CURE RATE AND DESTRUCTIVE CURE RATE MODELS UNDER PROPORTIONAL ODDS LIFETIME DISTRIBUTIONS

CURE RATE AND DESTRUCTIVE CURE RATE MODELS UNDER PROPORTIONAL ODDS LIFETIME DISTRIBUTIONS

BY

TIAN FENG, B.Sc., M.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS & STATISTICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

PH.D. OF MATHEMATICS

Ph.D. of MATHEMATICS (2018)	McMaster University
(Mathematics & Statistics)	Hamilton, Ontario, Canada

TITLE:	CURE RATE AND DESTRUCTIVE CURE RATE MODELS UNDER PROPORTIONAL ODDS LIFETIME DISTRIBUTIONS
AUTHOR:	TIAN FENG B.Sc., Actuarial Science University of Toronto, Toronto, Canada M.Sc., Statistics McMaster University, Hamilton, Canada
SUPERVISOR:	Professor N. Balakrishnan
NUMBER OF PAGES:	xvi, 204

To my husband Dr. Hon-yiu So

To my parents Mr. Shenghu Feng, Ms. Jingping Yuan

To my grandmother Ms. Xiumei Tian

Abstract

Cure rate models, introduced by Boag (1949), are very commonly used while modelling lifetime data involving long time survivors. Applications of cure rate models can be seen in biomedical science, industrial reliability, finance, manufacturing, demography and criminology. In this thesis, cure rate models are discussed under a competing cause scenario, with the assumption of proportional odds (PO) lifetime distributions for the susceptibles, and statistical inferential methods are then developed based on right-censored data.

In Chapter 2, a flexible cure rate model is discussed by assuming the number of competing causes for the event of interest following the Conway-Maxwell (COM) Poisson distribution, and their corresponding lifetimes of non-cured or susceptible individuals can be described by PO model. This provides a natural extension of the work of Gu *et al.* (2011) who had considered a geometric number of competing causes. Under right censoring, maximum likelihood estimators (MLEs) are obtained by the use of expectation-maximization (EM) algorithm. An extensive Monte Carlo simulation study is carried out for various scenarios, and model discrimination between some well-known cure models like geometric, Poisson and Bernoulli is also examined. The goodness-of-fit and model diagnostics of the model are also discussed. A cutaneous melanoma dataset example is used to illustrate the models as well as the inferential methods.

Next, in Chapter 3, the destructive cure rate models, introduced by Rodrigues *et al.*

(2011), are discussed under the PO assumption. Here, the initial number of competing causes is modelled by a weighted Poisson distribution with special focus on exponentially weighted Poisson, length-biased Poisson and negative binomial distributions. Then, a damage distribution is introduced for the number of initial causes which do not get destroyed. An EM-type algorithm for computing the MLEs is developed. An extensive simulation study is carried out for various scenarios, and model discrimination between the three weighted Poisson distributions is also examined. All the models and methods of estimation are evaluated through a simulation study. A cutaneous melanoma dataset example is used to illustrate the models as well as the inferential methods.

In Chapter 4, frailty cure rate models are discussed under a gamma frailty wherein the initial number of competing causes is described by a Conway-Maxwell (COM) Poisson distribution in which the lifetimes of non-cured individuals can be described by PO model. The detailed steps of the EM algorithm are then developed for this model and an extensive simulation study is carried out to evaluate the performance of the proposed model and the estimation method. A cutaneous melanoma dataset as well as a simulated data are used for illustrative purposes.

Finally, Chapter 5 outlines the work carried out in the thesis and also suggests some problems of further research interest.

KEY WORDS: Cure rate models; Mixture model; Long-term survivors; COM-Poisson distribution; Weighted Poisson distribution; EM algorithm; Right censoring; Non-informative censoring; Profile likelihood; Asymptotic variances and covariances; Maximum likelihood estimation; Likelihood-ratio test; Exponential distribution; Proportional odds model; Weibull distribution; Log-logistic distribution; Gamma distribution; Mixture of chi-square;

Akaike Information Criterion (AIC); Bayesian Information Criterion (BIC); Model discrimination; Monte Carlo simulations; Goodness-of-fit test; Cutaneous melanoma.

Acknowledgements

I will take this opportunity to thank my extraordinary supervisor, Dr. Narayanaswamy Balakrishnan. It was a wonderful experience to work under his supervision. Also, I want to thank the supervisory committee members, Dr. Roman Viveros-Aguilera, and Dr. Aaron Childs for their support, recommendations, and suggestions, and Professors Hans U. Boden and Bartosz Protas for their suggestions and help.

I am always grateful to my parents Jingping Yuan and Shenghu Feng, my husband Hon-yiu So for their encouragement, and my friends Suvra Pal, Sandip Barui, Xin Zhang, He Yi, Katherine Davies, Tao Tan, Kai Liu, Mu He, Jing Cai, Yanqiao Wang and all the other research students for their kind assistance.

Notation and abbreviations

- AIC... Akaike Information Criterion
- BIC... Bayesian Information Criterion
- cdf...cumulative distribution function
- COM-Poisson...Conway-Maxwell Poisson
- CP...coverage probability
- i.i.d. ...independent and identically distributed
- MCMC...Markov chain Monte Carlo
- pdf...probability density function
- PO...proportional odds
- RMSE...root mean square error
- SE...standard error
- sf...survival function

Contents

List of Figures

Chapter 1

Introduction

1.1 Cure rate model

Due to the significant progress in technology and medical science in the past several decades, it is highly likely that many patients can be cured and do not experience the disease again over a long period of time following a good prognosis and efficient treatment. So, the cured proportion of the overall population should not be ignored and becomes an important indicator for the effect of the treatment. Recently, the methodology that accounts for cure has received considerable wide attention in the survival analysis literature. The patients who are cured are called long-time survivors or immune individuals, while the others, who still experience the disease or relapse or recurrence of the disease, are called susceptibles. Statistical models accommodating such a surviving fraction are generally referred to as cure rate models.

The basic cure rate model can be seen as a two-component mixture model. Suppose we have an indicator variable I such that $I = 0$ if the subject is immune/cured (belongs to set I_0) with probability p_0 and $I = 1$ if the subject is susceptible (belongs to set I_1) with probability $1 - p_0$, where p_0 is the cure rate. The cumulative distribution function of the overall population can be viewed as a mixture of the two populations as follows:

$$\begin{aligned} F_p(y) &= P[Y \leq y | I = 0] P(I = 0) + P[Y \leq y | I = 1] P(I = 1) \\ &= F_s(y)(1 - p_0); \end{aligned} \tag{1.1}$$

thus, the survival function of the overall population is then evidently

$$S_p(y) = p_0 + (1 - p_0) S_s(y), \tag{1.2}$$

where p_0 is the probability of cure, $S_s(y)$ is the survival function of the non-cured or susceptible individuals in the population, and $F_s(y)$ is the corresponding distribution function of the susceptibles. It is to be noted that $S_p(y)$ is an improper survival function since it approaches p_0 as y goes to ∞. The improper density function can be easily obtained as $f_p(y) = -\frac{\partial S_p(y)}{\partial y}$. Moreover, a proper survival function of the susceptible population can be found from the overall population and the cure rate as

$$S_s(y) = \frac{S_p(y) - p_0}{1 - p_0}, \; y > 0. \tag{1.3}$$

More generally, a cure model can be approached through a competing risks set-up as follows. Suppose M is an unobservable random variable denoting the number of competing causes related to the occurrence of an event of interest. Let W_j, $j = 1, \ldots, m$, be the random variable denoting the time-to-event for the jth competing cause. Given $M = m$, $W_1, \ldots, W_m$ are assumed to be independent and identically distributed (i.i.d.) with a common cumulative distribution function (c.d.f.) $F(w) = 1 - S(w)$. Then, the population time-to-event or lifetime is given by

$$Y = \min\{W_0, W_1, \dots, W_m\}, \tag{1.4}$$

where W_0 is corresponding to the individual who are not susceptible to the event occurrence (namely, with infinite lifetime); that is, $P(w_0 = \infty) = 1$. This leads to a proportion of the cured group, known as cure rate. The survival function for the entire population is then given by [Rodrigues *et al.* (2009)]

$$S_p(y) = \sum_{m=0}^{\infty} P(M = m)[S(y)]^m = A_M(S(y)), \tag{1.5}$$

where $A_M(.)$ is the probability generating function (p.g.f.) of M.

