On some Likelihood Inference for the Destructive COM-Poisson Cure Rate Model with Generalised Gamma Lifetime

Jacob Majakwara

Supervisor: Dr. Suvra Pal

School of Statistics and Actuarial Science

A Thesis submitted to the Faculty of Science in fulfillment of the requirements for the degree of

Doctor of Philosophy

June 2018

Declaration

I, Jacob Majakwara, declare that the contents of this thesis are original and have not been submitted for any other degree or qualification in any university. This thesis is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg.

.....
Signature of candidate

28th day of June 2018

Jacob Majakwara
June 2018

Abstract

In this thesis, based on a competing causes scenario, the destructive Conway-Maxwell Poisson (COM-Poisson) cure rate model is studied. The model assumes the occurrence of the event under study to undergo a destructive process of the initial competing causes and only the undamaged portion relating to the initial number of competing causes is recorded. This provides a real and practical way of interpreting the occurrence of the event under study in a biological system. This research assumes the distribution of competing causes to be COM-Poisson which includes some of the commonly used discrete distributions as its particular cases. Furthermore, we propose to model the lifetime by the generalised gamma distribution, which includes some of the commonly used lifetime distributions as its particular cases. The main contribution of this thesis is to develop the expectation maximisation (EM) algorithm to determine the maximum likelihood estimates (MLEs) of the model parameters and to carry out the likelihood inference assuming the data to be right censored. Model discrimination within the COM-Poisson and generalised gamma families are carried out to select a parsimonious distribution for the competing cause and the lifetime that jointly provides an adequate fit to the data. We develop the estimation procedure using both profile likelihood and complete likelihood approaches and make a comparison between the two techniques through the EM algorithm. The performance of the proposed method of inference is demonstrated by carrying out a comprehensive Monte Carlo simulation study. The flexibilities of the COM-Poisson and generalised gamma families are utilised to carry out a two-way model discrimination using the likelihood-and information-based methods. The proposed estimation technique is then applied to a real melanoma data for illustrative purpose.

The results show that both the COM-Poisson and generalised gamma distributions provide additional flexibility in modelling survival data with surviving fraction. We have also shown how covariates can influence the cure rate. Most importantly, the model fits the data better when the destructive mechanism is taken into account.

KEYWORDS: COM-Poisson distribution; Competing cause scenario; EM algorithm; Maximum likelihood estimates (MLEs); Profile likelihood; complete likelihood; Model discrimination; Long-term survivors; Generalised gamma distribution; Melanoma data

... to all those oppressed by the greediness of others, remember “all life is an experiment. The more experiments you make, the better.” Ralph Waldo Emerson.
... to those entrusted with power of knowledge, remember “the ability to simplify means to eliminate the unnecessary so that the necessary may speak” Hans Hoffmann.

Acknowledgements

The journey to a doctoral degree requires discipline. Discipline is doing what you know needs to be done, even though you do not want to. My journey is dedicated to all those individuals who provided motivation and guidance.

I would like to express my deepest gratitude primarily to my Ph.D. supervisor, Dr. Suvra Pal, for introducing me to this area of destructive cure rate modelling. His enormous knowledge, constant guidance, inspiration, influence, infinite patience and support have made my journey much safe and easy. This helped me a lot to navigate through this rigorous process for the last four years. My colleagues in the school of Statistics deserve many thanks as they encouraged me to complete this journey. Lastly, I would want to extend my sincere gratitude to my family members for supporting me during this period.

Outputs from the thesis

Papers submitted to journals

- Pal, S., Majakwara, J., & Balakrishnan, N. (2018). An EM algorithm for the destructive COM-Poisson regression cure rate model. *Metrika*, Vol. 81, Issue 2, pp. 143-171.
- Majakwara, J. & Pal, S. On some inferential issues for the destructive COM-Poisson-generalized gamma regression cure rate model (Under peer review).

Conference presentations

- Destructive COM-Poisson Cure Rate Model and Associated Likelihood Inference under Lognormal Lifetime. SASA 58th Annual Conference. 28 November - 1 December 2016, Capetown, South Africa.
- Likelihood Inference on destructive COM-Poisson Gamma Cure Rate Model. SASA 59th Annual Conference. 27 - 30 November 2017, Bloemfontein, South Africa.

The first journal paper is based on materials covered in Chapter 2 and the second paper is based on material covered in Chapter 5. The SASA conference presentations were made based on materials covered in Chapter 3 and 4.

Table of Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Introduction

Models for survival data with a surviving fraction play an important role in survival and reliability analysis as well as in areas such as economics, financial studies, demography and criminology. In standard survival analysis, generally it is anticipated that all subjects should experience the event of interest if complete follow up is possible, however, there are situations where some subjects may not experience the event at the end of the follow up period. In some instances, some subjects may not even experience the event under study even if the follow up period is extended. This leads to a proportion of subjects getting cured permanently. Due to progress in cancer treatment, more research has focused on cure models, which are models for survival data consisting of a fraction cure. Those subjects in medical and epidemiological studies who do not experience the event of interest are termed immunes or long-term survivors. Heterogeneity between diseased patients who are immunes and those who are not can be investigated using cure rate models. Cure rate models are still being developed as pointed out by Othus et al. (2012). As per them, they state that though cure models have been actively developed for the past two decades, they have not yet been used in some clinical literature areas.

INTRODUCTION

Cure models as part of mixture models assume that the population of interest is made up of two groups, the cured and the non-cured, and they are quite appealing from both statistical and biological points of view. Interest lies in whether an event of interest can occur and if it can occur, when it will occur. These models can only be used if there is strong scientific evidence that the subjects come from two separate groups. For this, subjects need to be followed up for a long time period and the data should show heavy censoring towards the end of the follow up period (see Maller and Zhou, 1992). The survival function curve for such a population should level off to a non-zero proportion indicating the presence of cure. In these models, the aim is to estimate the cured proportion and the survival probability of the uncured group, as well as determining whether covariates have any effect on these quantities of interest. Estimating the cure rate is crucial because it acts as an important marker to measure the efficacy of a treatment or therapy. The book of Maller and Zhou (1996) provides an excellent introduction on analysis of survival data with immunes.

In cancer clinical trials, patients are generally heterogeneous leading to various responses from treatment. Some will respond favourably to the treatment and this group of patients is considered to be cured if they show no signs or symptoms of the disease for a sufficiently long period of time. For the other group of patients, the disease will persist leading to relapse. Using mixture models, it is possible to estimate the proportion of immunes and also to study the causes of failure for the susceptible group of patients. This cured fraction is an important and useful measure as it allows to observe the trend in the survival of patients suffering from a particular disease. Cure rate models have two advantages; "they enrich our ability to interpret survival analysis in terms of characteristics that have clear biomedical meaning; and they lead to more general regression models, thereby extending our ability to describe actual data" (Tsodikov et al., 2003, p. 1063).

This research will focus on a flexible type of cure rate model, named as the destructive cure rate model, using the expectation maximisation (EM) algorithm to determine the maximum likelihood estimates (MLEs) of the parameters of the proposed model. This is an extension of the work of Rodrigues et al. (2011) where the authors proposed the destructive weighted

Poisson cure rate model for describing the elimination of initiated cells after some initial treatment, which presents a realistic and practical interpretation of the biological mechanism of the occurrence of an event in a competing cause scenario. It assumes that the original number of competing causes produced by nature undergoes a destructive process and what is recorded is only from the undamaged portion of the initial number of competing causes (Pal and Balakrishnan, 2015).

This work will rely on the EM algorithm as an optimisation technique for maximum likelihood estimation. Optimisation of an objective function is commonly encountered in computational statistics and the objective function can be sums of squares, log-likelihood, a penalised log-likelihood or a log posterior function. EM algorithm is the most effective algorithm for maximisation when dealing with incomplete data because it simply transfers maximisation from a complex function to a simple surrogate function (Becker et al., 1997). The notion of missing or incomplete data lies at the heart of the EM algorithm. McLachlan and Krishnan (2007, p. 2) said that "the situations where the EM algorithm can be applied include not only evidently incomplete-data situations, where there are missing data, truncated distributions, or censored or grouped observations, but also a whole variety of situations where the incompleteness of the data is not all that natural or evident. These include statistical models such as random effects, mixtures, convolutions, log linear models, and latent class and latent variable structures". The incomplete data situation will be considered in this work and missing-ness comes from the set of censored observations as Cooner et al. (2007) pointed out that because of finite monitoring time, cure can never be 'observed'.

1.1.1 Background

Boag (1949) introduced the mixture-based approach to cure rate models and later Berkson and Gage (1952) modified it. It consist of a mixture of a certain proportion p_0 of the population being long-term survivors and $1-p_0$ being susceptibles such that $S_{\text{pop}}(y) = p_0 + (1-p_0)S_1(y)$, where $S_1(y)$ is the survival function for the susceptibles and

$S_{\text{pop}}(y)$ is the survival function of the entire population. This model consists of the incidence model (for p_0) and the latent model (for the event time). Mixture models offer natural models for unobserved population heterogeneity and, therefore, are considered as a special type of the frailty model. The frailty model deals with addition of random effects in survival models, therefore, some of the well known identifiability issues in frailty models (see Hougaard, 1995) will need to be considered when dealing with mixture cure model.

The other approach to cure rate models is due to Yakovlev and Tsodikov (1996) who pioneered the stochastic tumour models and then Chen et al. (1999) extended them to a hierarchical framework. The event of interest occurs due to competing causes, which is a sub-discipline of survival analysis where in addition to the survival time W corresponding to each cause, the number of causes $M \in \{0, 1, 2, ...\}$ are regarded as latent variables in the modelling. The cure fraction (surviving fraction) is what makes cure rate models different from regular survival models like Cox regression model, which assumes that all subjects should experience the event under study if follow up is extended.

In clinical trials, there can be a number of causes competing to give rise to the event under study. Generally, the number of competing causes is not observable (latent variable) and recently, different authors have used different distributions to model this latent number of causes. For instance, Cooner et al. (2007) extended the competing cause cure rate model, where the event of interest T was considered as the r^{th} order statistic of the $W_1, W_2, \ldots, W_M$, and the number of causes M was modelled using some well known discrete distributions. Borges et al. (2012) used a generalised power series distribution for modelling the number of competing causes. Rodrigues et al. (2009) modelled the number of competing causes by Conway-Maxwell Poisson (COM-Poisson), whereas Rodrigues et al. (2011) used compound weighted Poisson distribution.

Rodrigues et al. (2011) introduced the destructive cure rate model, which gives a more realistic approach to cure rate models as they assume that the initial number of competing causes undergo a process of destruction in a competing cause scenario. According to

Rodrigues et al. (2011), let the initial number of competing causes M be an unobservable weighted Poisson random variable and be related to the occurrence of an event of interest with probability mass function (pmf) $p_m = P[M = m]$ for $m = 0, 1, \ldots$. Using the notation used by Pal and Balakrishnan (2017b); for $M = m$, let X_j, $j = 1, 2, \ldots, M$, be independent Bernoulli random variables, independent of M, with probability of success p indicating the presence of the j^{th} competing cause. Let D represent the total number of competing causes, among the M initial competing causes, which are not destroyed after a specific treatment. Then,

$$D = \begin{cases} X_1 + X_2 + \cdots + X_M, & \text{if } M > 0 \\ 0, & \text{if } M = 0. \end{cases} \tag{1.1}$$

Destruction or damage implies $D \leq M$. According to Rodrigues et al. (2011), given $M = m$, the distribution of D is termed the damaged distribution. One of the main interest is to investigate the characteristics of the distribution of the random variables M and D.

For a particular $D = d$, let random variables W_i (for $i = 1, 2, \ldots, d$) denote the survival time due to the i^{th} active competing cause. We assume that the W_i's are independent, they must also be distributed independently of D, and that they share a common cumulative distribution function $F(y) = 1 - S(y)$, where $S(y)$ denotes a proper survival function. When dealing with competing causes, both the number of active competing causes (D) and their corresponding lifetimes (W_i) are unobservable, and for all the active competing causes only the minimum lifetime is observed. However, the survival data includes subjects who are long-term survivors, and for them to be accommodated, the lifetime is given as

$$Y = \min\{W_0, W_1, \ldots, W_D\}, \tag{1.2}$$

where $P(W_0 = \infty) = 1$, which leads to the cure rate p_0, a fraction of the population who are immunes (Pal and Balakrishnan, 2015).

1.2 Aims and objectives

This research focuses on the computational aspects of the destructive COM-Poisson cure rate model by employing the EM algorithm so as to obtain the MLEs of the model parameters and using several particular distributions for the lifetimes as well as nesting them under the wider generalised gamma distribution. Assuming right censored observations, the steps of the EM algorithm for determining the MLEs of the model parameters are developed in detail. The generalised gamma distribution is a family of distributions whose particular cases are the Weibull, gamma, lognormal and exponential distributions. Model discrimination is carried out within the wider class of generalised gamma distribution to obtain the best distribution for the lifetime that provides the best fit to the data. For the number of competing causes, the COM-Poisson distribution is used and its different special cases are explored in order to obtain a parsimonious competing cause distribution. The particular cases of COM-Poisson distribution are Poisson, geometric and Bernoulli distributions and are obtained when the dispersion parameter (of COM-Poisson) takes the value 1, 0 and infinity, respectively.

1.2.1 Aims of the research

To develop likelihood inference for the destructive COM-Poisson cure rate model by employing the EM algorithm to estimate the model parameters assuming several lifetime distributions which are nested under the wider class of generalised gamma distribution.

1.2.2 Objectives of the research

1. To estimate the cure rate of patients based on a destructive competing cause scenario using a generalised family of distributions for the lifetime and the number of competing causes.

2. To find a parsimonious lifetime distribution along with a suitable competing causes

distribution that jointly provides adequate goodness-of-fit.

3. To assess the performance of the proposed method of inference by carrying out a comprehensive Monte Carlo (MC) simulation study and illustrate its performance using real data.

4. To make a comparison between two different estimation techniques, profile likelihood and complete likelihood, in terms of precision and accuracy of MLEs.

5. Study if there are any other factors (covariates) that affect the cure rate and survival probability.

6. To study the cause of failure for the uncured group of patients.

1.2.3 Hypotheses and questions

1. To test for a suitable competing cause distribution within the COM-Poisson family that can address goodness-of-fit.

2. To test for a suitable lifetime distribution within the generalised gamma family that can address goodness-of-fit.

3. To test the performance of the proposed method of inference using an extensive simulation study through the calculation of accuracy and precision of estimates.

4. Through the use of real data, test the performance of the method of inference developed.

5. To test for the importance of the destructive mechanism and show that its inclusion improves the fit.

1.3 Literature review

The assumption of the equality of the variance and mean of Poisson model makes it too restrictive for the majority of data which tend to be over-dispersed. The extra variability can be accounted for by having a Poisson model with an additional parameter. Generalisation of the Poisson distribution gives rise to the COM-Poisson distribution . "The COM-Poisson distribution is a flexible distribution that generalises several classical distributions (namely, the Poisson, geometric and Bernoulli) via its dispersion parameter" (Sellers et al., 2012, p. 114).

Two component mixture cure rate models allow for the effect of covariates to be different on susceptibles and immunes. Variations of the mixture cure rate models include mixture model of Kuk and Chen (1992), which combines logistic regression with proportional hazards regression, semi-parametric transformation cure models by Lu and Ying (2004), and semi-parametric accelerated failure time cure model by Li and Taylor (2002), among others. Under a competing cause scenario and a piecewise constant assumption for the baseline hazard function, Larson and Dinse (1985) developed the Cox proportional hazard cure rate model, whilst Fang et al. (2005) studied the Cox proportional hazard cure rate model where the properties of the maximum likelihood (ML) estimators were discussed. Lu (2008) used a non-parametric approach for estimating the parameters of the model developed by Larson and Dinse (1985). Zeng et al. (2006) developed semi-parametric transformation models which have the proportional odds cure rate and the proportional hazard models as particular cases. A similar approach on Cox proportional hazard cure rate model based on EM likelihood inference was developed by Liu et al. (2006) for time-to-event data obtained through interval mapping. From a Bayesian perspective, Zhao et al. (2014) estimated the parameters of the Cox proportional hazard cure rate model.

It is difficult to distinguish a censored subject in the susceptible group from an immune one (Farewell, 1986). This presents a problem in trying to distinguish "models with high incidence of susceptibles and long tails of the latency from low incidence of susceptibles and

short tails of the latency distribution" (Li et al., 2001, p. 390). Because of this, identifiability of cure rate models is an issue which needs to be addressed when dealing with cure rate models in general.

Cure rate models proposed by Berkson and Gage (1952); Chen et al. (1999) and Yakovlev and Tsodikov (1996) were later extended by Rodrigues et al. (2009) using the weighted Poisson distribution. Chen et al. (1999) derived the expressions of the mixture cure rate models and the properties of the hazard function. They further discussed the properties of the hazard function and Bayesian inference which were not discussed by Yakovlev and Tsodikov (1996). The use of maximum likelihood estimation of parameters in the COM-Poisson cure rate model was discussed by Rodrigues et al. (2009). These authors managed to unify the commonly used cure rate models. Using the model proposed by Rodrigues et al. (2009), Balakrishnan and Pal (2012, 2013, 2015, 2016) implemented the EM algorithm for estimating the model parameters.

Cancho et al. (2012) extended the promotion cure rate model by generalising it to the COM-Poisson cure rate model as a way of unifying the theory of cure rate models. Cancho et al. (2012) used Markov Chain Monte Carlo techniques to develop the COM-Poisson cure rate model through Bayesian procedures and the Weibull distribution was used as the lifetime distribution.

The cure rate models discussed so far do not take into account the possible elimination or destruction of the initial number of competing causes. In this regard, the destructive weighted Poisson cure rate model was proposed by Rodrigues et al. (2011), which describes the destruction of initiated cells after a specific treatment. This enables, in a competitive scenario, a real and practical way of interpreting the biological mechanism of the occurrence of an event. Borges et al. (2012) extended the work of Rodrigues et al. (2011) by incorporating association between the initiated cells, using an extended generalised power series distribution with an inflation parameter. The inflation parameter captures the association between the initiated cells. Cancho et al. (2013) through a latent activation scheme developed the

destructive negative binomial cure rate model and the maximum likelihood approach was used for parameter estimation. Through the EM algorithm, Pal and Balakrishnan (2015; 2016; 2017a; 2017b) studied some particular cases of the destructive cure rate models.

Although the destructive cure rate model has recently received some attention, the authors have considered simple and convenient-to-use distributions to represent the original number of competing causes, which in turn results in simple distributions for the number of active causes after the destruction. This simplicity, however, does not facilitate goodness-of-fit using as a formal test of hypothesis. Similarly, to study different destructive cure rate models, authors have considered the commonly used lifetime distributions to model the time taken by each active cause to produce the event of interest. In this research, we study a more general destructive cure rate model by assuming the family of COM-Poisson distribution to model the initial number of competing causes (and then deriving the distribution of the number of active causes) and the generalised gamma distribution to model the time taken by each active cause to produce the event. This two-way generalisation, one corresponding to the number of active causes and the other corresponding to the lifetime of each active cause, allows us to carry out a two-way model selection to select a parsimonious distribution for the active causes together with a suitable distribution for the lifetime associated with each active cause that would jointly address goodness-of-fit. Furthermore, model discrimination for the destructive cure rate model has not been addressed although such model discrimination has been addressed by Balakrishnan and Pal (2015) for the usual cure rate model without the destructive process.

1.3.1 COM-Poisson cure rate model

The COM-Poisson distribution was first introduced by Conway and Maxwell (1962) and was revived by Shmueli et al. (2005) as a flexible discrete distribution that accounts for over-dispersion and under-dispersion. Its a two parameter distribution which generalises the Poisson distribution where the additional parameter accounts for both under-dispersion

and over-dispersion. This makes it more elegant and flexible to model count data. The COM-Poisson distribution has the probability mass function (pmf) given by

$$P(M=m)=\frac{\eta^m}{(m!)^\phi}\frac{1}{Z(\eta,\phi)}, \qquad m=0,1,2,\ldots, \tag{1.3}$$

where

$$Z(\eta,\phi)=\sum_{j=0}^{\infty}\frac{\eta^j}{(j!)^\phi}.$$

It is undefined when $\phi=0$ and $\eta\geq 1$ because $Z(\eta,\phi)$ does not converge. The dispersion parameter, ϕ, is such that $\phi<1$ represents over-dispersion and $\phi>1$ represents under-dispersion. The COM-Poisson is a generalisation of some commonly used discrete distributions.

- For $\phi=1$, $Z(\eta,\phi)=e^\eta$, and the distribution in (1.3) reduces to the equi-dispersed Poisson distribution with parameter η.
- If $\phi=0$ and $\eta<1$, $Z(\eta,\phi)=\sum_{j=0}^{\infty}\eta^j=\frac{1}{1-\eta}$, which is a geometric distribution with parameter $1-\eta$ and is over-dispersed.
- As $\phi\to\infty$, $Z(\eta,\phi)\to 1+\eta$ and the distribution is under-dispersed Bernoulli with $P(M=1)=\frac{\eta}{1+\eta}$.

Inference and model discrimination can be developed using COM-Poisson distribution's hierarchical structure because it nests the three most commonly used discrete distributions.

The moment generating function of COM-Poisson distribution is

$$M_M(s)=E(e^{Ms})=\frac{Z(\eta e^s,\phi)}{Z(\eta,\phi)}$$

and its probability generating function is

$$A_p(s)=E(s^M)=\sum_{m=0}^{\infty}p(m;\eta,\phi)s^m=\frac{Z(\eta s,\phi)}{Z(\eta,\phi)}, \quad 0<s<1.$$

The survival function and density function of the COM-Poisson cure rate model are respectively expressed as

$$S_{\text{pop}}(y) = A_p(S(y)) = \frac{Z(\eta S(y), \phi)}{Z(\eta, \phi)} \tag{1.4}$$

and

$$f_{\text{pop}}(y) = -\frac{d}{dy} S_{\text{pop}}(y) = \frac{S(y)}{Z(\eta, \phi) S(y)} \sum_{j=0}^{\infty} \frac{j(\eta S(y))^j}{(j!)^{\phi}}, \tag{1.5}$$

where $f(y)$ is the pmf of the lifetimes in (1.2). The cured proportion is expressed as $p_0 = \lim_{y\to\infty} S_{\text{pop}}(y) = \frac{1}{Z(\eta,\,\phi)}$. "This model encompasses the promotional time cure model when $\phi = 1$ (Chen et al., 1999; Yakovlev and Tsodikov, 1996) and the mixture cure model when $\phi \to \infty$ (Berkson and Gage, 1952; Boag, 1949)" (Rodrigues et al., 2009, p. 3608). This flexibility makes it quite appealing when dealing with survival data with immunes. In the promotion time cure model, the proportion of immunes is $p_0 = \frac{1}{Z(\eta,1)} = e^{-\eta}$ and is such that the cured fraction (p_0) increases with decreasing η (Cancho et al., 2012).

The COM-Poisson distribution is a member of the weighted Poisson distribution with $w(m;\phi) = (m!)^{1-\phi}$, being the weight function (Rodrigues et al., 2011). According to Kokonendji et al. (2008), the weighted Poisson distributions are widely used for modelling partially recorded data, which are mostly encountered in survival analysis.

1.3.2 Destructive COM-Poisson cure rate model

Poisson distribution cannot handle over-dispersed or under-dispersed count data. According to Kokonendji et al. (2008, p. 1287), they "arise due to one or more possible causes such as heterogeneity and aggregation for over-dispersion and repulsion for under-dispersion". Weighted Poisson distribution can handle both under-dispersion and over-dispersion in a unified way. A lot of studies have been done on weighted distributions; Fisher (1934) introduced them using method of ascertainment, Gupta and Kirmani (1990) as well as Patil (2002) made further studies on them. Let M be a standard Poisson random variable with pmf

$p(m;\eta) = P(M = m;\eta)$, where η is the canonical parameter. For $M = m$, let the probability of ascertaining it be $w(m,\phi)$. Then, the pmf of M is given by

$$p(m;\eta,\phi) = \frac{w(m,\phi)p(m;\eta)}{E_\eta[w(M,\phi)]} = P(M = m;\eta,\phi), \quad m = 0,1,2,\ldots, \tag{1.6}$$

where $E_\eta(\cdot)$ denotes the mean value depending on η with respect to the random variable M, $w(m;\phi)$ is a weight function such that $\phi \geq 0$ and $0 < E_\eta(w(M,\phi)) < \infty$. Note that the weighted Poisson distribution in (1.6) can have different formats of the weight function with the simplest one being $w(m;\phi) = m$, and is called the size-biased Poisson distribution of the random variable M (see Patil and Rao, 1978). This kind of distribution can only capture under-dispersion. If the weight function in (1.6) is $w(m;\phi) = (m!)^{1-\phi}$, the distribution becomes the COM-Poisson distribution.

The lifetime Y in (1.2) then has a destructive or compound weighted cure survival function (Rodrigues et al., 2011) of the form

$$S_{\text{pop}}(y) = P(Y \geq y) = \sum_{m=0}^{\infty} P[D = m](S(y))^m = A_D(S(y)), \tag{1.7}$$

where $S(y)$ is the survival function of the lifetimes in (1.2), $S_{\text{pop}}(y)$ is the population survival function and $A_D(\cdot)$ is the probability generating function (pgf) of the compound random variable D.

"In order to be more clear about the damaged distribution, what we mean is that there is a positive probability of the value m of M to be reduced to some value j of D or, after some destructive process, $m - j$ of the m initial risk factors remain active in a competitive scenario" (Rodrigues et al., 2011, p. 337). The limit, $\lim_{y\to\infty} S_{\text{pop}}(y;\eta,\phi,p) = p_0$ is the proportion of immunes (long-term survivors), which suggests that $S_{\text{pop}}(\cdot)$ is not a proper survival function.

According to Rodrigues et al. (2011), for an integer ϕ, the total damaged random variable D has the weight function given by

$$\begin{aligned} w_p(j;\eta,\phi,p) &= E_{\eta(1-p)}\left(\{(j+U)!\}^{1-\phi}\right) \\ &= \exp\{-\eta(1-p)\}(j!)^{1-\phi}\,{}_1F_\phi(j+1;j+1,\ldots,j+1;\eta(1-p)), \end{aligned}$$

where $\eta(1-p)$ is the parameter of a Poisson random variable $U = M - j$, and the generalised hyper-geometric function is defined by

$${}_uF_v(a_1,\ldots,a_u;b_1,\ldots,b_v;x) = \sum_{m=0}^{\infty} \frac{(a_1)_m(a_2)_m \times \cdots \times (a_u)_m}{(b_1)_m(b_2)_m \times \cdots \times (b_v)_m} \frac{x^m}{m!},$$

with $(a)_m = a(a+1) \times \cdots \times (a+m-1)$. The pdf of the total damaged variable and survival function of the destructive COM-Poisson model are, respectively, given by

$$P(D = j;\eta,\phi,p) = e^{-\eta p}\,\frac{{}_1F_\phi(1;j+1,\ldots,j+1;\eta(1-p))}{{}_1F_\phi(1;1,\ldots,1;\eta)}, \quad j = 0,1,2\ldots. \tag{1.8}$$

$$S_{\text{pop}}(y;\eta,\phi,p) = \exp\{-\eta pF(y)\}\,\frac{{}_1F_\phi(1;1,\ldots,1;\eta\{1-pF(y)\})}{{}_1F_\phi(1;1,\ldots,1;\eta)}. \tag{1.9}$$

The cure rate is then obtained as $p_0 = \lim_{y\to\infty} S_{\text{pop}}(y) = \frac{Z(\eta(1-p),\,\phi)}{Z(\eta,\,\phi)}$. Table 1.1 gives the expressions for $S_{\text{pop}}(y)$, $f_{\text{pop}}(y)$ and p_0 for the destructive COM-Poisson model and some of its particular cases.

Table 1.1 Expressions for $S_{\text{pop}}(y)$, $f_{\text{pop}}(y)$ and p_0 for the destructive COM-Poisson model and some of its particular cases

Model	$S_{\text{pop}}(y)$	$f_{\text{pop}}(y)$	p_0
Bernoulli	$\frac{1+\eta(1-pF(y))}{1+\eta}$	$\frac{\eta p f(y)}{1+\eta}$	$\frac{1+\eta(1-p)}{1+\eta}$
Poisson	$e^{-\eta pF(y)}$	$\eta p f(y) S_{\text{pop}}(y)$	$e^{-\eta p}$
Geometric	$\frac{1-\eta}{1-\eta(1-pF(y))}$	$\frac{\eta p f(y) S_{\text{pop}}(y)}{1-\eta(1-pF(y))}$	$\frac{1-\eta}{1-\eta(1-p)}$
COM-Poisson	$\frac{Z(\eta(1-pF(y)),\phi)}{Z(\eta,\phi)}$	$\frac{pf(y)}{Z(\eta,\phi)(1-pF(y))}\sum_{j=1}^{\infty}\frac{j\{\eta(1-pF(y))\}^j}{(j!)^\phi}$	$\frac{Z(\eta(1-p),\phi)}{Z(\eta,\phi)}$

1.3.3 Lifetime distributions

Numerous parametric models are used for analysing lifetime data but a few are useful in a wide range of situations. These are the Weibull, gamma and log-normal distributions. The generalised gamma distribution is also very useful since its a family of distributions and includes the aforementioned distributions as particular cases.

Exponential distribution

The exponential distribution is the simplest of all lifetime distributions. It is a special case of both Weibull and gamma distributions. Its pdf is expressed as

$$f(y;\gamma) = \gamma \exp(-\gamma y), \qquad y > 0,\ \gamma > 0. \tag{1.10}$$

The mean is given by $E(Y) = \frac{1}{\gamma}$ and the variance is given by Var(Y) $= \frac{1}{\gamma^2}$. The survival function takes the following form

$$S(y;\gamma) = e^{-\gamma y}, \quad y > 0.$$

A constant hazard function characterizes the exponential distribution and thus limits its application to practical situations.

Weibull distribution

The Weibull distribution plays an important role (because of its flexibility) in lifetime data analysis, reliability studies and life testing. The pdf of a Weibull distribution is expressed as

$$f(y;\gamma_t) = \frac{1}{\gamma_1 y}(\gamma_2 y)^{\frac{1}{\gamma_1}} \exp(-(\gamma_2 y)^{\frac{1}{\gamma_1}}), \quad y > 0,\ \gamma_1 > 0,\ \gamma_2 > 0 \tag{1.11}$$

where $\gamma_t = (\gamma_1, \gamma_2)'$ with shape parameter $\frac{1}{\gamma_1}$ and scale parameter $\frac{1}{\gamma_2}$. The mean and variance are, respectively, given as

$$E(Y) = \frac{1}{\gamma_2}\Gamma(1+\gamma_1) \qquad \text{and} \qquad \text{Var}(Y) = \frac{1}{\gamma_2^2}[\Gamma(1+2\gamma_1) - \{\Gamma(1+\gamma_1)\}^2],$$

where

$$\Gamma(a) = \int_0^\infty e^{-x}x^{a-1}dx$$

is the gamma function. The survival function is expressed by

$$S(y;\gamma_t) = \exp(-(\gamma_2 y)^{\frac{1}{\gamma_1}}), \qquad y > 0. \tag{1.12}$$

The hazard function of the Weibull distribution can be constant, increasing or decreasing, showing the flexibility of the distribution and thus making it the most commonly used distribution in survival analysis.

Gamma distribution

Gamma distribution is also useful to model lifetime data. Its pdf is expressed as

$$f(y;\gamma_t) = \frac{(\frac{\gamma_2}{\gamma_1^2})^{\frac{1}{\gamma_1^2}}}{\Gamma(\frac{1}{\gamma_1^2})}\exp\left(-\frac{\gamma_2 y}{\gamma_1^2}\right)y^{\frac{1}{\gamma_1^2}-1}, \qquad y > 0,\ \gamma_1 > 0,\ \gamma_2 > 0 \tag{1.13}$$

where $\gamma_t = (\gamma_1, \gamma_2)'$ with shape parameter $\frac{1}{\gamma_1^2}$ and scale parameter $\frac{\gamma_1^2}{\gamma_2}$, where $\Gamma(\cdot)$ is the gamma function defined before. The mean and variance of the gamma distribution are respectively expressed as

$$E(Y) = \frac{1}{\gamma_2} \qquad \text{and} \qquad \text{Var}(Y) = \frac{\gamma_1^2}{\gamma_2^2}.$$

The survival function is

$$S(y;\gamma_t) = \frac{\Gamma(\frac{1}{\gamma_1^2}, \frac{\gamma_2 y}{\gamma_1^2})}{\Gamma(\frac{1}{\gamma_1^2})}, \qquad y > 0, \tag{1.14}$$

where

$$\Gamma(a,b) = \int_b^\infty e^{-x} x^{a-1} dx$$

is the upper incomplete gamma function.

Lognormal distribution

The lognormal distribution is another important distribution used in lifetime modelling and has been found to be a serious competitor to the Weibull distribution. Its pdf is written as

$$f(y;\gamma_t) = \frac{1}{\sqrt{2\pi}\gamma_1 y}\exp\left(-\frac{1}{2}\left(\frac{\log(\gamma_2 y)}{\gamma_1}\right)^2\right), \quad y>0,\ \gamma_1>0,\ \gamma_2>0 \tag{1.15}$$

where $\gamma_t = (\gamma_1, \gamma_2)'$ with mean $\exp\left(\frac{\gamma_1^2}{2} - \log(\gamma_2)\right)$ and variance $\exp(\gamma_1^2 - 2\log(\gamma_2))(\exp(\gamma_1^2) - 1)$. The survival function of the log-normal distribution is

$$S(y;\gamma_t) = 1 - \Phi\left(\frac{\log(\gamma_2 y)}{\gamma_1}\right), \quad y>0, \tag{1.16}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal random variable.

Generalised gamma distribution

This distribution introduced by Stacy (1962) is a flexible distribution and it includes all the aforementioned distributions as its particular cases. Its pdf is is expressed as

$$f(y;\gamma_t) = \begin{cases} q(q^{-2})^{q^{-2}}(\lambda y)^{q^{-2}(q/\sigma)}\exp\left(-q^{-2}(\lambda y)^{q/\sigma}\right)/\left(\Gamma(q^{-2})\sigma y\right), & q>0 \\ \left(\sqrt{2\pi}\sigma y\right)^{-1}\exp(-(\log(\lambda y))^2/(2\sigma^2)), & q=0, \end{cases} \tag{1.17}$$

where $\gamma_t = (q,\ \sigma,\ \lambda)'$ with $q \geq 0$ and $\sigma > 0$ being the shape parameters and $\lambda > 0$ being the scale parameter.

The generalised gamma family has special cases distributions such as exponential, Weibull, gamma and lognormal and because of its flexibility, the shape of the uncured patients' survival distribution is easily captured (Yu et al., 2004). The particular cases of generalised gamma distribution can be obtained by changing the shape parameters as follows:

- $q = 1$, it reduces to the Weibull distribution;
- $q/\sigma = 1$, it becomes a gamma distribution;
- $q = 0$ (which is also the limit of (1.17) as $q \to 0$), it reduces to the lognormal distribution; and
- $q = \sigma = 1$, it becomes an exponential distribution.

This flexibility of the generalised gamma as a lifetime distribution enables model discrimination to be carried out within its family in order to select a particular lifetime distribution that has the potential to provide the best fit to a given data set.

The mean and variance, respectively, are expressed as

$$E(Y) = \frac{(q^2)^{\sigma/q}}{\lambda\Gamma(1/q^2)}\Gamma\left(\frac{\sigma}{q}+\frac{1}{q^2}\right); \quad \text{Var}(Y) = \frac{q^{4\sigma/q}}{\lambda^2}\frac{1}{\Gamma(1/q^2)}\left[\Gamma\left(\frac{2\sigma}{q}+\frac{1}{q^2}\right) - \frac{\Gamma^2(\frac{\sigma}{q}+\frac{1}{q^2})}{\Gamma(1/q^2)}\right].$$

The generalised gamma distribution's survival function involves the incomplete gamma function and is given by

$$S(y;\gamma_t) = \begin{cases} \Gamma(q^{-2}, q^{-2}(\lambda y)^{q/\sigma})/\Gamma(q^{-2}), & q > 0 \\ 1 - \Phi\left(\frac{\log(\lambda y)}{\sigma}\right), & q = 0, \end{cases} \tag{1.18}$$

where

$$\Gamma(a,b) = \int_b^\infty e^x x^{a-1} dx$$

is the upper incomplete gamma function.

1.4 Data and the likelihood function

In reality, lifetimes of all subjects may not be completely observed due to censoring. In this research, we will only consider right censoring which occurs when for some subjects exact lifetimes are known and for the rest the lifetimes are only known to exceed certain values.

Denoting the censoring time as C_i and actual lifetime as defined in (1.2), the observed individual lifetime is defined as $T_i = \min(Y_i, C_i)$, for $i = 1, 2, \ldots, n$. Let the right censoring indicator be $\delta_i = I\,(Y_i \leq C_i)$, where $\delta_i = 1$ if the lifetime is observed and $\delta_i = 0$ if the lifetime is right censored. The censoring is assumed to be non-informative, that is, Y_i and C_i are independent. In the parametric destructive COM-Poisson cure rate model, link functions are utilised to connect covariates with the fraction cure. The parameters to be used are η and p. In this research, we will link the parameters η and p to covariates $\mathbf{x}_1$ and $\mathbf{x}_2$ by the log-linear link function $\eta_i = \exp(\mathbf{x}_{1i}^{'}\beta_1)$ and the logistic link function $p_i = \frac{\exp(\mathbf{x}_{2i}'\beta_2)}{1+\exp(\mathbf{x}_{2i}'\beta_2)}$, respectively. In these link functions, β_1 and β_2 represent the vector of regression coefficients. Thus, we deal with a more practical situation wherein we introduce the covariate effect and this would allow us to study in detail how the cure rate is affected by the covariates.

Let us introduce a cured status indicator variable J, which is defined as follows

$$J = \begin{cases} 1, & \text{if a subject is susceptible} \\ 0, & \text{if a subject is cured.} \end{cases}$$

It follows that $P(J = 0) = p_0$ (which represents the fraction cure) and $P(J = 1) = 1 - p_0$, which represents the proportion of susceptibles. Let $F_{\text{pop}}(y)$ be the cumulative distribution function (cdf) of the overall population and $F_1(y)$ be the same for the susceptibles only. Then, for finite $y > 0$, we have $P(Y \leq y|J = 0) = 0$ and $P(Y \leq y|J = 1) = F_1(y)$. Note that $F_1(y)$ is a proper cdf but the cdf for the overall population is not. $F_{\text{pop}}(y)$ is given by (see Balakrishnan and Pal, 2012)

$$\begin{aligned} F_{\text{pop}}(y) &= P(Y \leq y) \\ &= P(Y \leq y|J=0)P(J=0) + P(Y \leq y|J=1)P(J=1) \\ &= (1-p_0)F_1(y). \end{aligned}$$

Thus, the overall population survival function is written as

$$S_{\text{pop}}(y) = 1 - F_{\text{pop}}(y) = p_0 + (1-p_0)S_1(y),$$

where $S_1(\cdot)$ represents the proper survival function of susceptibles. Note that $S_{\text{pop}}(y)$ is not a proper survival function. It should also be noted that if subjects belong to the set $I_0 = \{i : \delta_i = 0\}$, J is unknown, whereas J takes the value 1 if subjects belong to the set $I_1 = \{i : \delta_i = 1\}$. "This actaully introduces the missing data and motivates us to employ the EM algorithm for the determination of the MLEs of the model parameters" (Balakrishnan and Pal, 2015, p. 157).

Assuming non-informative censoring and for a given sample of size n, the observed data likelihood function is expressed as

$$\begin{aligned} L(\theta; \mathbf{t}, \mathbf{x}_1, \mathbf{x}_2, \delta) &\propto \prod_{i=1}^{n} \{f_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\}^{\delta_i} \{S_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\}^{1-\delta_i} \\ &= \prod_{I_1} f_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta) \prod_{I_0} S_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta), \end{aligned}$$

where $\theta = (\phi, \beta_1', \beta_2', \gamma')'$, $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$, and $\mathbf{x}_1$ and $\mathbf{x}_2$ denote the vectors of covariates $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$, respectively.

1.5 Likelihood-based inference

Inferential statistics is mainly concerned with estimating parameters of a model and for testing model adequacy. One of the most commonly used principles is the likelihood principle, where the likelihood function is used to convey information about unknown quantities. It is the likelihood function that captures all the information in the data about a parameter and as

pointed out by Yudi (2001, p. 30), "the likelihood is a tool for an objective reasoning with data, especially for dealing with the uncertainty due to the limited amount of information contained in the data". The MLE will therefore provide a point estimate of interest that maximises the likelihood function. The likelihood is the joint distribution of the data given the parameter values. The maximum likelihood estimation provides properties such as consistency, minimal sufficiency, efficiency, asymptotic unbiasedness and asymptotic normality for its estimates.

Likelihood based techniques are mainly computational and require numerical methods. Since lifetime data are incomplete, the EM algorithm appears to be a natural choice in such problems. The EM algorithm is one of the most versatile algorithms in modern statistics and is used for maximum likelihood (or maximum a posteriori) estimation when some latent variables are involved. When dealing with the family of mixture distributions, they should be identifiable in order to have a meaningful estimation procedure.

1.6 EM algorithm

The EM algorithm is a useful method of estimation of parameters when there is a many-to-one mapping from the underlying distribution to the distribution governing the observation. "Thus, in the EM algorithm, one augments the observed data with latent data such that one complicated maximisation is replaced by an iterative series of simple maximisations" (Tanner, 1996, p. 3). The likelihood is easy to optimise if it is expressed in terms of latent data. "The fitting of mixture models by maximum likelihood is a classic example of a problem that is simplified considerably by the EM's conceptual unification of maximum likelihood estimation from data that can be viewed as being incomplete" (McLachlan and Krishnan, 2007, p. 13). The standard errors are calculated using the method of Louis (1982), which involves computing the observed information matrix within the EM frame work. In the iterations of the EM algorithm, computation of the gradient (score) of complete data as well as the second order derivatives matrix are required. Inverting the second order derivatives will give the observed information matrix.

EM algorithm is applied as an iterative computation of MLEs where Newton-Raphson is difficult to use. The EM algorithm has the advantage that its iteration always increases the likelihood, though this does not guarantee convergence. EM algorithm usually handles parameter constraints automatically, that is, each maximisation step produces a MLE-type estimate. Each iteration of EM algorithm has two steps; these are the expectation step (E-step) and maximisation step (M-step). "The E-step consists in manufacturing data for the complete-data problem, using the observed data set of the incomplete-data problem and the current value of the parameters, so that the simpler M-step computation can be applied to this 'complete' data set" (McLachlan and Krishnan, 2007, p. 2). The maximisation step is achieved by estimating the parameters to this 'complete' data set.

Let $(t_i, \delta_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}, J_i)$, $i = 1, 2, \cdots, n$, denote the complete data which includes the observed data and the unobserved J_i's. The complete data likelihood function is then given by (see Balakrishnan and Pal, 2013)

$$\begin{aligned} L_c(\theta; \mathbf{t}, \mathbf{x}_1, \mathbf{x}_2, \delta, \mathbf{J}) \quad \propto \quad & \prod_{I_1} \{f_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\}^{J_i} \times \\ & \prod_{I_0} \{p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i})\}^{1-J_i} \{(1 - p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i})) S_1(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\}^{J_i}, \end{aligned}$$

where $\mathbf{J}$ denote the vector of J_i values.