1.2 A brief literature review

The cure rate model was first introduced by Boag (1949) followed by Berkson and Gage (1952), and have been subsequently studied by many authors. Primarily, cure data have been analyzed in the literature by the structure of the underling survival model of the non-cured individuals $S_s(t)$ as proportional hazards (PH) mixture cure model [Peng and Dear (2000), Liu *et al.* (2012), and Sy and Taylor (2000)], accelerated failure time (AFT) mixture cure rate model [Zhang and Peng (2007), Lu (2010), and Li and Taylor (2002)], accelerated hazards (AH) mixture cure rate model [Zhang and Peng (2009)], and proportional odds (PO) mixture cure rate model [Gu *et al.* (2011), and Mao and Wang (2010)]. Later, Chen *et al.* (1999) pointed out that the mixture cure rate model has several drawbacks from both frequentist and Bayesian perspective, and so introduced a promotion time cure model from

Bayesian point of view.

Rodrigues *et al.* (2009) proposed the Conway-Maxwell-Poisson cure rate model by extending the model in Chen *et al.* (1999) and also including the models of Boag (1949) and Berkson and Gage (1952). Subsequently, these authors developed a destructive cure rate model which assumes the initial number of competing causes to undergo a destructive process due to a treatment, such as chemotherapy or radiation. This model has a more realistic interpretation of the underlying biological mechanism than the usual proportion time cure model.

Vaupel *et al.* (1979) introduced a frailty model which provides a convenient way to accommodate unobserved covariates and/or heterogeneity in survival data in the form of a frailty term. The frailty model was then studied by Clayton (1978) for bivariate data and Greenwood and Yule (1920) for recurrent data. Common choices for the frailty distribution is the gamma distribution and positive stable distribution [Hougaard (1986)] as they all have simple Laplace transforms and are therefore convenient to use. Balakrishnan and Peng (2006) introduced a generalized gamma frailty model, which includes gamma, Weibull and log-normal frailty models all as special cases.

Applications of cure rate models are not limited to biomedical studies, and can also be seen in many other fields such as industrial reliability, finance, manufacturing, demography, and criminology. For a detailed research in these applications, interested readers may refer to Klebanov *et al.* (1993), Ibrahim *et al.* (2005), and Hoggart and Griffin (2001).

1.3 COM-Poisson cure rate model

The Conway-Maxwell (COM) Poisson distribution, introduced by Conway and Maxwell (1962), is a generalization of the Poisson distribution allowing for under- and over-dispersion, depending on the value of the dispersion parameter. It thus provides a continuous bridge between under- and over-dispersion.

Suppose the number of the competing causes M follows a COM-Poisson distribution. The probability mass function (p.m.f.) of M is given by

$$P(M = m; \eta, \phi) = \frac{1}{Z(\eta, \phi)} \frac{\eta^m}{(m!)^\phi},\ m = 0, 1, 2, \ldots, \tag{1.6}$$

where the normalization constant is given by,

$$Z(\eta, \phi) = \sum_{j=0}^{\infty} \frac{\eta^j}{(j!)^\phi}, \tag{1.7}$$

with $\phi \geq 0$ and $\eta > 0$. Here, the proportion of cured individuals from the whole population is the probability of no competing risk, that is,

$$p_0 = P(M = 0; \eta, \phi) = (Z(\eta, \phi))^{-1}. \tag{1.8}$$

As a weighted Poisson random varaiable (r.v.), M leads to a Poisson r.v. with mean equal to η when $\phi = 1$, and M leads to a under- or over-dispersion if $\phi > 1$ or $\phi < 1$ [see Shmueli *et al.* (2005), Kadane *et al.* (2006), and Kokonendji *et al.* (2008, 2009)]. For example, M approaches the Bernoulli r.v. with parameter $\frac{1}{1+\eta}$ when $\phi \to \infty$ and $Z(\eta, \phi) \to 1 + \eta$, and M reduces to a Geometric r.v. with parameter $1 - \eta$ if $\phi = 0$, $\eta < 1$ and $Z(\eta, \phi) = \frac{1}{1-\eta}$. Note that M is undefined for $\eta \geq 1$ and $\phi = 0$. The population survival function and density function of the time-to-event Y is then

$$S_p(y) = \frac{Z(\eta S(y), \phi)}{Z(\eta, \phi)}, \tag{1.9}$$

$$f_p(y) = \frac{1}{Z(\eta, \phi)} \frac{f(y)}{S(y)} \sum_{j=1}^{\infty} \frac{j(\eta S(y))^j}{(j!)^{\phi}}, \tag{1.10}$$

Note that as $y \to \infty$, $S_p(y) \to p_0 > 0$. Hence, $S_p(y)$ is not a proper survival function.

For a detailed discussion, interested readers may refer to Balakrishnan and Pal (2012, 2013, 2014, 2016), Balakrishnan *et al.* (2017), and Rodrigues *et al.* (2009).

1.4 Destructive weighted Poisson cure rate models

Suppose M is an unobservable random variable representing the initial number of competing causes. Given $M = m$, let X_j $j = 1, \ldots, m$, be independent Bernoulli variables with success probability $p \in (0, 1)$, independently of M. Define the following "damage" random variable D for the number of initial competing causes which are still not destroyed:

$$D = \begin{cases} X_1 + X_2 + \cdots + X_M, & M > 0 \\ 0, & M = 0. \end{cases} \tag{1.11}$$

The distribution of D, conditional on $M = m$, is known as the "damaged" distribution. It is distributed binomially with parameters m and p; see Rodrigues *et al.* (2011) for details. Note that $D \leq M$. Also, $p = 1$ means no damage or destruction occurs. Given the undamaged number of initial competing causes $D = d$, let W_j (latent) be the time-to-event associated with the j-th competing cause. In the competing causes scenario, the number of competing causes D and the lifetimes W associated with these causes are not observable (latent variables). Given $M = m$, $W_1, \ldots, W_D$ are assumed to be independent and identically distributed with common cumulative distribution function $F(w)$ and survival function $S(w)$. The observed lifetime is then given by

$$Y = \min\{W_0, W_1, \ldots, W_D\}, \tag{1.12}$$

where $P(w_0 = \infty) = 1$ as before.

A random variable X follows a weighted Poisson distribution if its p.m.f. is given by

$$p(x; \phi, \lambda) = \frac{\omega(x; \tau) p(x; \lambda)}{E_\lambda(\omega(X; \tau))}, \tag{1.13}$$

where $\omega(X; \tau)$ is a non-negative weight function with parameter vector τ, $p(x; \lambda)$ is the p.m.f. of a standard Poisson distribution with parameter $\lambda > 0$, and E_λ denotes the expectation with respect to $p(x; \lambda)$, i.e.,

$$E_\lambda(\omega(X; \tau)) = \sum_{x=0}^{\infty} \omega(x; \tau) p(x; \lambda). \tag{1.14}$$

From the above weighted Poisson model, we focus here in this thesis on three special cases.