Then, the log-likelihood function (excluding the constant) is expressed as

$$\begin{aligned} l_c(\theta; \mathbf{t}, \mathbf{x}_1, \mathbf{x}_2, \delta, \mathbf{J}) \quad = \quad & \sum_{I_1} J_i \log\{f_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\} + \sum_{I_0} (1 - J_i) \log\{(p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i})\} \\ & + \sum_{I_0} J_i \log\{(1 - p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i})) S_1(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\}. \end{aligned} \tag{1.19}$$

E-Step:

In this step, we calculate the expectation of $l_c(\theta; \mathbf{t}, \mathbf{x}_1, \mathbf{x}_2, \delta, \mathbf{J})$ with respect to the distribution of the unobserved J_i's, given the current values of the parameter and the observed data $\mathbf{O}$, where $\mathbf{O} = \{\text{observed } J_i's, (t_i, \delta_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}); i = 1, 2, \ldots, n\}$. The complete data log-likelihood in (1.19) is linear in Bernoulli random variables J_i, and therefore we need to calculate

$\pi_i^{(k)} = E(J_i|\theta^{(k)}, \mathbf{O}_i)$, $i = 1, 2, \ldots, n$, where $\theta^{(k)}$ stands for the current parameter value at the k^{th} iteration step. For the i^{th} censored observation, we have (see Balakrishnan and Pal, 2016)

$$\begin{aligned}
\pi_i^{(k)} &= E(J_i|\theta^{(k)}, O_i) \\
&= P(J_i = 1|T_i > t_i; \theta^{(k)}) \\
&= \frac{P(T_i > t_i|J_i = 1)P(J_i = 1)}{P(T_i > t_i)}\Big|_{\theta=\theta^{(k)}} \\
&= \frac{(1 - p_{0i})S_1(t_i)}{S_{\text{pop}}(t_i)}\Big|_{\theta=\theta^{(k)}}.
\end{aligned} \tag{1.20}$$

For $i \in I_1$, we simply have $\pi_i^{(k)} = J_i = 1$. In the expectation step, the J_is in (1.19) are replaced by $\pi_i^{(k)}$ if $i \in I_0$ and by 1 if $i \in I_1$. Let $Q(\theta, \pi^{(k)})$ denote the conditional expectation of the complete data log-likelihood function, and is written as

$$\begin{aligned}
Q(\theta, \pi^{(k)}) \quad = \quad & \sum_{I_1} \log\{f_{\text{pop}}(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\} + \sum_{I_0} (1 - \pi_i^{(k)}) \log\{(p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i})\} \\
& + \sum_{I_0} \pi_i^{(k)} \log\{(1 - p_0(\theta, \mathbf{x}_{1i}, \mathbf{x}_{2i}))S_1(t_i, \mathbf{x}_{1i}, \mathbf{x}_{2i}; \theta)\},
\end{aligned}$$

where $\pi^{(k)}$ is the vector of $\pi_i^{(k)}$ values.

M-Step:

In the maximisation step, the function $Q(\theta, \pi^{(k)})$ is maximised with respect to θ over the corresponding parameter space Θ to obtain an improved estimate of θ as

$$\theta^{(k+1)} = \arg\max_{\theta \in \Theta} Q(\theta, \pi^{(k)}).$$

The E-step and M-step are then iteratively continued until convergence to obtain the MLE of the parameter θ. In our case, the maximisation step is carried out using the EM gradient algorithm (Lange, 1995) because the MLEs do not have explicit expressions. Explicit expressions of the E-step and M-step are presented in the Appendices A.1 and B.1 for several flexible destructive cure rate models with different lifetime distributions.

1.6.1 Standard errors

Generally, the observed information matrix is used to estimate the covariance matrix rather than the expected information matrix since it does not require an expectation to be computed. Therefore, an asymptotic covariance matrix of $\hat{\theta}$ is approximated by inverting the observed information matrix of θ. Using $l_c(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})$ from the EM algorithm, we obtain the first order and second order derivatives of the observed data log-likelihood function. Through the use of Louis' principle, the observed information matrix can be calculated as follows (see Louis, 1982).

$$\begin{aligned} I(\theta) &= E[B(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})] - E[S(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})S^T(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})] \\ &\quad + E[S(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})]E[S^T(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})], \end{aligned}$$

where $I(\theta)$ is the information on θ, $B(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})$ and $S(\theta;\mathbf{t},\mathbf{x}_1,\mathbf{x}_2,\delta,\mathbf{J})$ denote the negative of the matrix of second derivatives and the gradient vector of (1.19), respectively. The standard errors of MLEs are then calculated by taking the square roots of the diagonal elements of $I^{-1}(\theta)$. The asymptotic (large sample) distribution of the MLE $(\hat{\theta})$ is multivariate normal with mean θ and variance $I^{-1}(\theta)$, and therefore we can obtain the asymptotic confidence intervals of the parameters. This allows us to study the coverage probabilities of the parameters through a simulation study.

1.7 Simulation study

There are times when Monte Carlo (MC) simulation study is needed in statistical applications, especially when there is no closed form solution to a statistical problem or to compare results of an inference under different assumptions. According to (Mooney, 1997, p. 20), MC simulations are useful to: "make inferences when weak statistical theory exists for an estimator; test null hypotheses under a variety of plausible conditions; assess the quality of an inference method; assess the robustness of parametric inference to assumption violations; and compare estimator's properties".

Monte Carlo simulation is utilised in this research for the purpose of assessing the quality of an inference method used and also its robustness.

1.7.1 Model fitting

We evaluate the performance of the proposed method of inference through an extensive MC simulation study. Varying samples sizes are considered in the simulation study in order to assess the behaviour of the model under small and relatively large samples.

In this research, data are simulated from several destructive cure rate models and the method of estimation discussed in Section 1.6 is used to estimate the parameters of the model. There is need to assess precision and accuracy of estimates in any simulation study. Precision specifies how finely an estimate of a parameter is while accuracy specifies how likely the estimate is to be the true value of the parameter. Precision measures to be used are standard error, root mean square error (RMSE) and power of the test while accuracy measures to be used are bias, RMSE and coverage probability. Note that mean square error (MSE) combines bias and variance, and therefore RMSE measures both accuracy and precision, and hence gives an overall sense of the quality of the estimator. It is also of interest to study the empirical bias and RMSE of the cure fraction under model mis-specification. R software is used for all simulations and the results are based on 1000 MC runs.

1.7.2 Model discrimination

"Specifying a correct model is important for a consistent inference and efficiency" (Yudi, 2001, p. 375). In order to achieve this, for each given data set, suitable methods should be utilised for comparing several competing models and then using some criterion the one that best describes the data is to be selected. Since the competing cause and lifetime distributions used in our modeling are flexible and includes other distributions as special cases, two approaches can be used to select the best model, the likelihood based method and the information-based

method.

Likelihood-based method

The likelihood approach normally leads to the best possible inference. Since the likelihood approach requires complete specification of the probability structure, it becomes difficult to quantify the uncertainty associated with model selection. So, sensitivity analysis will help to gauge the correctness of the assumed model. The models for the COM-Poisson family of distributions are nested and therefore the likelihood ratio test (LRT) can be utilised. The LRT's performance is studied by testing that the null hypothesis follows one of the particular cases of COM-Poisson distribution versus that its a member of the COM-Poisson family besides the one stated in the null hypothesis. This is true for a fixed lifetime distribution, and the same can be carried out for a fixed competing cause distribution where we test for the suitability of a lifetime distribution within the family of generalised gamma.

The LRT is closely related to the maximum likelihood estimation and it is defined as $\Lambda = -2(\hat{l}_0 - \hat{l}_1)$, where $\hat{l}_1$ and $\hat{l}_0$ denote values of the unrestricted and restricted maximized log-likelihood function, respectively. Under the standard likelihood theory, the asymptotic null hypothesis of Λ is a central chi-square distribution with one degree of freedom. If the parameter lies on the boundary of parameter space (the likelihood will not be regular), the asymptotic null hypothesis of Λ is a 50-50 mixture of a point mass at zero and a chi-square distribution with one degree of freedom, that is $P(\Lambda \leq \lambda) = \frac{1}{2} + \frac{1}{2}P(\chi_1^2 \leq \lambda)$ (see Self and Liang, 1987)

The LRT statistic of the fitted particular cases of destructive COM-Poisson cure rate model against the fitted destructive COM-Poisson cure rate model is calculated for each simulated data set. The observed significance levels and powers of the LRT are then computed based on 1000 data sets for each situation and at a 5% nominal level of significance.

Information-based methods

The Kullback-Leibler distance measures how well a good model fits the data. The Kullback-Leibler distance is $D(g,f) = E_g\left[\log\frac{g(X)}{f(X)}\right]$, where $f(\cdot)$ is an assumed model density that should fit the data and $g(\cdot)$ is the pdf of the true model. Working with several models necessitate choosing the one that is closest to the true distribution, that is, minimising the Kullback-Leibler distance to the true distribution. Minimising the Kullback-Leibler distance is the same as maximising the likelihood. We cannot compare the maximised likelihood for competing models if the number of parameters vary between the models because the one with more parameters will yield larger log-likelihood value. To correct the bias analytically the Akaike information criterion (AIC) is used. Minimising the AIC is the same as minimising the Kullback-Leibler distance. The AIC is used to compare different models $f_k(\mathbf{x},\theta_k)$ given the data $\mathbf{x}$. The models need not necessarily be nested, thus making it a general purpose model selection tool. AIC $= -2l + 2p$, where p is the number of parameters and l is the maximised log-likelihood value of the fitted model. A model that minimises the AIC is the best fitted model. "We can interpret the first term in the AIC as a measure of data fit and the second as a penalty" (Yudi, 2001, p. 375). This makes the comparison of several models on equal footing though not necessarily containing the true model. The possible models are easily ranked from best to worst using the AIC given the empirical data. Similarly, we can use the Bayesian information criterion (BIC) defined as BIC $= -2l + p\log(n)$, where n is the sample size. The model with the minimum value of BIC is the best one. The BIC has an added advantage over the AIC because it penalises bigger samples more.

Based on 1000 datasets in each situation and for different competing models, the AIC and BIC values are computed and the observed selection rates are recorded.

1.8 Real melanoma data

To illustrate the method of parameter estimation developed here, we applied our method to a real malignant melanoma data. Melanoma survival data can be found in the 'timereg' package (Scheike and Zhang, 2011) of R software (R Development Core Team, 2010). These data correspond to patients with a type of malignant cancer evaluated for postoperative treatment performance. The data set includes 205 patients observed after the removal of malignant melanoma operation for the period $1962 - 1977$ at Odense university hospital. The minimum and maximum time in years are 0.0274 and 15.25, respectively, with mean and standard deviation of 5.9 and 3.1 years, respectively. 72% of the observations were censored (showing heavy censoring) and therefore care has to be taken to prevent identifiability issue. Patients were followed until 1977 and times recorded refer to the time until a patient died or the censoring time, whichever occurred first. There were 14 patients who died from other causes and were treated as censored observations. Two covariates of interest are ulceration status (whether present or absent) and tumour thickness measured in millimeters.

Chapter 2

Destructive COM-Poisson model with Weibull lifetime

2.1 Model identifiability

Estimation of model parameters for cure rate models generally depends critically on model identifiability, that is, "no two sets of parameters may generate the same model output. Thus, identifiability is a structural property of the model, ..., for if a model is non-identifiable then estimation of its parameters from a sample of model outputs is impossible and, if attempted, would lead to instability of the estimates" (Hanin and Huang, 2014, p. 261). Note that the destructive COM-Poisson model suffers from the identifiability issue. The model can be made identifiable by modeling p as "a logistic regression with non-constant covariates (Li et al., 2001). So, it is necessary to include some covariates in the cure fraction to ensure identifiability" (Ortega et al., 2012, p. 708). To eliminate this problem of non-identifiability, we relate the parameter η to covariates $\mathbf{x}_1$ by the log-linear link function $\eta = \exp(\mathbf{x}_1'\beta_1)$ and the parameter p to covariates $\mathbf{x}_2$ by the logistic link function $p = \frac{\exp(\mathbf{x}_2'\beta_2)}{1+\exp(\mathbf{x}_2'\beta_2)}$ such that the covariates $\mathbf{x}_1$ and $\mathbf{x}_2$ do not share common elements and either β_1 or β_2 do not include an intercept. Note that this approach is along the lines of Rodrigues et al. (2011).

Let γ denote the parameter vector corresponding to the distribution of the lifetime in (1.2). For n subjects, the observed data likelihood function is expressed as

$$L(\theta) \propto \prod_{i=1}^{n} \left\{ f_{\text{pop}}(t_i) \right\}^{\delta_i} \left\{ S_{\text{pop}}(t_i) \right\}^{1-\delta_i}, \tag{2.1}$$

where $\theta = (\phi, \beta_1', \beta_2', \gamma')'$ is the vector of model parameters to be estimated. A fully parametric approach is considered in this thesis and in this chapter we assume the lifetime W in (1.2) to follow a two parameter Weibull distribution with

$$S(t) = e^{-(\gamma_2 t)^{\frac{1}{\gamma_1}}} \quad \text{and} \quad f(t) = \frac{(\gamma_2 t)^{\frac{1}{\gamma_1}}}{\gamma_1 t} S(t), \tag{2.2}$$

for $t > 0$, $\gamma_1 > 0$ and $\gamma_2 > 0$. Note that the choice of link function for η does not guarantee η to be less than unity and as such it violates the condition required for the destructive geometric cure rate model ($\phi = 0$). Therefore, the destructive geometric cure rate model will not be discussed in this research.

2.2 Parameter estimation

The expressions of the first- and second-order derivatives of $Q(\theta, \pi^{(k)})$ with respect to θ together with explicit expressions of the $Q(\theta, \pi^{(k)})$ function are presented in Appendix A.1 for the destructive COM-Poisson model and some of its particular cases. The E- and M-steps of the EM algorithm are repeated iteratively until a stopping criterion is satisfied to obtain the MLE of the parameter θ. Note that the M-step involves the computation of the infinite series $Z(\eta, \phi)$ and other related infinite series, as presented in Appendix A.1. The stopping criterion to be used is $|(\theta^{(k)} - \theta^{(k-1)})/\theta^{(k-1)}| < \tau$ where τ is a small tolerance value (value of $\tau = 0.001$ has been used in this thesis). The infinite series $Z(\eta, \phi)$ can be computed by truncating it as follows. Since the ratio of successive terms $j-1$ and j is η/j^{ϕ}, for $j \geq 1$, by choosing a small fixed ε, we can take the smallest k such that $k > (\eta)^{1/\phi}$ and

$Z(\eta,\phi) \approx \sum_{j=0}^{k} \eta/(j!)^{\phi}$. The other infinite series can be truncated in a similar way. For computational efficiency, we have chosen a small value for ε as 10^{-10} and observed that the number of terms k required to truncate a series increases as the COM-Poisson shape parameter ϕ gets close to zero.

2.2.1 Approaches for estimating parameters

Model parameters can be estimated either by utilising the profile likelihood or complete likelihood approaches within the EM algorithm framework.

Profile likelihood approach

In this approach, the ϕ parameter of the COM-Poisson distribution is profiled out by fixing values of ϕ which are admissible and distinct. For a fixed value of ϕ, the rest of the parameters of the model are estimated using the EM algorithm and the log-likelihood function is calculated. The shape parameter ϕ is then estimated by taking the value of ϕ for which the log-likelihood function is the maximum. Noting the problem when $\phi = 0$, as discussed in Section 2.1, the grid for ϕ to apply the profile likelihood was selected as $\{0.2, 0.3, \ldots, 2.3\}$.

Complete likelihood

Unlike the profile likelihood approach where ϕ is kept fixed, in the complete likelihood approach all the model parameters are simultaneously estimated using the EM algorithm by maximising with respect to all the parameters of the model. The advantage of the complete likelihood approach is that there is no need to fix ϕ, instead, ϕ is introduced in the estimation procedure and all the parameters are simultaneously estimated from the complete log-likelihood function. This makes the estimation procedure more efficient and avoid errors

incurred in standard error calculation when employing the profile likelihood approach where the uncertainty of ϕ is not taken into account (because ϕ is kept fixed). Note that this approach may fail to converge in some instances because of the flatness of the log-likelihood function with respect to ϕ and in some instances may lead to high standard error of the shape parameter.

2.2.2 Standard errors

An approximate variance-covariance matrix of the MLE is obtained by inverting the observed information matrix of ϕ, β_1, β_2 and γ. The first and negative second-order derivatives of the observed log-likelihood function with respective to ϕ, β_1, β_2 and γ gives the components of the score function and observed information matrix, respectively. Explicit expressions of the components of the observed information matrix for different destructive cure rate models are presented in Appendix A.2.

2.3 Simulation study

The performance of the proposed estimation method is evaluated through an extensive MC simulation study. We considered varying sample sizes of $n = 200$ and $n = 500$ to demonstrate the small and large sample behaviours. To generate data, we mimicked the real melanoma data analysed later in Section 2.5. As seen in the melanoma data, ulceration status (1: present and 0: absent) and tumour thickness (min $= 0.1$ mm, max $= 17.42$ mm) are the two covariates of interest with 44% of subjects having ulcer. To generate the covariate ulceration status, we generated a Uniform (0, 1) random variable. If the value of the variable was greater than 0.44, ulceration status was taken as 0, otherwise, it was taken as 1. The covariate tumour thickness was generated from a Uniform (0.1, 17.42) distribution. The parameter η was linked to ulceration status without the intercept and p to tumour thickness. Note that the absence of ulcer implies η taking the value 1 and hence if ulcer is present η should be

greater than 1 since on an average the number of competing causes is expected to be more. To obtain a true value of the regression parameter β_1 corresponding to η, we chose a value of η as 1.5 which gave the value of β_1 as 0.405, after solving $\exp(\beta_1) = 1.5$. For a choice of the regression parameters (β_{02} and β_{12}) corresponding to the parameter p, we chose two values of p as 0.2 and 0.99 and equated them to the logistic link functions with the values of tumour thickness as 0.1 and 17.42, respectively (that is, $1+\exp(\beta_{02}+0.1\beta_{12}) = 0.2$ and $1+\exp(\beta_{02}+17.42\beta_{12}) = 0.99$). After solving the two equations simultaneously, we obtained the true values of the regression parameters corresponding to p as $\beta_{02} = -1.421$ and $\beta_{12} = 0.345$. The censoring times (C) were generated from an exponential distribution with rate 0.2. To generate the lifetime, we followed the following steps:

i) Generate a random variable M from a COM-Poisson distribution with parameter η, for a fixed choice of ϕ.

ii) If $M = 0$, the observed time (T) is taken as the censoring time (C), i.e, $T = C$. Otherwise, generate a random variable D from a binomial distribution with parameters M and p.

iii) From (ii), if $D = 0$, set $T = C$. Otherwise, generate D Weibull random variables $(W_1, W_2, \ldots, W_D)$ with shape $= \frac{1}{\gamma_1}$ and scale $= \frac{1}{\gamma_2}$, for a specific choice of γ_1 and γ_2, and then take $T = \min\{\min\{W_1, W_2, \ldots, W_D\}, C\}$.

iv) If $T = C$, the censoring indicator δ is taken as 0, otherwise, it is taken as 1.

The choice of the lifetime parameters γ_1 and γ_2 was obtained after equating the theoretical mean and variance of the considered Weibull distribution to some fixed values. For this, we considered two different choices of the variance as 0.15 and 0.30 and one choice of the mean as 2.5. This gave us two different choices of (γ_1, γ_2) as (0.131, 0.376) and (0.191, 0.368). The proportion of censored observations, on average, was approximately 72% which is similar to the real melanoma data. Initial values of the parameters are to be provided for the iterative procedure to start. For this purpose, discrete four-dimensional grid search was done on the parameters (ϕ, $\beta_1, \beta_{02}, \beta_{12}$), and an interval for the lifetime parameter (γ_1, γ_2)

was created by taking 20% deviation off its true value. Then, we independently sampled the regression and lifetime parameter values which maximised the log-likelihood function to get the initial values.

In this simulation study, we examined the performance of the proposed estimation method through the empirical bias and RMSE of the estimates. The coverage probabilities were also computed based on two nominal levels of 5% and 10%. All simulation results were based on 1000 MC runs.

Tables 2.1 and 2.2 present the results for the destructive Bernoulli and destructive Poisson cure rate models. It is clear that the estimates obtained from the EM algorithm are close to the true parameter values. Both the standard error and RMSE decreases as sample size increases and hence satisfies the large sample properties. The coverage probabilities of the asymptotic confidence intervals are also close to the nominal levels used. It should be noted that the relative bias is high for the regression parameters as compared to the lifetime parameters. In Tables 2.3 - 2.5, we present the results for the destructive COM-Poisson model where two different true values of ϕ were chosen as 0.75 and 1.3. In this case, two approaches were employed in estimating ϕ; first, ϕ was estimated using the profile likelihood approach and then using the complete likelihood approach. Table 2.3 shows the results obtained using the profile likelihood approach. The bias in the estimate of ϕ decreases as sample size increases, as expected. Also, since standard errors are calculated for a fixed ϕ, it results in underestimation problem because the uncertainty of ϕ is not taken into account. This underestimation problem is observed in the coverage probabilities where all the model parameters suffer from under coverage and is more pronounced for β_1. Tables 2.4 and 2.5 shows the results obtained when using the complete likelihood approach for different values of ϕ. A significant improvement is noticed in the coverage probabilities as compared to the profile likelihood approach. Once again, the bias, standard errors and RMSEs decreases as sample size increases.

Table 2.1 Accuracy and precision of estimates for the destructive Bernoulli cure rate model

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
200	(0.191, 0.368)	β_1	0.514 (0.571)	0.109	27	0.514	0.946	0.974
		β_{02}	−1.715 (0.974)	-0.294	21	1.008	0.929	0.971
		β_{12}	0.419 (0.229)	0.074	21	0.250	0.869	0.903
		γ_1	0.189 (0.021)	-0.002	-1	0.019	0.912	0.960
		γ_2	0.369 (0.011)	0.001	0.5	0.010	0.904	0.955
500	(0.191, 0.368)	β_1	0.453 (0.318)	0.048	12	0.327	0.925	0.957
		β_{02}	−1.532 (0.571)	-0.111	8	0.613	0.903	0.958
		β_{12}	0.376 (0.133)	0.031	9	0.147	0.875	0.923
		γ_1	0.189 (0.013)	-0.002	-1	0.013	0.898	0.945
		γ_2	0.369 (0.007)	0.001	0.5	0.007	0.915	0.964
200	(0.131, 0.376)	β_1	0.484 (0.515)	0.079	20	0.470	0.941	0.976
		β_{02}	−1.782 (1.016)	-0.361	25	0.975	0.945	0.980
		β_{12}	0.454 (0.252)	0.109	32	0.232	0.941	0.961
		γ_1	0.129 (0.014)	-0.002	-2	0.013	0.897	0.948
		γ_2	0.376 (0.007)	0.000	0	0.007	0.913	0.953
500	(0.131, 0.376)	β_1	0.445 (0.309)	0.040	10	0.312	0.926	0.958
		β_{02}	−1.547 (0.579)	-0.126	9	0.608	0.914	0.957
		β_{12}	0.384 (0.137)	0.039	11	0.142	0.912	0.955
		γ_1	0.130 (0.009)	-0.001	-0.8	0.009	0.897	0.942
		γ_2	0.376 (0.005)	0.000	0	0.005	0.920	0.962

2.4 Model discrimination

The COM-Poisson distribution is flexible and it includes some of the discrete distributions as its particular cases. This makes it suitable as a tool to select a simple distribution within its family that provides an adequate fit. In this simulation study, we evaluate the performance of the LRT in testing the null hypothesis that the competing cause can be modelled by one of the Bernoulli, Poisson, COM-Poisson ($\phi = 0.3$) and COM-Poisson ($\phi = 2.0$) distributions versus the alternative hypothesis that the competing cause can be modelled by a member of the COM-Poisson family other than the one stated in the null hypothesis. The asymptotic null distribution of the LRT statistic (Λ) under the null hypothesis is a chi-square distribution with one degree of freedom. However, when testing for the Bernoulli distribution, the asymptotic null distribution of Λ is such that $P(\Lambda \leq \lambda) = \frac{1}{2} + \frac{1}{2}P(\chi_1^2 \leq \lambda)$, where χ_1^2 is a

Table 2.2 Accuracy and precision of estimates for the destructive Poisson cure rate model

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
200	(0.191, 0.368)	β_1	0.429 (0.220)	0.024	6	0.224	0.910	0.956
		β_{02}	−1.620 (0.809)	-0.199	14	0.888	0.927	0.971
		β_{12}	0.418 (0.227)	0.073	21	0.280	0.879	0.912
		γ_1	0.189 (0.017)	-0.002	-1	0.016	0.893	0.939
		γ_2	0.369 (0.010)	0.001	0.3	0.010	0.889	0.937
500	(0.191, 0.368)	β_1	0.410 (0.137)	0.005	1	0.140	0.889	0.945
		β_{02}	1.519 (0.480)	-0.098	7	0.519	0.917	0.964
		β_{12}	0.379 (0.131)	0.034	1	0.163	0.896	0.932
		γ_1	0.190 (0.011)	-0.001	-0.5	0.011	0.891	0.932
		γ_2	0.368 (0.006)	0.000	0	0.007	0.888	0.940
200	(0.131, 0.376)	β_1	0.423 (0.217)	0.018	4	0.221	0.901	0.950
		β_{02}	−1.670 (0.846)	-0.249	17	0.859	0.941	0.980
		β_{12}	0.447 (0.248)	0.102	30	0.260	0.939	0.963
		γ_1	0.129 (0.012)	-0.002	-2	0.011	0.879	0.933
		γ_2	0.376 (0.007)	0.000	0	0.007	0.895	0.937
500	(0.131, 0.376)	β_1	0.408 (0.137)	0.003	0.7	0.141	0.888	0.946
		β_{02}	−1.523 (0.483)	-0.102	7	0.497	0.921	0.971
		β_{12}	0.383 (0.133)	0.038	11	0.152	0.924	0.959
		γ_1	0.130 (0.007)	-0.001	-0.8	0.008	0.894	0.935
		γ_2	0.376 (0.004)	0.000	0	0.005	0.885	0.933

chi-square random variable with one degree of freedom (Self and Liang, 1987). Considering the same parameter settings used in the simulation study and for each generated data, we calculated the Λ values of the fitted Bernoulli, Poisson, COM-Poisson ($\phi = 0.3$) and COM-Poisson ($\phi = 2.0$) models versus the fitted general COM-Poisson model. Using 5% level of significance and based on 1000 MC runs, we calculated the observed levels (in bold) and power values of the LRT (determined by the rejection rate of null hypothesis) and present them in Table 2.6. First, it is noted that when the true model is either Poisson ($\phi = 1.0$) or COM-Poisson ($\phi = 2.0$), the observed levels are close to the true nominal level used. When the true model is Bernoulli ($\phi \to \infty$), the observed levels are liberal as they go beyond the nominal levels used, however, they improve as sample size increases. When the true model is COM-Poisson ($\phi = 0.3$), the observed levels are conservative, however, they improve as sample size increases. When COM-Poisson ($\phi = 0.3$) is the true model, the LRT has a very high power to reject the Bernoulli model, but, it has moderate power to reject the

Table 2.3 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 0.75$ using profile likelihood approach

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
200	(0.191, 0.368)	β_1	0.423 (0.170)	0.018	4	0.295	0.624	0.719
		β_{02}	−1.873 (0.813)	-0.452	23	1.131	0.849	0.920
		β_{12}	0.487 (0.274)	0.142	41	0.592	0.780	0.810
		γ_1	0.189 (0.016)	-0.002	-1	0.017	0.863	0.915
		γ_2	0.368 (0.010)	0.000	0	0.013	0.799	0.874
		ϕ	0.825 (−)	0.055	7	0.617	-	-
500	(0.191, 0.368)	β_1	0.406 (0.109)	0.001	0.2	0.193	0.645	0.723
		β_{02}	−1.608 (0.459)	-0.187	13	0.493	0.854	0.912
		β_{12}	0.391 (0.136)	0.046	13	0.214	0.811	0.854
		γ_1	0.190 (0.010)	-0.001	-0.5	0.011	0.862	0.917
		γ_2	0.368 (0.007)	0.000	0	0.009	0.785	0.864
		ϕ	0.775 (−)	0.025	3	0.397	-	-
200	(0.131, 0.376)	β_1	0.426 (0.174)	0.021	5	0.287	0.694	0.798
		β_{02}	−1.808 (0.802)	-0.387	27	0.952	0.879	0.936
		β_{12}	0.475 (0.268)	0.130	38	0.461	0.805	0.837
		γ_1	0.129 (0.011)	-0.002	-2	0.012	0.859	0.914
		γ_2	0.376 (0.007)	0.000	0	0.009	0.810	0.878
		ϕ	0.836 (−)	0.086	11	0.599	-	-
500	(0.131, 0.376)	β_1	0.401 (0.153)	-0.004	-1	0.218	0.736	0.803
		β_{02}	−1.617 (0.493)	-0.196	14	0.528	0.853	0.933
		β_{12}	0.382 (0.134)	0.037	11	0.200	0.802	0.847
		γ_1	0.130 (0.008)	-0.001	-0.8	0.008	0.852	0.929
		γ_2	0.376 (0.004)	0.000	0	0.006	0.767	0.852
		ϕ	1.314 (−)	0.014	2	0.623	-	-

Poisson and COM-Poisson ($\phi = 2.0$) models. When the true model is Poisson, the LRT only has moderate power to reject the Bernoulli. Similarly, when the true model is either COM-Poisson ($\phi = 2.0$) or Bernoulli, the test has low power to reject any of the competing models. Also, as expected, note that the power to reject the wrong model increases as sample size increases.

Table 2.4 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 0.75$ using complete likelihood approach

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.191, 0.368)	β_1	0.485 (0.355)	0.080	0.302	0.956	0.969
		β_{02}	-1.843 (0.944)	-0.422	0.932	0.955	0.983
		β_{12}	0.530 (0.328)	0.185	0.492	0.975	0.991
		γ_1	0.189 (0.018)	-0.002	0.016	0.923	0.959
		γ_2	0.371 (0.015)	0.003	0.012	0.960	0.986
		ϕ	0.944 (0.775)	0.194	0.566	0.981	0.989
500	(0.191, 0.368)	β_1	0.429 (0.194)	0.024	0.176	0.940	0.973
		β_{02}	-1.596 (0.521)	-0.175	0.481	0.923	0.968
		β_{12}	0.413 (0.171)	0.068	0.186	0.940	0.976
		γ_1	0.190 (0.011)	-0.001	0.011	0.908	0.949
		γ_2	0.370 (0.009)	0.002	0.008	0.952	0.984
		ϕ	0.826 (0.398)	0.076	0.318	0.963	0.979
200	(0.131, 0.376)	β_1	0.439 (0.325)	0.034	0.295	0.924	0.946
		β_{02}	-1.888 (0.938)	-0.467	0.900	0.910	0.970
		β_{12}	0.517 (0.320)	0.172	0.444	0.984	0.997
		γ_1	0.130 (0.012)	-0.001	0.011	0.913	0.955
		γ_2	0.377 (0.010)	0.001	0.009	0.945	0.976
		ϕ	0.902 (0.709)	0.152	0.578	0.937	0.963
500	(0.131, 0.376)	β_1	0.416 (0.187)	0.011	0.180	0.913	0.939
		β_{02}	-1.632 (0.515)	-0.211	0.466	0.907	0.957
		β_{12}	0.412 (0.167)	0.067	0.166	0.960	0.984
		γ_1	0.130 (0.008)	-0.001	0.007	0.920	0.960
		γ_2	0.376 (0.006)	0.000	0.006	0.917	0.957
		ϕ	0.802 (0.381)	0.052	0.352	0.927	0.953

Table 2.5 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 1.3$ using complete likelihood approach

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.191, 0.368)	β_1	0.436 (0.452)	0.031	0.339	0.959	0.976
		β_{02}	-1.953 (1.093)	-0.532	1.089	0.941	0.984
		β_{12}	0.487 (0.308)	0.142	0.418	0.956	0.981
		γ_1	0.190 (0.021)	-0.001	0.017	0.964	0.993
		γ_2	0.371 (0.016)	0.003	0.011	0.976	0.993
		ϕ	1.887 (2.228)	0.587	0.954	0.993	0.994
500	(0.191, 0.368)	β_1	0.443 (0.270)	0.038	0.217	0.952	0.989
		β_{02}	-1.576 (0.594)	-0.155	0.597	0.945	0.976
		β_{12}	0.411 (0.180)	0.066	0.244	0.928	0.977
		γ_1	0.190 (0.013)	-0.001	0.012	0.921	0.965
		γ_2	0.372 (0.010)	0.004	0.008	0.944	0.977
		ϕ	1.844 (1.328)	0.544	0.768	0.997	0.998
200	(0.131, 0.376)	β_1	0.405 (0.397)	0.000	0.307	0.922	0.957
		β_{02}	-2.049 (1.094)	-0.628	1.063	0.927	0.965
		β_{12}	0.529 (0.338)	0.184	0.404	0.979	0.996
		γ_1	0.130 (0.014)	-0.001	0.013	0.924	0.960
		γ_2	0.376 (0.011)	0.000	0.008	0.954	0.981
		ϕ	1.562 (1.571)	0.262	0.844	0.957	0.971
500	(0.131, 0.376)	β_1	0.418 (0.243)	0.013	0.221	0.913	0.957
		β_{02}	-1.645 (0.576)	-0.224	0.534	0.913	0.966
		β_{12}	0.408 (0.166)	0.063	0.197	0.928	0.972
		γ_1	0.131 (0.009)	0.000	0.008	0.914	0.971
		γ_2	0.376 (0.007)	0.000	0.006	0.925	0.967
		ϕ	1.471 (0.928)	0.171	0.712	0.938	0.965

Table 2.6 Observed levels and rejection rates of the likelihood ratio test ($\gamma_1 = 0.191$ and $\gamma_2 = 0.368$)

	Fitted destructive	True destructive COM-Poisson model			
n	COM-Poisson model	$\phi = 0.3$	$\phi = 1.0$	$\phi = 2.0$	$\phi \to \infty$
200	$\phi = 0.3$	**0.002**	0.163	0.201	0.189
	$\phi = 1.0$	0.428	**0.047**	0.058	0.070
	$\phi = 2.0$	0.663	0.178	**0.035**	0.014
	$\phi \to \infty$	0.976	0.567	0.216	**0.071**
500	$\phi = 0.3$	**0.015**	0.400	0.526	0.522
	$\phi = 1.0$	0.764	**0.048**	0.148	0.282
	$\phi = 2.0$	0.939	0.294	**0.036**	0.040
	$\phi \to \infty$	1.000	0.856	0.330	**0.065**

2.5 Melanoma data analysis

The estimation procedure was applied to a well-known data set on malignant melanoma (see Section 1.8). Figure 2.1 shows the Kaplan-Meier curve stratified by ulceration status and the curves level off at 0.4 and 0.85 showing that immunes are present in both the ulcer and non-ulcer groups.

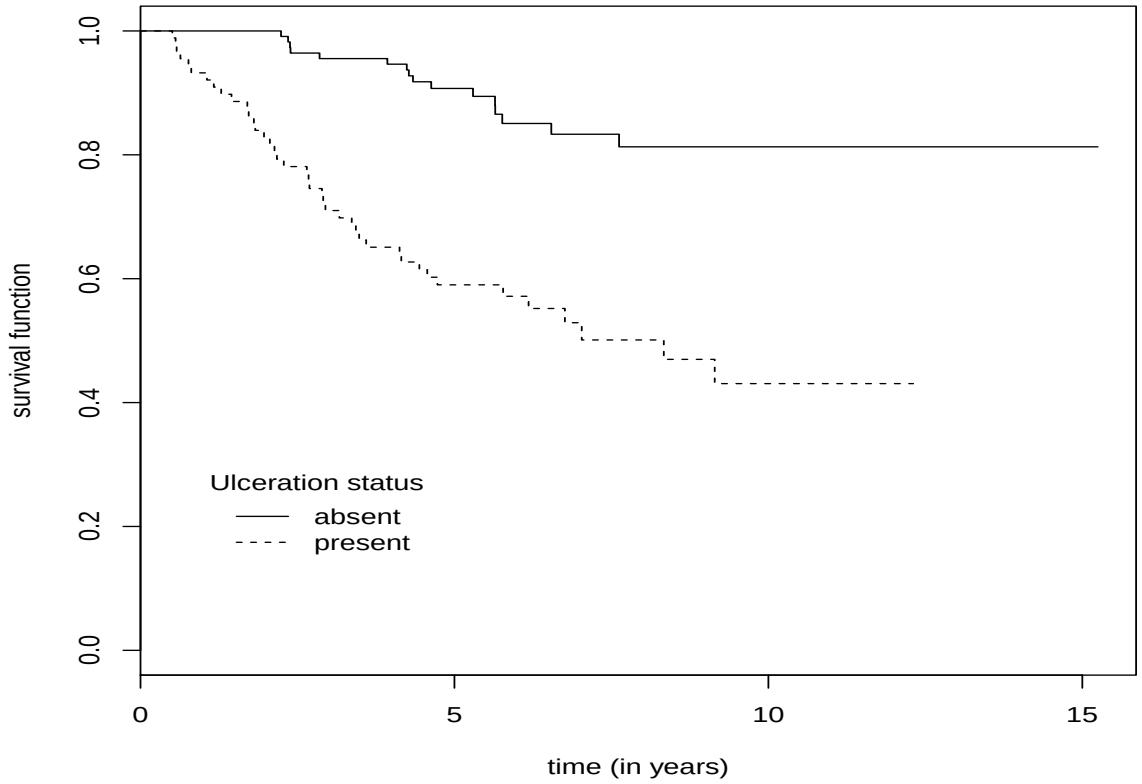

Fig. 2.1 Kaplan-Meier curve of melanoma data stratified by ulceration status

To circumvent the identifiability issue, as discussed in Section 2.1, we can consider the following four scenarios in practice.

Setting 1: η linked to ulceration status without an intercept and p linked to tumor thickness with an intercept; Setting 2: η linked to tumor thickness without an intercept and p linked to ulceration status with an intercept; Setting 3: η linked to ulceration status with an intercept

and p linked to tumor thickness without an intercept; Setting 4: η linked to tumor thickness with an intercept and p linked to ulceration status without an intercept.

The EM algorithm was run for all the four settings and we present the maximised log-likelihood values together with the AIC and BIC values in Table 2.7.

Table 2.7 Maximized log-likelihood values (l) and information criteria values

Setting	l	AIC	BIC
1	-206.799	425.598	445.536
2	-207.475	426.950	446.888
3	-207.932	427.864	447.802
4	-209.372	430.744	450.682

By comparing the maximised log-likelihood, AIC and BIC values of Table 2.7, it is clear that Setting 1, i.e., η linked to ulceration status without an intercept and p linked to tumor thickness with an intercept, provides the best fit to the data. For Setting 1, the MLE of ϕ, turned out to be ≈ 1. So, we considered testing for the destructive Poisson model ($H_0 : \phi = 1$). The LRT statistic turned out to be ≈ 0 and hence the p-value turned out to be ≈ 1, which gives us very strong evidence to conclude that the destructive Poisson model should be our working model. Thus, for the given data set the destructive COM-Poisson model reduces to the simple destructive Poisson model. We also considered testing for the suitability of the destructive Bernoulli model ($H_0 : \phi \to \infty$) and in this case the LRT statistic turned out to be 2.958 with the corresponding p-value 0.043. Note that the assumption of Bernoulli distribution gets rejected at 5% significance level. In Table 2.8, we present the MLEs and standard errors of the destructive COM-Poisson model obtained using the profile likelihood and complete likelihood approaches along with the results for the destructive Poisson model.

Table 2.8 shows that when profile likelihood approach is utilised, all the parameters are significant in the model at 5% level. When all the parameters are estimated simultaneously, all the regression and lifetime parameters are significant at 5% level except β_1, which is significant at 20% and ϕ, which is significant at 15% level. The high standard error of ϕ is

Table 2.8 MLEs, standard errors (se) and p- values for the destructive COM-Poisson (Poisson) model using both the profile and complete likelihood approaches

Model	Approach		ϕ	β_1	β_{02}	β_{12}	γ_1	γ_2
COM-Poisson	Complete Likelihood	MLE	0.983	0.536	-2.578	0.957	0.611	0.158
		se	0.935	0.407	0.710	0.435	0.099	0.052
		p- value	0.146	0.188	0.000	0.024	0.000	0.001
	Profile Likelihood	MLE	0.98	0.535	-2.579	0.957	0.611	0.158
		se	-	0.257	0.654	0.404	0.188	0.046
		p- value	-	0.037	0.000	0.018	0.000	0.000
Poisson		MLE	-	0.542	-2.579	0.964	0.612	0.158
		se	-	0.259	0.655	0.405	0.098	0.045
		p- value	-	0.036	0.000	0.017	0.000	0.000

induced by the flatness of the log-likelihood function, that is, ϕ can take different values but the maximised log-likelihood function remains almost the same. Note that the standard errors and hence the p-values for destructive Poisson model and the destructive COM-Poisson model obtained using the profile likelihood approach are similar. All parameters are significant for the destructive Poisson model.

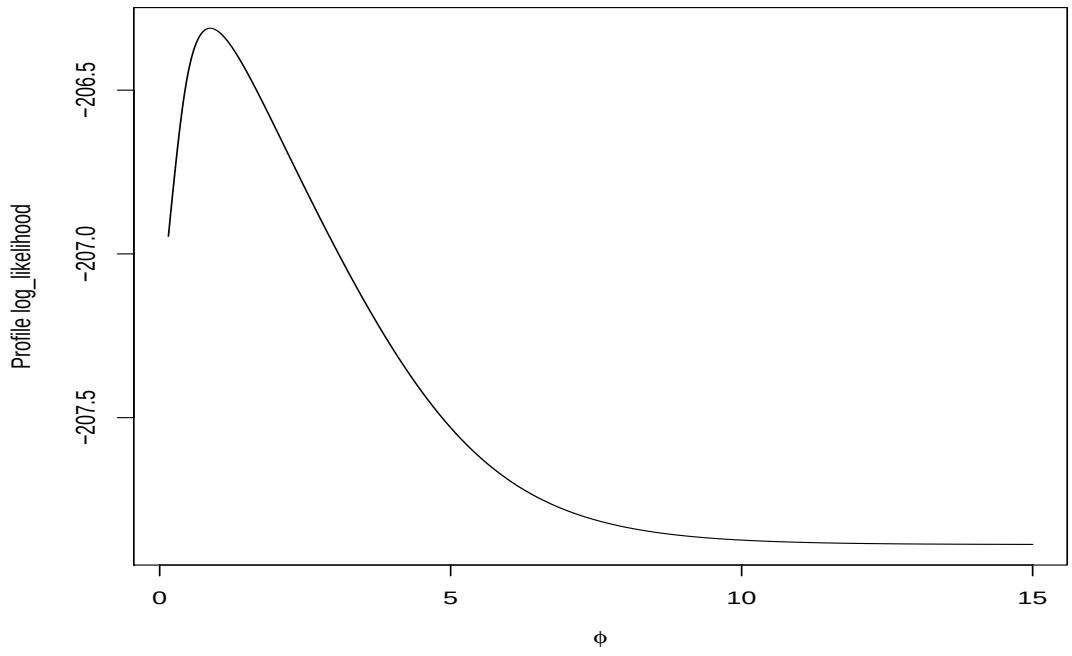

Fig. 2.2 Profile log-likelihood plot for different values of ϕ

We also fitted a particular case of the destructive Poisson model, namely, the Poisson model ($p = 1$). This relates to the case when the destructive mechanism is absent and to preserve the same number of parameters, the parameter η was linked to both covariates with an intercept term. In this case, the maximized log-likelihood value turned out to be -207.865, providing clear evidence of the fact that the destructive mechanism when taken into account gives a better fit to the data (note that the maximized log-likelihood value for the destructive Poisson model is -206.799). Finally, we also tested for the suitability of the exponential distribution for the lifetime ($H_0 : \gamma_1 = 1$) and the p-value corresponding to the LRT turned out to be 0.001, showing overwhelming evidence against the exponential distribution. In Figure 2.2, we present the profile log-likelihood plot for different values of ϕ under Setting 1, and it can be seen that the log-likelihood surface is flat with respect to ϕ and that the log-likelihood is maximum for $\phi \approx 1$, providing evidence for the destructive Poisson model.