1.4.1 Destructive negative binomial cure rate model

Suppose M follows a weighted Poisson distribution with parameter $\frac{\eta\phi}{1+\eta\phi}$ and weighted function $\Gamma(m + \frac{1}{\phi})$ with $\eta > 0$ and $\phi > 0$, and with p.m.f. as,

$$P[M = m; \eta, \phi] = \frac{\Gamma(m + \frac{1}{\phi})}{\Gamma(\frac{1}{\phi}) m!} \left(\frac{\eta\phi}{1 + \eta\phi} \right)^m \left(\frac{1}{1 + \eta\phi} \right)^{\frac{1}{\phi}}, m = 0, 1, 2, 3, \ldots \tag{1.15}$$

It can be seen that (1.15) is a negative binomial distribution with parameters $\eta > 0$ and $\phi > 0$. If we set $\phi = 1$, (1.15) becomes

$$P[M = m; \eta, \phi] = \left(\frac{\eta}{1+\eta}\right)^m \frac{1}{1+\eta}, \quad m = 0, 1, 2, 3, \ldots, \tag{1.16}$$

a geometric distribution with probability of success $\frac{1}{1+\eta}$. The mass function of the damage distribution D in (1.11) in this case is

$$\begin{aligned} P[D = d; \eta, \phi, p] &= \sum_{m=d}^{\infty} P(M = m; \eta, \phi) P(D = d | M = m) \\ &= \frac{\Gamma(d + \frac{1}{\phi})}{\Gamma(\frac{1}{\phi}) d!} \left(\frac{\eta\phi p}{1+\eta\phi p}\right)^d \left(\frac{1}{1+\eta\phi p}\right)^{\frac{1}{\phi}}, \\ & d = 0, 1, 2, 3, \ldots, m = 0, 1, 2, 3, \ldots, d \leq m, \end{aligned} \tag{1.17}$$

which is also a negative binomial distribution. The cure rate then is,

$$P(D = 0; \eta, \phi, p) = \left(\frac{1}{1+\eta\phi p}\right)^{\frac{1}{\phi}}. \tag{1.18}$$

The overall population survival function of the time-to-event Y is given by

$$S_p(y) = P[Y > y] = \sum_{d=0}^{\infty} P(D = d)[S(y)]^d = (1 + \eta\phi p F(y))^{-\frac{1}{\phi}}.$$

It is to be noticed that the population function $S_p(y)$ defined above is not a proper survival function. The improper density function can be easily obtained as

$$f_p(y) = -\frac{\delta S_p(y)}{\delta y} = \frac{\eta p f(y) S_p(y)}{1 + \eta\phi p F(y)}. \tag{1.19}$$

1.4.2 Destructive exponentially weighted Poisson cure rate model

Suppose M follows a weighted Poisson distribution with parameter $\eta > 0$ and exponential weight function $e^{\phi m}$, where $\phi \in \Re$, i.e.,

$$P[M = m; \eta, \phi] = e^{-\eta e^{\phi}} \frac{(\eta e^{\phi})^m}{m!},\ m = 0, 1, 2, 3, \dots \tag{1.20}$$

It can be seen that (1.20) is a Poisson distribution with parameter ηe^{ϕ}. If $\phi = 0$, (1.20) becomes

$$P[M = m; \eta, \phi] = e^{-\eta} \frac{\eta^m}{m!},\ m = 0, 1, 2, 3, \dots \tag{1.21}$$

a Poisson distribution parameter η. The p.m.f. of D is,

$$\begin{aligned} P[D = d; \eta, \phi, p] &= e^{-\eta p e^{\phi}} \frac{(\eta p e^{\phi})^d}{d!}, \\ &\qquad d = 0, 1, 2, 3, \dots, m = 0, 1, 2, 3, \dots, d \leq m. \end{aligned} \tag{1.22}$$

The cure rate in this case is,

$$P(D = 0; \eta, \phi, p) = e^{-\eta p e^{\phi}}. \tag{1.23}$$

The overall population survival function of the time-to-event Y is given by

$$S_p(y) = P[Y > y] = \sum_{d=0}^{\infty} P(D = d)[S(y)]^d = \sum_{d=0}^{\infty} e^{-\eta p e^{\phi}} \frac{(\eta p e^{\phi} S(y))^d}{d!} = e^{-\eta p e^{\phi} F(y)}.$$

The corresponding density function is given by

$$f_p(y) = S_p(y)f(y)\eta p e^{\phi} \tag{1.24}$$

1.4.3 Destructive length-biased Poisson cure rate model

Suppose M follows a weighted Poisson distribution with parameter $\eta > 0$ and weight function m, i.e.,

$$P[M = m; \eta, \phi] = \frac{m\frac{e^{-\eta}\eta^m}{m!}}{\sum_{m=0}^{\infty} m\frac{e^{-\eta}\eta^m}{m!}} = e^{-\eta}\frac{\eta^{m-1}}{(m-1)!}, m = 0, 1, 2, 3, \ldots \tag{1.25}$$

It is to be noted that D is a weighted Poisson random variable with parameter ηp and a weighted function $d + \eta(1-p)$. The p.m.f. of D is as follows,

$$\begin{aligned} P[D = d; \eta, p] &= \frac{e^{-\eta p}(\eta p)^d}{d!}(1 - p + \frac{d}{\eta}), \\ & d = 0, 1, 2, 3, \ldots, m = 0, 1, 2, 3, \ldots, d \leq m, \end{aligned} \tag{1.26}$$

The cure rate in this case is,

$$p_0 = P(D = 0; \eta, p) = (1-p)e^{-\eta p}. \tag{1.27}$$

The overall population survival function of the time-to-event Y is given by

$$S_p(y) = P[Y > y] = \sum_{d=0}^{\infty} P(D = d)[S(y)]^d = e^{-\eta p F(y)}(1 - pF(y)),$$

and the corresponding p.d.f. is,

$$f_p(y) = p\left\{\eta + \frac{1}{1 - pF(y)}\right\} S_p(y) f(y) \tag{1.28}$$

For a detailed discussion on the three special cases of the weighted Poisson model, interested readers may refer to Balakrishnan and Pal (2013, 2016), and Pal and Balakrishnan (2016, 2017, 2018).

1.5 Proportional odds model for lifetime data

In this thesis, we assume a proportional odds model for the distribution of W_j with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w) = \theta O_0(w), \tag{1.29}$$

where $O(w) = S(w)/F(w)$ is the odds of survival up to w, the proportionality term θ is linked to covariates as $e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2}$ with $\boldsymbol{x}_c = (x_1, \ldots, x_p)'$ being a vector of p covariates, $\boldsymbol{\gamma}_2 = (\gamma_{21}, \ldots, \gamma_{2p})'$ is the proportional odds regression coefficients, and $O_0(w)$ is the baseline odds function. We can further obtain the survival function of W_j as

$$S(w) = [1 + e^{-\boldsymbol{x}_c' \boldsymbol{\gamma}_2}(S_0(w)^{-1} - 1)]^{-1}, \tag{1.30}$$

with the corresponding probability density function (p.d.f.)

$$f(w) = f_0(w) e^{-\boldsymbol{x}_c' \boldsymbol{\gamma}_2} [(1 - S_0(w)) e^{-\boldsymbol{x}_c' \boldsymbol{\gamma}_2} + S_0(w)]^{-2}. \tag{1.31}$$

where S_0 and f_0 are the baseline survival function and baseline p.d.f., respectively.

In this thesis, two different baseline distributions for the proportional odds survival model corresponding to the time-to-event random variable are considered, namely, Weibull and log-logistic distributions. It should also be noted that the log-logistic distribution in fact processes the proportional odds property, while the Weibull distribution does not.

1.5.1 Proportional odds model with Weibull baseline

The survival function of W under a proportional odds model with Weibull baseline is

$$S(w, \boldsymbol{\gamma}) = [1 + e^{-\boldsymbol{x}_c' \boldsymbol{\gamma}_2}(e^{(\gamma_1 w)^{1/\gamma_0}} - 1)]^{-1}, w > 0, \tag{1.32}$$

and the corresponding p.d.f. is

$$f(w, \boldsymbol{\gamma}) = (\gamma_1 w)^{1/\gamma_0} e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2 - (\gamma_1 w)^{1/\gamma_0}} [e^{-(\gamma_1 w)^{1/\gamma_0}} (e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2} - 1) + 1]^{-2} / (\gamma_0 w), w > 0, \tag{1.33}$$

where $\gamma_0 > 0$ and $\gamma_1 > 0$ are the shape and scale parameters, respectively.

1.5.2 Proportional odds model with log-logistic baseline

Suppose the baseline distribution is a log-logistic distribution with $\gamma_0 > 0$ and $\gamma_1 > 0$ as the scale and shape parameters, respectively. Then, the corresponding odds function of W_i is given by

$$O(w, ; \boldsymbol{x}_c', \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1}}{w^{\gamma_1}} e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2} = O_0(w, ; \gamma_0, \gamma_1) e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2}. \tag{1.34}$$

We observe that W_i still follows a two-parameter log-logistic distribution ($\gamma_0, \gamma_1 > 0$) with shape parameter γ_1 and scale parameter $\gamma_0 e^{-\boldsymbol{x}_c' \boldsymbol{\gamma}_2/\gamma_1}$, and with corresponding survival function

$$S(w, \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2}}{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c' \boldsymbol{\gamma}_2} + w^{\gamma_1}}, w > 0. \tag{1.35}$$

Note that the mean does not exist if $\gamma_1 < 1$ and the variance does not exist if $\gamma_1 < 2$.

1.6 Gamma frailty proportional odds model

Frailty models provide a convenient way to accommodate unobserved covariates and/or heterogeneity in survival data in the form of a frailty term. Assume a proportional odds model with frailty term for the distribution of W_j ($j = 1, 2, \ldots$), with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w|r) = r\theta O_0(w) \tag{1.36}$$

where $O(w) = S(w)/F(w)$ is the odds of survival up to time w, the probability term θ is linked to covariates as $e^{\boldsymbol{\alpha}' \boldsymbol{x}_c}$ with $\boldsymbol{x}_c = (x_1, \ldots, x_p)'$ is a vector of p covariates, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ is the proportional odds regression coefficients, $O_0(w)$ is the baseline odds function, and r is the frailty parameter following a gamma distribution with shape $k > 0$ and scale $\xi > 0$. The mean and variance of r is k/ξ and k/ξ^2, respectively. Here, we set mean equal to 1 to avoid non-identifiability in the model so that $k = \xi$. Therefore, the p.d.f.

of r can be written as

$$f(r) = \frac{r^{\xi-1}\xi^{\xi}e^{-r\xi}}{\Gamma(\xi)},\ r \geq 0,\ \xi > 0. \tag{1.37}$$

We can further obtain the survival function of W_j, through unconditioning, as

$$S(t_i) = \int_0^{\infty} \frac{r_i S_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_c}}{r_i S_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_c} + F_0(t_i)} f_r(r_i) dr_i,\ t_i > 0, \tag{1.38}$$

with the corresponding probability density function (p.d.f.) as

$$f(t_i) = \int_0^{\infty} \frac{r_i f_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_c}}{(r_i S_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_c} + F_0(t_i))^2} f_r(r_i) dr_i,\ t_i > 0. \tag{1.39}$$

1.6.1 Gamma frailty proportional odds model with Weibull baseline

The survival function of W under a gamma frailty proportional odds model with Weibull baseline is given by

$$S(w, \boldsymbol{\gamma}) = \int_0^{\infty} [1 + e^{-\boldsymbol{\alpha}'\boldsymbol{x}_c}/r_i(e^{(\gamma_1 w)^{1/\gamma_0}} - 1)]^{-1} f_r(r_i) dr_i, w > 0, \tag{1.40}$$

and the corresponding p.d.f. is

$$f(w, \boldsymbol{\gamma}) = \int_0^{\infty} (\gamma_1 w)^{1/\gamma_0} r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_c - (\gamma_1 w)^{1/\gamma_0}} [e^{-(\gamma_1 w)^{1/\gamma_0}}(r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_c} - 1) + 1]^{-2}/(\gamma_0 w) f_r(r_i) dr_i, w > 0, \tag{1.41}$$

where $\gamma_0 > 0$ and $\gamma_1 > 0$ are the shape and scale parameters, respectively.

1.6.2 Gamma frailty proportional odds model with log-logistic baseline

Suppose the baseline distribution is a log-logistic distribution with $\gamma_0 > 0$ and $\gamma_1 > 0$ as the scale and shape parameters, respectively. Then, the corresponding odds function of W_i is given by

$$O(w_i; \boldsymbol{x}_c', \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1}}{w_i^{\gamma_1}} e^{\boldsymbol{\alpha}'\boldsymbol{x}_c} = r_i O_0(w_i; \gamma_0, \gamma_1) e^{\boldsymbol{\alpha}'\boldsymbol{x}_c}. \tag{1.42}$$

We observe that W_i still follows a two-parameter log-logistic distribution $(\gamma_0, \gamma_1 > 0)$ with shape parameter γ_1 and scale parameter $\gamma_0 e^{-\boldsymbol{\alpha}'\boldsymbol{x}_c/\gamma_1}$, and with corresponding survival function

$$S(w_i, \boldsymbol{\gamma}) = \int_0^\infty \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{\alpha}'\boldsymbol{x}_c}}{\gamma_0^{\gamma_1} e^{\boldsymbol{\alpha}'\boldsymbol{x}_c} + w_i^{\gamma_1}} f_r(r_i) dr_i, \ w_i > 0. \tag{1.43}$$

Note that the mean does not exist if $\gamma_1 < 1$ and the variance does not exist if $\gamma_1 < 2$.

1.7 Data and the likelihood

Suppose the time-to-event is not completely observed and is subject to non-informative right censoring, which means that the data above a certain value is not observed. Therefore, the observation time T_i, for the ith subject, would be the minimum of the censoring time C_i and the actual lifetime Y_i, i.e.,

$$T_i = \min\{Y_i, C_i\}, i = 1, \ldots, n. \tag{1.44}$$

We define an indicator function $\delta_i = I(Y_i \leq C_i)$ for the i-th subject such that $\delta_i = 1$ if the lifetime is observed while $\delta_i = 0$ if the lifetime is right censored, Δ_0 and Δ_1 are sets with all the i's equal to 0 and 1, respectively, and set Δ^* contains all the i's. It is to be noted that the cure rate $p_0 = Z(\eta, \phi)^{-1}$ is purely a function of η for a fixed value of ϕ. The range of $1/p_0$ is from 1 to infinity and it is monotone in η. Therefore, it is natural to use a logistic regression model $H_\phi(\eta) = 1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$ to link the covariate $\boldsymbol{x}_i$ to the cured proportion p_{0i}, i.e.,

$$p_{0i} = p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) = Z(\eta, \phi)^{-1} = H_\phi(\eta)^{-1} = (1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}, \tag{1.45}$$

where p_{0i} is the cured proportion for the ith category, $\boldsymbol{x}_i = (1, \boldsymbol{x}_{ic}')' = (1, x_{i1}, \ldots, x_{ip})'$ is a vector of $p+1$ covariates, and $\boldsymbol{\beta}$ is the vector of regression coefficients. Under this link function, η would equal $H^{-1}(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$, i.e., η can be calculated from the inverse function of $H_\phi(.)$ analytically for the Geometric, Poisson and Bernoulli distributions, and by using numerical method for the general COM-Poisson distribution.