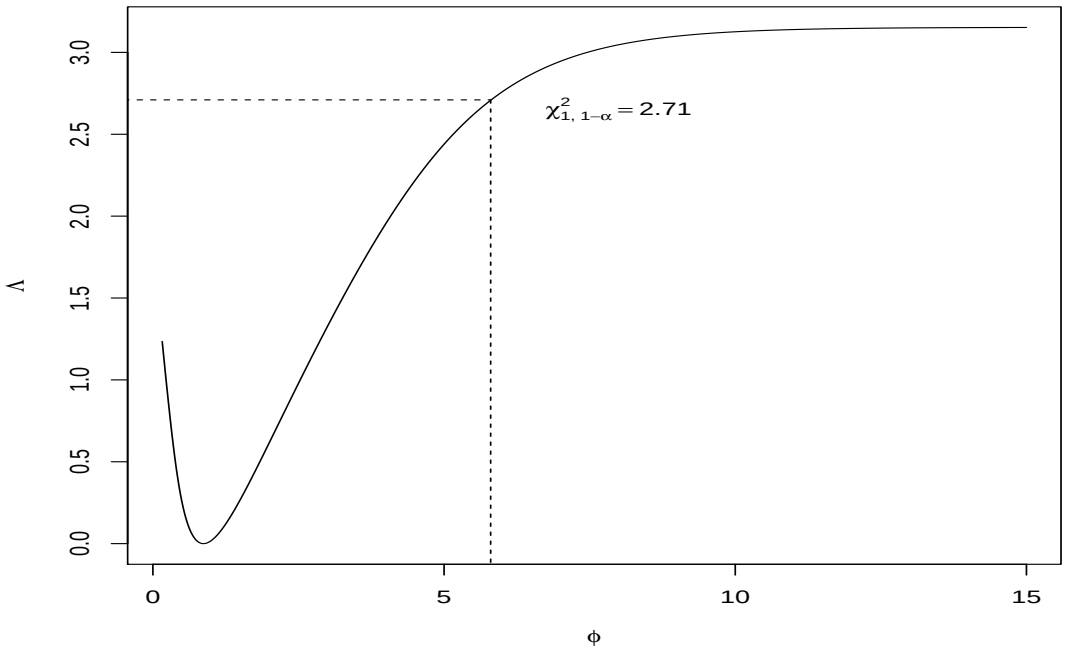

Fig. 2.3 Values of ϕ against the corresponding Λ-statistic

Figure 2.3 represents the values of ϕ plotted against the corresponding Λ statistic. We would want to look for an acceptable range of ϕ. For any value of ϕ, we reject the test at $\alpha = 10\%$

significance level if the corresponding Λ statistic value is greater than $\chi^2_{1,\,1-\alpha} = 2.71$. From the plot, an acceptable range of ϕ turns out to be $(0, 6.5)$, which clearly explains why the destructive Bernoulli model $(\phi \to \infty)$ got rejected and not the destructive Poisson model $(\phi = 1)$.

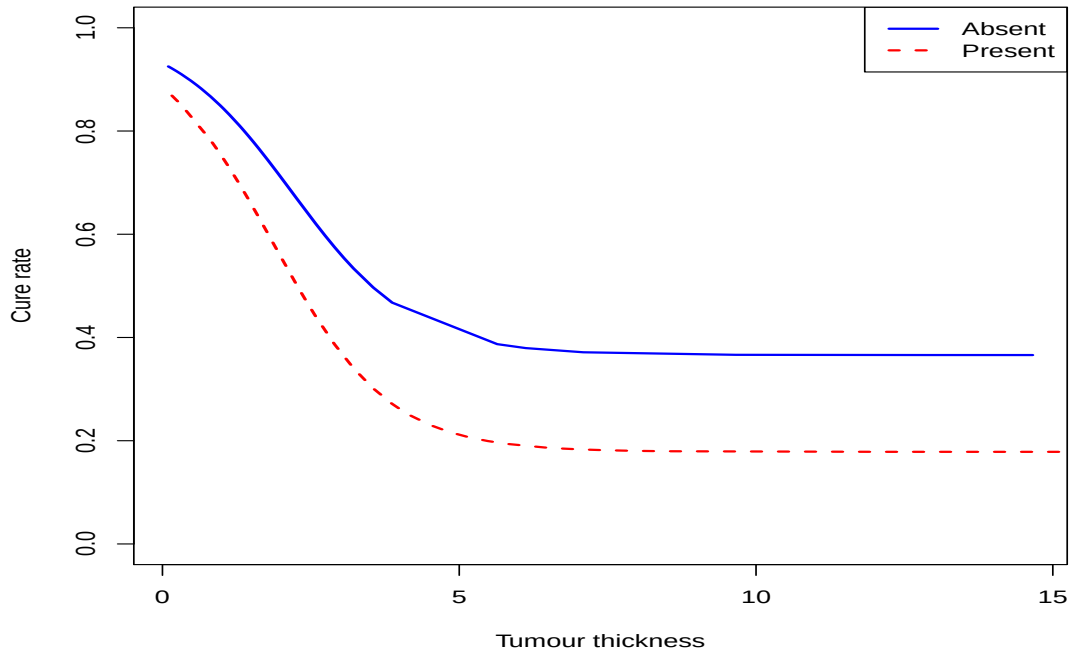

Fig. 2.4 Cure rate versus tumour thickness stratified by ulceration status

2.5.1 Effect of covariates on cure rate

To assess the covariate effect on the cure rate, note that the estimate of β_1 is positive. This implies that the presence of ulcer will have a higher mean number of competing causes, thus leading to a decrease in cure rate, as expected (the cure rate being a decreasing function of η). On the other hand, the estimate of β_{12} being positive implies that the initial competing causes will remain active with a higher probability with an increase in tumor thickness, which in turn implies a decrease in cure rate (the cure rate being a decreasing function of p). Figure 2.4 shows the cure rate plotted against the tumour thickness and stratified by ulceration status.

Clearly, an increase in tumour thickness reduces the cure rate, as the curve decreases more rapidly for thicker tumours and levels off for tumour thickness of 5 mm. The two lines are almost parallel for tumour thickness greater than 5 mm, and the cure rate levels off around 0.2 and 0.4 for patients with and without ulcers, respectively. Patients without ulcer have a higher chance of being long-term survivors as compared to those with ulcers.

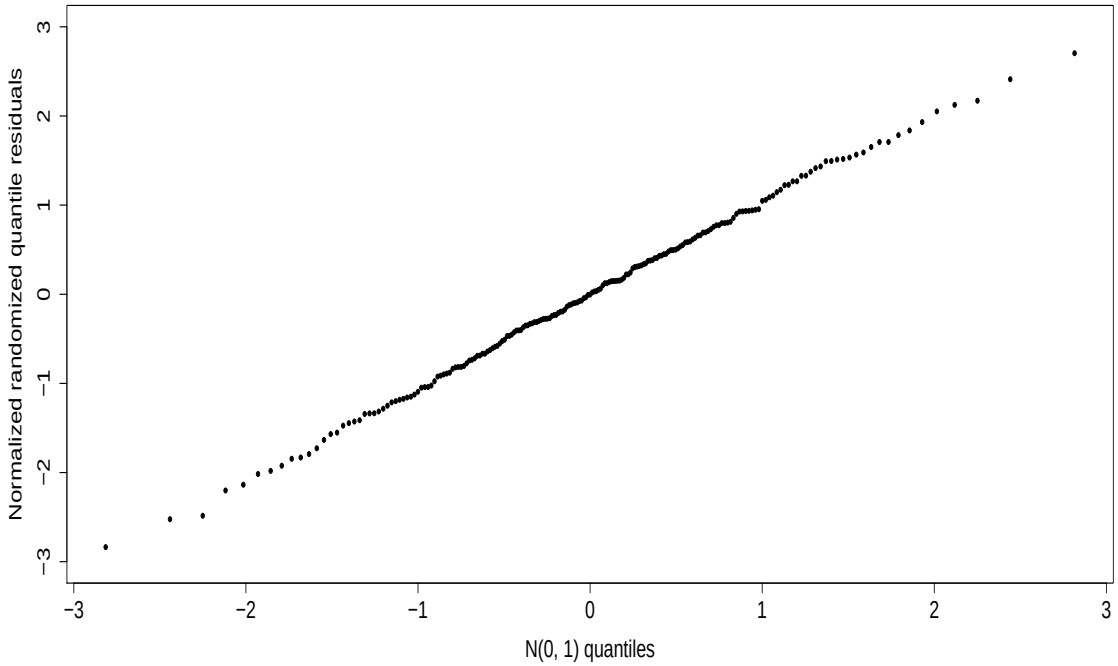

Fig. 2.5 Q-Q plot of the normalised randomised quantile residuals for the melanoma data

2.5.2 Model diagnosis

The adequacy of the destructive COM-Poisson (Poisson) model was verified using the normalised randomised quantile residuals (see Dunn and Smyth, 1996; Rigby and Stasinopoulos, 2005). Under the assumption that $1 - S_{\text{pop}}(y)$ are uniformly distributed between 0 and 1, the quantile residuals are defined by $q_i = \Phi^{-1}(1 - S_{\text{pop}}(y))$, where Φ^{-1} is the standard normal distribution quantile function. The QQ plot presented in Figure 2.5 is obtained by plotting the median of five sets of ordered residuals. As the plot is linear, it suggests that the destructive COM-Poisson model fits the data well. A formal test

for normality of residuals was carried out using the Kolmogorov-Smirnov test. The test gave a p-value of 0.997 implying very strong evidence for the goodness-of-fit.

2.6 Conclusion

In this chapter, we have developed the likelihood inference based on the EM algorithm for the flexible destructive COM-Poisson cure rate model when the lifetime follows a Weibull distribution. This model is more realistic than the usual cure rate model (where the destructive mechanism is absent) as it takes into account the biological mechanism of the occurrence of an event of interest. Right censoring introduces the missing data as in the right censored group cured subjects cannot be distinguished from the susceptibles. This motivates us to develop the EM algorithm for the estimation of the model parameters, which forms the primary contribution of this work. The simulation study results show that the proposed EM algorithm works very well. The profile likelihood approach to estimate the COM-Poisson shape parameter ϕ also performs well, however, it results in underestimation problem. The coverage probabilities are close to the nominal values and gets better as sample size increases. However, when using the complete likelihood approach, there is no problem of underestimation. The other major contribution of this work is in evaluating the flexibility of the COM-Poisson distribution to carry out a model discrimination so as to select a proper competing cause distribution that provides an adequate fit to the data. In the real data analysis, we have clearly illustrated this flexibility of the COM-Poisson distribution where the LRT rejects the assumption of the Bernoulli distribution and the Poisson model turned out to provide the best fit to the data. We have also shown that the destructive mechanism when taken into account provides a better fit to the data.

Chapter 3

Likelihood inference with lognormal lifetime

3.1 Introduction

The lognormal distribution is one of the most under utilised lifetime distribution in time to event data. A key feature of this distribution is that the hazard is not monotone, it rises quickly from zero to the peak and falls gradually. Boag (1949) used a lognormal survival model in a cure rate set up but not many authors have considered parametric cure rate models using the lognormal distribution as opposed to the Weibull distribution. In this chapter, the likelihood inference of the destructive COM-Poisson cure rate model is carried out using the lognormal as the lifetime distribution.

Assuming a two parameter lognormal distribution for the lifetime W as in (1.15) with survival function as given in (1.16), expressions for the $Q(\theta,\ \pi^{(k)})$ function are as follows.

3.2 Expressions of the Q function for destructive cure rate models

Destructive Bernoulli cure rate model

The $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned} Q(\theta,\pi^{(k)}) &= \sum_{I_1}\left\{\log\left(\frac{\eta_i p_i f(t_i)}{\sqrt{2\pi}\gamma_1 t_i(1+\eta_i)}\right)-\frac{1}{2}\left(\frac{\log(\gamma_2 t_i)}{\gamma_1}\right)^2\right\} \\ &\quad +\sum_{I_0}(1-\pi_i^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{\eta_i p_i S(t_i)}{1+\eta_i}\right), \end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=\frac{\eta_i p_i S(t_i)}{1+\eta_i(1-p_i F(t_i))}\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation, $f(t_i)$ and $S(t_i)$ are as given in (1.11) and (1.16), respectively.

Destructive Poisson cure rate model

The $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned} Q(\theta,\pi^{(k)}) &= \sum_{I_1}\{\log(\eta_i p_i f(t_i))-\eta_i p_i F(t_i)\}-\sum_{I_0}(1-\pi_i^{(k)})\eta_i p_i \\ &\quad +\sum_{I_0}\pi_i^{(k)}\log\left[\exp(-\eta_i p_i F(t_i))-\exp(-\eta_i p_i)\right], \end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=1-\exp(-\eta_i p_i S(t_i))\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation, and $F(t)=\Phi\left(\frac{\log(\gamma_2 t_i)}{\gamma_1}\right)$.

Destructive COM-Poisson cure rate model

The $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned}\mathbf{Q}(\theta,\pi^{(k)}) &= \sum_{I_1}\left\{\log\left(p_i f(t_i)\right)+\log z_{f1}-\log(z(1-p_i F(t_i)))\right\}\\ &\quad+\sum_{I_0}(1-\pi_i^{(k)})\log\left(\frac{z_p}{z}\right)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{z_f-z_p}{z}\right),\end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=\frac{z_f-z_p}{z-z_p}\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation, where $z_f=\sum_{j=0}^{\infty}\frac{[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}$, $z_p=\sum_{j=0}^{\infty}\frac{[\eta_i(1-p_i)]^j}{(j!)^\phi}$, $z=\sum_{j=0}^{\infty}\frac{\eta_i^j}{(j!)^\phi}$ and $z_{f1}=\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}$.

Explicit expressions of the first- and second-order derivatives of $Q(\theta,\pi^{(k)})$ with respect to θ are presented in Appendix B.1 for the destructive COM-Poisson model, destructive Poisson model and destructive Bernoulli model.

3.3 Standard errors

By inverting the Hessian matrix of ϕ, β_1, β_2 and γ, an approximate variance-covariance matrix of MLEs can be obtained. Explicit expressions of the score function and observed information matrix for different destructive cure rate models are presented in Appendix B.2.

3.4 Simulation study

In this simulation study, we studied the effects of different sample sizes, different censoring proportions as well as lifetime parameters in order to assess the model's performance and robustness. Two varying sample sizes of $n=200$ and $n=400$ were considered to demonstrate the small and moderately large sample behaviours. The same regression parameter setting as done in Section 2.3 (Scenario 1) is considered here and we also considered the following setting (Scenario 2): we chose a value of η as 3 which gave the true value of β_1 as 1.098, using the log-linear link function. For the true values of regression parameters β_{02} and β_{12}, p was equated to 0.33 and 0.99 using the logistic link function. The true values of β_{02}

and β_{12} were then obtained as $\beta_{02} = -0.739$ and $\beta_{12} = 0.306$. The censoring times were generated from an exponential distribution with rate 0.15. These two different settings gave the censoring proportion of approximately 70% for Scenario 1 and 50% for Scenario 2.

The choice of the lifetime parameters γ_1 and γ_2 were obtained after equating the theoretical mean and variance of the considered lognormal distribution to some fixed values. For this, we considered two different choices of mean as 1.5 and 2.5 and one value of variance as 1.1, which gave two different choices of $(\gamma_1,\ \gamma_2)$ as (0.225, 0.41) and (0.333, 0.705).

Table 3.1 presents the results for the destructive Bernoulli and destructive Poisson cure rate models. It is clear that the estimates obtained from the EM algorithm are close to the true parameter values. The bias, standard error and RMSE decrease as sample size increases and hence satisfies the large sample properties. The coverage probabilities are also close to the nominal levels used. Using Scenario 1, Tables 3.2 and 3.3 present the results for the general destructive COM-Poisson model for two different true values of ϕ as 0.7 and 1.4, respectively. In this case, ϕ was also estimated simultaneously with the rest of model parameters using the complete likelihood approach. The profile likelihood approach works very well in estimating ϕ with small bias. Furthermore, both the bias and RMSE of ϕ decreases as sample size increases. In Setting 1, the estimates of the lifetime parameters have small bias, standard error and RMSE for both the approaches. The coverage probabilities for both approaches are similar for the lifetime parameters. Under-coverage is noticeable in the regression parameters when employing the profile likelihood approach while in the complete likelihood approach there is over-coverage. Results of profile likelihood approach for Setting 2 are similar to Setting 1 and are therefore omitted. Both settings yielded similar results, the high standard error in ϕ (using complete likelihood approach) seems to affect the precision and accuracy of other regression parameters. Generally, the bias, standard error and RMSE decreases as sample size increases and the coverage probability improves when the complete likelihood approach is used.

Scenario 2 results are presented in Table 3.4 and once again the profile and complete

likelihood approaches are employed. Generally, the bias, standard error, and RMSE decreases as sample size increases. Note the heavy under-coverage for β_1 when the profile likelihood approach is used as compared to Scenario 1. Also, the complete likelihood approach has large bias, standard error and RMSE of ϕ.

Table 3.1 Accuracy and precision of estimates for the destructive Poisson and destructive Bernoulli cure rate models

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
			Destructive Poisson					
200	(0.225, 0.41)	β_1	0.440 (0.237)	0.035	9	0.239	0.904	0.941
		β_{02}	-1.571 (0.887)	-0.150	11	0.858	0.932	0.978
		β_{12}	0.431 (0.264)	0.086	25	0.297	0.880	0.916
		γ_1	0.227 (0.022)	0.002	0.9	0.019	0.944	0.975
		γ_2	0.407 (0.014)	-0.003	-0.7	0.011	0.944	0.977
400	(0.225, 0.41)	β_1	0.431 (0.163)	0.026	6	0.165	0.893	0.945
		β_{02}	-1.525 (0.581)	-0.104	7	0.597	0.915	0.967
		β_{12}	0.394 (0.164)	0.049	14	0.182	0.895	0.932
		γ_1	0.228 (0.016)	0.003	1	0.013	0.947	0.981
		γ_2	0.408 (0.010)	-0.002	-0.5	0.008	0.942	0.982
200	(0.333, 0.705)	β_1	0.438 (0.205)	0.033	9	0.211	0.891	0.943
		β_{02}	-1.635 (0.818)	-0.214	15	0.885	0.904	0.959
		β_{12}	0.464 (0.269)	0.119	34	0.324	0.895	0.932
		γ_1	0.336 (0.029)	0.003	0.9	0.025	0.939	0.978
		γ_2	0.697 (0.030)	-0.008	-1	0.024	0.942	0.977
400	(0.333, 0.705)	β_1	0.438 (0.145)	0.033	9	0.141	0.902	0.954
		β_{02}	-1.513 (0.527)	-0.092	6	0.549	0.908	0.957
		β_{12}	0.397 (0.158)	0.052	15	0.173	0.918	0.948
		γ_1	0.336 (0.020)	0.003	0.9	0.018	0.941	0.976
		γ_2	0.697 (0.021)	-0.008	-1	0.018	0.915	0.964
			Destructive Bernoulli					
200	(0.225, 0.41)	β_1	0.558 (0.723)	0.153	38	0.582	0.943	0.966
		β_{02}	-1.593 (1.044)	-0.172	12	0.945	0.966	0.991
		β_{12}	0.401 (0.259)	0.056	16	0.232	0.912	0.935
		γ_1	0.226 (0.026)	0.001	0.4	0.022	0.954	0.981
		γ_2	0.408 (0.015)	-0.002	-0.5	0.013	0.937	0.964
400	(0.225, 0.41)	β_1	0.490 (0.421)	0.085	21	0.412	0.938	0.974
		β_{02}	-1.578 (0.744)	-0.157	11	0.738	0.916	0.968
		β_{12}	0.403 (0.193)	0.058	17	0.209	0.907	0.940
		γ_1	0.228 (0.019)	0.003	1	0.016	0.957	0.985
		γ_2	0.407 (0.010)	-0.003	-0.7	0.009	0.904	0.956
200	(0.333, 0.705)	β_1	0.527 (0.578)	0.122	30	0.482	0.945	0.970
		β_{02}	-1.759 (1.046)	-0.338	24	1.121	0.941	0.982
		β_{12}	0.458 (0.287)	0.113	33	0.313	0.920	0.950
		γ_1	0.334 (0.034)	0.001	0.3	0.030	0.941	0.976
		γ_2	0.699 (0.033)	-0.006	-0.9	0.028	0.935	0.969
400	(0.333, 0.705)	β_1	0.486 (0.370)	0.081	20	0.358	0.940	0.972
		β_{02}	-1.561 (0.671)	-0.140	10	0.698	0.915	0.966
		β_{12}	0.406 (0.183)	0.061	18	0.210	0.905	0.942
		γ_1	0.336 (0.024)	0.003	1	0.021	0.943	0.982
		γ_2	0.697 (0.023)	-0.008	-1	0.021	0.909	0.952

Table 3.2 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 0.7$ using both the profile and complete likelihood approaches

Likelihood	n	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
Profile	200	$\beta_1 = 0.405$	0.407 (0.166)	0.002	0.306	0.589	0.701
Setting 1		$\beta_{02} = -1.421$	-1.956 (0.811)	-0.535	0.890	0.847	0.916
		$\beta_{12} = 0.345$	0.456 (0.241)	0.111	0.327	0.785	0.824
		$\gamma_1 = 0.225$	0.225 (0.020)	0.000	0.020	0.884	0.947
		$\gamma_2 = 0.410$	0.407 (0.014)	-0.003	0.017	0.835	0.902
		$\phi = 0.700$	0.722 ($-$)	0.022	0.565	-	-
	400	$\beta_1 = 0.405$	0.399 (0.121)	-0.006	0.206	0.647	0.718
		$\beta_{02} = -1.421$	-1.722 (0.553)	-0.301	0.588	0.846	0.921
		$\beta_{12} = 0.345$	0.418 (0.167)	0.073	0.223	0.820	0.866
		$\gamma_1 = 0.225$	0.224 (0.014)	-0.001	0.015	0.884	0.938
		$\gamma_2 = 0.410$	0.409 (0.009)	-0.001	0.013	0.791	0.863
		$\phi = 0.700$	0.722 ($-$)	0.022	0.406	-	-
Complete	200	$\beta_1 = 0.405$	0.490 (0.349)	0.085	0.323	0.951	0.973
Setting 1		$\beta_{02} = -1.421$	-2.024 (1.089)	-0.603	1.064	0.931	0.970
		$\beta_{12} = 0.345$	0.655 (0.435)	0.310	0.397	0.977	0.996
		$\gamma_1 = 0.225$	0.223 (0.021)	-0.002	0.020	0.902	0.957
		$\gamma_2 = 0.410$	0.412 (0.019)	0.002	0.017	0.926	0.968
		$\phi = 0.700$	0.971 (0.737)	0.271	0.600	0.967	0.978
	400	$\beta_1 = 0.405$	0.445 (0.200)	0.040	0.196	0.912	0.950
		$\beta_{02} = -1.421$	-1.660 (0.559)	-0.239	0.536	0.913	0.966
		$\beta_{12} = 0.345$	0.469 (0.201)	0.124	0.207	0.966	0.989
		$\gamma_1 = 0.225$	0.224 (0.013)	-0.001	0.012	0.918	0.964
		$\gamma_2 = 0.410$	0.412 (0.012)	0.002	0.012	0.908	0.967
		$\phi = 0.700$	0.830 (0.399)	0.130	0.401	0.953	0.977
Complete	200	$\beta_1 = 0.405$	0.553 (0.436)	0.148	0.374	0.984	0.997
Setting 2		$\beta_{02} = -1.421$	-1.895 (0.932)	-0.474	0.915	0.930	0.984
		$\beta_{12} = 0.345$	0.627 (0.364)	0.282	0.347	0.969	0.995
		$\gamma_1 = 0.333$	0.327 (0.027)	-0.006	0.027	0.869	0.929
		$\gamma_2 = 0.705$	0.715 (0.044)	0.010	0.039	0.925	0.972
		$\phi = 0.700$	1.059 (0.878)	0.359	0.730	0.993	0.995
	400	$\beta_1 = 0.405$	0.487 (0.187)	0.082	0.147	0.975	0.996
		$\beta_{02} = -1.421$	-1.562 (0.493)	-0.141	0.446	0.935	0.977
		$\beta_{12} = 0.345$	0.466 (0.179)	0.121	0.166	0.960	0.986
		$\gamma_1 = 0.333$	0.329 (0.017)	-0.004	0.017	0.889	0.947
		$\gamma_2 = 0.705$	0.714 (0.027)	0.009	0.024	0.924	0.967
		$\phi = 0.700$	0.914 (0.378)	0.214	0.303	0.998	0.998

Table 3.3 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 1.4$ using both the profile and complete likelihood approaches

Approach	n	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
Profile	200	$\beta_1 = 0.405$	0.385 (0.251)	-0.020	0.347	0.721	0.800
Likelihood		$\beta_{02} = -1.421$	-1.936 (0.898)	-0.515	0.983	0.837	0.905
Setting 1		$\beta_{12} = 0.345$	0.425 (0.249)	0.080	0.346	0.760	0.791
		$\gamma_1 = 0.225$	0.223 (0.022)	-0.002	0.022	0.899	0.944
		$\gamma_2 = 0.410$	0.406 (0.014)	-0.004	0.017	0.804	0.882
		$\phi = 1.400$	1.303 $(-)$	-0.097	0.860	-	-
	400	$\beta_1 = 0.405$	0.379 (0.181)	-0.026	0.251	0.714	0.783
		$\beta_{02} = -1.421$	-1.775 (0.601)	-0.354	0.689	0.818	0.900
		$\beta_{12} = 0.345$	0.397 (0.164)	0.052	0.255	0.762	0.821
		$\gamma_1 = 0.225$	0.225 (0.016)	0.000	0.016	0.894	0.942
		$\gamma_2 = 0.410$	0.408 (0.010)	-0.002	0.013	0.754	0.830
		$\phi = 1.400$	1.346 $(-)$	-0.054	0.753	-	-
Complete	200	$\beta_1 = 0.405$	0.417 (0.418)	0.012	0.325	0.923	0.962
Likelihood		$\beta_{02} = -1.421$	-2.121 (1.241)	-0.700	1.212	0.946	0.992
Setting 1		$\beta_{12} = 0.345$	0.620 (0.418)	0.275	0.400	0.992	0.998
		$\gamma_1 = 0.225$	0.223 (0.022)	-0.002	0.023	0.878	0.934
		$\gamma_2 = 0.410$	0.410 (0.020)	0.000	0.016	0.949	0.984
		$\phi = 1.400$	1.737 (1.890)	0.337	1.007	0.961	0.980
	400	$\beta_1 = 0.405$	0.407 (0.290)	0.002	0.239	0.918	0.942
		$\beta_{02} = -1.421$	-1.849 (0.732)	-0.428	0.741	0.893	0.961
		$\beta_{12} = 0.345$	0.518 (0.248)	0.173	0.269	0.949	0.984
		$\gamma_1 = 0.225$	0.223 (0.016)	-0.002	0.016	0.897	0.938
		$\gamma_2 = 0.410$	0.410 (0.014)	0.000	0.013	0.939	0.966
		$\phi = 1.400$	1.669 (1.268)	0.269	0.827	0.958	0.975
Complete	200	$\beta_1 = 0.405$	0.504 (0.417)	0.099	0.290	0.981	0.993
Likelihood		$\beta_{02} = -1.421$	-1.927 (1.068)	-0.506	0.951	0.953	0.983
Setting 2		$\beta_{12} = 0.345$	0.607 (0.371)	0.262	0.363	0.972	0.995
		$\gamma_1 = 0.333$	0.327 (0.029)	-0.006	0.030	0.858	0.936
		$\gamma_2 = 0.705$	0.714 (0.047)	0.009	0.038	0.948	0.986
		$\phi = 1.400$	2.074 (2.281)	0.674	1.071	0.988	0.993
	400	$\beta_1 = 0.405$	0.499 (0.262)	0.094	0.214	0.972	0.991
		$\beta_{02} = -1.421$	-1.625 (0.572)	-0.204	0.575	0.931	0.967
		$\beta_{12} = 0.345$	0.494 (0.196)	0.149	0.215	0.947	0.982
		$\gamma_1 = 0.333$	0.328 (0.018)	-0.005	0.018	0.886	0.933
		$\gamma_2 = 0.705$	0.717 (0.030)	0.012	0.025	0.928	0.975
		$\phi = 1.400$	2.057 (1.431)	0.657	0.897	0.996	0.996

Table 3.4 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model using both profile and complete likelihood approaches with $\phi = 0.7$ and $\phi = 1.4$

Approach	n	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
Profile Likelihood	200	$\beta_1 = 1.098$	1.016 (0.238)	-0.082	0.577	0.454	0.524
		$\beta_{02} = -0.739$	-1.295 (0.783)	-0.556	0.876	0.759	0.830
		$\beta_{12} = 0.306$	0.385 (0.243)	0.079	0.340	0.732	0.772
		$\gamma_1 = 0.225$	0.233 (0.021)	0.008	0.017	0.948	0.978
		$\gamma_2 = 0.410$	0.400 (0.014)	-0.010	0.019	0.716	0.800
		$\phi = 1.400$	1.219 $(-)$	-0.181	0.836	-	-
	400	$\beta_1 = 1.098$	1.072 (0.185)	-0.026	0.463	0.471	0.543
		$\beta_{02} = -0.739$	-1.053 (0.560)	-0.314	0.645	0.787	0.857
		$\beta_{12} = 0.306$	0.370 (0.186)	0.064	0.283	0.789	0.820
		$\gamma_1 = 0.225$	0.228 (0.014)	0.003	0.012	0.962	0.980
		$\gamma_2 = 0.410$	0.406 (0.009)	-0.004	0.014	0.708	0.789
		$\phi = 1.400$	1.375 $(-)$	-0.025	0.749	-	-
Complete Likelihood	200	$\beta_1 = 1.098$	1.341 (0.717)	0.243	0.516	0.981	0.992
		$\beta_{02} = -0.739$	-1.142 (0.916)	-0.403	0.839	0.908	0.962
		$\beta_{12} = 0.306$	0.478 (0.345)	0.172	0.303	0.976	0.991
		$\gamma_1 = 0.225$	0.224 (0.018)	-0.001	0.018	0.906	0.944
		$\gamma_2 = 0.410$	0.411 (0.019)	0.001	0.016	0.936	0.966
		$\phi = 1.400$	1.903 (1.604)	0.503	1.238	0.972	0.987
	400	$\beta_1 = 1.098$	1.290 (0.479)	0.192	0.452	0.963	0.987
		$\beta_{02} = -0.739$	-0.996 (0.605)	-0.257	0.554	0.905	0.956
		$\beta_{12} = 0.306$	0.438 (0.222)	0.132	0.248	0.973	0.990
		$\gamma_1 = 0.225$	0.224 (0.013)	-0.001	0.013	0.892	0.939
		$\gamma_2 = 0.410$	0.411 (0.013)	0.001	0.013	0.910	0.960
		$\phi = 1.400$	1.812 (1.051)	0.412	1.012	0.968	0.978
Complete Likelihood	200	$\beta_1 = 1.098$	1.426 (0.664)	0.328	0.546	0.983	0.988
		$\beta_{02} = -0.739$	-1.009 (0.945)	-0.270	0.844	0.945	0.969
		$\beta_{12} = 0.306$	0.632 (0.475)	0.326	0.354	0.976	0.995
		$\gamma_1 = 0.225$	0.220 (0.017)	-0.005	0.016	0.867	0.933
		$\gamma_2 = 0.410$	0.418 (0.018)	0.008	0.016	0.916	0.964
		$\phi = 0.700$	1.085 (0.688)	0.385	0.600	0.988	0.995
	400	$\beta_1 = 1.098$	1.327 (0.413)	0.229	0.377	0.976	0.994
		$\beta_{02} = -0.739$	-0.886 (0.613)	-0.147	0.527	0.951	0.976
		$\beta_{12} = 0.306$	0.547 (0.312)	0.241	0.310	0.955	0.982
		$\gamma_1 = 0.225$	0.222 (0.012)	-0.003	0.011	0.894	0.931
		$\gamma_2 = 0.410$	0.416 (0.013)	0.006	0.011	0.913	0.961
		$\phi = 0.700$	0.969 (0.411)	0.269	0.350	0.990	0.996

3.5 Model discrimination

3.5.1 Likelihood-based method

In this simulation study, we evaluate the performance of the LRT in testing the null hypothesis that the competing cause can be modelled by one of the COM-Poisson ($\phi = 0.2$), Poisson, COM-Poisson ($\phi = 1.5$) and Bernoulli distributions against an alternative where the competing cause can be modelled by a member of the COM-Poisson family besides the one specified in the null hypothesis. Considering the parameter settings of Scenario 1, we calculated the Λ values of the fitted COM-Poisson ($\phi = 0.2$), Poisson, COM-Poisson ($\phi = 1.5$) and Bernoulli models versus the fitted general COM-Poisson model. Based on 1000 MC runs and using 5% level of significance, we calculated the observed levels (in bold) and power values of the LRT and present them in Table 3.5. Note that when the true model is either Poisson ($\phi = 1$) or COM-Poisson ($\phi = 1.5$), the observed levels are close to the true nominal level used. When the true model is COM-Poisson ($\phi = 0.2$), the observed levels are conservative. When the true model is Bernoulli ($\phi \to \infty$), as was observed in Section 2.4, the observed levels are liberal, however, they improve as sample size increases. When COM-Poisson ($\phi = 0.2$) is the true model, the likelihood ratio test has very high power to reject any of the competing models. When the true model is either Poisson or COM-Poisson ($\phi = 1.5$), the LRT only has moderate power to reject the Bernoulli, and when the true model is Bernoulli, the test has low power to reject any of the competing models. Also, as expected, note that the power to reject the wrong model increases as sample size increases.

3.5.2 Information-based criteria

The performances of AIC and BIC are investigated in selecting the correct model within the COM-Poisson family. Table 3.6 gives the AIC values (BIC values were identical to AIC values) and it is clear that AIC performs very well in selecting the correct model when the

Table 3.5 Observed levels and rejection rates of the likelihood ratio test ($\gamma_1 = 0.225$ and $\gamma_2 = 0.410$)

n	Fitted destructive COM-Poisson model	True destructive COM-Poisson model			
		$\phi = 0.2$	$\phi = 1.0$	$\phi = 1.5$	$\phi \to \infty$
200	$\phi = 0.2$	**0.015**	0.328	0.192	0.169
	$\phi = 1.0$	0.939	**0.036**	0.030	0.021
	$\phi = 1.5$	0.972	0.093	**0.038**	0.003
	$\phi \to \infty$	1.000	0.633	0.354	**0.110**
400	$\phi = 0.2$	**0.021**	0.373	0.393	0.384
	$\phi = 1.0$	0.986	**0.038**	0.054	0.122
	$\phi = 1.5$	1.000	0.118	**0.039**	0.036
	$\phi \to \infty$	1.000	0.844	0.538	**0.094**

true model is COM-Poisson ($\phi = 0.2$). When the true model is either Bernoulli, Poisson or COM-Poisson ($\phi = 1.5$), the correct selection rate is conservative, however, it improves as sample size increases.

Table 3.6 Observed selection rates based on AIC ($\gamma_1 = 0.225$ and $\gamma_2 = 0.410$)

n	Fitted destructive COM-Poisson model	True destructive COM-Poisson model			
		$\phi = 0.2$	$\phi = 1.0$	$\phi = 1.5$	$\phi \to \infty$
200	$\phi = 0.2$	0.959	0.328	0.242	0.197
	$\phi = 1.0$	0.041	0.372	0.293	0.151
	$\phi = 1.5$	0.000	0.211	0.284	0.251
	$\phi \to \infty$	0.000	0.089	0.181	0.401
400	$\phi = 0.2$	0.984	0.224	0.149	0.079
	$\phi = 1.0$	0.016	0.461	0.311	0.109
	$\phi = 1.5$	0.000	0.267	0.396	0.290
	$\phi \to \infty$	0.000	0.048	0.144	0.522

3.6 Analysis of melanoma data

Here, the estimation procedure is applied to the malignant melanoma data described before. To circumvent the identifiability issue, as discussed in Section 1.3, four scenarios are considered, as done in Section 2.5 (see also Section 5.4). The scenario with η linked to ulceration status without an intercept and p linked to tumor thickness with an intercept turned out to be the best in terms of log-likelihood, AIC and BIC values. The log-likelihood

value turned out to be -205.039, which is even larger compared to the model under Weibull lifetime (see Section 2.5). The MLE of ϕ turned out to be 0.23 (see Figure 3.1).

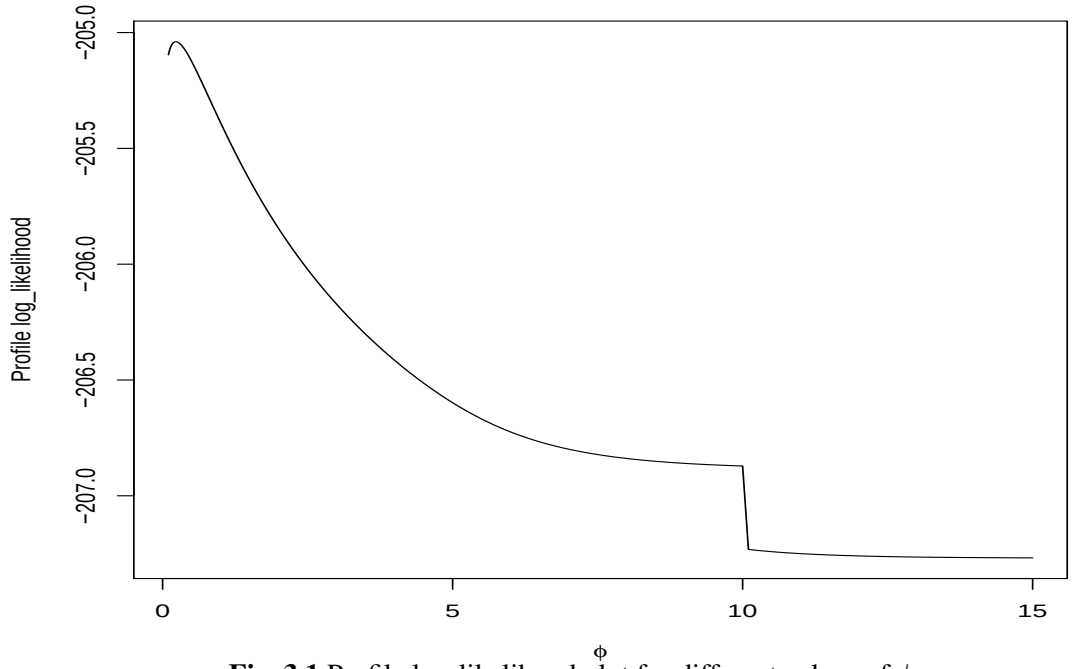

Fig. 3.1 Profile log-likelihood plot for different values of ϕ

Based on LRT, the suitability of the destructive COM-Poisson ($H_0 : \phi = 0.23$), destructive Poisson ($H_0 : \phi = 1$) and destructive Bernoulli($H_0 : \phi \to \infty$) models were tested and the p-values obtained were ≈ 1.0, 0.403 and 0.025, respectively. The destructive Bernoulli model got rejected at 5% significance level. We adopted the destructive COM-Poisson ($\phi = 0.23$) as our working model and Table 3.7 presents the MLEs, standard errors and the p- values.

Table 3.7 MLEs, standard errors (se) and p- values for the destructive COM-Poisson model using profile likelihood approach

	ϕ	β_1	β_{02}	β_{12}	γ_1	γ_2
MLE	0.23	0.320	-2.526	1.011	1.349	0.051
se	-	0.107	0.721	0.432	0.214	0.029
p- value	-	0.003	0.000	0.019	0.000	0.039

From Table 3.7, note that all the regression parameters are significant at 5% level of significance.

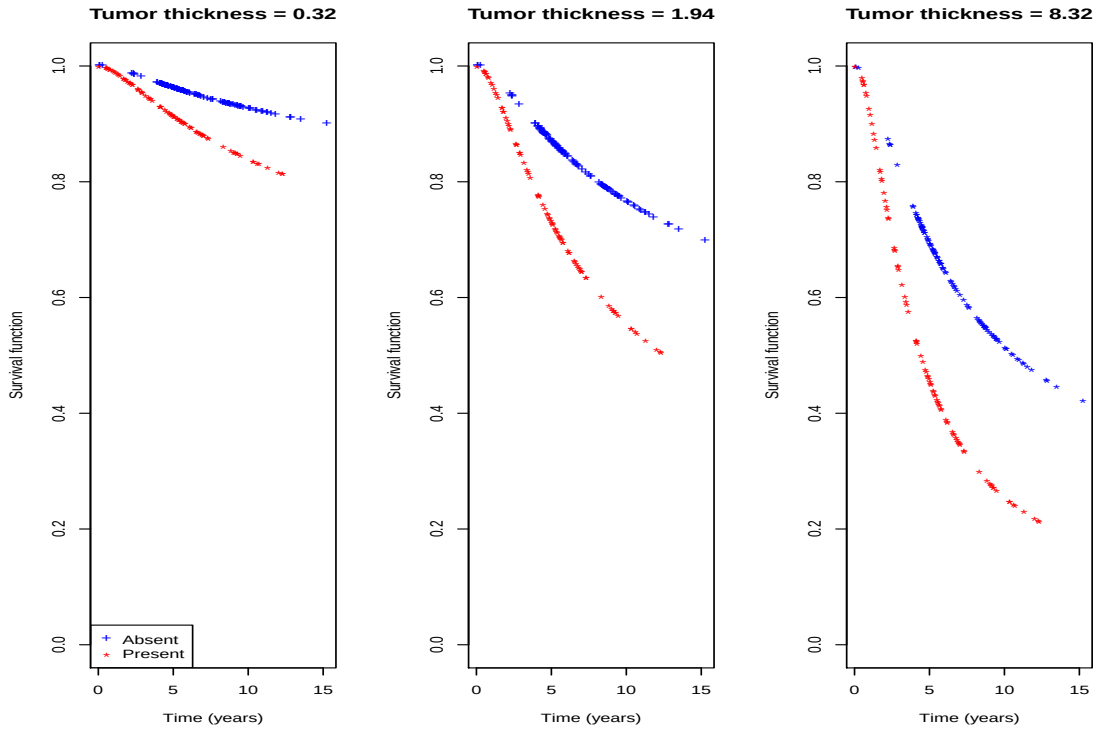

Fig. 3.2 Surviving function under destructive COM-Poisson ($\phi = 0.23$) model stratified by ulceration status (upper: absent, lower: present) for patients with different tumor thickness

3.6.1 Effect of covariates on cure rate

It is important to investigate the effect of covariates on the cure rate. Note that the cure rate depends only on the regression parameters (through the link functions for η and p). Since the estimate of $\beta_1 > 0$, it means that the presence of ulceration status results in higher mean number of competing causes and thus leading to a decrease in cure rate. Also, with $\beta_{12} > 0$, the initial competing causes will remain active with a higher probability with an increase in tumor thickness, which in turn implies a decrease in cure rate. Figure 3.2 clearly shows the effect of tumour thickness on survival. As tumour thickness increase, survival probability decreases. Also subjects without ulcer have a higher probability of being long-term survivors.

3.6.2 Profile likelihood plot for ϕ

Figure 3.3 shows the plot for different values of ϕ against the corresponding LRT (Λ) statistic. For any value of ϕ, the test is rejected at α significance level if the corresponding $\Lambda > \chi^2_{1,\,1-\alpha}$. From the figure, an acceptable range of ϕ turns out to be $(0,\ 4)$ for $\alpha = 10\%$ (when $\alpha = 5\%$, acceptable range of ϕ turns out to be $0,\ 10$), which clearly explains why the destructive Bernoulli model ($\phi \to \infty$) got rejected and not the destructive Poisson model ($\phi = 1$).