For n pairs of observations $(\boldsymbol{t}, \boldsymbol{\delta})$=$\{(t_1,\ \delta_1),\ldots,(t_n,\ \delta_n)\}$, the observed data likelihood function under the non-informative censoring is then given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \{f_p(t_i; \boldsymbol{\theta})\}^{\delta_i} \{S_p(t_i; \boldsymbol{\theta})\}^{1-\delta_i}, \tag{1.46}$$

where $\boldsymbol{\theta}$ is the set of parameters $(\phi, \boldsymbol{\beta}', \boldsymbol{\gamma}')$, which is equivalent to

$$\begin{aligned} L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto & \prod_{i=1,\{\delta_i=1\}} f_p(t_i; \boldsymbol{\theta}) \prod_{i=1,\{\delta_i=0\}} S_p(t_i; \boldsymbol{\theta}) \\ = & \prod_{i \in \Delta_1} f_p(t_i; \boldsymbol{\theta}) \prod_{i \in \Delta_0} \{p_0 + (1 - p_0) S_s(t_i; \boldsymbol{\theta})\}. \end{aligned} \tag{1.47}$$

1.8 Likelihood inference

1.8.1 EM algorithm

We develop here an Expectation-Maximization (EM) algorithm for obtaining the MLE of $\boldsymbol{\theta}$, and a profile likelihood approach for the estimation of the dispersion patameter ϕ. It is well-known that EM is an effective technique for finding the MLEs of unknown parameters of a model involving unobserved variables (for further discussion, refer to McLachlan and Krishnan (2007)). In our model, the random variable I_i's are observed for i in the set Δ_1 and unobserved for i in the set Δ_0, where $I_i = 1$ if the individual is susceptible and $I_i = 0$ if the individual is cured. Let us denote the set of complete data by ($\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{I}$)=$\{(t_1, \delta_1, \boldsymbol{x}_1, I_1), \ldots, (t_n, \delta_n, \boldsymbol{x}_n, I_n)\}$. The complete data likelihood function is then

$$L_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I}) \propto \prod_{i\in\Delta_1} f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) \prod_{i\in\Delta_0} p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)^{1-I_i}[(1 - p_0(\boldsymbol{\beta}, \boldsymbol{x}_i))S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta})]^{I_i}, \tag{1.48}$$

where $\mathbf{I} = (I_1, \ldots, I_n)'$, $\boldsymbol{x}_{ic} = (x_{i1}, \ldots, x_{ip})'$ and $\boldsymbol{x}_i = (1, \boldsymbol{x}'_{ic})'$. The corresponding complete log-likelihood function

$$\begin{aligned} l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I}) = \text{constant} &+ \sum_{i\in\Delta_1} \log f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) + \sum_{i\in\Delta_0} (1 - I_i)\log p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) \\ &+ \sum_{i\in\Delta_0} I_i \log[1 - p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)] + \sum_{i\in\Delta_0} I_i \log S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}). \end{aligned} \tag{1.49}$$

1.8.2 E-step

The expectation step is achieved by calculating the expected value of the complete data log-likelihood function with respect to the conditional distribution of the unobserved I_i's ($i \in \Delta_0$), given the observed data $\boldsymbol{O} = \{(t_i, \delta_i, \boldsymbol{x}_i), i \in \Delta_1\}$ and the current estimates of the parameters $\boldsymbol{\theta}^{(k)} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ for a fixed value of ϕ. Let us denote this function as

$$Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) = E(l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I})|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}), \tag{1.50}$$

at the k-th iteration step. In our model, I_i's are Bernoulli random variables and we can easily find the conditional expectation if the ith individual is susceptible as

$$\begin{aligned}
\boldsymbol{\pi}_i^{(k)} &= \mathrm{E}(I_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}) \\
&= P(I_i = 1|T > t) \\
&= \frac{(1 - p_0(\boldsymbol{\beta}^{(k)}, \boldsymbol{x}_i))S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}^{(k)})}{S_p(t_i, \boldsymbol{x}_i; \boldsymbol{\theta}^{(k)})}|_{\theta=\theta^{(k)}}.
\end{aligned} \tag{1.51}$$

Now, for a fixed value of ϕ, the Q function is given by

$$\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) = &\sum_{I \in \Delta_1} \log f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) + \sum_{i \in \Delta_0} (1 - \boldsymbol{\pi}_i^{(k)}) \log p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) \\
&+ \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log[1 - p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)] + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}).
\end{aligned} \tag{1.52}$$

1.8.3 M-step

The M-step is done by maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function in (1.52) in order to obtain the improved estimate of $\boldsymbol{\theta}$, i.e.,

$$\boldsymbol{\theta}^{*(k+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}). \tag{1.53}$$

The MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not have explicit expressions. Here, the numerical maximization is carried out by Newton-Raphson method.

For a fixed value of ϕ, the E-step and M-step are alternated until the parameter estimate converges to a desired level of accuracy. The parameter ϕ is determined by using the profile likelihood technique. For this purpose, we consider a range of ϕ with small increment, and then for each value of ϕ, the MLEs of other parameters are found, and the estimate with the largest likelihood is chosen as the final estimate. The explicit forms of the first- and second-order derivatives of the Q function as well as the update function for the case of COM-Poisson distribution and the special cases are presented in the Appendix.

1.8.4 Estimation of standard errors

We may approximate the asymptotic variance-covariance matrix of the MLEs $(\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}')'$ by inverting the observed Fisher information matrix of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for a fixed value of ϕ. The components of the observed Fisher information matrix can then be calculated from the negative of the second-order derivatives of the complete data likelihood function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ (for detailed information, refer to Louis (1982)). Thus, we can obtain the standard errors of the estimates and then construct corresponding asymptotic confidence intervals for the parameters.

1.8.5 Estimation of the cure rate and its standard error

Suppose $\hat{\boldsymbol{\beta}}$ is the MLE of the regression coefficient $\boldsymbol{\beta}$. The estimated cure rate for the corresponding group i is a function of $\hat{\boldsymbol{\beta}}$. The standard error of $\hat{p}_{0i}$ can then be found through delta method as

$$\mathrm{sd}(\hat{p_{0i}}) = \sqrt{\left(\frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_0}, \ldots, \frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_p}\right) \mathrm{var}\hat{\boldsymbol{\beta}} \left(\frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_0}, \ldots, \frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_p}\right)'}. \tag{1.54}$$

1.9 Simulation study and real data analysis

The accuracy, precision, and robustness of the models and the estimation techniques are studied and validated using extensive Monte Carlo simulations. Different scenarios are considered in the simulation studies by varying sample sizes, censoring proportions, cure rates, and the parameters in different distributions. The estimates were calculated through EM method. The empirical Bias, standard errors(SE), root Mean Square Error (RMSE), and 95% coverage probabilities (CPs) are all reported. In addition, the estimated cure rate with corresponding SEs and 95% CPs are also computed. The codes corresponding to the different methods are written in R software with version 3.2.2. The McMaster math department computer Bayes, Anatolius and the shared Hierarchical Academic Research Computing Network (SHARCNET) were used to compile all the R-codes.

For model validation and model discrimination, for the COM-Poisson model, the asymptotic distribution of the test statistic Λ, under $H_0 : \phi = 1$ follows a χ^2 distribution with one degree of freedom. However, the distribution of the test statistic Λ in the boundary cases when $\phi = 0$ (Geometric) and $\phi \longrightarrow \infty$ (Bernoulli) has a mixture distribution of χ^2_0 and χ^2_1 distributions such that $P(\Lambda \le \lambda) = \frac{1}{2} + \frac{1}{2}P(\Lambda_1 \le \lambda_1)$, where $\Lambda_1 \sim \chi^2_1$, χ^2_0 is chi-square distribution with 0 degrees of freedom and χ_1^2 is the chi-square distribution with one degree of freedom.

In this thesis, we consider a malignant cancer data, a cutaneous melanoma data, to illustrate the performance of the proposed model and the fitting methodology. These data were first introduced by Kirkwood *et al.* (2000), and subsequently studied by many authors including Balakrishnan and Pal (2012, 2014, 2015, 2016), Balakrishnan *et al.* (2017), and Rodrigues *et al.* (2009). These data were taken from Ibrahim *et al.* (2005), and were originally used to detect the prospective treatment performance on the high-dose interferon alfa-2b therapy in order to prevent the recurrence of the disease. The study included 427 patients in total from years 1991 to 1995 and followed up until year 1998. Among them, 10 patients were removed in our analysis due to the missingness of the tumor thickness data. The overall percentage of censored observations is 55.6%. The mean and standard deviation of the observed lifetimes are 3.18 and 1.69 in years, respectively. We choose the nodule categories based on the tumor thickness as the only covariate. The subjects were therefore divided into four different nodule categories ($x = 0, 1, 2, 3$), with corresponding sample sizes $n_1 = 111$, $n_2 = 137$, $n_3 = 87$, and $n_4 = 82$. The percentage of censored observations for the group were 67.57%, 61.31%, 52.87%, 32.93%. See Figure 1.1 for a plot of the lifetimes of susceptibles for these four nodule categories.

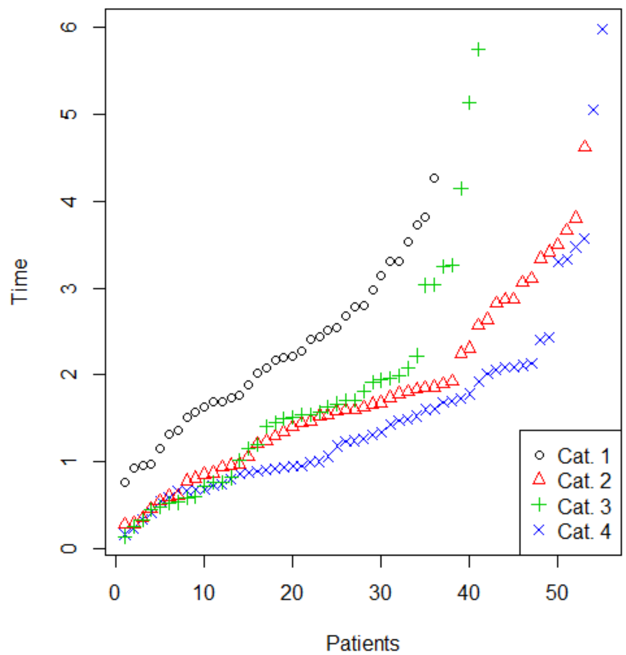

Figure 1.1: Lifetime of the susceptibles of the four nodule categories in the cutaneous melanoma data

Another dataset considered in this thesis is the one available in the “timereg” package in R software [see Andersen *et al.* (2017), and Drzewiecki *et al.* (1980)], which is a well-known dataset in the literature and studied by many authors including Rodrigues *et al.* (2011). It includes 205 patients (refer to Figure 1.2) corresponding to the survival of individuals after operation for malignant melanoma and collected at Odense University Hospital by K.T. Drzewiecki during the period 1962 to 1977 and followed until 1977. Among them, there were 57 and 14 patients died from melanoma and other causes, respectively. All rest of them were alive at the end of the study and thus censored.

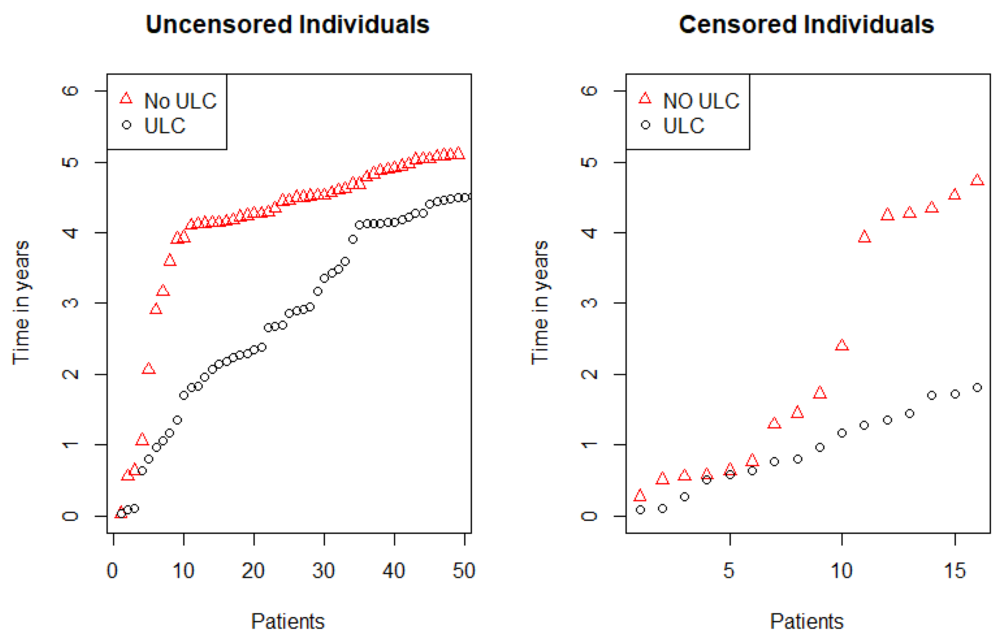

Figure 1.2: Lifetimes of patients in the cutaneous melanoma data with and without ulceration (ULC)

1.10 Scope of the thesis

The following chapters present the detailed choice of link function, specific likelihood function, Q function, step by step EM algorithm, simulation study results and the data analysis results for each described cure rate model. In Chapter 2, a flexible cure rate model is presented by assuming the number of competing causes for the event of interest to follow the COM-Poisson distribution. The lifetimes of susceptible individuals follow a PO model. In Chapter 3, we introduce a flexible destructive cure rate model for lifetime data. We assume the number of competing causes of the event of interest to follow the Weighted Poisson distribution (including LBP, NB, and EWP models) and the lifetimes of the non-cured individuals to follow a proportional odds survival model. The baseline odds distribution is considered to be either Weibull or Log-logistic distribution. A damage distribution is introduced due to the fact that some of the competing causes may not remain active following a treatment. In Chapter 4, we introduce a gamma frailty cure rate model for lifetime data by assuming the number of competing causes for the event of interest to follow the COM-Poisson distribution and the lifetimes of the non-cured individuals to follow a proportional odds model. The baseline distribution is considered to be either Weibull or Log-logistic distribution. Statistical inference is then developed under right censored data.