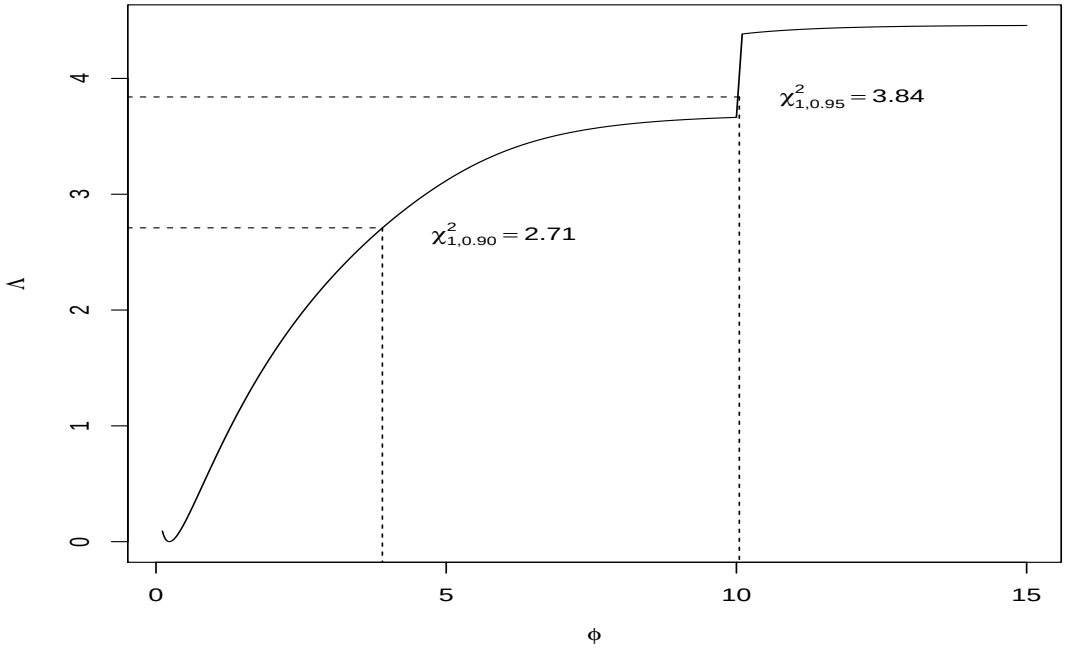

Fig. 3.3 Values of ϕ against the corresponding Λ-statistic

A particular case of the destructive COM-Poisson model, namely, the COM-Poisson model ($p = 1$) was also fitted. In this way, the destructive mechanism is absent and to preserve the same number of parameters, the parameter η was linked to both the covariates with an intercept term. In this case, the maximized log-likelihood value turned out to be -207.603 with p-value of 0.022. This provides clear evidence that a better fit to the data is realised when the destructive mechanism is taken into account.

3.7 Conclusion

The likelihood inference for destructive COM-Poisson cure rate model using lognormal as the lifetime distribution was considered. Two scenarios were considered with different censoring proportions. The profile likelihood approach performed well in estimating ϕ with small bias, though there was under-coverage for all the parameters. Generally, the bias of ϕ when using complete likelihood approach was higher compared to the profile likelihood approach, but the rest of the parameters were estimated with small bias. In this chapter, it has been shown that the lognormal distribution is a serious competitor to the Weibull distribution as it has performed better when applied to the melanoma data set. Also, it has been shown that the destructive mechanism provides a significantly better fit, and hence should not be ignored.

Chapter 4

Likelihood inference with gamma lifetime

4.1 Introduction

In this chapter, the lifetime W in (1.2) is assumed to follow a two parameter gamma distribution with pdf as given in (1.13). The gamma distribution has a constant, decreasing or increasing hazard function depending on whether $\gamma_1 < 1$, $\gamma_1 > 1$ or $\gamma_1 = 1$, respectively.

By letting $H(\gamma) = \frac{\left(\frac{\gamma_2}{\gamma_1^2}\right)^{\frac{1}{\gamma_1^2}}}{\Gamma\left(\frac{1}{\gamma_1^2}\right)}$, the expressions for $Q(\theta, \pi^{(k)})$ function are obtained as follows.

4.2 Expressions of the Q function for destructive cure rate models

Destructive Bernoulli cure rate model

Expression for the $Q(\theta,\ \pi^{(k)})$ function is

$$\begin{aligned} Q(\theta,\pi^{(k)}) &= \sum_{I_1}\left\{\log\left(\eta_i p_i H(\gamma)\right)-\frac{\gamma_2 t}{\gamma_1^2}+\left(\frac{1}{\gamma_1^2}-1\right)\log t_i-\log(1+\eta_i)\right\} \\ &\quad +\sum_{I_0}(1-\pi_i^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{\eta_i p_i\Gamma\left(\frac{1}{\gamma_1^2},\frac{\gamma_2 y}{\gamma_1^2}\right)}{(1+\eta_i)\Gamma\left(\frac{1}{\gamma_1^2}\right)}\right), \end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=\frac{\eta_i p_i S(t_i)}{1+\eta_i(1-p_i F(t_i))}\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation, and $S(t_i)$ is as given in (1.14).

Destructive Poisson cure rate model

The $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned} Q(\theta,\pi^{(k)}) &= \sum_{I_1}\left\{\log\left(\eta_i p_i H(\gamma)\right)-\frac{\gamma_2 t}{\gamma_1^2}+\left(\frac{1}{\gamma_1^2}-1\right)\log t_i-\eta_i p_i F(t_i)\right\} \\ &\quad -\sum_{I_0}(1-\pi_i^{(k)})\eta_i p_i+\sum_{I_0}\pi_i^{(k)}\log\left[\exp(-\eta_i p_i F(t_i))-\exp(-\eta_i p_i)\right], \end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=1-\exp(-\eta_i p_i S(t_i))\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation, and $F(t)=1-S(t_i)$.

Destructive COM-Poisson cure rate model

The $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned}\mathbf{Q}(\theta,\pi^{(k)}) &= \sum_{I_1}\left\{\log\left(\frac{p_iH(\gamma)}{z(1-p_iF(t_i))}\right)-\frac{\gamma_2 t}{\gamma_1^2}+\left(\frac{1}{\gamma_1^2}-1\right)\log t_i+\log z_{f1}\right\}\\ &\quad -\sum_{I_1}\log\{z(1-p_iF(t_i))\}+\sum_{I_0}(1-\pi_i^{(k)})\,(\log z_p-\log z)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{z_f-z_p}{z}\right),\end{aligned}$$

for $i=1,2,\ldots,n$ and $\pi_i^{(k)}=\frac{z_f-z_p}{z-z_p}|_{\theta=\theta^{(k)}}$ for the i-th censored observation, where $z_f=\sum_{j=0}^{\infty}\frac{[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}$, $z_p=\sum_{j=0}^{\infty}\frac{[\eta_i(1-p_i)]^j}{(j!)^\phi}$, $z=\sum_{j=0}^{\infty}\frac{\eta_i^j}{(j!)^\phi}$ and $z_{f1}=\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}$.

The first-order and second-order derivatives of the $Q(\theta,\pi^{(k)})$ function with respect to θ are all calculated numerically. The first and the negative second order derivatives of the observed log-likelihood function with respect to θ gives the components of the score function and observed information matrix, respectively. Note that the components of the score function can also be obtained within the EM framework and using the method of Louis (see Louis, 1982) as follows:

$$\frac{\partial}{\partial\theta}l(\theta)=\frac{\partial}{\partial\theta}\log L(\theta)=\frac{\partial}{\partial\theta}l_c(\theta)\Big|_{I_i=\pi_i\,\forall i},$$

where

$$\pi_i=\frac{(1-p_{0i})S_1(t_i)}{S_{\text{pop}}(t_i)}$$

if the i-th observation is censored and $\pi_i=1$ if the i-th observation is uncensored. The R code for calculating the MLEs and standard errors is given in Appendix C.

4.3 Simulation study

Two varying sample sizes of $n=200$ and $n=500$ were considered. The regression parameter values were obtained in the same way as in Section 2.3 except for β_1, which was calculated as 0.693 after assuming a value of η as 2.0.

The choice of the lifetime parameters γ_1 and γ_2 were obtained after equating the theoretical mean and variance of the considered gamma distribution to some fixed values. For this, we

chose two different values of variance as 0.06 and 0.12 and two different values of mean as 0.2 and 0.4, which gave us four different choices of (γ_1, γ_2) as $(0.612, 2.5)$, $(0.866, 2.5)$, $(1.225, 5)$ and $(1.732, 5)$. These choices of (γ_1, γ_2) enables us to study the behaviour of the model when the lifetime distribution has the same variance but different mean and vice-versa.

Tables 4.1 and 4.2 present the results for the destructive Bernoulli and destructive Poisson cure rate models. It is clear that the estimates obtained from the EM algorithm are close to the true parameter values. Both the standard error and RMSE decreases as sample size increases. The coverage probabilities are also close to the nominal levels used. Note that when γ_1 increases, no significant change is noticeable in the bias, standard errors, RMSE and coverage probability of the parameters. But, when both γ_1 and γ_2 increases, there is a significant increase in bias, standard error and RMSE for γ_2. In Table 4.3, we present the results for the general destructive COM-Poisson model with true value of ϕ as 0.5. The complete likelihood approach works very well in estimating ϕ with small bias. Furthermore, the bias, standard error and RMSE of ϕ decreases as sample size increases. There is slight over-coverage for the regression parameters and ϕ, but, the coverage probability improves as sample size increases. Table 4.4 present the estimation results when using both profile likelihood approach and complete likelihood approach. Note that the profile likelihood approach works very well in estimating parameters with low bias and standard error. The estimate of ϕ is close to the true value and gets better as sample size increases. The profile likelihood approach leads to under-coverage for all the parameters, and is more pronounced for β_1, β_{12} and γ_1. The complete likelihood approach gives better estimates as compared to the profile approach as the biases are smaller and the coverage probabilities are close to the nominal levels used. Table 4.6 presents the complete likelihood approach results for different values of the lifetime parameters with $\phi = 1.5$ (under-dispersed data). The approach works very well in recovering the true values. Once again, an increase in sample size leads to a reduction in the bias, standard error and RMSE. The coverage probabilities for both regression and lifetime parameters are close to the nominal levels.

From Tables 4.1 - 4.6, note that when only γ_1 increases, there is no difference in the precision

and accuracy of the estimates, however, when both γ_1 and γ_2 increases, a significant increase is noticeable in the bias, standard error and RMSE of ϕ and γ_2. When the parameter γ_2 increases (see Tables 4.5 and 4.6), the precision and accuracy of the parameters are lowered.

Table 4.1 Accuracy and precision of estimates for the destructive Poisson cure rate model

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.612, 2.5)	β_1	0.728 (0.175)	0.035	0.175	0.891	0.938
		β_{02}	-1.545 (0.593)	-0.124	0.583	0.935	0.977
		β_{12}	0.387 (0.168)	0.042	0.219	0.874	0.911
		γ_1	0.609 (0.036)	-0.003	0.036	0.887	0.941
		γ_2	2.513 (0.179)	0.013	0.181	0.897	0.939
500	(0.612, 2.5)	β_1	0.703 (0.109)	0.010	0.107	0.914	0.963
		β_{02}	-1.466 (0.361)	-0.045	0.365	0.902	0.955
		β_{12}	0.363 (0.098)	0.018	0.108	0.890	0.929
		γ_1	0.612 (0.023)	0.000	0.022	0.902	0.949
		γ_2	2.495 (0.113)	-0.005	0.105	0.913	0.964
200	(0.866, 2.5)	β_1	0.712 (0.173)	0.019	0.175	0.887	0.943
		β_{02}	-1.594 (0.636)	-0.173	0.743	0.913	0.970
		β_{12}	0.429 (0.201)	0.084	0.356	0.874	0.911
		γ_1	0.864 (0.047)	-0.002	0.047	0.905	0.962
		γ_2	2.517 (0.270)	0.017	0.276	0.898	0.953
500	(0.866, 2.5)	β_1	0.703 (0.108)	0.010	0.113	0.891	0.939
		β_{02}	-1.479 (0.361)	-0.058	0.359	0.902	0.952
		β_{12}	0.365 (0.098)	0.020	0.103	0.887	0.926
		γ_1	0.861 (0.030)	-0.005	0.030	0.899	0.948
		γ_2	2.511 (0.169)	0.011	0.168	0.899	0.961
200	(1.225, 5.0)	β_1	0.714 (0.168)	0.021	0.171	0.889	0.945
		β_{02}	-1.578 (0.601)	-0.157	0.649	0.910	0.964
		β_{12}	0.413 (0.176)	0.068	0.236	0.882	0.920
		γ_1	1.217 (0.060)	-0.008	0.059	0.901	0.939
		γ_2	5.075 (0.766)	0.075	0.766	0.895	0.954
500	(1.225, 5.0)	β_1	0.701 (0.106)	0.008	0.107	0.902	0.951
		β_{02}	-1.497 (0.353)	-0.076	0.357	0.926	0.966
		β_{12}	0.369 (0.096)	0.024	0.111	0.896	0.940
		γ_1	1.223 (0.038)	-0.002	0.039	0.890	0.948
		γ_2	5.036 (0.483)	0.036	0.486	0.898	0.953
200	(1.732, 5.0)	β_1	0.717 (0.169)	0.024	0.178	0.888	0.937
		β_{02}	-1.597 (0.618)	-0.176	0.825	0.902	0.960
		β_{12}	0.422 (0.194)	0.077	0.520	0.862	0.915
		γ_1	1.725 (0.079)	-0.007	0.080	0.893	0.940
		γ_2	5.235 (1.230)	0.235	1.326	0.882	0.939
500	(1.732, 5.0)	β_1	0.704 (0.106)	0.011	0.105	0.897	0.950
		β_{02}	-1.482 (0.350)	-0.061	0.358	0.908	0.956
		β_{12}	0.364 (0.094)	0.019	0.100	0.893	0.939
		γ_1	1.726 (0.050)	-0.006	0.048	0.909	0.956
		γ_2	5.119 (0.760)	0.119	0.768	0.902	0.951

Table 4.2 Accuracy and precision of estimates for the destructive Bernoulli cure rate model

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.612, 2.5)	β_1	0.785 (0.513)	0.092	0.495	0.934	0.959
		β_{02}	-1.606 (0.751)	-0.185	0.827	0.920	0.967
		β_{12}	0.400 (0.182)	0.055	0.221	0.895	0.935
		γ_1	0.609 (0.047)	-0.003	0.046	0.887	0.947
		γ_2	2.521 (0.183)	0.021	0.185	0.892	0.944
500	(0.612, 2.5)	β_1	0.732 (0.279)	0.039	0.290	0.925	0.962
		β_{02}	-1.500 (0.443)	-0.079	0.478	0.906	0.949
		β_{12}	0.372 (0.104)	0.027	0.115	0.879	0.923
		γ_1	0.610 (0.029)	-0.002	0.029	0.899	0.948
		γ_2	2.508 (0.115)	0.008	0.109	0.918	0.967
200	(0.866, 2.5)	β_1	0.798 (0.529)	0.105	0.523	0.930	0.966
		β_{02}	-1.694 (0.769)	-0.273	0.882	0.921	0.964
		β_{12}	0.423 (0.191)	0.078	0.272	0.897	0.928
		γ_1	0.862 (0.063)	-0.004	0.064	0.896	0.936
		γ_2	2.541 (0.269)	0.041	0.259	0.915	0.963
500	(0.866, 2.5)	β_1	0.744 (0.282)	0.051	0.284	0.924	0.967
		β_{02}	-1.495 (0.440)	-0.074	0.469	0.892	0.949
		β_{12}	0.364 (0.102)	0.019	0.112	0.883	0.930
		γ_1	0.862 (0.040)	-0.004	0.039	0.910	0.953
		γ_2	2.506 (0.168)	0.006	0.168	0.894	0.943
200	(1.225, 5.0)	β_1	0.819 (0.502)	0.126	0.489	0.928	0.958
		β_{02}	-1.647 (0.760)	-0.226	1.044	0.913	0.960
		β_{12}	0.410 (0.188)	0.065	0.308	0.895	0.929
		γ_1	1.220 (0.082)	-0.005	0.080	0.900	0.944
		γ_2	5.098 (0.744)	0.098	0.753	0.908	0.948
500	(1.225, 5.0)	β_1	0.735 (0.274)	0.042	0.292	0.919	0.949
		β_{02}	-1.494 (0.430)	-0.073	0.454	0.900	0.952
		β_{12}	0.368 (0.101)	0.023	0.119	0.891	0.934
		γ_1	1.221 (0.052)	-0.004	0.050	0.913	0.957
		γ_2	5.055 (0.467)	0.055	0.475	0.896	0.953
200	(1.732, 5.0)	β_1	0.803 (0.527)	0.110	0.522	0.926	0.960
		β_{02}	-1.701 (0.769)	-0.280	1.028	0.921	0.970
		β_{12}	0.427 (0.193)	0.082	0.315	0.895	0.935
		γ_1	1.722 (0.109)	-0.010	0.105	0.914	0.951
		γ_2	5.231 (1.150)	0.231	1.193	0.894	0.956
500	(1.732, 5.0)	β_1	0.754 (0.281)	0.061	0.301	0.909	0.955
		β_{02}	-1.497 (0.429)	-0.076	0.451	0.900	0.955
		β_{12}	0.364 (0.100)	0.019	0.117	0.878	0.926
		γ_1	1.730 (0.069)	-0.002	0.067	0.910	0.958
		γ_2	5.086 (0.711)	0.086	0.737	0.890	0.945

Table 4.3 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 0.5$ using complete likelihood approach

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.612, 2.5)	β_1	0.834 (0.342)	0.141	0.257	0.985	0.997
		β_{02}	-1.555 (0.632)	-0.134	0.582	0.951	0.979
		β_{12}	0.460 (0.233)	0.115	0.267	0.976	0.989
		γ_1	0.611 (0.032)	-0.001	0.033	0.887	0.942
		γ_2	2.624 (0.274)	0.124	0.239	0.931	0.969
		ϕ	0.660 (0.364)	0.160	0.241	0.997	1.000
500	(0.612, 2.5)	β_1	0.772 (0.196)	0.079	0.141	0.979	0.996
		β_{02}	-1.446 (0.364)	-0.025	0.335	0.937	0.976
		β_{12}	0.391 (0.126)	0.046	0.122	0.963	0.987
		γ_1	0.612 (0.020)	0.000	0.021	0.892	0.938
		γ_2	2.558 (0.166)	0.058	0.137	0.930	0.968
		ϕ	0.584 (0.198)	0.084	0.119	0.994	0.999
200	(0.866 2.5)	β_1	0.804 (0.334)	0.111	0.326	0.958	0.983
		β_{02}	-1.576 (0.644)	-0.155	0.615	0.943	0.979
		β_{12}	0.481 (0.251)	0.136	0.361	0.972	0.994
		γ_1	0.866 (0.042)	0.000	0.042	0.900	0.956
		γ_2	2.590 (0.398)	0.090	0.385	0.929	0.960
		ϕ	0.637 (0.357)	0.137	0.373	0.983	0.987
500	(0.866 2.5)	β_1	0.716 (0.183)	0.023	0.168	0.933	0.971
		β_{02}	-1.491 (0.357)	-0.070	0.336	0.911	0.961
		β_{12}	0.383 (0.121)	0.038	0.123	0.935	0.978
		γ_1	0.863 (0.026)	-0.003	0.027	0.888	0.940
		γ_2	2.521 (0.241)	0.021	0.239	0.905	0.956
		ϕ	0.528 (0.179)	0.028	0.165	0.941	0.975
200	(1.225, 5.0)	β_1	0.936 (0.351)	0.243	0.292	0.961	0.997
		β_{02}	-1.558 (0.672)	-0.137	0.658	0.964	0.992
		β_{12}	0.570 (0.298)	0.225	0.434	0.966	0.989
		γ_1	1.241 (0.058)	0.016	0.056	0.899	0.964
		γ_2	5.383 (1.134)	0.383	1.080	0.933	0.966
		ϕ	0.759 (0.372)	0.259	0.266	0.997	1.000
500	(1.225, 5.0)	β_1	0.866 (0.198)	0.173	0.164	0.905	0.954
		β_{02}	-1.422 (0.373)	-0.001	0.353	0.939	0.975
		β_{12}	0.444 (0.140)	0.099	0.151	0.933	0.975
		γ_1	1.241 (0.036)	0.016	0.034	0.899	0.951
		γ_2	5.243 (0.694)	0.243	0.732	0.887	0.939
		ϕ	0.683 (0.205)	0.183	0.153	0.942	0.997
200	(1.732 5.0)	β_1	0.949 (0.351)	0.256	0.298	0.973	0.991
		β_{02}	-1.540 (0.706)	-0.119	0.693	0.944	0.976
		β_{12}	0.596 (0.347)	0.251	0.651	0.956	0.982
		γ_1	1.764 (0.082)	0.032	0.077	0.909	0.947
		γ_2	5.587 (1.744)	0.587	1.955	0.888	0.956
		ϕ	0.790 (0.382)	0.290	0.311	0.988	1.000
500	(1.732 5.0)	β_1	0.841 (0.191)	0.148	0.158	0.926	0.976
		β_{02}	-1.465 (0.368)	-0.044	0.370	0.897	0.944
		β_{12}	0.440 (0.140)	0.095	0.132	0.934	0.981
		γ_1	1.756 (0.050)	0.024	0.046	0.910	0.952
		γ_2	5.027 (0.995)	0.027	0.980	0.894	0.955
		ϕ	0.651 (0.194)	0.151	0.154	0.947	0.989

Table 4.4 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model using both the profile and complete likelihood approaches

Approach	n	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
Profile Likelihood	200	$\beta_1 = 0.693$	0.579 (0.197)	-0.114	-16	0.300	0.623	0.695
		$\beta_{02} = -1.421$	-1.821 (0.667)	-0.400	28	0.761	0.867	0.922
		$\beta_{12} = 0.345$	0.399 (0.206)	0.054	16	0.427	0.784	0.826
		$\gamma_1 = 0.612$	0.606 (0.039)	-0.006	-1	0.040	0.888	0.937
		$\gamma_2 = 2.500$	2.445 (0.177)	-0.055	-2	0.228	0.758	0.839
		$\phi = 1.500$	1.288 $(-)$	-0.212	-14	0.725	-	-
	500	$\beta_1 = 0.693$	0.649 (0.134)	-0.044	-6	0.208	0.688	0.768
		$\beta_{02} = -1.421$	-1.549 (0.374)	-0.128	9	0.414	0.848	0.921
		$\beta_{12} = 0.345$	0.357 (0.095)	0.012	3	0.122	0.791	0.851
		$\gamma_1 = 0.612$	0.608 (0.024)	-0.004	-0.7	0.024	0.894	0.935
		$\gamma_2 = 2.500$	2.474 (0.111)	-0.026	-1	0.153	0.756	0.838
		$\phi = 1.500$	1.432 $(-)$	-0.068	-5	0.581	-	-
Complete Likelihood	200	$\beta_1 = 0.693$	0.658 (0.403)	-0.035	-5	0.340	0.878	0.904
		$\beta_{02} = -1.421$	-1.734 (0.770)	-0.313	22	0.632	0.923	0.974
		$\beta_{12} = 0.345$	0.403 (0.201)	0.058	17	0.240	0.914	0.945
		$\gamma_1 = 0.612$	0.607 (0.040)	-0.005	-0.8	0.039	0.904	0.949
		$\gamma_2 = 2.500$	2.469 (0.266)	-0.031	-1	0.232	0.913	0.950
		$\phi = 1.500$	1.583 (1.348)	0.083	6	0.984	0.860	0.889
	500	$\beta_1 = 0.693$	0.704 (0.257)	0.011	2	0.241	0.909	0.946
		$\beta_{02} = -1.421$	-1.515 (0.433)	-0.094	7	0.425	0.897	0.960
		$\beta_{12} = 0.345$	0.364 (0.117)	0.019	6	0.124	0.895	0.946
		$\gamma_1 = 0.612$	0.609 (0.025)	-0.003	-0.5	0.025	0.902	0.936
		$\gamma_2 = 2.500$	2.495 (0.172)	-0.005	-0.2	0.162	0.910	0.963
		$\phi = 1.500$	1.495 (0.875)	-0.005	-0.3	0.767	0.883	0.917

Table 4.5 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model using complete likelihood approach

Approach	n	Parameters	MLEs (se)	Bias		RMSE	CP	
				Real	%		90%	95%
Seting 1	200	$\beta_1 = 0.693$	0.667 (0.398)	-0.026	-4	0.374	0.883	0.918
		$\beta_{02} = -1.421$	−1.706 (0.722)	-0.285	20	0.619	0.923	0.971
		$\beta_{12} = 0.345$	0.400 (0.195)	0.055	16	0.221	0.923	0.960
		$\gamma_1 = 0.612$	0.619 (0.039)	0.007	1	0.033	0.947	0.976
		$\gamma_2 = 5.000$	4.839 (0.503)	-0.161	-3	0.367	0.926	0.955
		$\phi = 1.500$	1.485 (1.299)	-0.015	-1	1.023	0.871	0.905
	500	$\beta_1 = 0.693$	0.678 (0.236)	-0.015	-2	0.223	0.900	0.931
		$\beta_{02} = -1.421$	−1.544 (0.414)	-0.123	9	0.407	0.893	0.950
		$\beta_{12} = 0.345$	0.363 (0.114)	0.018	5	0.120	0.893	0.944
		$\gamma_1 = 0.612$	0.615 (0.025)	0.003	0.5	0.022	0.940	0.974
		$\gamma_2 = 5.000$	4.911 (0.321)	-0.089	-2	0.261	0.927	0.960
		$\phi = 1.500$	1.488 (0.757)	-0.012	-0.8	0.701	0.860	0.902
Setting 2	200	$\beta_1 = 0.693$	0.706 (0.399)	0.013	2	0.353	0.917	0.941
		$\beta_{02} = -1.421$	−1.778 (0.781)	-0.357	25	0.944	0.914	0.963
		$\beta_{12} = 0.345$	0.483 (0.270)	0.138	40	0.850	0.931	0.965
		$\gamma_1 = 0.866$	0.863 (0.054)	-0.003	-0.3	0.051	0.913	0.960
		$\gamma_2 = 5.000$	5.029 (0.734)	0.029	0.6	0.652	0.943	0.975
		$\phi = 1.500$	1.676 (1.470)	0.176	12	1.056	0.898	0.936
	500	$\beta_1 = 0.693$	0.709 (0.248)	0.016	2	0.242	0.905	0.938
		$\beta_{02} = -1.421$	−1.505 (0.413)	-0.084	6	0.403	0.915	0.949
		$\beta_{12} = 0.345$	0.361 (0.113)	0.016	5	0.111	0.913	0.954
		$\gamma_1 = 0.866$	0.862 (0.034)	-0.004	-0.5	0.034	0.904	0.951
		$\gamma_2 = 5.000$	5.007 (0.466)	0.007	0.1	0.431	0.922	0.965
		$\phi = 1.500$	1.607 (0.869)	0.107	7	0.811	0.909	0.938

Table 4.6 Accuracy and precision of estimates for the destructive COM-Poisson cure rate model with $\phi = 1.5$ using complete likelihood approach

n	(γ_1, γ_2)	Parameters	MLEs (se)	Bias	RMSE	CP	
						90%	95%
200	(0.866, 2.5)	β_1	0.690 (0.418)	-0.003	0.376	0.905	0.928
		β_{02}	-1.788 (0.790)	-0.367	0.785	0.918	0.966
		β_{12}	0.446 (0.234)	0.101	0.415	0.904	0.960
		γ_1	0.862 (0.055)	-0.004	0.051	0.918	0.959
		γ_2	2.491 (0.390)	-0.009	0.354	0.916	0.961
		ϕ	1.607 (1.477)	0.107	1.093	0.869	0.902
500	(0.866, 2.5)	β_1	0.703 (0.256)	0.010	0.257	0.895	0.922
		β_{02}	-1.525 (0.431)	-0.104	0.411	0.912	0.951
		β_{12}	0.360 (0.115)	0.015	0.118	0.901	0.949
		γ_1	0.862 (0.035)	-0.004	0.034	0.910	0.951
		γ_2	2.496 (0.248)	-0.004	0.239	0.912	0.962
		ϕ	1.598 (0.920)	0.098	0.888	0.895	0.931
200	(1.225, 5.0)	β_1	0.726 (0.414)	0.033	0.368	0.900	0.937
		β_{02}	-1.699 (0.750)	-0.278	0.711	0.928	0.971
		β_{12}	0.439 (0.222)	0.094	0.312	0.923	0.961
		γ_1	1.217 (0.075)	-0.008	0.069	0.916	0.960
		γ_2	5.100 (1.057)	0.100	0.978	0.913	0.956
		ϕ	1.778 (1.662)	0.278	1.159	0.900	0.937
500	(1.225, 5.0)	β_1	0.712 (0.243)	0.019	0.231	0.907	0.955
		β_{02}	-1.528 (0.416)	-0.107	0.404	0.907	0.961
		β_{12}	0.377 (0.119)	0.032	0.129	0.912	0.954
		γ_1	1.224 (0.047)	-0.001	0.046	0.915	0.952
		γ_2	5.038 (0.659)	0.038	0.647	0.899	0.955
		ϕ	1.656 (0.861)	0.156	0.801	0.912	0.946
200	(1.732, 5.0)	β_1	0.727 (0.407)	0.034	0.383	0.919	0.946
		β_{02}	-1.720 (0.735)	-0.299	0.763	0.918	0.960
		β_{12}	0.451 (0.220)	0.106	0.394	0.924	0.968
		γ_1	1.717 (0.104)	-0.015	0.095	0.919	0.961
		γ_2	5.173 (1.509)	0.173	1.428	0.908	0.953
		ϕ	1.742 (1.569)	0.242	1.158	0.912	0.940
500	(1.732, 5.0)	β_1	0.701 (0.239)	0.008	0.231	0.904	0.948
		β_{02}	-1.512 (0.409)	-0.091	0.382	0.925	0.962
		β_{12}	0.369 (0.114)	0.024	0.125	0.897	0.952
		γ_1	1.728 (0.066)	-0.004	0.063	0.919	0.955
		γ_2	5.173 (0.962)	0.173	0.957	0.913	0.957
		ϕ	1.619 (0.812)	0.119	0.771	0.905	0.942

4.4 Likelihood-based model discrimination

The performance of the LRT is evaluated in testing the null hypothesis that the competing cause can be modelled by one of the Bernoulli, Poisson, COM-Poisson ($\phi = 0.3$) and COM-Poisson ($\phi = 2.0$) distributions against the alternative that the competing cause can be modelled by a member of the COM-Poisson family besides the one stated in the null hypothesis. The Λ values are calculated for each of the fitted particular model versus the fitted general COM-Poisson model. Based on 1000 MC runs and using 5% level of significance, we calculate the observed levels (in bold) and power values of the LRT and present them in Table 4.7. When the true model is either Poisson ($\phi = 1.0$) or COM-Poisson ($\phi = 2.0$), the observed levels are close to the true nominal level used, provided the sample size is large. When the true model is Bernoulli ($\phi \to \infty$) and COM-Poisson ($\phi = 0.3$), the observed levels go beyond the nominal level. When COM-Poisson ($\phi = 0.3$) is the true model, LRT test has very high power to reject the COM-Poisson ($\phi = 2.0$) and Bernoulli models, but, it has moderate power to reject the Poisson model. When the true model is Poisson, the LRT has high power to reject the Bernoulli. When the true model is COM-Poisson ($\phi = 2.0$), the test has moderate power to reject COM-Poisson ($\phi = 0.3$) and Bernoulli, while if Bernoulli is the true model, it has moderate power to reject COM-Poisson ($\phi = 0.3$). Also, as expected, note that when the sample size increases the power to reject the wrong model increases as well.

Table 4.7 Observed levels and rejection rates of the likelihood ratio test ($\gamma_1 = 0.612$ and $\gamma_2 = 2.5$)

n	Fitted destructive COM-Poisson model	True destructive COM-Poisson model			
		$\phi = 0.3$	$\phi = 1.0$	$\phi = 2.0$	$\phi \to \infty$
200	$\phi = 0.3$	**0.181**	0.256	0.369	0.397
	$\phi = 1.0$	0.554	**0.027**	0.036	0.131
	$\phi = 2.0$	0.792	0.184	**0.034**	0.010
	$\phi \to \infty$	0.914	0.812	0.510	**0.076**
500	$\phi = 0.3$	**0.089**	0.576	0.685	0.771
	$\phi = 1.0$	0.609	**0.054**	0.065	0.441
	$\phi = 2.0$	0.900	0.471	**0.065**	0.103
	$\phi \to \infty$	0.910	0.979	0.798	**0.075**

4.5 Information-based model discrimination

In Table 4.8, we present the model discrimination results based on AIC (BIC values were identical to the AIC values). It is clear that AIC performs very well in selecting the correct model when the true model is COM-Poisson ($\phi = 0.3$) and the performance of AIC is moderate when the true model is either Poisson ($\phi = 1.0$) or Bernoulli. When the true model is COM-Poisson ($\phi = 2.0$), the correct selection rate is very conservative, however, it improves as sample size increases.

Table 4.8 Observed selection rates based on AIC ($\gamma_1 = 0.612$ and $\gamma_2 = 2.5$)

n	Fitted destructive COM-Poisson model	True destructive COM-Poisson model			
		$\phi = 0.3$	$\phi = 1.0$	$\phi = 2.0$	$\phi \to \infty$
200	$\phi = 0.3$	0.705	0.247	0.130	0.048
	$\phi = 1.0$	0.100	0.546	0.458	0.173
	$\phi = 2.0$	0.193	0.170	0.285	0.269
	$\phi \to \infty$	0.003	0.037	0.126	0.510
500	$\phi = 0.3$	0.782	0.151	0.040	0.009
	$\phi = 1.0$	0.123	0.702	0.534	0.082
	$\phi = 2.0$	0.086	0.142	0.387	0.282
	$\phi \to \infty$	0.009	0.005	0.038	0.627

4.6 Analysis of malignant melanoma data

In this section, the estimation procedure is applied to the melanoma data discussed in Section 1.8. The MLE of ϕ turned out to be 0.86 (see Figure 4.1) with the maximised log-likelihood value of -206.311.

The suitability of the destructive COM-Poisson model ($H_0 : \phi = 0.86$) was tested based on the LRT. The LRT statistic turned out to be ≈ 0 and hence the p-value turned out to be ≈ 1, which gives overwhelming evidence to conclude that the destructive COM-Poisson ($\phi = 0.86$) model should be the working model. The suitability of the destructive Bernoulli model ($H_0 : \phi \to \infty$) and destructive Poisson model ($\phi = 1$) were also tested, and in this case the LRT statistic turned out to be 4.460 and 0.700, respectively, with the corresponding

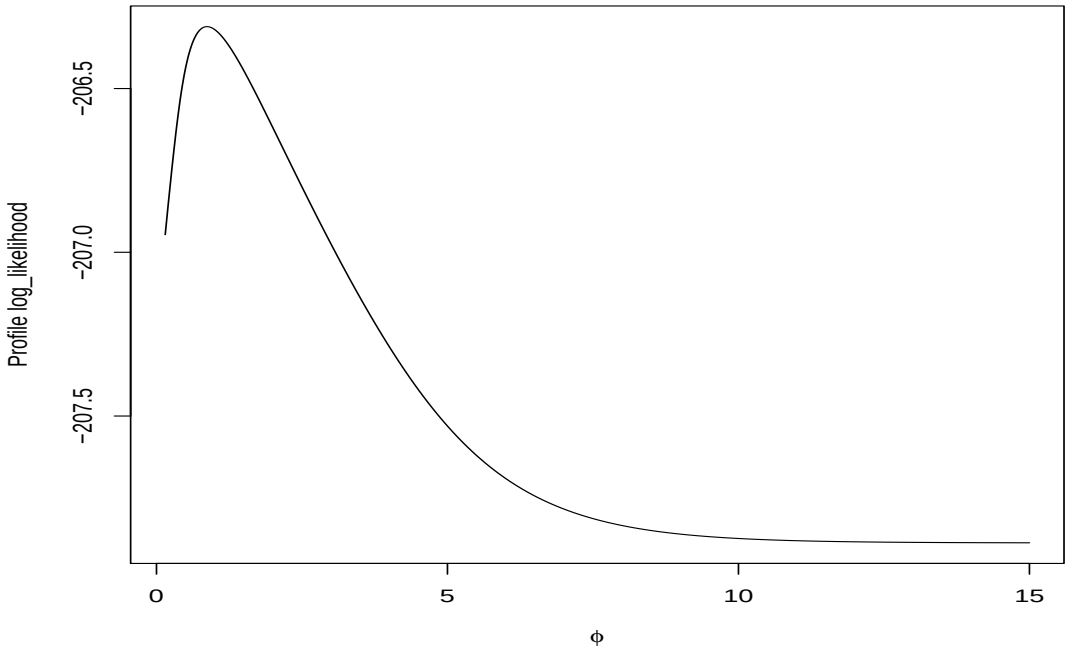

Fig. 4.1 Profile log-likelihood plot for different values of ϕ

p-values of 0.025 and 0.403, respectively. It is clear that the assumption of destructive Bernoulli model gets rejected at 5% level of significance while the assumption of destructive Poisson model fails to get rejected. The reason why the destructive Poisson model did not get rejected is because the value of $\phi = 0.86$ is not signifcantly different from $\phi = 1$, while for the destructive Bernoulli model, $\phi \to \infty$ is far away from $\phi = 0.86$. Table 4.9 presents the MLEs, standard errors and the p- values for the destructive COM-Poisson ($\phi = 0.86$) model. The standard error of ϕ is large due to the flatness of the log-likelihood surface (see Figure 4.1) and this caused the standard error of β_1 to be large as well. When the profile likelihood approach is utilised, all the parameters are significant in the model at 5% significance level.

Table 4.9 MLEs, standard errors (se) and p- values for the destructive COM-Poisson model using both the profile and complete likelihood approaches

Approach		ϕ	β_1	β_{02}	β_{12}	γ_1	γ_2
Profile	MLE	0.86	0.559	-2.552	0.980	0.727	0.143
Likelihood	se	-	0.248	0.657	0.393	0.083	0.049
	p- value	-	0.024	0.000	0.013	0.000	0.002
Complete	MLE	0.863	0.558	-2.555	0.978	0.728	0.142
Likelihood	se	0.922	0.442	0.722	0.433	0.087	0.066
	p- value	0.175	0.207	0.000	0.024	0.000	0.016

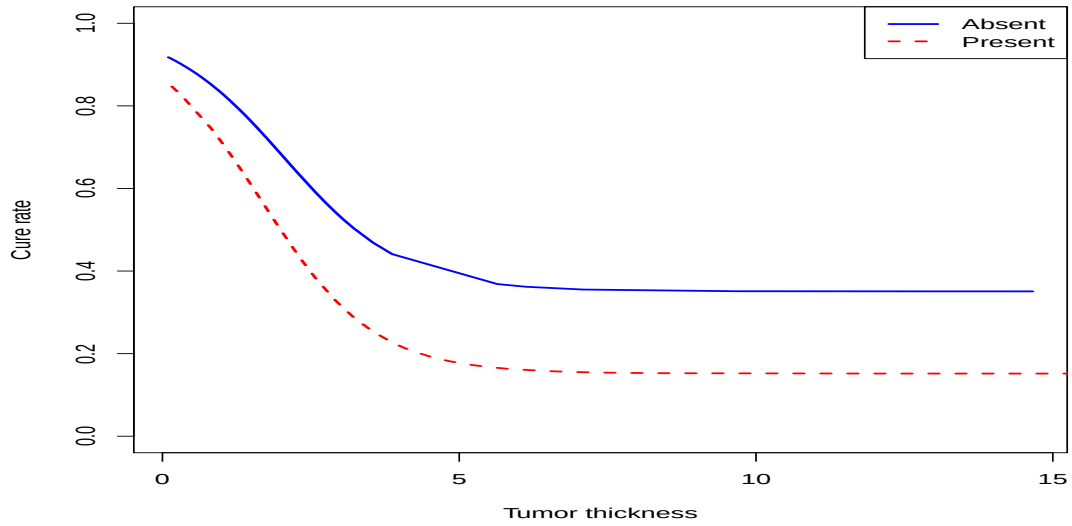

Fig. 4.2 Cure rate versus tumour thickness stratified by ulceration status

The estimates of β_1 and β_{12} being positive, Figure 4.2 shows the impact these estimates have on the cure rate. Initially, the cure rate decrease rapidly with increase in tumour thickness and then levels off for tumour thickness greater than 5 mm. The cure rate for patients without ulcer is higher compared to those with ulcers. Therefore, the presence of ulcer significantly reduces the chance of being a long-term survivor.

The COM-Poisson model was also fitted, a special case of the destructive COM-Poisson model. That is, we tested for $H_0 : p = 1$ versus $H_1 : p < 1$. The parameter η was linked to

both covariates with an intercept term in order to preserve the same number of parameters. In this case, the maximized log-likelihood value turned out to be -207.603 with a p-value of 0.022, providing clear evidence of the fact that a better fit to the data is realised when the destructive mechanism is taken into account.

Figure 4.3 shows the plot for different values of ϕ against the corresponding Λ statistic. For any value of ϕ, the test is rejected at $\alpha = 10\%$ significance level if the corresponding value of Λ statistic is greater than $\chi^2_{1,\,0.90} = 2.71$. From the plot, an acceptable range of ϕ turns out to be $(0,\ 5.8)$. Since the range include $\phi = 1$, this further suggests why the destructive Poisson model failed to get rejected.

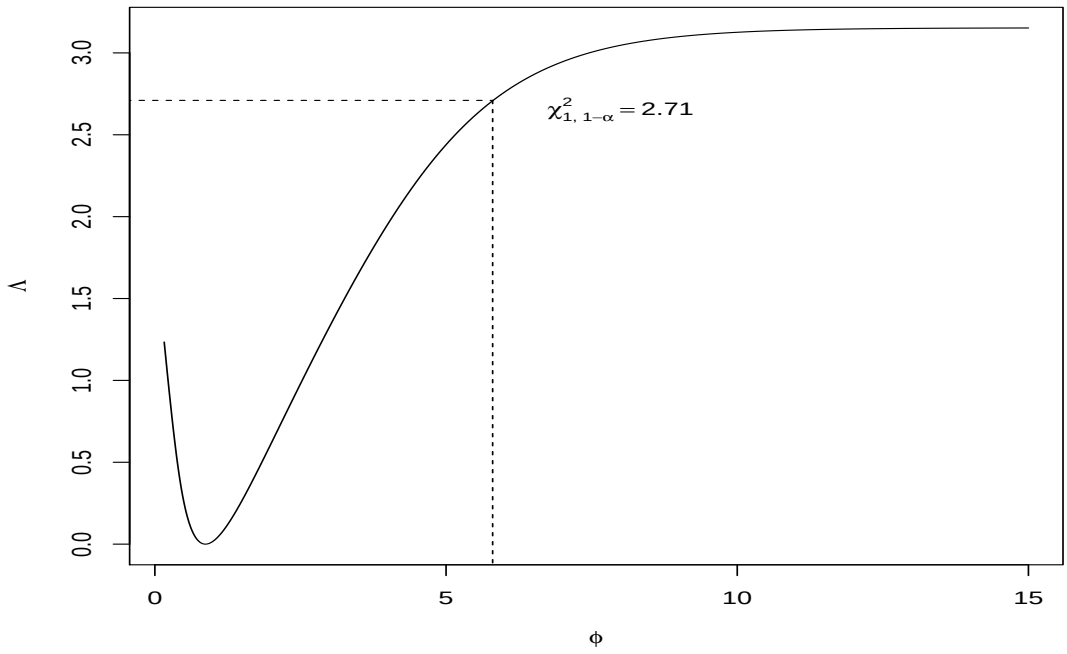

Fig. 4.3 Values of ϕ against the corresponding Λ-statistic

4.7 Conclusion

In this chapter, the EM algorithm has been developed for parameter estimation and standard error calculated using the Louis principle. Both the profile and complete likelihood approaches were utilised. Profile likelihood approach works well in estimating ϕ with small bias and the complete likelihood approach clearly worked very well in recovering the parameter values. It has been shown that changing only the parameter γ_1 does not affect the accuracy and precision of the parameters, but, when the parameter γ_2 increases, there is a reduction in the precision and accuracy of ϕ and γ_2. The application to the real melanoma data works well and the destructive mechanism improves the model fit.

Chapter 5

Destructive COM-Poisson cure rate model with generalised gamma lifetime

5.1 Introduction

The generalised gamma distribution was introduced by Stacy (1962) and includes many commonly used lifetime distributions as particular cases. The generalised gamma distribution defined in (1.17) has special cases as Weibull distribution (when $q = 1$), gamma distribution (when $q = \sigma$) and log-normal distribution (when $q = 0$, which is also the limit of (1.17) as $q \to 0$). The generalised gamma includes four most common types of hazard functions: monotonically increasing, monotonically decreasing, bathtub and arc-shaped hazards. The type of the hazard function is determined by the parameters q and σ together while λ simply act as a multiplicative factor of time. This flexibility of the generalised gamma as a lifetime distribution enables model discrimination to be carried out within its family to select a parsimonious lifetime distribution that adequately fits the data.

5.2 Likelihood and estimation of parameters

The first- and second-order derivatives of the $Q(\theta,\ \pi^{(k)})$ with respect to θ are computed numerically and so is the second order derivatives of the Hessian matrix. The code for data generation and EM algorithm for estimation of parameters for the destructive Poisson cure rate model with generalised gamma lifetime are given in Appendix D.1 and Appendix D.2, respectively.

Simplified expressions of the Q-function for some particular models are presented next.

Destructive Bernoulli cure rate model

In this case, the population density function can be expressed as $f_{\text{pop}}(t) = \frac{\eta p f(t)}{1+\eta}$, the population survival function as $S_{\text{pop}}(t) = \frac{1+\eta(1-pF(t))}{1+\eta}$, and the cure rate as $p_0 = \frac{1+\eta(1-p)}{1+\eta}$. Using these expressions, the $Q(\theta,\ \pi^{(k)})$ function can be simplified as

$$Q(\theta,\pi^{(k)}) = \sum_{I_1} \log\left(\frac{\eta_i p_i f(t_i)}{1+\eta_i}\right) + \sum_{I_0}(1-\pi^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right) + \sum_{I_0}\pi^{(k)}\log\left(\frac{\eta_i p_i S(t_i)}{1+\eta_i}\right),$$

where $\pi_i^{(k)} = \frac{\eta_i p_i S(t_i)}{1+\eta_i(1-p_i F(t_i))}\big|_{\theta=\theta^{(k)}}$ for the i-th censored observation. Note that the $Q(\theta,\ \pi^{(k)})$ function can be split into two parts, one containing the regression parameters $\beta = (\beta_1', \beta_2')'$ and the other containing the lifetime parameter γ. Thus, the maximisation step can be carried out by maximising $Q_1(\beta,\ \pi^{(k)})$ with respect to β and $Q_2(\gamma,\ \pi^{(k)})$ with respect to γ, separately, where

$$Q_1(\beta,\ \pi^{(k)}) = \sum_{I_1}\log\left(\frac{\eta_i p_i}{1+\eta_i}\right) + \sum_{I_0}(1-\pi^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right) + \sum_{I_0}\pi^{(k)}\log\left(\frac{\eta_i p_i}{1+\eta_i}\right)$$

and

$$Q_2(\gamma,\ \pi^{(k)}) = \sum_{I_1}\log f(t_i) + \sum_{I_0}\pi^{(k)}\log S(t_i).$$

Destructive Poisson cure rate model

In this case, the population density function can be expressed as $f_{\text{pop}}(t) = \eta p f(t) S_{\text{pop}}(t)$, the population survival function as $S_{\text{pop}}(t) = \exp(-\eta p F(t))$, and the cure rate as $p_0 = \exp(-\eta p)$. Using these expressions, the $Q(\theta,\ \pi^{(k)})$ function can be simplified as

$$\begin{aligned} Q(\theta, \pi^{(k)}) &= \sum_{I_1} [\log(\eta_i p_i f(t_i)) - \eta_i p_i F(t_i)] - \sum_{I_0} (1 - \pi_i^{(k)}) \eta_i p_i \\ &\quad + \sum_{I_0} \pi_i^{(k)} \log\left[\exp(-\eta_i p_i F(t_i)) - \exp(-\eta_i p_i)\right], \end{aligned}$$

where ${\pi_i}^{(k)} = 1 - \exp(-\eta_i p_i S(t_i))|_{\theta=\theta^{(k)}}$ for the i-th censored observation.

Destructive COM-Poisson cure rate model

In this case, the $Q(\theta,\ \pi^{(k)})$ function can be expressed as

$$\begin{aligned} Q(\theta, \pi^{(k)}) &= \sum_{I_1} \left\{\log[p_i f(t_i) z_{f1}] - \log[z(1 - p_i F(t_i))]\right\} + \sum_{I_0} (1 - \pi_i^{(k)}) \log\left(\frac{z_p}{z}\right) \\ &\quad + \sum_{I_0} \pi_i^{(k)} \log\left(\frac{z_f - z_p}{z}\right), \end{aligned}$$

where $\pi_i^{(k)} = \frac{z_f - z_p}{z - z_p}|_{\theta=\theta^{(k)}}$ for the i-th censored observation, $z_f = \sum_{j=0}^{\infty} \frac{[\eta_i(1-p_i F(t_i))]^j}{(j!)^\phi}$, $z_p = \sum_{j=0}^{\infty} \frac{[\eta_i(1-p_i)]^j}{(j!)^\phi}$, $z = \sum_{j=0}^{\infty} \frac{\eta_i^j}{(j!)^\phi}$, and $z_{f1} = \sum_{j=1}^{\infty} \frac{j[\eta_i(1-p_i F(t_i))]^j}{(j!)^\phi}$.

5.3 Simulation study

An extensive MC simulation study is carried out to evaluate whether the proposed inference methodology is able to satisfactorily recover the parameters values. Varying sample sizes of $n = 200$ and $n = 400$ are chosen to demonstrate the small and moderately large sample behaviours. Two covariates, ulceration status and tumor thickness, are considered. Ulceration status is generated from Uniform $(0, 1)$ distribution and if the value is greater than 0.44,

ulceration status is taken as 0, otherwise, it is taken as 1. Tumor thickness is generated from a Uniform $(0.1, 17.42)$ distribution, since in the melanoma data the minimum and maximum values of tumor thickness are 0.1 and 17.42 mm, respectively. For the simulation study, the parameters η and p are linked to ulceration status and tumor thickness, respectively. To come up with the true value of β_1 corresponding to η, we can choose a value of η as 1.4, which gives the true value of β_1 as 0.34, using the log-linear link function. This value of $\eta = 1.4$ corresponds to the presence of ulcer, and when the ulcer is absent, the value of η is expected to be lower, which is 1 in this case. For the true values of the regression parameters corresponding to tumor thickness (β_{02} and β_{12}), the logistic link function for p with the minimum and maximum tumor thickness values are equated to 0.1 and 0.9, respectively. The true values are calculated as $\beta_{02} = -2.223$ and $\beta_{12} = 0.254$ after solving the equations. Note that lower tumor thickness would imply lower activation probability p.

To generate the lifetimes, we first generate a COM-Poisson random variable M with parameter η, for a fixed choice of ϕ. If the value of M is zero, the observed time (T) is taken as the censoring time (C). Otherwise, if $M = m\,(> 0)$, a random variable D is generated from Binomial(m, p) distribution. If $D = 0$, the observed time (T) is once gain taken as the censoring time (C). Otherwise, D generalised gamma random variables are generated $(G_1, G_2, \ldots, G_D)$ with density function as in (1.17), for specific choices of λ, σ, and q. Then, T is taken as the $\min\{\min(G_1, G_2, \cdots, G_D), C\}$. Finally, if $T = C$, the censoring indicator δ is taken as 0, otherwise, it is taken as 1. The censoring times (C) are generated from an exponential distribution with rate 0.3. The R package "compoisson" is used to generate data from the COM-Poisson distribution. To generate data from the generalised gamma distribution, we made use of the following result: If G follows a generalised gamma distribution as in (1.17), then, $V = G^{q/\sigma}$ follows a gamma distribution with shape = $\frac{1}{q^2}$ and scale = $\frac{q^2}{\lambda^{q/\sigma}}$. Hence, $V^{\sigma/q}$ is a generalised gamma variable of our interest.

The true values of the lifetime parameters λ and σ, for a chosen value of q, are obtained by equating the mean $\frac{(q^2)^{\sigma/q}}{\lambda\Gamma(1/q^2)}\Gamma(\frac{\sigma}{q} + \frac{1}{q^2})$ and the variance $\frac{q^{4\sigma/q}}{\lambda^2\Gamma(1/q^2)}\left\{\Gamma(\frac{2\sigma}{q} + \frac{1}{q^2}) - \frac{\Gamma^2(\sigma/q+1/q^2)}{\Gamma(1/q^2)}\right\}$ of the underlying generalised gamma distribution in

(1.17) to some fixed values. We consider the true values of q to be 0.8 and 1.5, and using 0.3 and 0.05 as the mean and variance values, respectively, we obtain $(\lambda,\ \sigma)$ as $(2.454,\ 0.745)$ and $(2.293,\ 0.684)$. To start the iterative EM algorithm, initial value of $\gamma_1 = (\lambda, \sigma)'$ can be obtained by drawing a random number from the interval $(0.8\gamma_1, 1.2\gamma_1)$. With this initial value for γ_1, a discrete four-dimensional grid search on $(\beta_1,\ \beta_{02},\ \beta_{12},\ q)$ can be applied, using a suitable range for each parameter, for a particular value of ϕ. The log-likelihood value needs to be calculated for each possible combination and the combination that results in the maximum log-likelihood value can be taken as the initial value. All the simulations are based on 1000 MC runs. For the destructive COM-Poisson-generalised gamma cure rate model, the log-likelihood surface turned out to be flat with respect to the COM-Poisson shape parameter ϕ and the generalised gamma shape parameter q. Hence, along with complete likelihood approach, we also employ a profile-likelihood technique, where the EM algorithm is first run for several chosen fixed values of ϕ and q. Then, the values of ϕ and q that results in the maximum log-likelihood value is taken as their MLEs. The corresponding estimates of other model parameters are taken as their MLEs. Note that the profile likelihood approach does not allow us to compute the standard errors of ϕ and q since the information matrix is inverted on the assumption that ϕ and q are fixed.

Tables 5.1 and 5.2 present the results for the destructive Poisson and destructive Bernoulli generalised gamma regression cure rate models, where the parameters are estimated with and without using the profile likelihood approach. Under setting 1, when $q = 1.5$, the profile likelihood approach is applied with respect to the generalised gamma shape parameter q. For this, a discrete grid for q is selected as $\{0, 0.1, \ldots, 2.0\}$. The profile likelihood approach works very well in estimating q with small bias. Also, the bias and RMSE of q decreases as sample size increases. Since in the standard error calculation, the uncertainty of q is not taken into account, it results in under-estimation problem. This can be seen in the coverage probabilities where the under-coverage is clearly noticeable for the lifetime parameters σ and λ. A similar observation was made by Balakrishnan and Pal (2015). When all the parameters are estimated using the complete likelihood approach, there is no underestimation problem, as expected. However, the coverage probability of q is generally high because of the flatness

of the likelihood function with respect to q. The bias, RMSE and standard errors of the model parameters decreases as sample size increases and the coverage probabilities are close to the nominal values. Note that the MLEs, RMSEs and standard errors of the regression parameters obtained by both methods are similar, its only for the lifetime parameters that the profile likelihood approach results in smaller standard errors. Considering Setting 2, when $q = 0.8$ in this case, complete likelihood approach is used. The EM algorithm once again performs very well and satisfies the large sample properties. The profile likelihood approach results for Setting 2 are similar to Setting 1 and, therefore, are not included.

Table 5.1 Accuracy and precision of estimates for the destructive Poisson cure rate model with generalised gamma lifetime and using profile likelihood technique together with complete likelihood technique via the EM algorithm

Approach	n	Parameter	MLEs (se)	Bias	RMSE	CP	
						90%	95%
Profile	200	$\beta_1 = 0.340$	0.350 (0.219)	0.010	0.221	0.901	0.943
Likelihood		$\beta_{02} = -2.223$	-2.408 (0.551)	-0.185	0.577	0.916	0.964
(Setting 1)		$\beta_{12} = 0.254$	0.291 (0.086)	0.037	0.096	0.956	0.986
		$\sigma = 0.684$	0.659 (0.061)	-0.025	0.106	0.595	0.689
		$\lambda = 2.293$	2.372 (0.225)	0.079	0.413	0.640	0.735
		$q = 1.500$	1.555 $(-)$	0.055	0.378	-	-
	400	$\beta_1 = 0.340$	0.350 (0.156)	0.010	0.150	0.908	0.962
		$\beta_{02} = -2.223$	-2.280 (0.365)	-0.057	0.358	0.920	0.966
		$\beta_{12} = 0.254$	0.266 (0.055)	0.012	0.052	0.949	0.984
		$\sigma = 0.684$	0.670 (0.044)	-0.014	0.079	0.589	0.680
		$\lambda = 2.293$	2.326 (0.158)	0.033	0.282	0.659	0.729
		$q = 1.500$	1.534 $(-)$	0.034	0.280	-	-
Complete	200	$\beta_1 = 0.340$	0.336 (0.222)	-0.004	0.233	0.883	0.933
Likelihood		$\beta_{02} = -2.223$	-2.391 (0.550)	-0.168	0.591	0.910	0.958
(Setting 1)		$\beta_{12} = 0.254$	0.284 (0.086)	0.030	0.098	0.948	0.980
		$\sigma = 0.684$	0.648 (0.120)	-0.036	0.128	0.883	0.929
		$\lambda = 2.293$	2.360 (0.405)	0.067	0.423	0.891	0.951
		$q = 1.500$	1.661 (0.532)	0.161	0.617	0.922	0.971
	400	$\beta_1 = 0.340$	0.349 (0.157)	0.009	0.158	0.898	0.954
		$\beta_{02} = -2.223$	-2.258 (0.363)	-0.035	0.359	0.915	0.960
		$\beta_{12} = 0.254$	0.261 (0.054)	0.007	0.053	0.930	0.973
		$\sigma = 0.684$	0.664 (0.083)	-0.020	0.091	0.882	0.927
		$\lambda = 2.293$	2.314 (0.277)	0.021	0.291	0.889	0.940
		$q = 1.500$	1.584 (0.326)	0.084	0.428	0.915	0.961
Complete	200	$\beta_1 = 0.340$	0.351 (0.226)	0.011	0.229	0.893	0.949
Likelihood		$\beta_{02} = -2.223$	-2.342 (0.548)	-0.119	0.585	0.895	0.954
(Setting 2)		$\beta_{12} = 0.254$	0.274 (0.085)	0.020	0.094	0.902	0.948
		$\sigma = 0.745$	0.713 (0.097)	-0.032	0.098	0.880	0.932
		$\lambda = 2.454$	2.472 (0.375)	0.018	0.376	0.908	0.952
		$q = 0.800$	0.885 (0.349)	0.085	0.372	0.953	0.991
	400	$\beta_1 = 0.340$	0.343 (0.158)	0.003	0.150	0.910	0.964
		$\beta_{02} = -2.223$	-2.280 (0.370)	-0.057	0.368	0.918	0.960
		$\beta_{12} = 0.254$	0.264 (0.056)	0.010	0.056	0.918	0.968
		$\sigma = 0.745$	0.734 (0.066)	-0.011	0.070	0.888	0.928
		$\lambda = 2.454$	2.465 (0.259)	0.011	0.257	0.903	0.949
		$q = 0.800$	0.819 (0.221)	0.019	0.226	0.921	0.963
	600	$\beta_1 = 0.340$	0.344 (0.129)	0.004	0.128	0.905	0.954
		$\beta_{02} = -2.223$	-2.266 (0.298)	-0.043	0.308	0.897	0.955
		$\beta_{12} = 0.254$	0.261 (0.044)	0.007	0.047	0.903	0.948
		$\sigma = 0.745$	0.736 (0.054)	-0.009	0.053	0.895	0.943
		$\lambda = 2.454$	2.467 (0.209)	0.013	0.213	0.899	0.946
		$q = 0.800$	0.805 (0.175)	0.005	0.172	0.926	0.967

Table 5.2 Accuracy and precision of estimates for the destructive Bernoulli cure rate model with generalised gamma lifetime and using profile likelihood technique together with complete likelihood technique via the EM algorithm

Approach	n	Parameter	MLEs (se)	Bias	RMSE	CP	
						90%	95%
Profile	200	$\beta_1 = 0.340$	0.506 (0.695)	0.166	0.621	0.957	0.978
Likelihood		$\beta_{02} = -2.223$	-2.473 (0.707)	-0.250	0.777	0.917	0.966
(Setting 1)		$\beta_{12} = 0.254$	0.292 (0.098)	0.038	0.115	0.935	0.973
		$\sigma = 0.684$	0.659 (0.079)	-0.025	0.113	0.715	0.796
		$\lambda = 2.293$	2.450 (0.263)	0.157	0.492	0.659	0.771
		$q = 1.500$	1.556 $(-)$	0.056	0.426	-	-
	400	$\beta_1 = 0.340$	0.433 (0.416)	0.093	0.400	0.969	0.992
		$\beta_{02} = -2.223$	-2.354 (0.466)	-0.131	0.493	0.892	0.963
		$\beta_{12} = 0.254$	0.273 (0.063)	0.019	0.063	0.923	0.963
		$\sigma = 0.684$	0.660 (0.056)	-0.024	0.082	0.692	0.786
		$\lambda = 2.293$	2.333 (0.177)	0.040	0.298	0.672	0.758
		$q = 1.500$	1.534 $(-)$	0.034	0.311	-	-
Complete	200	$\beta_1 = 0.340$	0.455 (0.665)	0.115	0.626	0.940	0.964
Likelihood		$\beta_{02} = -2.223$	-2.441 (0.706)	-0.218	0.783	0.910	0.959
(Setting 1)		$\beta_{12} = 0.254$	0.291 (0.098)	0.037	0.115	0.936	0.978
		$\sigma = 0.684$	0.656 (0.133)	-0.028	0.130	0.895	0.946
		$\lambda = 2.293$	2.429 (0.498)	0.136	0.494	0.933	0.976
		$q = 1.500$	1.603 (0.627)	0.103	0.653	0.933	0.972
	400	$\beta_1 = 0.340$	0.385 (0.404)	0.045	0.397	0.957	0.979
		$\beta_{02} = -2.223$	-2.363 (0.471)	-0.140	0.499	0.891	0.959
		$\beta_{12} = 0.254$	0.275 (0.064)	0.021	0.064	0.925	0.963
		$\sigma = 0.684$	0.658 (0.093)	-0.026	0.102	0.896	0.935
		$\lambda = 2.293$	2.319 (0.330)	0.026	0.327	0.907	0.950
		$q = 1.500$	1.629 (0.423)	0.129	0.554	0.906	0.959
Complete	200	$\beta_1 = 0.340$	0.415 (0.679)	0.075	0.641	0.924	0.956
Likelihood		$\beta_{02} = -2.223$	-2.456 (0.745)	-0.233	0.852	0.935	0.977
(Setting 2)		$\beta_{12} = 0.254$	0.298 (0.108)	0.044	0.146	0.938	0.982
		$\sigma = 0.745$	0.699 (0.112)	-0.046	0.106	0.886	0.923
		$\lambda = 2.454$	2.454 (0.464)	0.000	0.373	0.947	0.970
		$q = 0.800$	0.934 (0.467)	0.134	0.424	0.980	0.997
	400	$\beta_1 = 0.340$	0.390 (0.431)	0.050	0.436	0.917	0.949
		$\beta_{02} = -2.223$	-2.334 (0.474)	-0.111	0.520	0.886	0.943
		$\beta_{12} = 0.254$	0.269 (0.064)	0.015	0.069	0.909	0.953
		$\sigma = 0.745$	0.728 (0.077)	-0.017	0.074	0.902	0.941
		$\lambda = 2.454$	2.468 (0.326)	0.014	0.298	0.930	0.966
		$q = 0.800$	0.846 (0.292)	0.046	0.271	0.962	0.986
	600	$\beta_1 = 0.340$	0.364 (0.329)	0.024	0.334	0.916	0.945
		$\beta_{02} = -2.223$	-2.284 (0.377)	-0.061	0.373	0.917	0.956
		$\beta_{12} = 0.254$	0.264 (0.051)	0.010	0.053	0.905	0.952
		$\sigma = 0.745$	0.732 (0.062)	-0.013	0.065	0.880	0.942
		$\lambda = 2.454$	2.464 (0.262)	0.010	0.252	0.905	0.954
		$q = 0.800$	0.830 (0.231)	0.030	0.230	0.918	0.972

5.3.1 Model discrimination

The COM-Poisson and generalised gamma families are flexible in the sense that they contain many commonly used distributions as special cases. This will aid in selecting a parsimonious joint distribution that fits the data adequately. This model discrimination can be carried out by using the LRT and information-based criteria.

Likelihood ratio test

We first evaluate the performance of the LRT in testing the null hypothesis that the lifetime can be modelled by one of the exponential $(H_0 : \ q = 1, \ \sigma = 1)$, gamma $(H_0 : \ q = \sigma)$, Weibull $(H_0 : \ q = 1)$, lognormal $(H_0 : \ q = 0)$ or generalized gamma $(H_0 : \ q = 0.8)$ distributions against the alternative that the lifetime can be modelled by any other member of the generalised gamma family. For this purpose, i.e., utilising the flexibility of the generalised gamma family, we assume the true competing cause distribution to be either Bernoulli or Poisson. Considering the same parameter settings as in the simulation study and for each generated data, we calculated the Λ values of the fitted exponential, gamma, lognormal, Weibull, and generalised gamma ($q = 0.8$) distributions versus the fitted generalised gamma distribution. We present the observed levels (in bold) and power values of the LRT, all determined by the rejection rate of the null hypothesis, in Table 5.3. For this, we choose the true level as 5% and all the results are based on 1000 MC runs. We first note that the observed levels are close to the true nominal level used, note that when the true model is lognormal, the observed levels are liberal for both destructive Poisson and Bernoulli models, however, they improve as sample size increases. This shows that the chi-square and the mixture chi-square distributions provide a good approximation to the null distribution of the LRT statistic. It is clearly seen that when the true lifetime distribution is either lognormal, Weibull, gamma or generalised gamma ($q = 0.8$), the test generally has high power to reject the exponential model. However, when the true lifetime distribution is exponential, the test only has power to reject the lognormal distribution and no power to reject the Weibull,

gamma, and generalised gamma ($q = 0.8$) models. When the true lifetime distribution is either Weibull, gamma, exponential or generalised gamma ($q = 0.8$), the test has high power to reject the lognormal model, which increases as sample size increases. Furthermore, when the true lifetime distribution is lognormal, the test once again has high power to reject the Weibull, gamma, exponential, and generalised gamma ($q = 0.8$) models (note that the power to reject the gamma model is a little conservative, however, it increases as sample size increases). Hence, we can say that the LRT can detect a difference in model fit between the lognormal model and the Weibull, gamma, exponential, and generalised gamma ($q = 0.8$) models. Since the test has very low power to reject the gamma (Weibull) model when the true lifetime is Weibull (gamma), we can say that the LRT cannot distinctly discriminate between the gamma and Weibull models. Hence, the gamma and the Weibull models are expected to provide similar fits.

Table 5.3 Model discrimination within the generalised gamma family using the likelihood ratio test

n	Fitted model	True generalized gamma model				
		Lognormal	$q = 0.8$	Weibull	Gamma	Exponential
		Destructive Poisson Model				
200	Lognormal	**0.061**	0.801	0.910	0.769	0.869
	$q = 0.8$	0.691	**0.056**	0.110	0.050	0.098
	Weibull	0.834	0.089	**0.042**	0.112	0.057
	Gamma	0.540	0.062	0.104	**0.058**	0.059
	Exponential	1.000	0.964	0.837	0.973	**0.061**
400	Lognormal	**0.054**	0.976	0.999	0.957	0.995
	$q = 0.8$	0.925	**0.055**	0.160	0.044	0.159
	Weibull	0.991	0.146	**0.057**	0.223	0.056
	Gamma	0.782	0.062	0.169	**0.045**	0.053
	Exponential	1.000	1.000	0.989	1.000	**0.057**
		Destructive Bernoulli Model				
200	Lognormal	**0.066**	0.653	0.798	0.579	0.801
	$q = 0.8$	0.492	**0.059**	0.077	0.057	0.100
	Weibull	0.656	0.078	**0.046**	0.083	0.040
	Gamma	0.391	0.050	0.082	**0.056**	0.050
	Exponential	0.992	0.769	0.628	0.800	**0.057**
400	Lognormal	**0.057**	0.877	0.960	0.820	0.944
	$q = 0.8$	0.786	**0.042**	0.124	0.057	0.139
	Weibull	0.943	0.085	**0.047**	0.143	0.050
	Gamma	0.643	0.049	0.119	**0.046**	0.046
	Exponential	1.000	0.975	0.906	0.993	**0.049**

A similar model discrimination is carried out within the COM-Poisson family for a fixed lifetime distribution and the results are presented in Table 5.4. In this case, we evaluate the performance of the LRT in testing the null hypothesis that the competing cause can

be modelled by one of the Bernoulli ($\phi \to \infty$), Poisson ($\phi = 1$), COM-Poisson ($\phi = 0.3$) and COM-Poisson ($\phi = 2$) distributions against the alternative that the competing cause can be modelled by any other member of the COM-Poisson family. For this, we assume the lifetimes to follow the generalised gamma ($q = 0.8$) distribution. From Table 5.4, we note that the observed levels are conservative except for the case when the true competing cause is COM-Poisson ($\phi = 0.3$). When the true competing cause distribution is Bernoulli ($\phi \to \infty$), the test has no power to reject Poisson ($\phi = 1$) and COM-Poisson ($\phi = 2$), however, it has very low power to reject COM-Poisson ($\phi = 0.3$). On the other hand, when the true competing cause is COM-Poisson ($\phi = 0.3$), the test has very high power to reject the Bernoulli. When the true competing cause is Poisson, the test only has moderate power to reject the Bernoulli, provided the sample size is large.

Table 5.4 Model discrimination within the COM-Poisson family for generalised gamma ($q = 0.8$) lifetimes using likelihood ratio test

n	Fitted destructive COM-Poisson model	True destructive COM-Poisson model			
		$\phi = 0.3$	$\phi = 1.0$	$\phi = 2.0$	$\phi \to \infty$
200	$\phi = 0.3$	**0.042**	0.189	0.153	0.133
	$\phi = 1.0$	0.079	**0.029**	0.047	0.042
	$\phi = 2.0$	0.638	0.020	**0.003**	0.004
	$\phi \to \infty$	0.975	0.241	0.048	**0.016**
400	$\phi = 0.3$	**0.043**	0.228	0.284	0.247
	$\phi = 1.0$	0.419	**0.033**	0.076	0.068
	$\phi = 2.0$	0.969	0.062	**0.007**	0.003
	$\phi \to \infty$	1.000	0.503	0.118	**0.017**

Information-based criteria

The performance of the AIC and BIC in selecting one of the lognormal, Weibull, gamma, exponential and generalised gamma ($q = 0.8$) as the lifetime distribution given a true lifetime distribution is investigated in this section. In this study, we considered the same parameter

settings as in Section 5.3 and we only considered the destructive Poisson and destructive Bernoulli cure rate models. Based on 1000 data sets in each simulation, we calculated the observed selection rates for each of the two selection criteria, and these results are presented in Table 5.5. When the true model is either lognormal or exponential, it is noted that both the selection criteria performs very well in selecting the correct model. When the true model is Weibull, the selection rates are little conservative. In all cases, the correct selection rate increases as sample size increases. When gamma is the true lifetime distribution, the information criteria cannot distinguish between the Weibull and gamma models. Also, a generalised gamma ($q = 0.8$) is closer to Weibull than gamma.

In this section, the flexibility of the generalised gamma family has been utilised to select a parsimonious lifetime distribution that provides the best fit to the data. The flexibilities of the COM-Poisson family and the generalised gamma family will both be utilised to select a pair of parsimonious distributions and the next section will illustrate this two-way model discrimination using the real melanoma data set.

Table 5.5 Observed selection rates based on AIC (BIC) for the destructive model within the generalised gamma family of lifetimes

n	Fitted model	True generalized gamma model				
		Lognormal	$q = 0.8$	Weibull	Gamma	Exponential
		Destructive Poisson Model				
200	Lognormal	0.814 (0.814)	0.083 (0.083)	0.031 (0.031)	0.106 (0.106)	0.043 (0.017)
	$q = 0.8$	0.009 (0.009)	0.169 (0.160)	0.137 (0.114)	0.181 (0.138)	0.078 (0.009)
	Weibull	0.003 (0.003)	0.418 (0.403)	0.620 (0.529)	0.361 (0.351)	0.073 (0.014)
	Gamma	0.173 (0.174)	0.328 (0.323)	0.192 (0.180)	0.351 (0.345)	0.057 (0.008)
	Exponential	0.000 (0.000)	0.003 (0.030)	0.020 (0.145)	0.000 (0.024)	0.749 (0.951)
400	Lognormal	0.919 (0.919)	0.023 (0.023)	0.001 (0.001)	0.017 (0.36)	0.003 (0.000)
	$q = 0.8$	0.001 (0.001)	0.236 (0.236)	0.153 (0.153)	0.306 (0.247)	0.134 (0.010)
	Weibull	0.000 (0.000)	0.335 (0.335)	0.705 (0.693)	0.217 (0.240)	0.070 (0.010)
	Gamma	0.080 (0.080)	0.406 (0.406)	0.141 (0.140)	0.461 (0.477)	0.076 (0.007)
	Exponential	0.000 (0.000)	0.000 (0.000)	0.000 (0.013)	0.000 (0.000)	0.718 (0.972)
		Destructive Bernoulli Model				
200	Lognormal	0.766 (0.766)	0.147 (0.136)	0.053 (0.047)	0.164 (0.155)	0.055 (0.006)
	$q = 0.8$	0.036 (0.036)	0.153 (0.106)	0.125 (0.074)	0.154 (0.109)	0.061 (0.006)
	Weibull	0.019 (0.019)	0.433 (0.342)	0.574 (0.402)	0.386 (0.317)	0.121 (0.022)
	Gamma	0.178 (0.178)	0.242 (0.225)	0.162 (0.132)	0.278 (0.258)	0.066 (0.019)
	Exponential	0.000 (0.000)	0.025 (0.191)	0.085 (0.345)	0.018 (0.161)	0.697 (0.948)
400	Lognormal	0.865 (0.865)	0.051 (0.054)	0.017 (0.017)	0.104 (0.103)	0.027 (0.005)
	$q = 0.8$	0.004 (0.004)	0.215 (0.202)	0.152 (0.137)	0.222 (0.217)	0.085 (0.012)
	Weibull	0.000 (0.000)	0.390 (0.374)	0.663 (0.581)	0.289 (0.286)	0.077 (0.002)
	Gamma	0.132 (0.132)	0.342 (0.344)	0.164 (0.148)	0.384 (0.382)	0.035 (0.000)
	Exponential	0.000 (0.000)	0.001 (0.026)	0.004 (0.116)	0.001 (0.012)	0.776 (0.980)

5.4 Analysis of malignant melanoma data

In this section, we illustrate an application of the proposed estimation procedure and the flexibilities of the COM-Poisson and generalised gamma families to the data set on malignant melanoma discussed in Section 1.8. Keeping in mind the identifiability issue as discussed earlier, we can consider four practical scenarios as discussed in Section 2.5.

Table 5.6 Maximized log-likelihood values (l) and information criteria values

Setting	l	AIC	BIC
1	-205.039	424.079	447.340
2	-207.375	428.750	452.011
3	-208.205	430.409	453.670
4	-211.956	437.913	461.172

The EM algorithm is run for all the four scenarios and the maximised log-likelihood values together with the AIC and BIC values are presented in Table 5.6.

Clearly, Scenario 1 provides the best fit as the log-likelihood value is the maximum and the AIC and BIC values are the minimum. Thus, we consider Scenario 1 as our working scenario. Two-way profile likelihood approach is employed for the parameters ϕ and q. Figure 5.1 shows the profile log-likelihood plot for different values of ϕ and q. Its clear that the log-likelihood surface is flat with respect to both higher ϕ and q values and the maximum is obtained when these values are 0.23 and 0, respectively. Thus, for the real data, the generalised gamma distribution reduces to a lognormal distribution. Table 5.7 presents the MLEs, standard errors and p-values of the destructive COM-Poisson model with generalised gamma ($q = 0$) lifetimes. Note that the profile and complete likelihood approaches produce the same MLEs and as noted in the simulation study, the standard errors corresponding to the complete likelihood approach are larger.

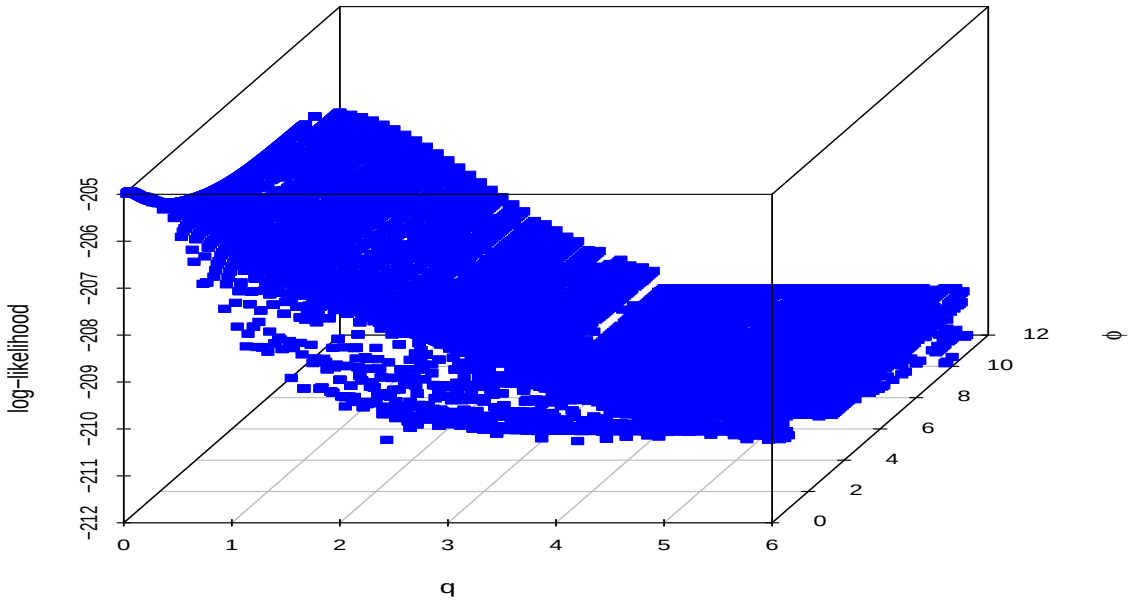

Fig. 5.1 Two-way profile log-likelihood plot for different values of ϕ and q

Table 5.7 MLEs, standard errors (se) and p-values for the destructive COM-Poisson model using profile and complete likelihood approaches with generalised gamma ($q = 0$) lifetimes

Approach		ϕ	β_1	β_{02}	β_{12}	σ	λ
Profile	MLE	0.23	0.319	-2.526	1.011	1.349	0.051
Likelihood	se	-	0.107	0.721	0.432	0.215	0.029
	p -value	-	0.003	0.000	0.019	0.000	0.039
Complete	MLE	0.229	0.319	-2.526	1.010	1.349	0.051
Likelihood	se	0.489	0.472	0.751	0.438	0.448	0.084
	p -value	0.320	0.500	0.000	0.021	0.001	0.272

Table 5.8 Log-likelihood value (l), Λ-statistic and p -values for different combinations of lifetime and competing cause distributions

Competing cause distribution		Lifetime distribution			
		Lognormal	Weibull	Gamma	Exponential
Poisson	l	-205.389	-206.799	-206.320	-210.906
	Λ	0.700	3.520	2.561	11.733
	p-value	0.352	0.172	0.278	0.008
Bernoulli	l	-207.269	-208.278	-207.906	-211.879
	Λ	4.460	6.477	5.734	13.678
	p-value	0.027	0.020	0.028	0.002
COM-Poisson ($\phi = 0.23$)	l	-205.039	-207.376	-206.857	-210.046
	Λ	≈ 0.000	4.673	3.635	10.012
	p-value	≈ 1.000	0.097	0.162	0.007

Next, we use the flexibilities of the COM-poisson and generalised gamma families to formally test for different combinations of competing cause and lifetime distributions. The results are presented in Table 5.8. The results show that the lognormal and COM-Poisson ($\phi = 0.23$) combination provides the best fit to the data. Also, the Poisson and any of the lifetime distributions except the exponential provides adequate fit to the data at 5% significance level. COM-Poisson ($\phi = 0.23$) and either Weibull or gamma lifetimes also provide adequate fit while the Bernoulli and any of the lifetime distributions does not provide good fit to the data at all. In Table 5.9, we present the maximised log-likelihood (l), AIC and BIC values for the combinations (from Table 5.8) that are not rejected at 5% level of significance. It is clear from Table 5.9 that the lognormal and COM-Poisson ($\phi = 0.23$) combination provides the best fit to the data and is thus our working model.

Table 5.9 AIC and BIC values for combinations of lifetime and competing cause distributions that provide an adequate fit

Competing cause	Lifetime distribution	l	AIC	BIC
COM-Poisson ($\phi = 0.23$)	Weibull	-207.376	428.752	452.013
Poisson	Weibull	-206.799	427.599	450.860
COM-Poisson ($\phi = 0.23$)	Gamma	-206.857	427.714	450.975
Poisson	Gamma	-206.320	426.640	449.901
COM-Poisson ($\phi = 0.23$)	Lognormal	-205.039	424.079	447.340
Poisson	Lognormal	-205.389	424.778	448.039

5.4.1 Effect of covariates on the cure rate

To determine the effect of covariates on the cure rate, it is important to note that the estimate of β_1 is positive. This implies higher mean number of competing causes with the presence of ulcer, thus leading to a decrease in cure rate, as expected. On the other hand, the estimate of β_{12} being positive implies that the initial competing causes will remain active with a higher probability with an increase in tumor thickness, which in turn implies a decrease in cure rate (the cure rate being a decreasing function of p). It is also of interest to fit a particular case of the working model when $p = 1$, namely, the COM-Poisson ($\phi = 0.23$) model. In this way, the destructive mechanism is absent and to preserve the same number of parameters, the parameter η is linked to both the covariates with an intercept term. In this case, the maximised log-likelihood value turned out to be -207.603 (p-value of 0.022), providing overwhelming evidence of the fact that the destructive mechanism when taken into account gives a better fit to the data.

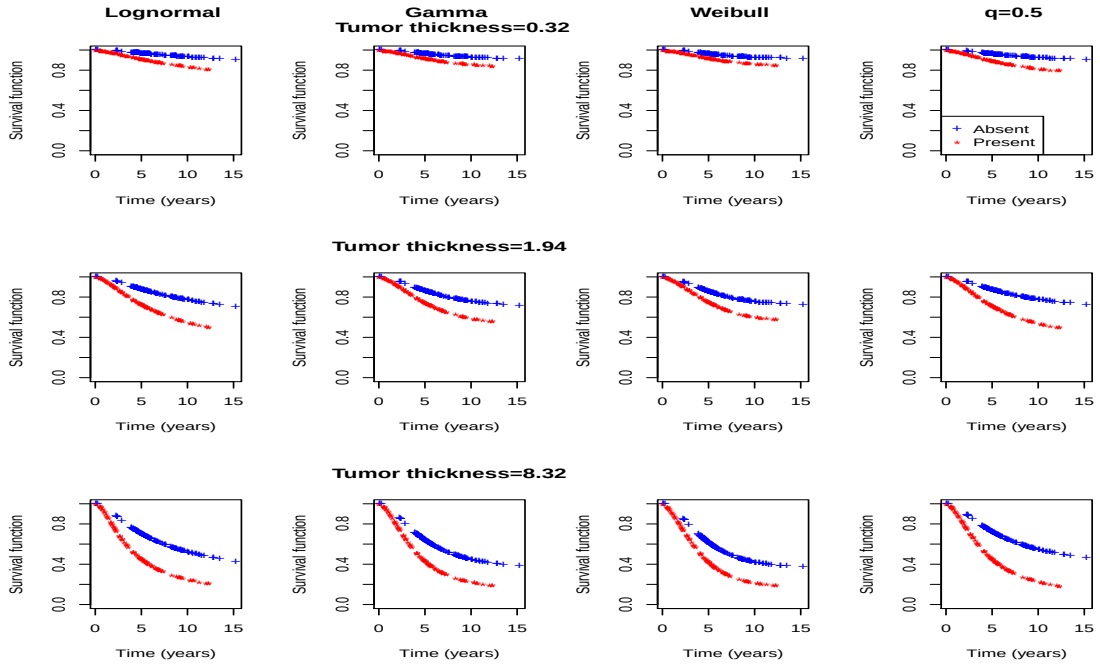

Fig. 5.2 Survival function under the destructive COM-Poisson model stratified by ulceration status for different lifetime distributions with different tumor thickness values

5.4.2 Profile likelihood surface for q and ϕ

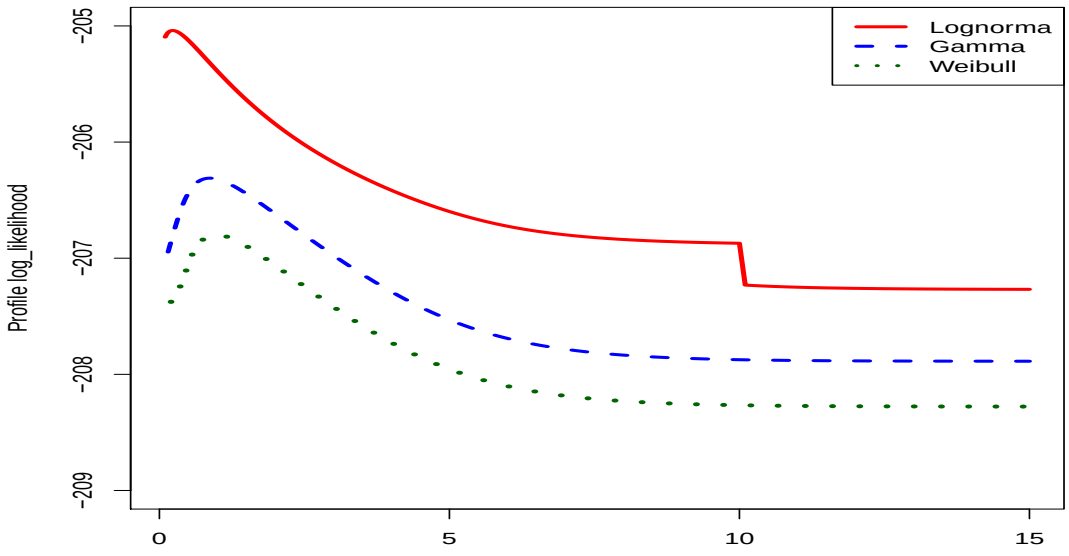

Fig. 5.3 Profile log-likelihood plot of ϕ for different lifetime distributions

Figure 5.3 shows the profile log-likelihood plot for different values of ϕ for special cases of the generalised gamma distribution. In order to find an acceptable range of ϕ for the data, we considered testing for different values of ϕ and Figure 5.4 shows the plot of the values of ϕ against the corresponding LRT statistic. We reject the test at $\alpha\%$ level of significance if the value of Λ exceeds $\chi^2_{1,1-\alpha}$. Figure 5.4 shows two values of $\chi^2_{1,1-\alpha}$ for $\alpha = 5\%$ and 10% for different lifetime distributions. For $\Lambda = 3.84$, the corresponding value of ϕ using lognormal as the lifetime distribution is 10.1 (from the plot) and hence an acceptable range of ϕ for the data is $(0, 10.1)$. From the figure, it is clear why the assumption of Bernoulli competing causes got rejected for all the lifetimes in Table 5.8.