For all the models considered, we derive the maximum likelihood estimators (MLEs) with the full usage of expectation maximization (EM) method for all the model parameters, except the dispersion parameter, which is estimated by a profile likelihood approach. Detailed expressions for the Q-functions, the first- and second-derivatives of the Q-functions and the observed information matrix, corresponding to different cure rate models discussed in each chapter, are all presented in Appendices A-C. The model discrimination among some well-known special cases are discussed using both likelihood- and information-based criteria. An extensive Monte Carlo simulation study is carried out by varying censoring rates, sample sizes, distributional parameters, and the cure rates to examine the performance of the proposed model as well as the inferential methods developed in the thesis. Analysis of the cutaneous melanoma data is also carried out for illustrative purposes.

Chapter 2

COM-Poisson Cure Rate Model under Proportional Odds Lifetimes

2.1 Introduction

In this Chapter, we assume a proportional odds model for the distribution of lifetime varaible of a susceptible individual W_j, with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w; \boldsymbol{x}_c) = \theta O_0(w), \tag{2.1}$$

where $O(w) = S(w)/F(w)$ is the odds of survival up to w, the proportionality term θ is linked to covariates as $e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}$ with $\boldsymbol{x}_c = (x_1, \ldots, x_p)'$ being a vector of p covariates, $\boldsymbol{\gamma}_2 = (\gamma_{21}, \ldots, \gamma_{2p})'$ is the vector of proportional odds regression coefficients, and $O_0(w)$

is the baseline odds function. We can then obtain the survival function of W_j as

$$S(w) = [1 + e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2}(S_0(w)^{-1} - 1)]^{-1}, \tag{2.2}$$

with the corresponding probability density function (p.d.f.) as

$$f(w) = f_0(w)e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2}[(1 - S_0(w))e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2} + S_0(w)]^{-2}. \tag{2.3}$$

The rest of this chapter proceeds as follows. Section 2.2 describes the data and the likelihood, while the estimation of the cure rate and associated inferential issues are discussed in Section 2.3. In Section 2.4, an extensive Monte Carlo simulation study is carried out. In Section 2.5, we discuss model discrimination using information- and likelihood-based methods. A data on cutaneous melanoma is analyzed in Section 2.6 finally for illustrative purpose.

2.2 Data and the likelihood

Suppose the time-to-event is not completely observed and is subject to non-informative right censoring, which means that the data above a certain value are not observed. Therefore, the observation time T_i for the ith subject, would be the minimum of the censoring time C_i and the actual lifetime Y_i, i.e.,

$$T_i = \min\{Y_i, C_i\}, i = 1, \ldots, n. \tag{2.4}$$

We define an indicator function $\delta_i = I(Y_i \leq C_i)$ for the i-th subject such that $\delta_i = 1$ if the lifetime is observed while $\delta_i = 0$ if the lifetime is right censored, Δ_0 and Δ_1 are sets with all the i's equal to 0 and 1, respectively, and set Δ^* contains all the i's. It is to be noted that the cure rate $p_0 = Z(\eta, \phi)^{-1}$ is purely a function of η for a fixed value of ϕ. The range of $1/p_0$ is from 1 to ∞ and it is monotone in η. Therefore, it is natural to use a logistic regression model $H_\phi(\eta) = 1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$ to link the covariate $x_1, \ldots, x_p$ to the cured proportion p_{0i}, i.e.,

$$p_{0i} = p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) = Z(\eta, \phi)^{-1} = H_\phi(\eta)^{-1} = (1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}, \tag{2.5}$$

where p_{0i} is the cured proportion for the ith category, $\boldsymbol{x}_i = (1, \boldsymbol{x}_{ic}')' = (1, x_{i1}, \ldots, x_{ip})'$ is a vector of $p+1$ covariates, and $\boldsymbol{\beta}$ is the vector of regression coefficients. Under this link, η would equal $H^{-1}(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$, i.e., η can be calculated from the inverse function of $H_\phi(.)$ analytically for the Geometric, Poisson and Bernoulli distributions, and by using numerical method for the general COM-Poisson distribution.

For n pairs of observations $(\boldsymbol{t}, \boldsymbol{\delta})$=$\{(t_1,\ \delta_1)$,...,$(t_n,\ \delta_n)\}$, the observed data likelihood function under the non-informative censoring is then given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \{f_p(t_i; \boldsymbol{\theta})\}^{\delta_i} \{S_p(t_i; \boldsymbol{\theta})\}^{1-\delta_i}, \tag{2.6}$$

where $\boldsymbol{\theta}$ is the set of parameters $(\phi, \boldsymbol{\beta}', \boldsymbol{\gamma}')$, which is equivalent to

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i \in \Delta_1} f_p(t_i; \boldsymbol{\theta}) \prod_{i \in \Delta_0} \{p_0 + (1 - p_0) S_s(t_i; \boldsymbol{\theta})\}. \tag{2.7}$$

Here, we consider two baseline distributions for the proportional odds survival model corresponding to the time-to-event random variable, namely, Weibull and log-logistic distributions. It should also be noted that log-logistic distribution in fact processes the proportional odds property, while the Weibull distribution does not. The survival function and p.d.f. of W under a Weibull baseline, for example, are

$$S(w,\boldsymbol{\gamma}) = [1 + e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2}(e^{(\gamma_1 w)^{1/\gamma_0}} - 1)]^{-1}, w > 0, \tag{2.8}$$

$$f(w,\boldsymbol{\gamma}) = (\gamma_1 w)^{1/\gamma_0} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 - (\gamma_1 w)^{1/\gamma_0}} [e^{-(\gamma_1 w)^{1/\gamma_0}}(e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2} - 1) + 1]^{-2}/(\gamma_0 w), \tag{2.9}$$

where $w > 0, \gamma_0 > 0$ and $\gamma_1 > 0$ are the shape and scale parameters of the baseline Weibull distribution, respectively. On the other hand, if we assume the baseline distribution to be a log-logistic distribution with $\gamma_0 > 0$ and $\gamma_1 > 0$ as the scale and shape parameters, respectively, then the corresponding odds function of W_i is given by

$$O(w,;\boldsymbol{x}_c',\boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1}}{w^{\gamma_1}} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2} = O_0(w,;\gamma_0,\gamma_1)e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}. \tag{2.10}$$

We observe that W_i still follows a two-parameter log-logistic distribution $(\gamma_0, \gamma_1 > 0)$ with shape parameter γ_1 and scale parameter $\gamma_0 e^{-\boldsymbol{x}_c',\boldsymbol{\gamma}_2/\gamma_1}$, and with corresponding survival function

$$S(w,\boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}}{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2} + w^{\gamma_1}}, w > 0. \tag{2.11}$$

Note that the mean does not exist if $\gamma_1 < 1$ and the variance does not exist if $\gamma_1 < 2$.

Results for the Special cases of COM-Poisson cure rate model

As mentioned earlier, the COM-Poisson distribution includes the Bernoulli, Poisson and Geometric distributions as special cases. Here, we detail the steps of the EM algorithm for the corresponding three special cure models.