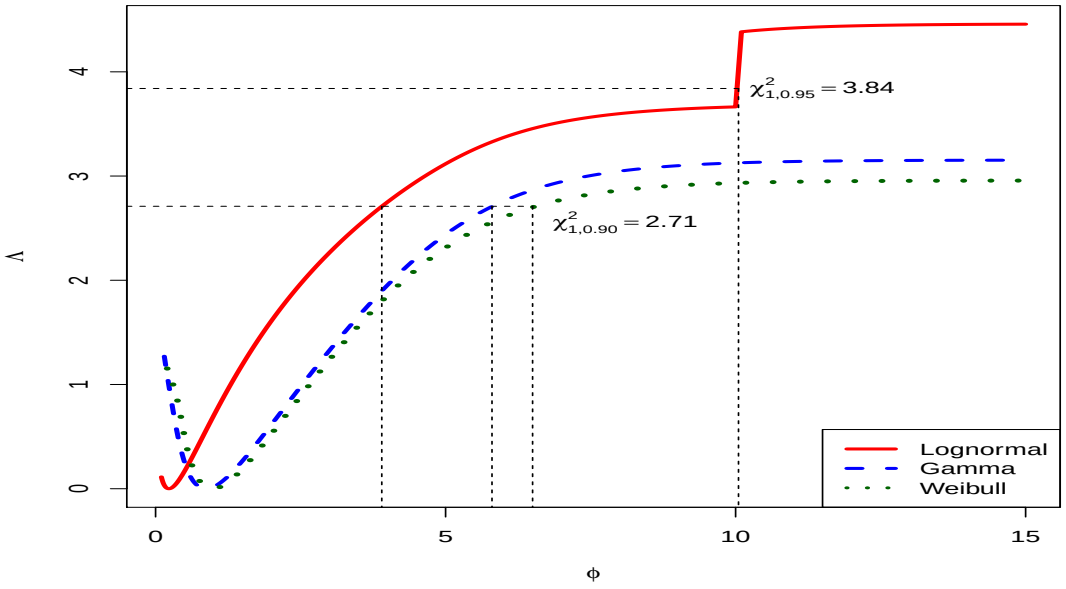

Fig. 5.4 Values of ϕ against the corresponding Λ-statistic for different lifetime distributions

5.4.3 Plots showing the survival function and effect of covariates on the cure rate

Figure 5.2 shows the surviving function for patients stratified by ulceration status with tumor thickness as 0.320, 1.94, and 8.32 mm, which corresponds to the 5%, 50%, and 95%

quantiles, respectively. The surviving probability decreases more rapidly for patients with thicker tumors. Furthermore, the survival probability is higher for patients without the ulcer. Figure 5.5 shows the effect of covariates on the cured fraction. The cure rate decreases with an increase in tumour thickness and then levels off for tumor thickness greater than 5 mm. As expected, the cure rate is significantly higher when ulceration status is absent. This clearly shows that subjects without ulcers have a higher probability of being long term survivors compared to those with ulcers.

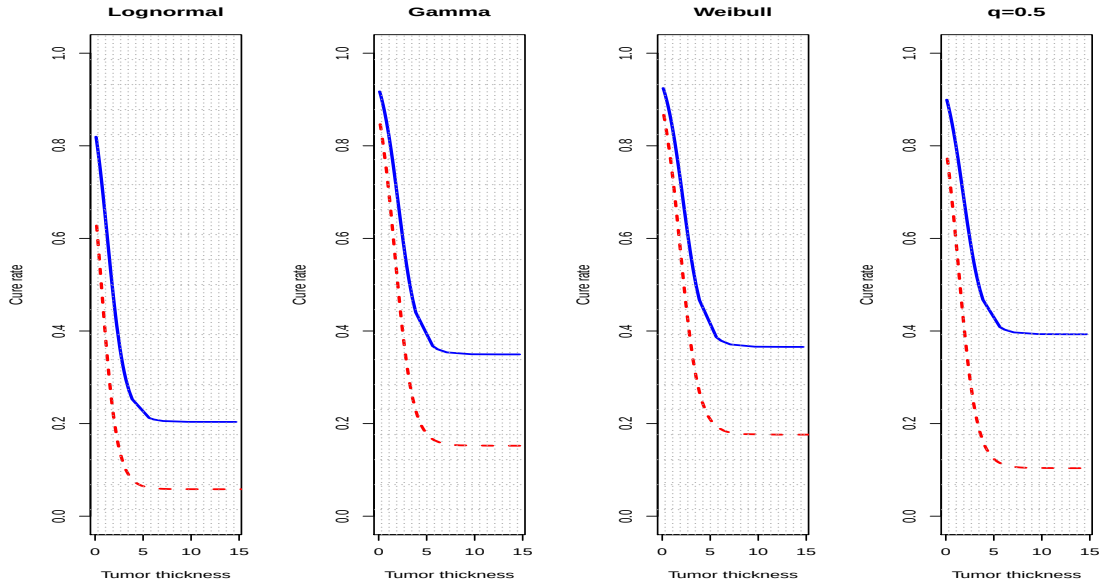

Fig. 5.5 Cure rate for the destructive COM-Poisson model stratified by ulceration status (upper: absent, lower: present) for different lifetime distributions

5.4.4 Goodness-of-fit

We also checked for the adequacy of the destructive COM-Poisson-generalised gamma cure rate model using the normalised randomised quantile residuals and the QQ plot is presented in Figure 5.6, where each point corresponds to the median of five sets of ordered residuals. The plot suggests that the model fits the data well. The p-value turned out to be 0.998 for the Kolmogorov-Smirnov test, which provides significant evidence that residuals are normally distribuited.

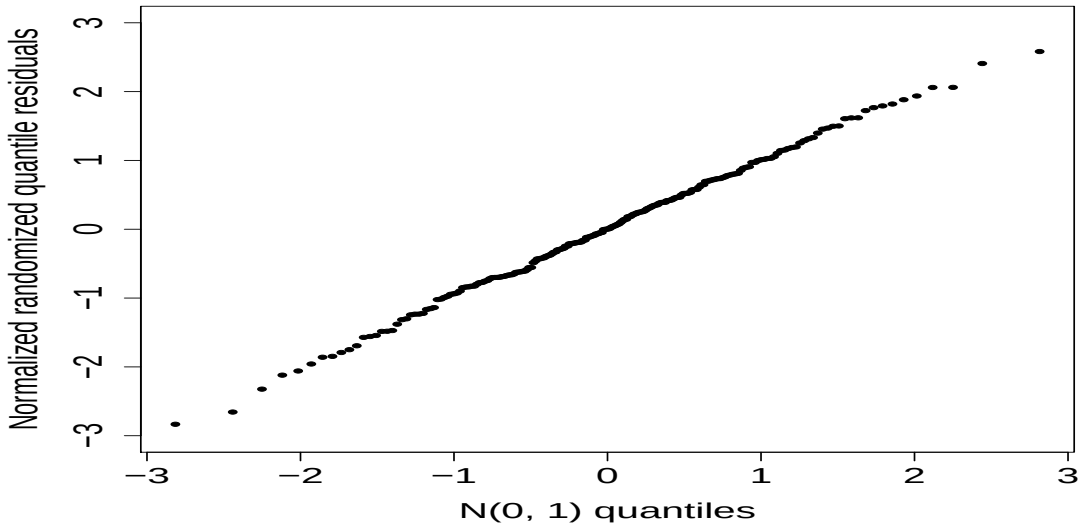

Fig. 5.6 Q-Q plot of the normalised randomised quantile residuals for the melanoma data

5.5 Conclusion

In this chapter, we have developed likelihood inference based on the EM algorithm for the flexible destructive COM-Poisson cure rate model assuming the lifetime to follow the flexible generalised gamma distribution. Our simulation study shows that the proposed EM algorithm works very well. Within the EM framework, we have illustrated both complete likelihood and profile likelihood approaches. The model discrimination results within the generalised gamma family shows that only the lognormal and exponential distributions are significantly different from the other lifetime distributions within this family. The main contributions of our work are in evaluating the flexibility of the the COM-Poisson family and generalised gamma family to carry out a two-way model discrimination to select a parsimonious pair of distributions that adequately fit the data and developing the EM algorithm for the estimation of model parameters. We have also shown through the real data that when the destructive mechanism is taken into account, the model provides a better fit to the data.

Chapter 6

Conclusion

6.1 Summary of research

A competitive scenario assumes that an event of interest can be due to several initial risk factors, but, in many situations it is conceivable that some of the original number of risks factors will go through destruction by intensive treatment. In this regard, Rodrigues et al. (2011) proposed the "destructive weighted Poisson cure rate model for describing the biological process of elimination of initiated cells after some specific treatment, but assuming biological independence of cells" (Borges et al., 2012), which makes a realistic interpretation of the biological mechanism of the occurrence of events in a competitive scenario. This research focuses on the development of methodology for the estimation of parameters of the destructive COM-Poisson cure rate model with generalised gamma lifetime through the EM algorithm. Data obtained from clinical trials are often right censored. In this research, the EM algorithm was used as an alternate efficient tool in determining the MLEs of the model parameters based on right censored data.

Several lifetime distributions were considered, namely, Weibull, lognormal, gamma and generalised gamma for the destructive COM-Poisson cure rate model. Also, special cases of

the COM-Poisson model were considered under each of the lifetime distributions. In each of these cases, the exact likelihood inference based on EM algorithm was developed. Extensive simulation studies were done to evaluate the performance of the proposed estimation methodology and to demonstrate the flexibilities of the COM-Poisson and generalised gamma families.

Under the Weibull lifetime, two approaches were implemented in estimating the model parameters of the destructive COM-Poisson model, using and without using the profile likelihood approach. The profile likelihood approach generally results in under-coverage which is more pronounced for the regression parameters. This is because the uncertainty of ϕ is not taken into account in the estimation procedure. When ϕ was estimated simultaneously from the model with the rest of the model parameters, it resulted in over-coverage. Generally, with an increase in sample size, there was a decrease in bias, standard error and RMSE. These large sample properties were satisfied for both under- and over-dispersed data. The simulation results show that the proposed EM algorithm works very well.

The lognormal and gamma lifetime distributions were then considered and in these cases, the parameters were once again estimated using the profile likelihood and complete likelihood approaches. The necessary steps of the EM algorithm were developed for the destructive COM-Poisson and some of its special cases. Again, the results showed that the EM algorithm converged to the true parameters accurately. The complete likelihood approach outperformed the profile likelihood approach in terms of precision and accuracy values and it worked very well especially under gamma lifetime. Model discrimination was carried out using both likelihood and information-based techniques. When the information based technique was utilised, the AIC and BIC gave similar results and generally gave better selection results compared to the likelihood ratio test.

The melanoma data was analysed assuming the aforementioned lifetime distributions for the destructive COM-Poisson model. In each case, the parameters were estimated using the two approaches, i.e, with and without the profile likelihood technique. Both approaches yielded

identical MLEs and log-likelihood values, with the profile likelihood approach having smaller standard errors. In each case, the destructive Poisson model provided adequate fit to the data while the destructive Bernoulli model was rejected. Lognormal as the lifetime distribution for the destructive COM-Poisson model had the largest log-likelihood value followed by gamma. A better fit to the data was realised when the destructive mechanism was taken into account.

Finally the generalised gamma as the lifetime distribution was considered. The EM algorithm was developed and simulation study was carried out to evaluate the performance of the estimation procedure. Again, the MLEs were obtained using both, with and without the profile likelihood approach. The methodology worked very well in estimating the parameters, but the profile likelihood approach resulted in under-coverage for the lifetime parameters. However, when all the parameters were estimated simultaneously from the model, the under-coverage was eradicated and the procedure worked well to recover the parameters. Over-coverage was noticed for the q parameter and this was caused due to the flatness of the log-likelihood function with respect to q. Model discrimination was carried out using both likelihood-based and information-based criteria, and both the criteria performed well in discriminating the special cases of the generalised gamma distribution.

Finally, the melanoma data was analysed in order to illustrate the two-way model discrimination within the COM-Poisson and generalised gamma families. It turned out that the lognormal as the lifetime distribution together with the COM-Poisson ($\phi = 0.23$) as the number of competing causes together resulted in the best fit. Inclusion of the destructive mechanism made the fit much better, as indicated by the QQ plot of the randomised quantile residuals.

6.2 Future research

The current research can be extended in the following ways:

- Considering other wider classes for the lifetime distribution, such as the generalised F distribution that nests the F, generalized gamma, log-logistic and Burr distributions; and generalised power series distribution for the number of competing causes, which contains many discrete distributions as special cases.

- Generally, the time to event and the censoring time are assumed to be independent, which is not always true or may not have sound biological basis. For example, in a clinical trial, patients might withdraw because their health condition deteriorates with the new treatment and hence in this case the right censoring is informative. It will be of interest to extend the current research to the case of informative censoring. In this case, a suitable copula function may be helpful to capture the dependency between the lifetime and censoring variables.

- Applying the EM algorithm estimation procedure to a more general type of censoring, such as interval censoring, which includes right censoring, as its special case.

- The current framework can be extended to a Bayesian approach, which requires choosing appropriate prior distributions. Inferences can be based on Markov chain Monte Carlo (MCMC) methods, especially the Gibbs sampler algorithm. In such a scenario, the results obtained from the frequentist and Bayesian approaches can be compared.

- Many data in clinical trials involve biomarkers collected repeatedly over several time points. As a result, joint modelling of longitudinal and survival data becomes important. Little work has been done in this regard that can incorporate a cured fraction. It will be of great interest to inroduce competing cause and destructive mechanism in such modelling.

Appendices

Appendix A

Under Weibull lifetime distribution

A.1 First- and second-order derivatives of the Q-function

Destructive Bernoulli cure rate model

$$
\begin{aligned}
Q(\theta,\pi^{(k)}) \quad \propto \quad & \sum_{I_1}\log\left(\frac{\eta_i p_i f(t_i)}{1+\eta_i}\right)+\sum_{I_0}(1-\pi_i^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right) \\
& +\sum_{I_0}\pi_i^{(k)}\log\left(\frac{\eta_i p_i S(t_i)}{1+\eta_i}\right), \text{ where } \pi_i^{(k)}=\frac{\eta_i p_i S(t_i)}{1+\eta_i(1-p_i F(t_i))}.
\end{aligned}
$$

$$
\frac{\partial Q}{\partial \beta_{1l}} = \sum_{I_1}\frac{x_{1il}}{1+\eta_i}-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}\eta_i p_i}{(1+\eta_i)\{1+\eta_i(1-p_i)\}}+\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}}{1+\eta_i},
$$

$$
\frac{\partial Q}{\partial \beta_{2k}} = \sum_{I_1}x_{2ik}(1-p_i)-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}\eta_i p_i(1-p_i)}{\{1+\eta_i(1-p_i)\}}+\sum_{I_0}\pi_i^{(k)}x_{2ik}(1-p_i),
$$

$$
\frac{\partial Q}{\partial \gamma_1} = \sum_{I_1}E_i+\sum_{I_0}\pi_i^{(k)}\frac{\log(\gamma_2 t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2},
$$

$$
\frac{\partial Q}{\partial \gamma_2} = \sum_{I_1}H_i-\sum_{I_0}\pi_i^{(k)}\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1\gamma_2},
$$

$$
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \beta_{1l'}} = -\sum_{I_1}\frac{x_{1il}x_{1il'}\eta_i}{(1+\eta_i)^2}-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}x_{1il'}\eta_i p_i\{1-\eta_i^2(1-p_i)\}}{(1+\eta_i)^2\{1+\eta_i(1-p_i)\}^2}-\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}x_{1il'}\eta_i}{(1+\eta_i)^2},
$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} &= -\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i) \\
&\quad -\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)^2\{1+\eta_i-\frac{p_i}{(1-p_i)}\}\{1+\eta_i(1-p_i)\}}{\{1+\eta_i(1-p_i)\}^2} \\
&\quad -\sum_{I_0}\pi_i^{(k)}x_{2ik}x_{2ik'}p_i(1-p_i),
\end{aligned}$$

$$\frac{\partial^2 Q}{\partial\gamma_1^2} = -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1}+\frac{1}{\gamma_1^2}+\frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4}\right\}-\sum_{I_0}\pi_i^{(k)}\frac{\log(\gamma_2 t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}\{\log(\gamma_2 t_i)+2\gamma_1\}}{\gamma_1^4},$$

$$\frac{\partial^2 Q}{\partial\gamma_2^2} = -\sum_{I_1}\left\{\frac{H_i}{\gamma_2}+\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2}\right\}+\sum_{I_0}\pi_i^{(k)}\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}(\gamma_1-1)}{\gamma_1^2\gamma_2^2},$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{2k}} = -\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}x_{2ik}\eta_i p_i(1-p_i)}{\{1+\eta_i(1-p_i)\}^2},$$

$$\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} = \sum_{I_1}\left\{\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}\log(\gamma_2 t_1)}{\gamma_1^2\gamma_2}-\frac{H_i}{\gamma_1}\right\}+\sum_{I_0}\pi_i^{(k)}\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2}\left(1+\frac{\log(\gamma_2 t_1)}{\gamma_1}\right),$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_s} = 0,$$

$$\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_s} = 0,$$

for $s=1,2$, where $E_i=\frac{\log(\gamma_2 t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2}-\frac{\log(\gamma_2 t_i)}{\gamma_1^2}-\frac{1}{\gamma_1}$ and $H_i=\frac{1-(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1\gamma_2}$.

Destructive Poisson cure rate model

$$\begin{aligned}
Q(\theta,\pi^{(k)}) &\propto \sum_{I_1}\{\log(\eta_i p_i f(t_i))-\eta_i p_i F(t_i)\}-\sum_{I_0}(1-\pi_i^{(k)})\eta_i p_i \\
&\quad +\sum_{I_0}\pi_i^{(k)}\log\{\exp(-\eta_i p_i F(t_i))-\exp(-\eta_i p_i)\}, \text{ where } \pi_i^{(k)}=1-\exp(-\eta_i p_i S(t_i)).
\end{aligned}$$

$$\frac{\partial Q}{\partial\beta_{1l}} = \sum_{I_1}x_{1il}(1-\eta_i p_i F(t_i))-\sum_{I_0}(1-\pi_i^{(k)})x_{1il}\eta_i p_i-\sum_{I_0}\pi_i^{(k)}x_{1il}\eta_i p_i B_i,$$

$$\begin{aligned}
\frac{\partial Q}{\partial\beta_{2k}} &= \sum_{I_1}x_{2ik}(1-p_i)(1-\eta_i p_i F(t_i))-\sum_{I_0}(1-\pi_i^{(k)})x_{2ik}\eta_i p_i(1-p_i) \\
&\quad -\sum_{I_0}\pi_i^{(k)}x_{2ik}\eta_i p_i(1-p_i)B_i,
\end{aligned}$$

Under Weibull lifetime distribution

$$\frac{\partial Q}{\partial \gamma_1} = \sum_{I_1}\left(E_i - \eta_i p_i F_1'(t_i)\right) - \sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i F_1'(t_i)}{C_i},$$

$$\frac{\partial Q}{\partial \gamma_2} = \sum_{I_1}\left(H_i - \eta_i p_i F_2'(t_i)\right) - \sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i F_2'(t_i)}{C_i},$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{1l'}} &= -\sum_{I_1}x_{1il}x_{1il'}\eta_i p_i F(t_i) - \sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{1il'}\eta_i p_i \\ &\quad + \sum_{I_0}\pi_i^{(k)}x_{1il}x_{1il'}\eta_i p_i\left\{\eta_i p_i\left(B_i^* - B_i^2\right) + B_i\right\},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} &= -\sum_{I_1}x_{2ik}x_{2ik'}p_i(1-p_i)\left\{\eta_i(1-p_i)F(t_i) + 1 - \eta_i p_i F(t_i)\right\} \\ &\quad + \sum_{I_0}(1-\pi_i^{(k)})x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)(2p_i-1) \\ &\quad + \sum_{I_0}\pi_i^{(k)}x_{1il}x_{2ik}\eta_i p_i(1-p_i)^2\left\{\eta_i p_i(B_i^* - B_i^2) + \frac{(2p_i-1)B_i}{1-p_i}\right\},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\gamma_1^2} &= -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1} + \frac{1}{\gamma_1^2} + \frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4} + \eta_i p_i F_{11}''(t_i)\right\} \\ &\quad - \sum_{I_0}\pi_i^{(k)}\left\{\frac{\eta_i p_i F_{11}''(t_i)}{C_i} + \frac{\left(\eta_i p_i F_1'(t_i)\right)^2(1-C_i)}{C_i^2}\right\},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\gamma_2^2} &= -\sum_{I_1}\left\{\frac{H_i}{\gamma_2} + \frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2} + \eta_i p_i F_{22}''(t_i)\right\} \\ &\quad - \sum_{I_0}\pi_i^{(k)}\left\{\frac{\eta_i p_i F_{22}''(t_i)}{C_i} + \frac{\eta_i^2 p_i^2\left(F_2'(t_i)\right)^2(1-C_i)}{C_i^2}\right\},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{2k}} &= -\sum_{I_1}x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i) - \sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{2ik}\eta_i p_i(1-p_i) \\ &\quad + \sum_{I_0}\pi_i^{(k)}x_{1il}x_{2ik}\eta_i p_i(1-p_i)\left\{\eta_i p_i\left(B_i^* - B_i^2\right) - B_i\right\},\end{aligned}$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_s} = -\sum_{I_1}x_{1il}\eta_i p_i F_s'(t_i) - \sum_{I_0}\pi_i^{(k)}\frac{x_{1il}\eta_i p_i F_s'(t_i)}{C_i}\left\{1 - \frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_s} = -\sum_{I_1}x_{2ik}\eta_i p_i(1-p_i)F_s'(t_i) - \sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}\eta_i p_i(1-p_i)F_s'(t_i)}{C_i}\left\{1 - \frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\begin{aligned}\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} &= \sum_{I_1}\left\{\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}\log(\gamma_2 t_i)}{\gamma_1^3\gamma_2}-\frac{H_i}{\gamma_1}-\eta_i p_i F_{12}''(t_i)\right\}-\sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i F_{12}''(t_i)}{C_i}\\ &\quad+\sum_{I_0}\pi_i^{(k)}\frac{\eta_i^2 p_i^2 F_1'(t_i)F_2'(t_i)\,(C_i-1)}{C_i^2},\end{aligned}$$

for $s=1,2$, where E_i and H_i are as defined before with $B_i=\frac{S_{\text{pop}}(t_i)F(t_i)-p_{0i}}{S_{\text{pop}}(t_i)-p_{0i}}$, $B_i^*=\frac{S_{\text{pop}}(t_i)F^2(t_i)-p_{0i}}{S_{\text{pop}}(t_i)-p_{0i}}$, $C_i=1-\exp(-\eta_i p_i S(t_i))$,

$F_1'(t_i)=\frac{\partial F(t_i)}{\partial\gamma_1}=-\frac{S(t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}\log(\gamma_2 t_i)}{\gamma_1^2}$, $\quad F_2'(t_i)=\frac{\partial F(t_i)}{\partial\gamma_2}=\frac{S(t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1\gamma_2}$,

$F_{11}''(t_i)=\frac{\partial^2 F(t_i)}{\partial\gamma_1^2}=F_1'(t_i)\left(E_i-\frac{1}{\gamma_1}\right)$, $\quad F_{22}''(t_i)=\frac{\partial^2 F(t_i)}{\partial\gamma_2^2}=F_2'(t_i)\left(H_i-\frac{1}{\gamma_2}\right)$,

$F_{12}''(t_i)=\frac{\partial^2 F(t_i)}{\partial\gamma_1\partial\gamma_2}=F_2'(t_i)E_i$.

Destructive COM-Poisson cure rate model

$$\begin{aligned}\mathbf{Q}(\theta,\pi^{(k)}) &\propto \sum_{I_1}\left[\log(p_i f(t_i))+\log z_{f10}-\log\{z(1-p_i F(t_i))\}\right]\\ &\quad+\sum_{I_0}(1-\pi_i^{(k)})\,(\log z_p-\log z)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{z_f-z_p}{z}\right),\\ &\quad\text{where } \pi_i^{(k)}=\tfrac{z_f-z_p}{z_f}.\end{aligned}$$

$$\begin{aligned}\frac{\partial Q}{\partial\beta_{1l}} &= \sum_{I_1}x_{1il}\left\{\frac{z_{f20}}{z_{f10}}-\frac{z_1}{z}\right\}+\sum_{I_0}(1-\pi_i^{(k)})x_{1il}\left\{\frac{z_{p10}}{z_p}-\frac{z_1}{z}\right\}\\ &\quad+\sum_{I_0}\pi_i^{(k)}x_{1il}\left\{\frac{z_{f10}-z_{p10}}{z_f-z_p}-\frac{z_1}{z}\right\},\\ \frac{\partial Q}{\partial\beta_{2k}} &= \sum_{I_1}x_{2ik}(1-p_i)\left\{\frac{1}{1-p_iF(t_i)}-p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)\right\}\\ &\quad-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}p_i(1-p_i)z_{p11}}{z_p}-\sum_{I_0}\pi_i^{(k)}x_{2ik}p_i(1-p_i)\left\{\frac{F(t_i)z_{f11}-z_{p11}}{z_f-z_p}\right\},\\ \frac{\partial Q}{\partial\gamma_1} &= \sum_{I_1}\left\{E_i-p_iF_1'(t_i)\left(\frac{z_{f21}}{z_{f10}}-\frac{1}{1-p_iF(t_i)}\right)\right\}-\sum_{I_0}\pi_i^{(k)}\frac{p_iF_1'(t_i)z_{f11}}{z_f-z_p},\\ \frac{\partial Q}{\partial\phi} &= \sum_{I_1}\left\{\frac{1}{z_{f10}}z_{hf10}-\frac{1}{z}z_h\right\}+\sum_{I_0}(1-\pi_i^{(k)})\left\{\frac{1}{z_p}z_{hp}-\frac{1}{z}z_h\right\}\\ &\quad+\sum_{I_0}\pi_i^{(k)})\left\{\frac{1}{z_f-z_p}\left[z_{hf}-z_{hp}\right]-\frac{1}{z}z_h\right\},\end{aligned}$$

$$
\begin{aligned}
\frac{\partial Q}{\partial \gamma_2} &= \sum_{I_1}\left\{H_i - p_i F_2'(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{1-p_iF(t_i)}\right)\right\} - \sum_{I_0}\pi_i^{(k)}\frac{p_iF_2'(t_i)z_{f11}}{z_f - z_p},\\
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{1l'}} &= \sum_{I_1}x_{1il}x_{1il'}\left\{\frac{z_{f30}}{z_{f10}} - \left(\frac{z_{f20}}{z_{f10}}\right)^2 + \left(\frac{z_1}{z}\right)^2 - \frac{z_2}{z}\right\}\\
&\quad + \sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{1il'}\left\{\frac{z_{p20}}{z_p} - \left(\frac{z_{p10}}{z_p}\right)^2 + \left(\frac{z_1}{z}\right)^2 - \frac{z_2}{z}\right\}\\
&\quad + \sum_{I_0}\pi_i^{(k)}x_{1il}x_{1il'}\left\{\frac{z_{f20}-z_{p20}}{z_f-z_p} - \left(\frac{z_{f10}-z_{p10}}{z_f-z_p}\right)^2 + \left(\frac{z_1}{z}\right)^2 - \frac{z_2}{z}\right\},\\
\frac{\partial^2 Q}{\partial\phi^2} &= \sum_{I_1}\left\{\frac{z_{hhf10}}{z_{f10}} - \left(\frac{z_{hf10}}{z_{f10}}\right)^2 - \frac{z_{hh}}{z} + \left(\frac{z_h}{z}\right)^2\right\}\\
&\quad + \sum_{I_0}(1-\pi_i^{(k)})\left\{\frac{z_{hhp}}{z_p} - \left(\frac{z_{hp}}{z_p}\right)^2 - \frac{z_{hh}}{z} + \left(\frac{z_h}{z}\right)^2\right\}\\
&\quad + \sum_{I_0}\pi_i^{(k)}\left\{\frac{z_{hhf}-z_{hhp}}{z_f-z_p} - \left(\frac{z_{hf}-z_{hp}}{z_f-z_p}\right)^2 - \frac{z_{hh}}{z} + \left(\frac{z_h}{z}\right)^2\right\},\\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} &= \sum_{I_1}\frac{x_{2ik}x_{2ik'}(1-p_i)^2}{(1-p_iF(t_i))^2}\left\{p_iF(t_i) - \frac{p_i(1-p_iF(t_i))}{1-p_i}\right\}\\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i^2F(t_i)(1-p_i)z_{f21}}{z_{f10}}\\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_iF(t_i)(1-p_i)^2}{z_{f10}}\left\{p_iF(t_i)\left(z_{f22} - \frac{z_{f21}^2}{z_{f10}}\right) - z_{f21}\right\}\\
&\quad + \sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}(1-p_i)^2p_i(p_iz_{p12}-z_{p11})}{z_p}\\
&\quad + \sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)z_{p11}}{z_p}\left\{1-(1-p_i)\frac{z_{p11}}{z_p}\right\}\\
&\quad - \sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}x_{2ik'}p_i(1-p_i)^2}{z_f-z_p}\left\{F(t_i)z_{f11} - z_{p11} - p_i(F^2(t_i)z_{f12} - z_{p12})\right\}\\
&\quad + \sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)(F(t_i)z_{f11}-z_{p11})}{z_f-z_p}\\
&\quad - \sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)^2(F(t_i)z_{f11}-z_{p11})^2}{(z_f-z_p)^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \gamma_1^2} &= -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1}+\frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4}+\frac{1}{\gamma_1^2}+p_i F_{11}''(t_i)\frac{z_{f21}}{z_{f10}}\right\} \\
&\quad +\sum_{I_1}\left[p_i^2(F_1'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}}-\left(\frac{z_{f21}}{z_{f10}}\right)^2\right\}+\frac{p_i}{1-p_iF(t_i)}\left\{F_{11}''+\frac{p_i(F_1'(t_i))^2}{1-p_iF(t_i)}\right\}\right] \\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{p_i}{z_f-z_p}\left\{F_{11}''(t_i)z_{f11}-p_i\left(F_1'(t_i)\right)^2\left(z_{f12}-\frac{z_{f11}^2}{z_f-z_p}\right)\right\}, \\
\frac{\partial^2 Q}{\partial \gamma_2^2} &= -\sum_{I_1}\left\{\frac{H_i}{\gamma_2}+\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2}+p_i F_{22}''(t_i)\frac{z_{f21}}{z_{f10}}\right\} \\
&\quad +\sum_{I_1}\left[p_i^2(F_2'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}}-\left(\frac{z_{f21}}{z_{f10}}\right)^2\right\}+\frac{p_i}{1-p_iF(t_i)}\left\{F_{22}''+\frac{p_i(F_2'(t_i))^2}{1-p_iF(t_i)}\right\}\right] \\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{p_i}{z_f-z_p}\left\{F_{22}''(t_i)z_{f11}-p_i\left(F_2'(t_i)\right)^2\left(z_{f12}-\frac{z_{f11}^2}{z_f-z_p}\right)\right\}, \\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \beta_{2k}} &= \sum_{I_1}\frac{x_{1il}x_{2ik}p_iF(t_i)(1-p_i)}{z_{f10}}\left(\frac{z_{f21}z_{f20}}{z_{f10}}-z_{f31}\right) \\
&\quad +\sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{2ik}p_i(1-p_i)\left\{\frac{z_{p11}z_{p10}}{z_p^2}-\frac{z_{p21}}{z_p}\right\} \\
&\quad +\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}x_{2ik}p_i(1-p_i)}{(z_f-z_p)^2}\left[z_{f10}\left\{F(t_i)z_{f11}-z_{p11}\right\}-z_{p10}\left\{F(t_i)z_{f11}-z_{p11}\right\}\right] \\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}x_{2ik}p_i(1-p_i)(F(t_i)z_{f21}-z_{p21})}{z_f-z_p}, \\
\\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \gamma_s} &= \sum_{I_1}x_{1il}p_iF_s'(t_i)\left\{(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{z_{f31}}{z_{f10}}\right\} \\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}p_i(1-p_i)F_s'(t_i)z_{f11}z_{p11}}{(z_f-z_p)^2}+\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}p_i(1-p_iF(t_i))F_s'(t_i)z_{f11}^2}{(z_f-z_p)^2} \\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}p_i(1-p_i)F_s'(t_i)z_{f21}}{z_f-z_p}, \\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \phi} &= \sum_{I_1}x_{1il}\left\{\frac{z_{hf20}}{z_{f10}}-\frac{z_{f20}z_{hf10}}{(z_{f10})^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\} \\
&\quad +\sum_{I_0}(1-w_i^{(k)})x_{1il}\left\{\frac{z_{hp10}}{z_p}-\frac{z_{p10}z_{hp}}{(z_p)^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\} \\
&\quad +\sum_{I_0}w_i^{(k)}x_{1il}\left\{\frac{z_{hf10}-z_{hp10}}{z_f-z_p}-\frac{(z_{f10}-z_{p10})(z_{hf}-z_{hp})}{(z_f-z_p)^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_s} &= -\sum_{I_1} x_{2ik}p_iF_s'(t_i)(1-p_i)\left\{\frac{z_{f21}}{z_{f10}} - p_iF(t_i)\frac{z_{f22}}{z_{f10}}\right\} \\
&\quad -\sum_{I_1} x_{2ik}p_iF_s'(t_i)(1-p_i)\left\{p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2 - \frac{1}{(1-p_iF(t_i))^2}\right\} \\
&\quad -\sum_{I_0} \pi_i^{(k)}\frac{x_{2ik}p_i^2F_s'(t_i)(1-p_i)}{(z_f-z_p)^2}\left\{F(t_i)z_{f11}^2 - z_{f11}z_{p11}\right\} \\
&\quad +\sum_{I_0} \pi_i^{(k)}\frac{x_{2ik}p_iF_s'(t_i)(1-p_i)}{z_f-z_p}\left\{p_iF(t_i)z_{f12} - z_{f11}\right\}, \\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\phi} &= -\sum_{I_1} x_{2ik}(1-p_i)p_iF(t_i)\left(\frac{z_{hf21}}{z_{f10}} - \frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right) \\
&\quad -\sum_{I_0} (1-w_i^{(k)})x_{2ik}p_i(1-p_i)\left(\frac{z_{hp11}}{z_p} - \frac{z_{p11}z_{hp}}{(z_p)^2}\right) \\
&\quad -\sum_{I_0} w_i^{(k)}x_{2ik}p_i(1-p_i)\left\{\frac{F(t_i)z_{hf11} - z_{hp11}}{z_f-z_p} - \frac{\left(F(t_i)z_{f11}-z_{p11}\right)\left(z_{hf}-z_{hp}\right)}{(z_f-z_p)^2}\right\}, \\
\frac{\partial^2 Q}{\partial\phi\partial\gamma_s} &= -\sum_{I_1} p_iF_s'(t_i)\left(\frac{z_{hf21}}{z_{f10}} - \frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right) \\
&\quad -\sum_{I_0} w_i^{(k)}p_iF_s'(t_i)\left(\frac{z_{hf11}}{z_f-z_p} - \frac{z_{f11}(z_{hf}-z_{hp})}{(z_f-z_p)^2}\right),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} &= \sum_{I_1}\left\{\frac{p_iF_{12}''(t_i)}{(1-p_iF(t_i))} - \frac{H_i}{\gamma_1}\right\} + \sum_{I_1}\left\{\frac{(\gamma_2t_i)^{\frac{1}{\gamma_1}}\log(\gamma_2t_i)}{\gamma_1^3\gamma_2} + \frac{p_i^2F_1'(t_i)F_2'(t_i)}{(1-p_iF(t_i))^2}\right\} \\
&\quad +\sum_{I_1} p_i^2F_1'(t_i)F_2'(t_i)\frac{z_{f22}}{z_{f10}} - \sum_{I_1} p_iF_{12}''(t_i)\frac{z_{f21}}{z_{f10}} - \sum_{I_1} p_i^2F_1'(t_i)F_2'(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2 \\
&\quad +\sum_{I_0} \pi_i^{(k)}\frac{p_i^2F_1'(t_i)F_2'(t_i)}{z_f-z_p}\left\{z_{f12} - \frac{z_{f11}^2}{z_f-z_p}\right\} - \sum_{I_0} \pi_i^{(k)}\frac{p_iF_{12}''(t_i)z_{f11}}{z_f-z_p},
\end{aligned}
$$

for $s = 1,2$, where E_i, H_i, $F_1'(t_i)$, $F_2'(t_i)$, $F_{11}''(t_i)$, $F_{12}''(t_i)$ and $F_{22}''(t_i)$ are all as defined before with

$$
\begin{aligned}
&z = Z(\eta_i,\phi), && z_p = Z[\eta_i(1-p_i),\phi], && z_f = Z[\eta_i(1-p_iF(t_i)),\phi], \\
&z_h = \textstyle\sum_{j=2}^{\infty}\frac{\log(1/j!)\eta_i^j}{(j!)^\phi}, && z_{hp} = \textstyle\sum_{j=2}^{\infty}\frac{\log(1/j!)[\eta_i(1-p_i)]^j}{(j!)^\phi}, && z_{hf} = \textstyle\sum_{j=2}^{\infty}\frac{\log(1/j!)[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}, \\
&z_{hh} = \textstyle\sum_{j=2}^{\infty}\frac{[\log(1/j!)]^2\eta_i^j}{(j!)^\phi}, && z_{hhp} = \textstyle\sum_{j=2}^{\infty}\frac{[\log(1/j!)]^2[\eta_i(1-p_i)]^j}{(j!)^\phi}, && z_{hhf} = \textstyle\sum_{j=2}^{\infty}\frac{[\log(1/j!)]^2[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}, \\
&z_1 = \textstyle\sum_{j=1}^{\infty}\frac{j\eta_i^j}{(j!)^\phi}, && z_{p10} = \textstyle\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_i)]^j}{(j!)^\phi}, && z_{f10} = \textstyle\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}, \\
&z_{h1} = \textstyle\sum_{j=2}^{\infty}\frac{\log(1/j!)j\eta_i^j}{(j!)^\phi}, && z_{hp10} = \textstyle\sum_{j=2}^{\infty}\frac{j\log(1/j!)[\eta_i(1-p_i)]^j}{(j!)^\phi}, && z_{hf10} = \textstyle\sum_{j=2}^{\infty}\frac{j\log(1/j!)[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}, \\
&z_2 = \textstyle\sum_{j=1}^{\infty}\frac{j^2\eta_i^j}{(j!)^\phi}, && z_{p20} = \textstyle\sum_{j=1}^{\infty}\frac{j^2[\eta_i(1-p_i)]^j}{(j!)^\phi}, && z_{p21} = \textstyle\sum_{j=1}^{\infty}\frac{j^2\eta_i^j(1-p_i)^{j-1}}{(j!)^\phi}, \\
&z_{p11} = \textstyle\sum_{j=1}^{\infty}\frac{j\eta_i^j(1-p_i)^{j-1}}{(j!)^\phi}, && z_{f11} = \textstyle\sum_{j=1}^{\infty}\frac{j\eta_i^j(1-p_iF(t_i))^{j-1}}{(j!)^\phi}, && z_{p12} = \textstyle\sum_{j=2}^{\infty}\frac{j(j-1)\eta_i^j(1-p_i)^{j-2}}{(j!)^\phi},
\end{aligned}
$$

$$z_{hp11} = \sum_{j=2}^{\infty} \frac{j\log(1/j!)\eta_i^j(1-p_i)^{j-1}}{(j!)^\phi}, \qquad z_{hf11} = \sum_{j=2}^{\infty} \frac{j\log(1/j!)\eta_i^j(1-p_iF(t_i))^{j-1}}{(j!)^\phi},$$
$$z_{hhf10} = \sum_{j=1}^{\infty} \frac{j[\log(1/j!)]^2\eta_i^j(1-p_iF(t_i))^j}{(j!)^\phi}, \qquad z_{f20} = \sum_{j=1}^{\infty} \frac{j^2[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi},$$
$$z_{f21} = \sum_{j=1}^{\infty} \frac{j^2\eta_i^j(1-p_iF(t_i))^{j-1}}{(j!)^\phi}, \qquad z_{f12} = \sum_{j=2}^{\infty} \frac{j(j-1)\eta_i^j(1-p_iF(t_i))^{j-2}}{(j!)^\phi},$$
$$z_{f22} = \sum_{j=2}^{\infty} \frac{j^2(j-1)\eta_i^j(1-p_iF(t_i))^{j-2}}{(j!)^\phi}, \qquad z_{f31} = \sum_{j=1}^{\infty} \frac{j^3\eta_i^j(1-p_iF(t_i))^{j-1}}{(j!)^\phi},$$
$$z_{f30} = \sum_{j=1}^{\infty} \frac{j^3[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi}, \qquad z_{hf20} = \sum_{j=2}^{\infty} \frac{j^2\log(1/j!)[\eta_i(1-p_iF(t_i))]^j}{(j!)^\phi},$$
$$z_{hf21} = \sum_{j=2}^{\infty} \frac{j^2\eta_i^j\log(1/j!)(1-p_iF(t_i))^{j-1}}{(j!)^\phi}.$$

A.2 Expressions for the components of the observed information matrix

Destructive Bernoulli cure rate model

$$l(\theta) \propto \sum_{I_1} \log\left(\frac{\eta_i p_i f(t_i)}{1+\eta_i}\right) + \sum_{I_0} \log S_{\text{pop}}(t_i).$$

$$\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{1l'}} = -\sum_{I_1} \frac{x_{1il}x_{1il'}\eta_i}{(1+\eta_i)^2} + \sum_{I_0} x_{1il}x_{1il'}\left[\frac{\eta_i(1-p_iF(t_i))}{\{1+\eta_i(1-p_iF(t_i))\}^2} - \frac{\eta_i}{(1+\eta_i)^2}\right],$$

$$\frac{\partial^2 l}{\partial\beta_{2k}\beta_{2k'}} = -\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i) - \sum_{I_0} \frac{x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)F(t_i)}{1+\eta_i(1-p_iF(t_i))}\left\{1-2p_i+\frac{\eta_i p_i(1-p_i)F(t_i)}{1+\eta_i(1-p_iF(t_i))}\right\},$$

$$\frac{\partial^2 l}{\partial\gamma_1^2} = -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1} + \frac{1}{\gamma_1^2} + \frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4}\right\} - \sum_{I_0} \eta_i p_i\left[\frac{F_{11}''(t_i)}{1+\eta_i(1-p_iF(t_i))} + \frac{\eta_i p_i(F_1'(t_i))^2}{\{1+\eta_i(1-p_iF(t_i))\}^2}\right],$$

$$\frac{\partial^2 l}{\partial\gamma_2^2} = -\sum_{I_1}\left\{\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2} + \frac{H_i}{\gamma_2}\right\} - \sum_{I_0} \eta_i p_i\left[\frac{F_{22}''(t_i)}{1+\eta_i(1-p_iF(t_i))} + \frac{\eta_i p_i(F_2'(t_i))^2}{\{1+\eta_i(1-p_iF(t_i))\}^2}\right],$$

$$\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{2k}} = -\sum_{I_0} \frac{x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i)}{\{1+\eta_i(1-p_iF(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_s} = -\sum_{I_0} \frac{x_{1il}\eta_i p_i F_s'(t_i)}{\{1+\eta_i(1-p_iF(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_s} = -\sum_{I_0} \frac{x_{2ik}\eta_i p_i(1-p_i)F_s'(t_i)\,(1+\eta_i)}{\{1+\eta_i(1-p_iF(t_i))\}^2},$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \gamma_1\partial \gamma_2} &= \sum_{I_1}\left\{\frac{\log(\gamma_2 t_1)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^3\gamma_2} - \frac{H_i}{\gamma_1}\right\} \\ &\quad -\sum_{I_0}\frac{\eta_i p_i}{1+\eta_i(1-p_iF(t_i))}\left\{F_{12}''(t_i) + \frac{\eta_i p_i F_1'(t_i)F_2'(t_i)}{1+\eta_i(1-p_iF(t_i))}\right\}.\end{aligned}$$

Destructive Poisson cure rate model

$$l(\theta) \propto \sum_{I_1}\{\log(\eta_i p_i f(t_i)) - \eta_i p_i F(t_i)\} - \sum_{I_0}\eta_i p_i F(t_i).$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \beta_{1l'}} = -\sum_{I_1} x_{1il}x_{1il'}\eta_i p_i F(t_i) - \sum_{I_0} x_{1il}x_{1il'}\eta_i p_i F(t_i),$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \beta_{2k}\partial \beta_{2k'}} &= -\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i)\{\eta_i(1-p_i)F(t_i) + 1 - \eta_i p_i F(t_i)\} \\ &\quad -\sum_{I_0} x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)(1-2p)F(t_i),\end{aligned}$$

$$\frac{\partial^2 l}{\partial \gamma_1^2} = -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1} + \frac{1}{\gamma_1^2} + \frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4} + \eta_i p_i F_{11}''(t_i\right\} - \sum_{I_0}\eta_i p_i F_{11}''(t_i),$$

$$\frac{\partial^2 l}{\partial \gamma_2^2} = -\sum_{I_1}\left\{\frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2} + \frac{H_i}{\gamma_2} + \eta_i p_i F_{22}''(t_i)\right\} - \sum_{I_0}\eta_i p_i F_{22}''(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \beta_{2k}} = -\sum_{I_1} x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i) - \sum_{I_0} x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_s} = -\sum_{I_1} x_{1il}\eta_i p_i F_s'(t_i) - \sum_{I_0} x_{1il}\eta_i p_i F_s'(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_s} = -\sum_{I_1} x_{2ik}\eta_i p_i(1-p_i)F_s'(t_i) - \sum_{I_0} x_{2ik}\eta_i p_i(1-p_i)F_s'(t_i),$$

$$\frac{\partial^2 l}{\partial\gamma_1\partial\gamma_2} = \sum_{I_1}\left\{\frac{\log(\gamma_2 t_i)(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^3\gamma_2} - \frac{H_i}{\gamma_1} - \eta_i p_i F_{12}''(t_i)\right\} - \sum_{I_0}\eta_i p_i F_{12}''(t_i).$$

Destructive COM-Poisson cure rate model

$$\begin{aligned}
l(\boldsymbol{\theta}) &\propto \sum_{I_1}\left[\log(p_i f(t_i)) + \log z_{f10} - \log\{z(1-p_i F(t_i))\}\right] + \sum_{I_0}\log S_{\text{pop}}(t_i).\\
\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{1l'}} &= \sum_{I_1} x_{1il}x_{1il'}\left\{\frac{z_{f30}}{z_{f10}} - \left(\frac{z_{f20}}{z_{f10}}\right)^2 + \left(\frac{z_1}{z}\right)^2 - \frac{z_2}{z}\right\}\\
&\quad + \sum_{I_0}\frac{x_{1il}x_{1il'}}{z}\left\{\frac{z_{f20}}{z_f} - \frac{z_2}{z} + \left(\frac{z_1}{z}\right)^2 - \left(\frac{z_{f10}}{z_f}\right)^2\right\},\\
\frac{\partial^2 l}{\partial\beta_{2k}\partial\beta_{2k'}} &= \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i(1-p_i)^2}{(1-p_i F(t_i))^2}\left\{F(t_i) - \frac{1-p_i F(t_i)}{1-p_i}\right\}\\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i^2 F(t_i)(1-p_i)z_{f21}}{z_{f10}}\\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i F(t_i)(1-p_i)^2}{z_{f10}}\left\{p_i F(t_i)\left(z_{f22} - \frac{z_{f21}^2}{z_{f10}}\right) - z_{f21}\right\}\\
&\quad + \sum_{I_0}x_{2ik}x_{2ik'}p_i(1-p_i)F(t_i)\left[p_i(1-p_i)F(t_i)\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{(1-2p_i)z_{f11}}{z_f}\right],\\
\frac{\partial^2 l}{\partial\gamma_1^2} &= -\sum_{I_1}\left\{\frac{2E_i}{\gamma_1} + \frac{(\log(\gamma_2 t_i))^2(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^4} + \frac{1}{\gamma_1^2} + p_i F_{11}''(t_i)\frac{z_{f21}}{z_{f10}}\right\}\\
&\quad + \sum_{I_1}\left[p_i^2(F_1'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}} - \left(\frac{z_{f21}}{z_{f10}}\right)^2\right\} + \frac{p_i}{1-p_i F(t_i)}\left(F_{11}'' + \frac{p_i(F_1'(t_i))^2}{1-p_i F(t_i)}\right)\right]\\
&\quad + \sum_{I_0}p_i\left[p_i\left(F_1'(t_i)\right)^2\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{F_{11}''(t_i)z_{f11}}{z_f}\right],\\
\frac{\partial^2 l}{\partial\gamma_2^2} &= -\sum_{I_1}\left\{\frac{H_i}{\gamma_2} + \frac{(\gamma_2 t_i)^{\frac{1}{\gamma_1}}}{\gamma_1^2\gamma_2^2} + p_i F_{22}''(t_i)\frac{z_{f21}}{z_{f10}}\right\}\\
&\quad + \sum_{I_1}\left[p_i^2(F_2'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}} - \left(\frac{z_{f21}}{z_{f10}}\right)^2\right\} + \frac{p_i}{1-p_i F(t_i)}\left(F_{22}'' + \frac{p_i(F_2'(t_i))^2}{1-p_i F(t_i)}\right)\right]\\
&\quad + \sum_{I_0}p_i\left[p_i\left(F_2'(t_i)\right)^2\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{F_{22}''(t_i)z_{f11}}{z_f}\right],
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 l}{\partial \phi^2} &= \sum_{I_1}\left\{\frac{z_{hhf10}}{z_{f10}}-\left(\frac{z_{hf10}}{z_{f10}}\right)^2-\frac{z_{hh}}{z}+\left(\frac{z_h}{z}\right)^2\right\} \\
&\quad+\sum_{I_0}\left\{\frac{z_{hhf}}{z_f}-\left(\frac{z_{hp}}{z_p}\right)^2-\frac{z_{hh}}{z}+\left(\frac{z_h}{z}\right)^2\right\}, \\
\frac{\partial l}{\partial \beta_{1l}\partial \beta_{2k}} &= \sum_{I_1}\frac{x_{1il}x_{2ik}p_iF(t_i)(1-p_i)}{z_{f10}}\left(\frac{z_{f21}z_{f20}}{z_{f10}}-z_{f31}\right) \\
&\quad+\sum_{I_0}x_{1il}x_{2ik}p_i(1-p_i)F(t_i)\left(\frac{z_{f10}z_{f11}}{z_f^2}-\frac{z_{f21}}{z_f}\right), \\
\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_s} &= -\sum_{I_1}x_{1il}p_iF_s'(t_i)\frac{z_{f31}}{z_{f10}}+\sum_{I_1}x_{1il}p_iF_s'(t_i)(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2 \\
&\quad+\sum_{I_0}x_{1il}p_iF_s'(t_i)\left(\frac{z_{f11}z_{f10}}{z_f^2}-\frac{z_{f21}}{z_f}\right), \\
\frac{\partial^2 l}{\partial \beta_{1l}\partial \phi} &= \sum_{I_1}x_{1il}\left\{\frac{z_{hf20}}{z_{f10}}-\frac{z_{f20}z_{hf10}}{(z_{f10})^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\} \\
&\quad+\sum_{I_0}x_{1il}\left\{\frac{z_{hf10}}{z_f}-\frac{z_{f10}z_{hf}}{(z_p)^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\}, \\
\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_s} &= -\sum_{I_1}x_{2ik}p_iF_s'(t_i)(1-p_i)\left\{\frac{z_{f21}}{z_{f10}}-p_iF(t_i)\frac{z_{f22}}{z_{f10}}\right\} \\
&\quad-\sum_{I_1}x_{2ik}p_iF_s'(t_i)(1-p_i)\left\{p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{1}{(1-p_iF(t_i))^2}\right\} \\
&\quad+\sum_{I_0}x_{2ik}p_i(1-p_i)F_s'(t_i)\left[p_iF(t_i)\left\{\frac{z_{f12}}{z_f}-\left(\frac{z_{f11}}{z_f}\right)^2\right\}-\frac{z_{f11}}{z_f}\right], \\
\frac{\partial^2 l}{\partial \beta_{2k}\partial \phi} &= -\sum_{I_1}x_{2ik}(1-p_i)p_iF(t_i)\left(\frac{z_{hf21}}{z_{f10}}-\frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right) \\
&\quad-\sum_{I_0}x_{2ik}p_i(1-p_i)\left(\frac{z_{hf11}}{z_f}-\frac{z_{f11}z_{hf}}{(z_f)^2}\right), \\
\frac{\partial^2 l}{\partial \phi\partial \gamma_s} &= -\sum_{I_1}p_iF_s'(t_i)\left(\frac{z_{hf21}}{z_{f10}}-\frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right) \\
&\quad-\sum_{I_0}p_iF_s'(t_i)\left(\frac{z_{hf11}}{z_f}-\frac{z_{f11}z_{hf}}{(z_f)^2}\right), \\
\frac{\partial^2 l}{\partial \gamma_1\partial \gamma_2} &= \sum_{I_1}\left\{\frac{p_iF_{12}''}{1-p_iF(t_i)}-\frac{H_i}{\gamma_1}+\frac{(\gamma_2t_i)^{\frac{1}{\gamma_1}}\log(\gamma_2t_i)}{\gamma_1^3\gamma_2}+\frac{p_i^2F_1'(t_i)F_2'(t_i)}{(1-p_iF(t_i))^2}\right\} \\
&\quad+\sum_{I_1}p_i^2F_1'(t_i)F_2'(t_i)\frac{z_{f22}}{z_{f10}}-\sum_{I_1}p_iF_{12}''\frac{z_{f21}}{z_{f10}}-\sum_{I_1}p_i^2F_1'(t_i)F_2'(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2 \\
&\quad+\sum_{I_0}p_i\left[p_iF_1'(t_i)F_2'(t_i)\left\{\frac{z_{f12}}{z_f}-\left(\frac{z_{f11}}{z_f}\right)^2\right\}-\frac{F_{12}''(t_i)z_{f11}}{z_f}\right].
\end{aligned}
$$

Appendix B

Under lognormal lifetime distribution

B.1 First- and second-order derivatives of the Q-function and

Destructive Bernoulli cure rate model

$$Q(\theta,\pi^{(k)}) \propto \sum_{I_1}\log\left(\frac{\eta_i p_i f(t_i)}{1+\eta_i}\right)+\sum_{I_0}(1-\pi_i^{(k)})\log\left(\frac{1+\eta_i(1-p_i)}{1+\eta_i}\right)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{\eta_i p_i S(t_i)}{1+\eta_i}\right).$$

$$\frac{\partial Q}{\partial \beta_{1l}} = \sum_{I_1}\frac{x_{1il}}{1+\eta_i}-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}\eta_i p_i}{(1+\eta_i)\{1+\eta_i(1-p_i))\}}+\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}}{1+\eta_i},$$

$$\frac{\partial Q}{\partial \beta_{2k}} = \sum_{I_1}x_{2ik}(1-p_i)-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}\eta_i p_i(1-p_i)}{1+\eta_i(1-p_i)}+\sum_{I_0}\pi_i^{(k)}x_{2ik}(1-p_i),$$

$$\frac{\partial Q}{\partial \gamma_1} = \sum_{I_1}\frac{a_i^2-1}{\gamma_1}-\sum_{I_0}\pi_i^{(k)}\frac{F_1'(t_i)}{S(t_i)},$$

$$\frac{\partial Q}{\partial \gamma_2} = -\sum_{I_1}\frac{a_i}{\gamma_1\gamma_2}-\sum_{I_0}\pi_i^{(k)}\frac{F_2'(t_i)}{S(t_i)},$$

$$\frac{\partial^2 Q}{\partial \beta_{1l}\partial \beta_{1l'}} = -\sum_{I_1}\frac{x_{1il}x_{1il'}\eta_i}{(1+\eta_i)^2}-\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}x_{1il'}\eta_i p_i\{1-\eta_i^2(1-p_i)\}}{(1+\eta_i)^2\{1+\eta_i(1-p_i)\}^2}-\sum_{I_0}\pi_i^{(k)}\frac{x_{1il}x_{1il'}\eta_i}{(1+\eta_i)^2},$$

Under lognormal lifetime distribution

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{2k}\beta_{2k'}} &= -\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i) \\
&\quad -\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)^2[1+\eta_i-\frac{p}{1-p}\{1+\eta_i(1-p_i)\}]}{\{1+\eta_i(1-p_i)\}^2} \\
&\quad -\sum_{I_0}\pi_i^{(k)}x_{2ik}x_{2ik'}p_i(1-p_i),
\end{aligned}$$

$$\frac{\partial^2 Q}{\partial\gamma_1^2} = -\sum_{I_1}\frac{3a_i^2-1}{\gamma_1^2} - \sum_{I_0}\pi_i^{(k)}\left\{\frac{F_{11}''(t_i)}{S(t_i)} + \frac{(F_1'(t_i))^2}{S^2(t_i)}\right\},$$

$$\frac{\partial^2 Q}{\partial\gamma_2^2} = -\sum_{I_1}\frac{1-a_i\gamma_1}{\gamma_1^2\gamma_2^2} - \sum_{I_0}\pi_i^{(k)}\left\{\frac{F_{22}''(t_i)}{S(t_i)} + \frac{(F_2'(t_i))^2}{S^2(t_i)}\right\},$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{2k'}} = -\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{1il}x_{2ik'}\eta_i p_i(1-p_i)}{\{1+\eta_i(1-p_i)\}^2},$$

$$\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} = \sum_{I_1}\frac{2a_i}{\gamma_1^2\gamma_2} - \sum_{I_0}\pi_i^{(k)}\left\{\frac{F_{12}''(t_i)}{S(t_i)} + \frac{F_1'(t_i)F_2'(t_i)}{S^2(t_i)}\right\}.$$

Destructive Poisson cure rate model

$$\begin{aligned}
Q(\theta,\pi^{(k)}) &\propto \sum_{I_1}\{\log(\eta_i p_i f(t_i)) - \eta_i p_i F(t_i)\} - \sum_{I_0}(1-\pi_i^{(k)})\eta_i p_i \\
&\quad +\sum_{I_0}\pi_i^{(k)}\log\left[\exp(-\eta_i p_i F(t_i)) - \exp(-\eta_i p_i)\right].
\end{aligned}$$

$$\frac{\partial Q}{\partial\beta_{1l}} = \sum_{I_1}x_{1il}(1-\eta_i p_i F(t_i)) - \sum_{I_0}(1-\pi_i^{(k)})x_{1il}\eta_i p_i - \sum_{I_0}\pi_i^{(k)}x_{1il}\eta_i p_i B_i,$$

$$\frac{\partial Q}{\partial\beta_{2k}} = \sum_{I_1}x_{2ik}(1-p_i)(1-\eta_i p_i F(t_i)) - \sum_{I_0}(1-\pi_i^{(k)})x_{2ik}\eta_i p_i(1-p_i) - \sum_{I_0}\pi_i^{(k)}x_{2ik}\eta_i p_i(1-p_i)B_i,$$

$$\frac{\partial Q}{\partial\gamma_1} = \sum_{I_1}\left(\frac{a_i^2-1}{\gamma_1} - \eta_i p_i F_1'(t_i)\right) - \sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i F_1'(t_i)}{C_i},$$

$$\frac{\partial Q}{\partial\gamma_2} = -\sum_{I_1}\left(\frac{a_i}{\gamma_1\gamma_2} + \eta_i p_i F_2'(t_i)\right) - \sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i F_2'(t_i)}{C_i},$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{1l'}} &= -\sum_{I_1}x_{1il}x_{1il'}\eta_i p_i F(t_i) - \sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{1il'}\eta_i p_i \\
&\quad -\sum_{I_0}\pi_i^{(k)}x_{1il}x_{1il'}\eta_i p_i\left\{\eta_i p_i(B_i^2-B_i^*)+B_i\right\},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} &= -\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i)\left\{\eta_i(1-p_i)F(t_i)+1-\eta_i p_i F(t_i)\right\} \\
&\quad +\sum_{I_0}(1-\pi_i^{(k)})x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)(2p_i-1) \\
&\quad +\sum_{I_0}\pi_i^{(k)}x_{1il}x_{2ik}\eta_i p_i(1-p_i)^2\left\{\eta_i p_i\left(B_i^*-B_i^2\right)+\frac{(2p_i-1)B_i}{1-p_i}\right\},
\end{aligned}$$

$$\frac{\partial^2 Q}{\partial\gamma_1^2} = -\sum_{I_1}\left(\frac{3a_i^2-1}{\gamma_1^2}+\eta_i p_i F_{11}''(t_i)\right)-\sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i}{C_i}\left\{F_{11}''(t_i)+\frac{\eta_i p_i\left(F_1'(t_i)\right)^2(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\gamma_2^2} = -\sum_{I_1}\left(\frac{1-a_i\gamma_1}{\gamma_1^2\gamma_2^2}+\eta_i p_i F_{22}''(t_i)\right)-\sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i}{C_i}\left\{F_{22}''(t_i)+\frac{\eta_i p_i\left(F_2'(t_i)\right)^2(1-C_i)}{C_i}\right\},$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{2k}} &= -\sum_{I_1}x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i)-\sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{2ik}\eta_i p_i(1-p_i) \\
&\quad -\sum_{I_0}\pi_i^{(k)}x_{1il}x_{2ik}\eta_i p_i(1-p_i)\left\{\eta_i p_i(B_i^2-B_i^*)+B_i\right\},
\end{aligned}$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_1} = -\sum_{I_1}x_{1il}\eta_i p_i F_1'(t_i)-\sum_{I_0}w_i^{(k)}\frac{x_{1il}\eta_i p_i F_1'(t_i)}{C_i}\left\{1-\frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_2} = -\sum_{I_1}x_{1il}\eta_i p_i F_2'(t_i)-\sum_{I_0}w_i^{(k)}\frac{x_{1il}\eta_i p_i F_2'(t_i)}{C_i}\left\{1-\frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_1} = -\sum_{I_1}x_{2ik}\eta_i p_i(1-p_i)F_1'(t_i)-\sum_{I_0}\frac{\pi_i^{(k)}x_{2ik}\eta_i p_i(1-p_i)F_1'(t_i)}{C_i}\left\{1-\frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_2} = -\sum_{I_1}x_{2ik}\eta_i p_i(1-p_i)F_2'(t_i)-\sum_{I_0}\frac{w^{(k)}x_{2ik}\eta_i p_i(1-p_i)F_2'(t_i)}{C_i}\left\{1-\frac{\eta_i p_i S(t_i)(1-C_i)}{C_i}\right\},$$

$$\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} = \sum_{I_1}\left(\frac{2a_i}{\gamma_1^2\gamma_2}-\eta_i p_i F_{12}''(t_i)\right)-\sum_{I_0}\pi_i^{(k)}\frac{\eta_i p_i}{C_i}\left\{F_{12}''(t_i)+\frac{\eta_i p_i F_1'(t_i)F_2'(t_i)(1-C_i)}{C_i}\right\}.$$

Destructive COM-Poisson cure rate model

$$\begin{aligned}
\mathbf{Q}(\theta,\pi^{(k)}) &\propto \sum_{I_1}\left[\log\left(p_i f(t_i)\right)+\log\left\{\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_i F(t_i))]^j}{(j!)^{\phi}}\right\}\right] \\
&\quad -\sum_{I_1}\log\{z(1-p_i F(t_i))\}+\sum_{I_0}(1-\pi_i^{(k)})\left(\log z_p-\log z\right)+\sum_{I_0}\pi_i^{(k)}\log\left(\frac{z_f-z_p}{z}\right).
\end{aligned}$$

Under lognormal lifetime distribution

$$
\begin{aligned}
\frac{\partial Q}{\partial \beta_{1l}} &= \sum_{I_1} x_{1il}\left(\frac{z_{f20}}{z_{f10}}-\frac{z_1}{z}\right)+\sum_{I_0}(1-\pi_i^{(k)})x_{1il}\left(\frac{z_{p10}}{z_p}-\frac{z_1}{z}\right)+\sum_{I_0} w_i^{(k)} x_{1il}\left(\frac{z_{f10}-z_{p10}}{z_f-z_p}-\frac{z_1}{z}\right),\\
\frac{\partial Q}{\partial \beta_{2k}} &= \sum_{I_1} x_{2ik}(1-p_i)\left\{\frac{1}{1-p_iF(t_i)}-p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)\right\}\\
&\quad -\sum_{I_0}\frac{(1-w_i^{(k)})x_{2ik}p_i(1-p_i)z_{p11}}{z_p}-\sum_{I_0} w_i^{(k)}x_{2ik}p_i(1-p_i)\left(\frac{F(t_i)z_{f11}-z_{p11}}{z_f-z_p}\right),\\
\frac{\partial Q}{\partial \gamma_1} &= \sum_{I_1}\left[\frac{a_i^2-1}{\gamma_1}-p_iF_1'(t_i)\left\{\frac{z_{f21}}{z_{f10}}-\frac{1}{(1-p_iF(t_i))}\right\}\right]-\sum_{I_0} w_i^{(k)}\frac{p_iF_1'(t_i)z_{f11}}{z_f-z_p},\\
\frac{\partial Q}{\partial \gamma_2} &= -\sum_{I_1}\left[\frac{a_i}{\gamma_1\gamma_2}+p_iF_2'(t_i)\left\{\frac{z_{f21}}{z_{f10}}-\frac{1}{(1-p_iF(t_i))}\right\}\right]-\sum_{I_0} w_i^{(k)}\frac{p_iF_2'(t_i)z_{f11}}{z_f-z_p},\\
\frac{\partial Q}{\partial \phi} &= \sum_{I_1}\left\{\frac{1}{z_{f10}}z_{hf10}-\frac{1}{z}z_h\right\}+\sum_{I_0}(1-\pi_i^{(k)})\left\{\frac{1}{z_p}z_{hp}-\frac{1}{z}z_h\right\}\\
&\quad +\sum_{I_0}\pi_i^{(k)})\left\{\frac{1}{z_f-z_p}\left[z_{hf}-z_{hp}\right]-\frac{1}{z}z_h\right\},\\
\frac{\partial^2 Q}{\partial \beta_{1il}\partial \beta_{1il'}} &= \sum_{I_1} x_{1il}x_{1il'}\left\{\frac{z_{f30}}{z_{f10}}-\left(\frac{z_{f20}}{z_{f10}}\right)^2+\left(\frac{z_1}{z}\right)^2-\frac{z_2}{z}\right\}\\
&\quad +\sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{1il'}\left\{\frac{z_{p20}}{z_p}-\left(\frac{z_{p10}}{z_p}\right)^2+\left(\frac{z_1}{z}\right)^2-\frac{z_2}{z}\right\}\\
&\quad +\sum_{I_0} w_i^{(k)}x_{1il}x_{1il'}\left\{\frac{z_{f20}-z_{p20}}{z_f-z_p}-\left(\frac{z_{f10}-z_{p10}}{z_f-z_p}\right)^2+\left(\frac{z_1}{z}\right)^2-\frac{z_2}{z}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \beta_{2ik}\partial \beta_{2ik'}} &= \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i(1-p_i)^2}{(1-p_iF(t_i))^2}\left(F(t_i)-\frac{1-p_iF(t_i)}{1-p_i}\right)\\
&\quad +\sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)F(t_i)z_{f21}}{z_{f10}}\\
&\quad +\sum_{I_1}\frac{x_{2ik}x_{2ik'}p_iF(t_i)(1-p_i)^2}{z_{f10}}\left\{p_iF(t_i)\left(z_{f22}-\frac{z_{f21}^2}{z_{f10}}\right)-z_{f21}\right\}\\
&\quad +\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}(1-p_i)^2p_i(p_iz_{p12}-z_{p11})}{z_p}\\
&\quad +\sum_{I_0}(1-\pi_i^{(k)})\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)z_{p11}}{z_p}\left\{1-(1-p_i)\frac{z_{p11}}{z_p}\right\}\\
&\quad -\sum_{I_0}\frac{\pi_i^{(k)}x_{2ik}x_{2ik'}p_i(1-p_i)^2}{z_f-z_p}\left[F(t_i)z_{f11}-z_{p11}-p_i\{F^2(t_i)z_{f12}-z_{p12}\}\right]\\
&\quad +\sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)(F(t_i)z_{f11}-z_{p11})}{z_f-z_p}\\
&\quad -\sum_{I_0}\pi_i^{(k)}\frac{x_{2ik}x_{2ik'}p_i^2(1-p_i)^2(F(t_i)z_{f11}-z_{p11})^2}{(z_f-z_p)^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \gamma_1^2} &= -\sum_{I_1}\left\{\frac{3a_i^2-1}{\gamma_1^2}+p_iF_{11}''(t_i)\left(\frac{z_{f21}}{z_{f10}}-\frac{1}{1-p_iF(t_i)}\right)\right\}\\
&\quad+\sum_{I_1}p_i^2(F_1'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}}-\left(\frac{z_{f21}}{z_{f10}}\right)^2+\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad-\sum_{I_0}w_i^{(k)}\left\{\frac{p_iF_{11}''(t_i)z_{f11}}{z_f-z_p}-\frac{p_i^2(F_1'(t_i))^2z_{f12}}{z_f-z_p}+\frac{p_i^2(F_1'(t_i))^2z_{f11}^2}{\left(z_f-z_p\right)^2}\right\},\\
\frac{\partial^2 Q}{\partial \gamma_2^2} &= -\sum_{I_1}\left\{\frac{1-a_i\gamma_1}{\gamma_1^2\gamma_2^2}+p_iF_{22}''(t_i)\left(\frac{z_{f21}}{z_{f10}}-\frac{1}{1-p_iF(t_i)}\right)\right\}\\
&\quad+\sum_{I_1}p_i^2(F_2'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}}-\left(\frac{z_{f21}}{z_{f10}}\right)^2+\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad-\sum_{I_0}w_i^{(k)}\left\{\frac{p_iF_{22}''(t_i)z_{f11}}{z_f-z_p}-\frac{p_i^2(F_2'(t_i))^2z_{f12}}{z_f-z_p}+\frac{p_i^2(F_2'(t_i))^2z_{f11}^2}{\left(z_f-z_p\right)^2}\right\},\\
\frac{\partial^2 Q}{\partial \phi^2} &= \sum_{I_1}\left\{\frac{z_{hhf10}}{z_{f10}}-\left(\frac{z_{hf10}}{z_{f10}}\right)^2-\frac{z_{hh}}{z}+\left(\frac{z_h}{z}\right)^2\right\}\\
&\quad+\sum_{I_0}(1-\pi_i^{(k)})\left\{\frac{z_{hhp}}{z_p}-\left(\frac{z_{hp}}{z_p}\right)^2-\frac{z_{hh}}{z}+\left(\frac{z_h}{z}\right)^2\right\}\\
&\quad+\sum_{I_0}\pi_i^{(k)}\left\{\frac{z_{hhf}-z_{hhp}}{z_f-z_p}-\left(\frac{z_{hf}-z_{hp}}{z_f-z_p}\right)^2-\frac{z_{hh}}{z}+\left(\frac{z_h}{z}\right)^2\right\},\\
\frac{\partial Q}{\partial \beta_{1l}\partial \beta_{2k}} &= \sum_{I_1}\frac{x_{1il}x_{2ik}p_iF(t_i)(1-p_i)}{z_{f10}}\left(\frac{z_{f21}z_{f20}}{z_{f10}}-z_{f31}\right)\\
&\quad+\sum_{I_0}(1-\pi_i^{(k)})x_{1il}x_{2ik}p_i(1-p_i)\left(\frac{z_{p11}z_{p10}}{z_p^2}-\frac{z_{p21}}{z_p}\right)\\
&\quad+\sum_{I_0}\frac{w_i^{(k)}x_{1il}x_{2ik}p_i(1-p_i)}{\left(z_f-z_p\right)^2}\left[z_{f10}\left\{F(t_i)z_{f11}-z_{p11}\right\}-z_{p10}\left\{F(t_i)z_{f11}-z_{p11}\right\}\right]\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{1il}x_{2ik}p_i(1-p_i)(F(t_i)z_{f21}-z_{p21})}{z_f-z_p},\\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \gamma_1} &= \sum_{I_1}x_{1il}p_iF_1'(t_i)\left\{(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{z_{f31}}{z_{f10}}\right\}\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_i)F_1'(t_i)z_{f11}z_{p11}}{\left(z_f-z_p\right)^2}+\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_iF(t_i))F_1'(t_i)z_{f11}^2}{\left(z_f-z_p\right)^2}\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_i)F_1'(t_i)z_{f21}}{z_f-z_p},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_2} &= \sum_{I_1} x_{1il}p_iF_2'(t_i)\left\{(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{z_{f31}}{z_{f10}}\right\}\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_i)F_2'(t_i)z_{f11}z_{p11}}{\left(z_f-z_p\right)^2}+\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_iF(t_i))F_2'(t_i)z_{f11}^2}{\left(z_f-z_p\right)^2}\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{1il}p_i(1-p_i)F_2'(t_i)z_{f21}}{z_f-z_p},\\
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\phi} &= \sum_{I_1} x_{1il}\left\{\frac{z_{hf20}}{z_{f10}}-\frac{z_{f20}z_{hf10}}{(z_{f10})^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\}\\
&\quad+\sum_{I_0}(1-\pi_i^{(k)})x_{1il}\left\{\frac{z_{hp10}}{z_p}-\frac{z_{p10}z_{hp}}{(z_p)^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\}\\
&\quad+\sum_{I_0}w_i^{(k)}x_{1il}\left\{\frac{z_{hf10}-z_{hp10}}{z_f-z_p}-\frac{(z_{f10}-z_{p10})(z_{hf}-z_{hp})}{(z_f-z_p)^2}-\frac{z_{h1}}{z}+\frac{z_1z_h}{z^2}\right\},\\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_1} &= -\sum_{I_1} x_{2ik}p_iF_1'(t_i)(1-p_i)\left\{\frac{z_{f21}}{z_{f10}}-p_iF(t_i)\frac{z_{f22}}{z_{f10}}\right\}\\
&\quad-\sum_{I_1} x_{2ik}p_iF_1'(t_i)(1-p_i)\left\{p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad-\sum_{I_0}\frac{w_i^{(k)}x_{2ik}p_i^2F_1'(t_i)(1-p_i)}{(z_f-z_p)^2}\left\{F(t_i)\left(z_{f11}\right)^2-z_{f11}z_{p11}\right\}\\
&\quad+\sum_{I_0}\frac{w_i^{(k)}x_{2ik}p_iF_1'(t_i)(1-p_i)}{(z_f-z_p)}\left\{p_iF(t_i)z_{f12}-z_{f11}\right\},\\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\phi} &= -\sum_{I_1} x_{2ik}(1-p_i)p_iF(t_i)\left(\frac{z_{hf21}}{z_{f10}}-\frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right)\\
&\quad-\sum_{I_0}(1-w_i^{(k)})x_{2ik}p_i(1-p_i)\left(\frac{z_{hp11}}{z_p}-\frac{z_{p11}z_{hp}}{(z_p)^2}\right)\\
&\quad-\sum_{I_0}w_i^{(k)}x_{2ik}p_i(1-p_i)\left\{\frac{F(t_i)z_{hf11}-z_{hp11}}{z_f-z_p}-\frac{\left(F(t_i)z_{f11}-z_{p11}\right)\left(z_{hf}-z_{hp}\right)}{(z_f-z_p)^2}\right\},\\
\frac{\partial^2 Q}{\partial\phi\partial\gamma_s} &= -\sum_{I_1} p_iF_s'(t_i)\left(\frac{z_{hf21}}{z_{f10}}-\frac{z_{f21}z_{hf10}}{(z_{f10})^2}\right)\\
&\quad-\sum_{I_0}w_i^{(k)}p_iF_s'(t_i)\left(\frac{z_{hf11}}{z_f-z_p}-\frac{z_{f11}(z_{hf}-z_{hp})}{(z_f-z_p)^2}\right),
\end{aligned}
$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_2} &= -\sum_{I_1} x_{2ik} p_i F_2'(t_i)(1-p_i)\left\{\frac{z_{f21}}{z_{f10}} - p_i F(t_i)\frac{z_{f22}}{z_{f10}}\right\} \\
&\quad -\sum_{I_1} x_{2ik} p_i F_2'(t_i)(1-p_i)\left\{p_i F(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2 - \frac{1}{(1-p_i F(t_i))^2}\right\} \\
&\quad -\sum_{I_0} \frac{w_i^{(k)} x_{2ik} p_i^2 F_2'(t_i)(1-p_i)}{(z_f - z_p)^2}\left\{F(t_i)\left(z_{f11}\right)^2 - z_{f11}z_{p11}\right\} \\
&\quad +\sum_{I_0} \frac{w_i^{(k)} x_{2ik} p_i F_2'(t_i)(1-p_i)}{(z_f - z_p)}\left\{p_i F(t_i) z_{f12} - z_{f11}\right\},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 Q}{\partial\gamma_1\partial\gamma_2} &= \sum_{I_1}\left\{\frac{2a_i}{\gamma_1^2\gamma_2} - p_i F_{12}''(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{1-p_i F(t_i)}\right)\right\} \\
&\quad +\sum_{I_1} p_i^2 F_1'(t_i) F_2'(t_i)\left\{\frac{z_{f22}}{z_{f10}} - \left(\frac{z_{f21}}{z_{f10}}\right)^2 + \frac{1}{(1-p_i F(t_i))^2}\right\} \\
&\quad -\sum_{I_0} w_i^{(k)}\left\{\frac{p_i F_{12}''(t_i) z_{f11}}{z_f - z_p} - \frac{p_i^2 F_2'(t_i) F_2'(t_i) z_{f12}}{z_f - z_p} + \frac{p_i^2 F_2'(t_i) F_2'(t_i) z_{f11}^2}{\left(z_f - z_p\right)^2}\right\},
\end{aligned}$$

where:

$a_i = \frac{log(\gamma_2 t_i)}{\gamma_1}$, $g(a_i) = \frac{\phi(a_i)}{S(t_i)}$, $F_1'(t_i) = \frac{\partial F(t_i)}{\partial\gamma_1} = -\frac{a_i\phi(a_i)}{\gamma_1}$, $F_2'(t_i) = \frac{\partial F(t_i)}{\partial\gamma_2} = \frac{\phi(a_i)}{\gamma_1\gamma_2}$,
$F_{11}''(t_i) = \frac{\partial^2 F(t_i)}{\partial\gamma_1^2} = F_1'(t_i)\frac{a_i^2-2}{\gamma_1}$ $F_{22}''(t_i) = \frac{\partial^2 F(t_i)}{\partial\gamma_2^2} = -F_2'(t_i)\frac{a_i+\gamma_1}{\gamma_1\gamma_2}$,
$F_{12}''(t_i) = \frac{\partial^2 F(t_i)}{\partial\gamma_1\partial\gamma_2} = F_2'(t_i)\frac{a_i^2-1}{\gamma_1}$.

B.2 Expressions for the components of the observed information matrix

Destructive Bernoulli cure rate model

$$l(\theta) \propto \sum_{I_1}\log\left(\frac{\eta_i p_i f(t_i)}{1+\eta_i}\right) + \sum_{I_0}\log\left(S_{pop}(t_i)\right).$$

$$\begin{aligned}
\frac{\partial l}{\partial\beta_{1l}} &= \sum_{I_1}\frac{x_{1il}}{1+\eta_i} + \sum_{I_0} x_{1il}\left\{\frac{\eta_i(1-p_i F(t_i))}{1+\eta_i(1-p_i F(t_i))} - \frac{\eta_i}{1+\eta_i}\right\}, \\
\frac{\partial l}{\partial\beta_{2k}} &= \sum_{I_1} x_{2ik}(1-p_i) - \sum_{I_0}\frac{x_{2ik}\eta_i p_i(1-p_i)F(t_i)}{1+\eta_i(1-p_i F(t_i))},
\end{aligned}$$

Under lognormal lifetime distribution

$$\frac{\partial l}{\partial \gamma_1} = \sum_{I_1} \frac{a_i^2 - 1}{\gamma_1} - \sum_{I_0} \frac{\eta_i p_i F_1'}{1 + \eta_i (1 - p_i F(t_i))},$$

$$\frac{\partial l}{\partial \gamma_2} = -\sum_{I_1} \frac{a_i}{\gamma_1 \gamma_2} - \sum_{I_0} \frac{\eta_i p_i F_2'}{1 + \eta_i (1 - p_i F(t_i))},$$

$$\frac{\partial^2 l}{\partial \beta_{1l} \partial \beta_{1l'}} = -\sum_{I_1} \frac{x_{1il} x_{1il'} \eta_i}{(1+\eta_i)^2} + \sum_{I_0} x_{1il} x_{1il'} \left[\frac{\eta_i (1 - p_i F(t_i))}{\{1 + \eta_i (1 - p_i F(t_i))\}^2} - \frac{\eta_i}{(1+\eta_i)^2} \right],$$

$$\frac{\partial^2 l}{\partial \beta_{2k} \beta_{2k'}} = -\sum_{I_1} x_{2ik} x_{2ik'} p_i (1 - p_i) - \sum_{I_0} \frac{x_{2ik} x_{2ik'} \eta_i p_i (1 - p_i) F(t_i)}{1 + \eta_i (1 - p_i F(t_i))} \left\{ 1 - 2p_i + \frac{\eta_i p_i (1 - p_i) F(t_i)}{1 + \eta_i (1 - p_i F(t_i))} \right\},$$

$$\frac{\partial^2 l}{\partial \gamma_1^2} = -\sum_{I_1} \frac{3a_i^2 - 1}{\gamma_1^2} - \sum_{I_0} \eta_i p_i \left[\frac{F_{11}''(t_i)}{1 + \eta_i (1 - p_i F(t_i))} + \frac{\eta_i p_i (F_1'(t_i))^2}{\{1 + \eta_i (1 - p_i F(t_i))\}^2} \right],$$

$$\frac{\partial^2 l}{\partial \gamma_2^2} = -\sum_{I_1} \frac{1 - a_i \gamma_1}{\gamma_1^2 \gamma_2^2} - \sum_{I_0} \eta_i p_i \left[\frac{F_{22}''(t_i)}{1 + \eta_i (1 - p_i F(t_i))} + \frac{\eta_i p_i (F_2'(t_i))^2}{\{1 + \eta_i (1 - p_i F(t_i))\}^2} \right],$$

$$\frac{\partial^2 l}{\partial \beta_{1l} \partial \gamma_1} = -\sum_{I_0} \frac{x_{1il} \eta_i p_i F_1'(t_i)}{\{1 + \eta_i (1 - p_i F(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{2k} \partial \gamma_1} = -\sum_{I_0} \frac{x_{2ik} \eta_i p_i (1 - p_i) (1 + \eta_i) F_1'(t_i)}{\{1 + \eta_i (1 - p_i F(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{1l} \partial \gamma_2} = -\sum_{I_0} \frac{x_{1il} \eta_i p_i F_2'(t_i)}{\{1 + \eta_i (1 - p_i F(t_i)))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{2k} \partial \gamma_2} = -\sum_{I_0} \frac{x_{2ik} \eta_i p_i (1 - p_i) (1 + \eta_i) F_2'(t_i)}{\{1 + \eta_i (1 - p_i F(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \beta_{1l} \partial \beta_{2k'}} = -\sum_{I_0} \frac{x_{1il} x_{2ik'} \eta_i p_i (1 - p_i) F(t_i)}{\{1 + \eta_i (1 - p_i F(t_i))\}^2},$$

$$\frac{\partial^2 l}{\partial \gamma_1 \partial \gamma_2} = \sum_{I_1} \frac{2a_i}{\gamma_1^2 \gamma_2} - \sum_{I_0} \frac{\eta_i p_i}{1 + \eta_i (1 - p_i F(t_i))} \left\{ F_{12}''(t_i) + \frac{\eta_i p_i F_1'(t_i) F_2'(t_i)}{1 + \eta_i (1 - p_i F(t_i))} \right\}.$$

Destructive Poisson cure rate model

$$l(\theta) = \sum_{I_1} \{\log(\eta_i p_i f(t_i)) - \eta_i p_i F(t_i)\} - \sum_{I_0} (\eta_i p_i F(t_i)).$$

$$\frac{\partial l}{\partial \beta_{1l}} = \sum_{I_1} x_{1il}(1-\eta_i p_i F(t_i)) - \sum_{I_0} x_{1il}\eta_i p_i F(t_i),$$

$$\frac{\partial l}{\partial \beta_{2k}} = \sum_{I_1} x_{2ik}(1-p_i)(1-\eta_i p_i F(t_i)) - \sum_{I_0} x_{2ik}\eta_i p_i (1-p_i) F(t_i),$$

$$\frac{\partial l}{\partial \gamma_1} = \sum_{I_1}\left(\frac{a_i^2-1}{\gamma_1} - \eta_i p_i F_1'(t_i)\right) - \sum_{I_0}\eta_i p_i F_1'(t_i),$$

$$\frac{\partial l}{\partial \gamma_2} = -\sum_{I_1}\left(\frac{a_i}{\gamma_1\gamma_2} + \eta_i p_i F_2'(t_i)\right) - \sum_{I_0}\eta_i p_i F_2'(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \beta_{1l'}} = -\sum_{I_1} x_{1il}x_{1il'}\eta_i p_i F(t_i) - \sum_{I_0} x_{1il}x_{1il'}\eta_i p_i F(t_i),$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \beta_{2k}\partial \beta_{2k'}} = &-\sum_{I_1} x_{2ik}x_{2ik'}p_i(1-p_i)\{\eta_i(1-p_i)F(t_i) + 1 - \eta_i p_i F(t_i)\} \\ &-\sum_{I_0} x_{2ik}x_{2ik'}\eta_i p_i(1-p_i)(1-2p)F(t_i),\end{aligned}$$

$$\frac{\partial^2 l}{\partial \gamma_1^2} = -\sum_{I_1}\left(\frac{3a_i^2-1}{\gamma_1^2} + \eta_i p_i F_{11}''(t_i)\right) - \sum_{I_0}\eta_i p_i F_{11}''(t_i),$$

$$\frac{\partial^2 l}{\partial \gamma_2^2} = -\sum_{I_1}\left(\frac{1-a_i\gamma_1}{\gamma_1^2\gamma_2^2} + \eta_i p_i F_{22}''(t_i)\right) - \sum_{I_0}\eta_i p_i F_{22}''(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \beta_{2k}} = -\sum_{I_1} x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i) - \sum_{I_0} x_{1il}x_{2ik}\eta_i p_i(1-p_i)F(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_1} = -\sum_{I_1} x_{1il}\eta_i p_i F_1'(t_i) - \sum_{I_0} x_{1il}\eta_i p_i F_1'(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_2} = -\sum_{I_1} x_{1il}\eta_i p_i F_2'(t_i) - \sum_{I_0} x_{1il}\eta_i p_i F_2'(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_1} = -\sum_{I_0} x_{2ik}\eta_i p_i(1-p_i)F_1'(t_i) - \sum_{I_0} x_{2ik}\eta_i p_i(1-p_i)F_1'(t_i),$$

$$\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_2} = -\sum_{I_0} x_{2ik}\eta_i p_i(1-p_i)F_2'(t_i) - \sum_{I_0} x_{2ik}\eta_i p_i(1-p_i)F_2'(t_i),$$

$$\frac{\partial^2 l}{\partial \gamma_1\partial \gamma_2} = \sum_{I_1}\left(\frac{2a_i}{\gamma_1^2\gamma_2} - \eta_i p_i F_{12}''(t_i)\right) - \sum_{I_0}\eta_i p_i F_{12}''(t_i).$$

Destructive COM-Poisson cure rate model

$$
\begin{aligned}
l(\theta) \quad \propto \quad & \sum_{I_1}\left[\log(p_i f(t_i)) + \log\left\{\sum_{j=1}^{\infty}\frac{j[\eta_i(1-p_iF(t_i))]^j}{(j!)^{\phi}}\right\} - \log\{z(1-p_iF(t_i))\}\right] \\
& + \sum_{I_0} S_{pop}(t_i).
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_{1l}} &= \sum_{I_1} x_{1il}\left(\frac{z_{f20}}{z_{f10}} - \frac{z_1}{z}\right) + \sum_{I_0} x_{1il}\left(\frac{z_{f10}}{z_f} - \frac{z_1}{z}\right), \\
\frac{\partial l}{\partial \beta_{2k}} &= \sum_{I_1} x_{2ik}(1-p_i)\left\{\frac{1}{1-p_iF(t_i)} - p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)\right\} \\
&\quad - \sum_{I_0} x_{2ik}p_i(1-p_i)F(t_i)\left(\frac{z_{f11}}{z_f}\right), \\
\frac{\partial l}{\partial \gamma_1} &= \sum_{I_1}\left\{\frac{a_i^2-1}{\gamma_1} - p_iF_1'(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{(1-p_iF(t_i))}\right)\right\} - \sum_{I_0}\frac{p_iF_1'(t_i)z_{f11}}{z_f}, \\
\frac{\partial l}{\partial \gamma_2} &= -\sum_{I_1}\left\{\frac{a_i}{\gamma_1\gamma_2} + p_iF_2'(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{(1-p_iF(t_i))}\right)\right\} - \sum_{I_0}\frac{p_iF_2'(t_i)z_{f11}}{z_f}, \\
\frac{\partial^2 l}{\partial \beta_{1il}\partial \beta_{1il'}} &= \sum_{I_1} x_{1il}x_{1il'}\left\{\frac{z_{f30}}{z_{f10}} - \left(\frac{z_{f20}}{z_{f10}}\right)^2 + \left(\frac{z_1}{z}\right)^2 - \frac{z_2}{z}\right\} \\
&\quad + \sum_{I_0}\frac{x_{1il}x_{1il'}}{z}\left\{\frac{z_{f20}}{z_f} - \frac{z_2}{z} + \left(\frac{z_1}{z}\right)^2 - \left(\frac{z_{f10}}{z_f}\right)^2\right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \beta_{2ik}\partial \beta_{2ik'}} &= \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i(1-p_i)^2}{(1-p_iF(t_i))^2}\left\{F(t_i) - \frac{1-p_iF(t_i)}{1-p_i}\right\} \\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_i^2F(t_i)(1-p_i)z_{f21}}{z_{f10}} \\
&\quad + \sum_{I_1}\frac{x_{2ik}x_{2ik'}p_iF(t_i)(1-p_i)^2}{z_{f10}}\left\{p_iF(t_i)\left(z_{f22} - \frac{z_{f21}^2}{z_{f10}}\right) - z_{f21}\right\} \\
&\quad + \sum_{I_0} x_{2ik}x_{2ik'}p_i(1-p_i)F(t_i)\left[p_i(1-p_i)F(t_i)\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{(1-2p_i)z_{f11}}{z_f}\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l}{\partial \gamma_1^2} &= -\sum_{I_1}\left\{\frac{3a_i^2-1}{\gamma_1^2} + p_iF_{11}''(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{1-p_iF(t_i)}\right)\right\} \\
&\quad + \sum_{I_1} p_i^2(F_1'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}} - \left(\frac{z_{f21}}{z_{f10}}\right)^2 + \frac{1}{(1-p_iF(t_i))^2}\right\} \\
&\quad + \sum_{I_0} p_i\left[p_i\left(F_1'(t_i)\right)^2\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{F_{11}''(t_i)z_{f11}}{z_f}\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l}{\partial \gamma_2^2} &= -\sum_{I_1}\left\{\frac{1-a_i\gamma_1}{\gamma_1^2\gamma_2^2}+p_iF_{22}''(t_i)\left(\frac{z_{f21}}{z_{f10}}-\frac{1}{1-p_iF(t_i)}\right)\right\}\\
&\quad+\sum_{I_1}p_i^2(F_2'(t_i))^2\left\{\frac{z_{f22}}{z_{f10}}-\left(\frac{z_{f21}}{z_{f10}}\right)^2+\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad+\sum_{I_0}p_i\left[p_i\left(F_2'(t_i)\right)^2\left\{\frac{z_{f12}}{z_f}-\left(\frac{z_{f11}}{z_f}\right)^2\right\}-\frac{F_{22}''(t_i)z_{f11}}{z_f}\right],\\
\frac{\partial l}{\partial \beta_{1l}\partial \beta_{2k}} &= \sum_{I_1}\frac{x_{1il}x_{2ik}p_iF(t_i)(1-p_i)}{z_{f10}}\left(\frac{z_{f21}z_{f20}}{z_{f10}}-z_{f31}\right)\\
&\quad+\sum_{I_0}x_{1il}x_{2ik}p_i(1-p_i)F(t_i)\left(\frac{z_{f10}z_{f11}}{z_f^2}-\frac{z_{f21}}{z_f}\right),\\
\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_1} &= \sum_{I_1}x_{1il}p_iF_1'(t_i)\left\{(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{z_{f31}}{z_{f10}}\right\}\\
&\quad+\sum_{I_0}x_{1il}p_iF_1'(t_i)\left(\frac{z_{f11}z_{f10}}{z_f^2}-\frac{z_{f21}}{z_f}\right),\\
\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_2} &= \sum_{I_1}x_{1il}p_iF_2'(t_i)\left\{(1-p_iF(t_i))\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{z_{f31}}{z_{f10}}\right\}\\
&\quad+\sum_{I_0}x_{1il}p_iF_2'(t_i)\left(\frac{z_{f11}z_{f10}}{z_f^2}-\frac{z_{f21}}{z_f}\right),\\
\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_1} &= \sum_{I_1}x_{2ik}p_iF_1'(t_i)(1-p_i)\left\{\frac{p_iF(t_i)z_{f22}-z_{f21}}{z_{f10}}\right\}\\
&\quad-\sum_{I_1}x_{2ik}p_iF_1'(t_i)(1-p_i)\left\{p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad+\sum_{I_0}x_{2ik}p_i(1-p_i)F_1'(t_i)\left[p_iF(t_i)\left\{\frac{z_{f12}}{z_f}-\left(\frac{z_{f11}}{z_f}\right)^2\right\}-\frac{z_{f11}}{z_f}\right],\\
\frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_2} &= \sum_{I_1}x_{2ik}p_iF_2'(t_i)(1-p_i)\left\{\frac{p_iF(t_i)z_{f22}-z_{f21}}{z_{f10}}\right\}\\
&\quad-\sum_{I_1}x_{2ik}p_iF_2'(t_i)(1-p_i)\left\{p_iF(t_i)\left(\frac{z_{f21}}{z_{f10}}\right)^2-\frac{1}{(1-p_iF(t_i))^2}\right\}\\
&\quad+\sum_{I_0}x_{2ik}p_i(1-p_i)F_2'(t_i)\left[p_iF(t_i)\left\{\frac{z_{f12}}{z_f}-\left(\frac{z_{f11}}{z_f}\right)^2\right\}-\frac{z_{f11}}{z_f}\right],
\end{aligned}
$$

Under lognormal lifetime distribution

$$\begin{aligned}\frac{\partial^2 l}{\partial\gamma_1\partial\gamma_2} &= \sum_{I_1}\left\{\frac{2a_i}{\gamma_1^2\gamma_2} - p_i F_{12}''(t_i)\left(\frac{z_{f21}}{z_{f10}} - \frac{1}{1-p_i F(t_i)}\right)\right\} \\ &\quad + \sum_{I_1} p_i^2 F_1'(t_i)F_2'(t_i)\left\{\frac{z_{f22}}{z_{f10}} - \left(\frac{z_{f21}}{z_{f10}}\right)^2 + \frac{1}{(1-p_i F(t_i))^2}\right\} \\ &\quad + \sum_{I_0} p_i\left[p_i F_1'(t_i)F_2'(t_i)\left\{\frac{z_{f12}}{z_f} - \left(\frac{z_{f11}}{z_f}\right)^2\right\} - \frac{F_{12}''(t_i)z_{f11}}{z_f}\right].\end{aligned}$$

Appendix C

Codes for EM algorithm with gamma lifetime

C.1 Data generation for COM-Poisson distribution

```
rm(list=ls())
library(MASS)
library(compoisson)
library(optimx)
library(survival)
library(timereg)
library(numDeriv)
library(zipfR)

Sim_data_COMPois<-function(n,beta11,beta02,beta12,gam1,gam2,theta1,ph1){
x1<-ifelse(runif(n,0,1)>0.44,0,1)
x2<-runif(n,0.1,17.42)
eta<-exp(beta11*x1)
p<-exp(beta02+beta12*x2)/(1+exp(beta02+beta12*x2))
   C<-rexp(n,theta1)
  d<-rep(NA,n)
D<-rep(NA,n)
M<-rep(NA,n)
for (i in 1:n){
M[i]<-rcom(1,lambda=eta[i],nu=ph1)
if (M[i]>0){
      D[i]<-rbinom(1,size=M[i],prob=p[i])
         if (D[i]>0){
         y<-min(rgamma(D[i],shape=1/(gam1^2),scale=(gam1^2)/gam2))
           T[i]<-min(y,C[i])} else { #when D[i]=0
           T[i]<-C[i]       }} else { #when M[i]=0
D[i]<-0
T[i]<-C[i]}
if (T[i]<C[i]){
             d[i]<-1 } else { #d is status
d[i]<-0 }
}
      data<-data.frame(cbind(T,d,x1,x2))
```

```
return(data)
}
#when generating from Poisson use M[i]<-rpois(1,lambda=eta[i]) instead of
#M[i]<-rcom(1,lambda=eta[i],nu=ph1)
#and from Bernoulli use M[i]<-rbinom(1,size=1,prob=eta[i]/(1+eta[i]))

n=500
ph1=1.5
beta11=0.693
beta02=-1.421
beta12=0.345
gam1=1.225
gam2=5
theta1=0.25

dcp=Sim_data_COMPois(n,beta11,beta02,beta12,gam1,gam2,theta1,ph1)
data_obs=subset(dcp,d==1)$T
data_cens=subset(dcp,d==0)$T
x1t=subset(dcp,d==1)$x1
x1c=subset(dcp,d==0)$x1
x2t=subset(dcp,d==1)$x2
x2c=subset(dcp,d==0)$x2
```

C.2 Estimation of parameters using complete likelihood approach

```
dest_COMPois<-function(data_obs,data_cens,x1t,x1c,x2t,x2c,tol,maxit,b11,b02,b12,
g1,g2,ph){

grad=matrix(NA,nrow=6,ncol=1)
hess=matrix(NA,nrow=6,ncol=6)
pnew=matrix(NA,ncol=1,nrow=6)
pold=matrix(NA,ncol=1,nrow=6)

pold[1,1]<-b11
pold[2,1]<-b02
pold[3,1]<-b12
pold[4,1]<-g1
pold[5,1]<-g2
pold[6,1]<-ph

continue=TRUE
iter=1
while(continue){
#Q function: E step:

func=function(par=c(bb11,bb02,bb12,gg1,gg2,pph)){

zz1<-Vectorize(function(theta,ph,tol1){v1=1;j=1;continue=TRUE
while(continue==T){
v0=v1
h<-function(j){
if(j==0){return(1)}else{v=((theta/(j^(ph)))*h(j-1))}
return(v)}
```

```
v1=v1+h(j)
if(is.infinite(h(j)) | is.na(h(j)) | is.nan(h(j)) | v1>=max_tol){continue=F;
v1==Inf}else if(abs(v1-v0)<tol1){continue=F}else{j=j+1}}
return(v1)})

zz2<-Vectorize(function(theta,ph,tol1){v1=0;j=1;continue=TRUE
while(continue==T){
v0=v1
h<-function(j){
if(j==1){return(theta)}else{v=((theta/(j^(ph)))*h(j-1))}
return(v)}
v1=v1+j*h(j)
if(is.infinite(h(j)) | is.na(h(j)) | is.nan(h(j)) | v1>=max_tol){continue=F;
v1==Inf}else if(abs(v1-v0)<tol1){continue=F}else{j=j+1}}
return(v1)})

etat=exp(par[1]*x1t)
etac=exp(par[1]*x1c)
pt=(exp(par[2]+par[3]*x2t))/(1+exp(par[2]+par[3]*x2t))
pc=(exp(par[2]+par[3]*x2c))/(1+exp(par[2]+par[3]*x2c))
St=(Igamma(1/par[4]^2,(par[5]*data_obs)/par[4]^2,lower=F))/gamma(1/par[4]^2)
Sc=(Igamma(1/par[4]^2,(par[5]*data_cens)/par[4]^2,lower=F))/gamma(1/par[4]^2)
Ft=1-St
Fc=1-Sc
ft=((par[5]/par[4]^2)^(1/par[4]^2))/gamma(1/par[4]^2)*exp(-
(par[5]*data_obs/par[4]^2))*(data_obs)^(1/par[4]^2-1)
fc=((par[5]/par[4]^2)^(1/par[4]^2))/gamma(1/par[4]^2)*exp(-
(par[5]*data_cens/par[4]^2))*(data_cens)^(1/par[4]^2-1)

#calculation of infinite sums
spt=zz1(theta=etat*(1-pt*Ft),ph=par[6],tol1)/zz1(theta=etat,ph=par[6],tol1)
spc=zz1(theta=etac*(1-pc*Fc),ph=par[6],tol1)/zz1(theta=etac,ph=par[6],tol1)
fpt=(pt*ft*zz2(theta=etat*(1-pt*Ft),ph=par[6],tol1))/(zz1(theta=
etat,ph=par[6],tol1)*(1-pt*Ft))
p0t=zz1(theta=etat*(1-pt),ph=par[6],tol1)/zz1(theta=etat,ph=par[6],tol1)
p0c=zz1(theta=etac*(1-pc),ph=par[6],tol1)/zz1(theta=etac,ph=par[6],tol1)
s1c=(spc-p0c)/(1-p0c)

etak=exp(bb11*x1c)
pk=(exp(bb02+bb12*x2c))/(1+exp(bb02+bb12*x2c))
Fk=1-(Igamma(1/gg1^2,(gg2*data_cens)/gg1^2,lower=F))/gamma(1/gg1^2)
spk=zz1(theta=etak*(1-pk*Fk),pph,tol1)/zz1(theta=etak,pph,tol1)
p0k=zz1(theta=etak*(1-pk),pph,tol1)/zz1(theta=etak,pph,tol1)
s1k=(spk-p0k)/(1-p0k)
wk=(spk-p0k)/spk

dam1c=spc-p0c
for (j in 1:length(dam1c)){
   if(dam1c[j]==0){dam1c[j]=.Machine$double.eps}}

result1=sum(log(fpt))+sum((1-wk)*log(p0c))+sum(wk*log(dam1c))
return(-result1)
}
bb11<-pold[1,1]
bb02<-pold[2,1]
bb12<-pold[3,1]
```

```
gg1<-pold[4,1]
gg2<-pold[5,1]
pph<-pold[6,1]

param=c(pold[1,1],pold[2,1],pold[3,1],pold[4,1],pold[5,1],pold[6,1])

result=optimx(par=param, fn=func,method="Nelder-Mead",itnmax=5000,hessian=T)
hess=as.matrix(attr(result,"details")[,"nhatend"][[1]])
result$p1->pnew[1,1]
result$p2->pnew[2,1]
result$p3->pnew[3,1]
result$p4->pnew[4,1]
result$p5->pnew[5,1]
result$p6->pnew[6,1]

if(hess[1,1]=="NaN" | hess[2,2]=="NaN" | hess[3,3]=="NaN" | hess[4,4]=="NaN" |
hess[5,5]=="NaN" |  hess[6,6]=="NaN" |
hess[1,2]=="NaN" | hess[1,3]=="NaN" | hess[1,4]=="NaN" | hess[1,5]=="NaN" |
hess[1,6]=="NaN" |
hess[2,3]=="NaN" | hess[2,4]=="NaN" | hess[2,5]=="NaN" |  hess[2,6]=="NaN" |
hess[3,4]=="NaN" | hess[3,5]=="NaN" | hess[3,6]=="NaN" |
hess[4,5]=="NaN" | hess[4,6]=="NaN" | hess[5,6]=="NaN" |
hess[1,1]==Inf | hess[2,2]==Inf | hess[3,3]==Inf | hess[4,4]==Inf |
hess[5,5]==Inf |  hess[6,6]==Inf |
hess[1,2]==Inf | hess[1,3]==Inf | hess[1,4]==Inf | hess[1,5]==Inf |
hess[1,6]==Inf |
hess[2,3]==Inf | hess[2,4]==Inf | hess[2,5]==Inf | hess[2,6]==Inf |
hess[3,4]==Inf | hess[3,5]==Inf | hess[3,6]==Inf |
hess[4,5]==Inf | hess[4,6]==Inf | hess[5,6]==Inf |
hess[1,1]<0 | hess[2,2]<0 | hess[3,3]<0 | hess[4,4]<0 | hess[5,5]<0 |
hess[6,6]<0 |  det(hess)==0 | rcond(hess)<(1e-6) |
pnew[4,1]<0 | pnew[5,1]<0 | pnew[6,1]<0 |iter==maxit){

pnew[1,1]=0
pnew[2,1]=0
pnew[3,1]=0
pnew[4,1]=0
pnew[5,1]=0
pnew[6,1]=0
continue=FALSE
}else{
iter=iter+1
continue=(abs((pnew[1,1]-pold[1,1])/pold[1,1])>tol |
abs((pnew[2,1]-pold[2,1])/pold[2,1])>tol |
abs((pnew[3,1]-pold[3,1])/pold[3,1])>tol |
abs((pnew[4,1]-pold[4,1])/pold[4,1])>tol |
abs((pnew[5,1]-pold[5,1])/pold[5,1])>tol |
abs((pnew[6,1]-pold[6,1])/pold[6,1])>tol) &(iter<=maxit)

pold[1,1]=pnew[1,1]
pold[2,1]=pnew[2,1]
pold[3,1]=pnew[3,1]
pold[4,1]=pnew[4,1]
pold[5,1]=pnew[5,1]
pold[6,1]=pnew[6,1]
}#end of else
}#end of while
```

```
output=c(pnew)

#Standard Errors calculation
if(output[1]!=0 & output[2]!=0 & output[3]!=0 & output[4]!=0 &
output[5]!=0 & output[6]!=0){

#Calculating log-likelihood value for converging samples
qfunc=function(param=c(bb11,bb02,bb12,gg1,gg2,pph)){
etat=exp(param[1]*x1t)
etac=exp(param[1]*x1c)
pt=(exp(param[2]+param[3]*x2t))/(1+exp(param[2]+param[3]*x2t))
pc=(exp(param[2]+param[3]*x2c))/(1+exp(param[2]+param[3]*x2c))
St=(Igamma(1/param[4]^2,(param[5]*data_obs)/param[4]^2,lower=
F))/gamma(1/param[4]^2)
Sc=(Igamma(1/param[4]^2,(param[5]*data_cens)/param[4]^2,lower=
F))/gamma(1/param[4]^2)
Ft=1-St
Fc=1-Sc
ft=((param[5]/param[4]^2)^(1/param[4]^2))/gamma(1/param[4]^2)*exp(-
(param[5]*data_obs/param[4]^2))*(data_obs)^(1/param[4]^2-1)
fc=((param[5]/param[4]^2)^(1/param[4]^2))/gamma(1/param[4]^2)*exp(-
(param[5]*data_cens/param[4]^2))*(data_cens)^(1/param[4]^2-1)

#calculation of infinite sums
ph<-param[6]
zz1<-Vectorize(function(theta,ph,tol1){v1=1;j=1;continue=TRUE
while(continue==T){
v0=v1
h<-function(j){
if(j==0){return(1)}else{v=((theta/(j^(ph)))*h(j-1))}
return(v)}
v1=v1+h(j)
if(is.infinite(h(j)) | is.na(h(j)) | is.nan(h(j)) | v1>=max_tol){continue=F;
v1==Inf}else if(v1-v0<tol1){continue=F}else{j=j+1}}
return(v1)})

zz2<-Vectorize(function(theta,ph,tol1){v1=0;j=1;continue=TRUE
while(continue==T){
v0=v1
h<-function(j){
if(j==1){return(theta)}else{v=((theta/(j^(ph)))*h(j-1))}
return(v)}
v1=v1+j*h(j)
if(is.infinite(h(j)) | is.na(h(j)) | is.nan(h(j)) | v1>=max_tol){continue=F;
v1==Inf}else if(v1-v0<tol1){continue=F}else{j=j+1}}
return(v1)})

spt=zz1(theta=etat*(1-pt*Ft),ph,tol1)/zz1(theta=etat,ph,tol1)
spc=zz1(theta=etac*(1-pc*Fc),ph,tol1)/zz1(theta=etac,ph,tol1)
fpt=(pt*ft*zz2(theta=etat*(1-pt*Ft),ph,tol1))/(zz1(theta=etat,ph,tol1)*(1-pt*Ft))
p0t=zz1(theta=etat*(1-pt),ph,tol1)/zz1(theta=etat,ph,tol1)
p0c=zz1(theta=etac*(1-pc),ph,tol1)/zz1(theta=etac,ph,tol1)
s1c=(spc-p0c)/(1-p0c)

for(ii in 1: length(spc)){
if(spc[ii]==0){
```

```
spc[ii]=.Machine$double.eps}}

result=sum(log(fpt))+sum(log(spc))
return(result)
}
bb11=output[1]
bb02=output[2]
bb12=output[3]
gg1=output[4]
gg2=output[5]
pph=output[6]
param=c(bb11, bb02, bb12, gg1, gg2,pph)
lik=qfunc(param)
hess=hessian(qfunc,param,method="Richardson")
se=sqrt(diag(solve(-1*hess)))
}else{ #values of standard errors and likelihood for non-converging samples
se=c(0,0,0,0,0,0)
lik="NaN"
}

return(c(lik,output,se))

}#end of Dest_COMPois function

tol1=1e-10
tol=0.001
max_tol=1e+200
maxit=500

b11=0.268
b02=-2.519
b12=0.145
g1=1.21
g2=6.487
ph=1.457

ww1=as.numeric(dest_COMPois(data_obs,data_cens,x1t,x1c,x2t,x2c,tol,maxit,b11,b02,
b12,g1,g2,ph))
ww1
```

Appendix D

Codes for EM algorithm with generalised gamma lifetime

D.1 Data generation from Poisson distribution

```
rm(list=ls())
library(MASS)
library(compoisson)
library(optimx)
library(survival)
library(timereg)
library(numDeriv)
library(zipfR)

set.seed(1974)

Sim_data_Pois<-function(n,beta11,beta02,beta12,gam1,gam2,q1,theta1){
x1<-ifelse(runif(n,0,1)>0.44,0,1)
x2<-runif(n,0.1,17.42)
eta<-exp(beta11*x1)
p<-exp(beta02+beta12*x2)/(1+exp(beta02+beta12*x2))
   C<-rexp(n,theta1)
  d<-rep(NA,n)
D<-rep(NA,n)
M<-rep(NA,n)
for (i in 1:n){
M[i]<-rpois(1,lambda=eta[i])
if (M[i]>0){
      D[i]<-rbinom(1,size=M[i],prob=p[i])
         if (D[i]>0){
  y<-min((rgamma(D[i],shape=1/(q1^2),scale=(q1^2)/gam2^(q1/gam1)))^(gam1/q1))
           T[i]<-min(y,C[i])} else { #when D[i]=0
           T[i]<-C[i]       }} else { #when M[i]=0
D[i]<-0
T[i]<-C[i]}
if (T[i]<C[i]){
             d[i]<-1 } else { #d is status
```

```
d[i]<-0 }
}
      data<-data.frame(cbind(T,d,x1,x2))
return(data)
}

n=200

max_tol=1e+200
beta11=0.34
beta02=-2.223
beta12=0.254
gam1=0.745 #sigma=g1
gam2=2.454 #lambda=g2
q1=0.8
theta1=0.3

dp=Sim_data_Pois(n,beta11,beta02,beta12,gam1,gam2,q1,theta1)
data_obs<-subset(dp,d==1)$T
data_cens<-subset(dp,d==0)$T
x1t<-subset(dp,d==1)$x1
x1c<-subset(dp,d==0)$x1
x2t<-subset(dp,d==1)$x2
x2c<-subset(dp,d==0)$x2
```

D.2 Estimation of parameters for destructive Poisson model

```
dest_Pois<-function(data_obs,data_cens,x1t,x1c,x2t,x2c,tol,maxit,b11,
b02,b12,g1,g2,q){

grad=matrix(NA,nrow=6,ncol=1)
hess=matrix(NA,nrow=6,ncol=6)
pnew=matrix(NA,ncol=1,nrow=6)
pold=matrix(NA,ncol=1,nrow=6)

pold[1,1]<-b11
pold[2,1]<-b02
pold[3,1]<-b12
pold[4,1]<-g1
pold[5,1]<-g2
pold[6,1]<-q

continue=TRUE
iter=1

while(continue){
#Q function: E step:
func=function(par=c(bb11,bb02,bb12,gg1,gg2,qq)){
etat=exp(par[1]*x1t)
etac=exp(par[1]*x1c)
pt=(exp(par[2]+par[3]*x2t))/(1+exp(par[2]+par[3]*x2t))
pc=(exp(par[2]+par[3]*x2c))/(1+exp(par[2]+par[3]*x2c))
St=(Igamma(1/par[6]^2,1/par[6]^2*(par[5]*data_obs)^(par[6]/
```

```
par[4]),lower=F))/gamma(1/par[6]^2)   #for q>0
Sc=(Igamma(1/par[6]^2,1/par[6]^2*(par[5]*data_cens)^(par[6]/
par[4]),lower=F))/gamma(1/par[6]^2)  #for q>0
Ft=1-St
Fc=1-Sc
ft=(par[6]/((par[6]^2)^(1/par[6]^2)))*((par[5]*data_obs)^(1/
par[6]^2))^(par[6]/par[4])*exp(-1/par[6]^2*(par[5]*data_obs)^(par[6]/
par[4]))/(gamma(1/par[6]^2)*par[4]*data_obs)
fc=(par[6]/((par[6]^2)^(1/par[6]^2)))*((par[5]*data_cens)^(1/
par[6]^2))^(par[6]/par[4])*exp(-1/par[6]^2*(par[5]*data_cens)^(par[6]/
par[4]))/(gamma(1/par[6]^2)*par[4]*data_cens)
p0c<-exp(-etac*pc)
spt<-exp(-etat*pt*Ft)
spc<-exp(-etac*pc*Fc)
fpt=etat*pt*ft*(exp(-etat*pt*Ft))
etak=exp(bb11*x1c)
pk=(exp(bb02+bb12*x2c))/(1+exp(bb02+bb12*x2c))
Fk=1-(Igamma(1/qq^2,1/qq^2*(gg2*data_cens)^(qq/gg1),lower=F))/gamma(1/qq^2)
p0k<-exp(-etak*pk)
spk<-exp(-etak*pk*Fk)

wk=(spk-p0k)/spk
dam1c=spc-p0c
for (j in 1:length(dam1c)){
   if(dam1c[j]==0){dam1c[j]=.Machine$double.eps}}

result1=sum(log(fpt))+sum((1-wk)*log(p0c))+sum(wk*log(dam1c))

return(result1)
}
bb11<-pold[1,1]
bb02<-pold[2,1]
bb12<-pold[3,1]
gg1<-pold[4,1]
gg2<-pold[5,1]
qq<-pold[6,1]

param=c(pold[1,1],pold[2,1],pold[3,1],pold[4,1],pold[5,1],pold[6,1])

score=grad(func,param,method="Richardson")
hessmat=hessian(func,param,method="Richardson")

grad[1,1]<-score[1]
grad[2,1]<-score[2]
grad[3,1]<-score[3]
grad[4,1]<-score[4]
grad[5,1]<-score[5]
grad[6,1]<-score[6]

hess[1,1]<-hessmat[1,1]
hess[2,2]<-hessmat[2,2]
hess[3,3]<-hessmat[3,3]
hess[4,4]<-hessmat[4,4]
hess[5,5]<-hessmat[5,5]
hess[6,6]<-hessmat[6,6]
hess[1,2]<-hessmat[1,2]
```

```
hess[2,1]<-hessmat[2,1]
hess[1,3]<-hessmat[1,3]
hess[3,1]<-hessmat[3,1]
hess[1,4]<-hessmat[1,4]
hess[4,1]<-hessmat[4,1]
hess[1,5]<-hessmat[1,5]
hess[5,1]<-hessmat[5,1]
hess[1,6]<-hessmat[1,6]
hess[6,1]<-hessmat[6,1]
hess[2,3]<-hessmat[2,3]
hess[3,2]<-hessmat[3,2]
hess[2,4]<-hessmat[2,4]
hess[4,2]<-hessmat[4,2]
hess[2,5]<-hessmat[2,5]
hess[5,2]<-hessmat[5,2]
hess[2,6]<-hessmat[2,6]
hess[6,2]<-hessmat[6,2]
hess[3,4]<-hessmat[3,4]
hess[4,3]<-hessmat[4,3]
hess[3,5]<-hessmat[3,5]
hess[5,3]<-hessmat[5,3]
hess[3,6]<-hessmat[3,6]
hess[6,3]<-hessmat[6,3]
hess[4,5]<-hessmat[4,5]
hess[5,4]<-hessmat[5,4]
hess[4,6]<-hessmat[4,6]
hess[6,4]<-hessmat[6,4]
hess[5,6]<-hessmat[5,6]
hess[6,5]<-hessmat[6,5]

if(hess[1,1]=="NaN" | hess[2,2]=="NaN" | hess[3,3]=="NaN" |
hess[4,4]=="NaN" | hess[5,5]=="NaN" | hess[6,6]=="NaN" |
hess[1,2]=="NaN" | hess[1,3]=="NaN" | hess[1,4]=="NaN" |
hess[1,5]=="NaN" |hess[1,6]=="NaN" | hess[2,3]=="NaN" |
hess[2,4]=="NaN" | hess[2,5]=="NaN" |  hess[2,6]=="NaN" |
hess[3,4]=="NaN" | hess[3,5]=="NaN" | hess[3,6]=="NaN" |
hess[4,5]=="NaN" | hess[4,6]=="NaN" | hess[5,6]=="NaN" |
hess[1,1]==Inf | hess[2,2]==Inf | hess[3,3]==Inf | hess[4,4]==Inf |
hess[5,5]==Inf |  hess[6,6]==Inf |hess[1,2]==Inf | hess[1,3]==Inf |
hess[1,4]==Inf | hess[1,5]==Inf |  hess[1,6]==Inf | hess[2,3]==Inf |
hess[2,4]==Inf | hess[2,5]==Inf | hess[2,6]==Inf | hess[3,4]==Inf |
hess[3,5]==Inf | hess[3,6]==Inf | hess[4,5]==Inf | hess[4,6]==Inf |
hess[5,6]==Inf |hess[1,1]>0 | hess[2,2]>0 | hess[3,3]>0 | hess[4,4]>0 |
hess[5,5]>0 | hess[6,6]>0 | det(-1*hess)==0 | rcond(-1*hess)<(1e-6)){

  pnew[1,1]="NaN"
pnew[2,1]="NaN"
pnew[3,1]="NaN"
pnew[4,1]="NaN"
pnew[5,1]="NaN"
pnew[6,1]="NaN"

}else{

#M-step:
pnew=pold+solve(-1*hess)%*%grad
}

if(pnew[1,1]=="NaN" | pnew[2,1]=="NaN" | pnew[3,1]=="NaN" |
pnew[4,1]=="NaN" | pnew[5,1]=="NaN"  | pnew[6,1]=="NaN"  |
```

```
pnew[4,1]<0 | pnew[5,1]<0 | pnew[6,1]<0 |iter==maxit){

pnew[1,1]=0
pnew[2,1]=0
pnew[3,1]=0
pnew[4,1]=0
pnew[5,1]=0
pnew[6,1]=0

continue=FALSE
}else{
iter=iter+1
continue=(abs((pnew[1,1]-pold[1,1])/pold[1,1])>tol |
abs((pnew[2,1]-pold[2,1])/pold[2,1])>tol |
abs((pnew[3,1]-pold[3,1])/pold[3,1])>tol |
abs((pnew[4,1]-pold[4,1])/pold[4,1])>tol |
abs((pnew[5,1]-pold[5,1])/pold[5,1])>tol |
abs((pnew[6,1]-pold[6,1])/pold[6,1])>tol) &(iter<=maxit)

pold[1,1]=pnew[1,1]
pold[2,1]=pnew[2,1]
pold[3,1]=pnew[3,1]
pold[4,1]=pnew[4,1]
pold[5,1]=pnew[5,1]
pold[6,1]=pnew[6,1]
}#end of else
}#end of while

output=c(pnew)

#Standard Errors calculation

if(output[1]!=0 & output[2]!=0 & output[3]!=0 & output[4]!=0 &
output[5]!=0 & output[6]!=0){

#Calculating log-likelihood value for converging samples
qfunc=function(param=c(bb11,bb02,bb12,gg1,gg2,qq)){
etat=exp(param[1]*x1t)
etac=exp(param[1]*x1c)
pt=(exp(param[2]+param[3]*x2t))/(1+exp(param[2]+param[3]*x2t))
pc=(exp(param[2]+param[3]*x2c))/(1+exp(param[2]+param[3]*x2c))
St=(Igamma(1/param[6]^2,1/param[6]^2*(param[5]*data_obs)^(param[6]/
param[4]),lower=F))/gamma(1/param[6]^2)   #for q>0
Sc=(Igamma(1/param[6]^2,1/param[6]^2*(param[5]*data_cens)^(param[6]/
param[4]),lower=F))/gamma(1/param[6]^2)  #for q>0
Ft=1-St
Fc=1-Sc
ft=(param[6]/((param[6]^2)^(1/param[6]^2)))*((param[5]*data_obs)^(1/
param[6]^2))^(param[6]/param[4])*exp(-1/param[6]^2*(param[5]*
data_obs)^(param[6]/param[4]))/(gamma(1/param[6]^2)*param[4]*data_obs)
fc=(param[6]/((param[6]^2)^(1/param[6]^2)))*((param[5]*data_cens)^(1/
param[6]^2))^(param[6]/param[4])*exp(-1/param[6]^2*(param[5]*
data_cens)^(param[6]/param[4]))/(gamma(1/param[6]^2)*param[4]*data_cens)

p0c<-exp(-etac*pc)
spt<-exp(-etat*pt*Ft)
```

```
spc<-exp(-etac*pc*Fc)
fpt=etat*pt*ft*(exp(-etat*pt*Ft))
for(ii in 1: length(spc)){
if(spc[ii]==0){
spc[ii]=.Machine$double.eps}}

result=sum(log(fpt))+sum(log(spc))
return(result)
}
bb11=output[1]
bb02=output[2]
bb12=output[3]
gg1=output[4]
gg2=output[5]
pph=output[6]
param=c(bb11, bb02, bb12, gg1, gg2,qq)
hess=hessian(qfunc,param,method="Richardson")
se=sqrt(diag(solve(-1*hess)))
lik=qfunc(param)
}else{
se=c(0,0,0,0,0,0)
lik="NaN"
}

return(c(lik,output,se))

}#end of Dest_COMPois function

tol=0.001
max_tol=1e+200
maxit=500

b11=0.4703616
b02=-2.3493935
b12=0.2230183
g1=0.6909148
g2=2.4441480
q=0.9551

ww=as.numeric(dest_Pois(data_obs,data_cens,x1t,x1c,x2t,x2c,tol,maxit,b11,
b02,b12,g1,g2,q))
ww
```

References

Balakrishnan, N. and Pal, S. (2012). EM algorithm-based likelihood estimation for some cure rate models. *Journal of Statistical Theory and Practice*, 6(4):698–724.

Balakrishnan, N. and Pal, S. (2013). Lognormal lifetimes and likelihood-based inference for flexible cure rate models based on COM-Poisson family. *Computational Statistics & Data Analysis*, 67:41–67.

Balakrishnan, N. and Pal, S. (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.

Balakrishnan, N. and Pal, S. (2016). Expectation maximization-based likelihood inference for flexible cure rate models with Weibull lifetimes. *Statistical Methods in Medical Research*, 25(4):1535–1563.

Becker, M. P., Yang, I., and Lange, K. (1997). EM algorithms without missing data. *Statistical Methods in Medical Research*, 6(1):38–54.

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

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.

Borges, P., Rodrigues, J., and Balakrishnan, N. (2012). Correlated destructive generalized power series cure rate models and associated inference with an application to a cutaneous melanoma data. *Computational Statistics & Data Analysis*, 56(6):1703–1713.

Cancho, V. G., Bandyopadhyay, D., Louzada, F., and Yiqi, B. (2013). The destructive negative binomial cure rate model with a latent activation scheme. *Statistical Methodology*, 13:48–68.

Cancho, V. G., de Castro, M., and Rodrigues, J. (2012). A Bayesian analysis of the Conway–Maxwell–Poisson cure rate model. *Statistical Papers*, 53(1):165–176.

Chen, M. H., Ibrahim, J. G., and Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, 94(447):909–919.

Conway, R. W. and Maxwell, W. L. (1962). A queuing model with state dependent service rates. *Journal of Industrial Engineering*, 12(2):132–136.

Cooner, F., Banerjee, S., and Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, 102(478):560–572.

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

Fang, H.-B., Li, G., and Sun, J. (2005). Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. *Scandinavian Journal of Statistics*, 32(1):59–75.

Farewell, V. T. (1986). Mixture models in survival analysis: Are they worth the risk? *Canadian Journal of Statistics*, 14(3):257–262.

Fisher, R. A. (1934). The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6(1):13–25.

Gupta, R. C. and Kirmani, S. (1990). The role of weighted distributions in stochastic modeling. *Communications in Statistics-Theory and Methods*, 19(9):3147–3162.

Hanin, L. and Huang, L.-S. (2014). Identifiability of cure models revisited. *Journal of Multivariate Analysis*, 130:261–274.

Hougaard, P. (1995). Frailty models for survival data. *Lifetime Data Analysis*, 1(3):255–273.

Kokonendji, C. C., Mizere, D., and Balakrishnan, N. (2008). Connections of the Poisson weight function to overdispersion and underdispersion. *Journal of Statistical Planning and Inference*, 138(5):1287–1296.

Kuk, A. Y. C. and Chen, C. H. (1992). A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79(3):531–541.

Lange, K. (1995). A gradient algorithm locally equivalent to the EM algorithm. *Journal of the Royal Statisticsal Society: Series B*, 57:425–437.

Larson, M. G. and Dinse, G. E. (1985). A mixture model for the regression analysis of competing risks data. *Applied Statistics*, 34(3):201–211.

Li, C. S., Taylor, J. M., and Sy, J. P. (2001). Identifiability of cure models. *Statistics & Probability Letters*, 54(4):389–395.

Li, C. S. and Taylor, J. M. G. (2002). A semi-parametric accelerated failure time cure model. *Statistics in Medicine*, 21(21):3235–3247.

Liu, M., Lu, W., and Shao, Y. (2006). Mixture cure model with an application to interval mapping of quantitative trait loci. *Lifetime Data Analysis*, 12(4):421–440.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):226–233.

Lu, W. (2008). Maximum likelihood estimation in the proportional hazards cure model. *Annals of the Institute of Statistical Mathematics*, 60(3):545–574.

Lu, W. and Ying, Z. (2004). On semiparametric transformation cure models. *Biometrika*, 91(2):331–343.

Maller, R. A. and Zhou, S. (1992). Estimating the proportion of immunes in a censored sample. *Biometrika*, 79(4):731–739.

Maller, R. A. and Zhou, X. (1996). *Survival analysis with long-term survivors*. John Wiley & Sons.

McLachlan, G. and Krishnan, T. (2007). *The EM algorithm and extensions*. John Wiley & Sons, second edition.

Mooney, C. Z. (1997). *Monte Carlo simulation*. Sage Publications.

Ortega, E. M., Cordeiro, G. M., and Lemonte, A. J. (2012). A log-linear regression model for the β-Birnbaum–Saunders distribution with censored data. *Computational Statistics & Data Analysis*, 56(3):698–718.

Othus, M., Barlogie, B., LeBlanc, M. L., and Crowley, J. J. (2012). Cure models as a useful statistical tool for analyzing survival. *Clinical Cancer Research*, 18(14):3731–3736.

Pal, S. and Balakrishnan, N. (2015). Likelihood inference based on EM algorithm for the destructive length-biased Poisson cure rate model with Weibull lifetime. *Communications in Statistics-Simulation and Computation*, (to appear) DOI: 10.1080/03610918.2015.1053918.

Pal, S. and Balakrishnan, N. (2016). Destructive negative binomial cure rate model and EM-based likelihood inference under Weibull lifetime. *Statistics & Probability Letters*, 116:9–20.

Pal, S. and Balakrishnan, N. (2017a). An EM type estimation procedure for the destructive exponentially weighted Poisson regression cure model under generalized gamma lifetime. *Journal of Statistical Computation and Simulation*, 87(6):1107–1129.

Pal, S. and Balakrishnan, N. (2017b). Likelihood inference for the destructive exponentially weighted Poisson cure rate model with Weibull lifetime and an application to melanoma data. *Computational Statistics*, 32(2):429–449.

Patil, G. P. (2002). *Weighted distributions*. Wiley Online Library.

Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, 34(2):179–189.

R Development Core Team (2010). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3):507–554.

Rodrigues, J., de Castro, M., Balakrishnan, N., and Cancho, V. G. (2011). Destructive weighted Poisson cure rate models. *Lifetime Data Analysis*, 17(3):333–346.

Rodrigues, J., de Castro, M., Cancho, V. G., and Balakrishnan, N. (2009). COM-Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference*, 139(10):3605–3611.

Scheike, T. H. and Zhang, M.-J. (2011). Analyzing competing risk data using the R timereg package. *Journal of statistical software*, 38(2).

Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.

Sellers, K. F., Borle, S., and Shmueli, G. (2012). The COM-Poisson model for count data: a survey of methods and applications. *Applied Stochastic Models in Business and Industry*, 28(2):104–116.

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S., and Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(1):127–142.

Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33(3):1187–1192.

Tanner, M. A. (1996). *Tools for Statistical inference*. Springer, third edition.

Tsodikov, A. D., Ibrahim, J. G., and Yakovlev, A. Y. (2003). Estimating cure rates from survival data. *Journal of the American Statistical Association*, 98(464):1063–1078.

Yakovlev, A. Y. and Tsodikov, A. D. (1996). *Stochastic models of tumor latency and their biostatistical applications*. OECD Publishing, first edition.

Yu, B., Tiwari, R. C., Cronin, K. A., and Feuer, E. J. (2004). Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in Medicine*, 23(11):1733–1747.

Yudi, P. (2001). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press.

Zeng, D., Yin, G., and Ibrahim, J. G. (2006). Semiparametric transformation models for survival data with a cure fraction. *Journal of the American Statistical Association*, 101(474):670–684.

Zhao, L., Feng, D., Bellile, E. L., and Taylor, J. M. (2014). Bayesian random threshold estimation in a cox proportional hazards cure model. *Statistics in Medicine*, 33(4):650–661.