Bernoulli cure rate model

Let the competing cause random variable M follow a Bernoulli distribution with probability of success $\eta/(1+\eta)$. The probability density function for the whole population can then be expressed as

$$f_p(t_i, \boldsymbol{\theta}) = \frac{\eta}{1+\eta} f(t_i; \boldsymbol{\gamma}). \tag{2.12}$$

The survival function for the susceptible group is just the survival function for the time to event W, i.e., $S_s(t_i; \boldsymbol{\theta}) = S(t_i; \boldsymbol{\gamma})$. The inverse of the cure rate under this setting is $1/p_0 = 1 + \eta$. We, therefore, have $H_\phi(\eta) = 1 + \eta$ under the logistic link with a fixed value of ϕ, which implies $\eta = e^{\boldsymbol{x}_i' \boldsymbol{\beta}}$. The $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ function is then given by

$$\sum_{i \in \Delta_1} \boldsymbol{x}_i' \boldsymbol{\beta} + \sum_{i \in \Delta_1} \log f(t_i; \boldsymbol{x}_i, \boldsymbol{\gamma}) - \sum_{i \in \Delta *} \log(1 + e^{\boldsymbol{x}_i' \boldsymbol{\beta}}) + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i' \boldsymbol{\beta} + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma}). \tag{2.13}$$

It is readily seen that some of the terms in the Q function are only corresponding to $\boldsymbol{\beta}$ while the others are only corresponding to $\boldsymbol{\gamma}$. So, it can be split into two parts as follows:

$$Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)}) = Q_1(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) + Q_2(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}), \tag{2.14}$$

$$Q_1(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}) = \sum_{I \in \Delta_1} \log f(t_i; \boldsymbol{x}_i, \boldsymbol{\gamma}) + \sum_{I_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma}), \tag{2.15}$$

$$Q_2(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) = \sum_{i \in \Delta_1} \boldsymbol{x}_i' \boldsymbol{\beta} - \sum_{i \in \Delta *} \log(1 + e^{\boldsymbol{x}_i' \boldsymbol{\beta}}) + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i' \boldsymbol{\beta}, \tag{2.16}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i;\boldsymbol{\gamma}^{(k)})}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i;\boldsymbol{\gamma}^{(k)})} \tag{2.17}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\begin{aligned}
\frac{\partial Q}{\partial \gamma_j} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j}, \\
\frac{\partial Q}{\partial \beta_l} &= \sum_{i\in\Delta_1} x_{il} - \sum_{i\in\Delta*} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} x_{il}, \\
\frac{\partial^2 Q}{\partial \beta_l \partial \beta_{l'}} &= -\sum_{i\in\Delta*} \frac{x_{il} x_{il'} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}, \\
\frac{\partial^2 Q}{\partial \gamma_j \partial \gamma_{j'}} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}},
\end{aligned}$$

for $l, l' = 0, \ldots, p$, $j, j' = 1, 2, h = 21, \ldots, 2p$, $i = 1, \ldots, n$.

Poisson cure rate model

Let the competing cause random variable M follow a Poisson distribution. The probability density function for the whole population in this case can be expressed as

$$f_p(t_i;\boldsymbol{\theta}) = [\log(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] f(t_i;\boldsymbol{\gamma})(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{(S(t_i;\boldsymbol{\gamma})-1)}, \tag{2.18}$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = [(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{\gamma})} - 1]e^{-\boldsymbol{x}_i\boldsymbol{\beta}}. \tag{2.19}$$

The cure rate is $p_0 = e^{-\eta}$. We would then have $H_\phi(\eta) = e^\eta$ under the logistic link function with a fixed value of ϕ, which implies that $\eta = \ln(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$. The $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ function is then given by

$$\begin{aligned} Q = &\sum_{I\in\Delta_1} \log[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] + \sum_{I\in\Delta_1} \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma}) + \sum_{I\in\Delta_1} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) \\ &- \sum_{i\in\Delta*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)}\log((1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})} - 1) \end{aligned} \tag{2.20}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = 1 - (1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}})^{-S(t_i;\boldsymbol{\gamma}^{(k)})} \tag{2.21}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\begin{aligned} \frac{\partial Q}{\partial \beta_l} = &\sum_{i\in\Delta_1} \frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} - \sum_{i\in\Delta*} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} \\ &+ \sum_{i\in\Delta_0} \frac{\boldsymbol{\pi}^{(k)} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}, \\ \frac{\partial Q}{\partial \gamma_j} = &\sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_1} \frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) \\ &+ \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}} \frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j}, \end{aligned}$$

$$\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} = \sum_{i\in\Delta_1}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}\{\frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}[1-\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}]+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\} - \sum_{i\in\Delta*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}[1-\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}-1}],$$
$$\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_j} = \sum_{i\in\Delta_1}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}$$
$$+\sum_{i\in\Delta_0}\frac{\boldsymbol{\pi}^{(k)}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\frac{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}(1+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}))}{[1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},$$
$$\frac{\partial^2 Q}{\partial\gamma_j\partial\gamma_{j'}} = \sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_1}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}$$
$$-\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}[\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})]^2}{[1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_{j'}},$$

for $l, l' = 0, \ldots, p$, $j, j' = 1, 2$, $h = 21, \ldots, 2p$, $i = 1, \ldots, n$.

Geometric cure rate model

Let the competing cause random variable M follow a Geometric distribution. The probability density function for the whole population in this case can be expressed as

$$f_p(t_i;\boldsymbol{\theta}) = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}f(t_i,\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})^2}, \tag{2.22}$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = \frac{S(t_i;\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})}, \tag{2.23}$$

where $R_G(t_i, \boldsymbol{\theta}) = 1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}} - e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma})$. The cure rate under this setting is $p_0 = 1 - \eta$, and under the logistic link function with a fixed value of ϕ, we would have $H_\phi(\eta) = (1-\eta)^{-1}$, which implies that $\eta = e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}$. The $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ function is then given by

$$Q = \sum_{I \in \Delta_1} \boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{I \in \Delta_1} \log f(t_i, \boldsymbol{x}_{ic}, \boldsymbol{\gamma}) - \sum_{I \in \Delta_1} 2\log R_G(t_i, \boldsymbol{\theta}) - \sum_{i \in \Delta_0} \log(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} \boldsymbol{x}_i'\boldsymbol{\beta}$$
$$+ \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} \log S(t_i; \boldsymbol{\gamma}) - \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} \log R_G(t_i, \boldsymbol{\theta}), \tag{2.24}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{S(t_i; \boldsymbol{\gamma}^{(k)}) e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}} \tag{2.25}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\frac{\partial Q}{\partial \beta_l} = \sum_{i \in \Delta_1} x_{il} - 2 \sum_{i \in \Delta_1} \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l} - \sum_{i \in \Delta_0} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} x_{il} - \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l},$$
$$\frac{\partial Q}{\partial \gamma_j} = \sum_{i \in \Delta_1} \left(\frac{\partial \log f(t_i, \boldsymbol{x}_{ic}, \boldsymbol{\gamma})}{\partial \gamma_j} - 2 \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \gamma_j} \right) + \sum_{i \in \Delta_0} \boldsymbol{\pi}^{(k)} \left(\frac{\partial \log S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma})}{\partial \gamma_j} - \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \gamma_j} \right),$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} &= -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}} - \sum_{i\in\Delta_0}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}},\\
\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_j} &= -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j},\\
\frac{\partial Q^2}{\partial\gamma_j\partial\gamma_{j'}} &= \sum_{i\in\Delta_1}\left(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} - 2\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}\right) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\left(\frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}\right),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j} &= \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}, \quad \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} = \frac{x_{ih}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})},\\
\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} &= \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_{j'}},\\
\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_l} &= \frac{x_{il}x_{il'}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}(1-S(t_i,\boldsymbol{\theta})), \quad \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} = -\frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j},
\end{aligned}
$$

for $l, l' = 0,\dots,p$, $j, j' = 1, 2$, $h = 21,\dots,2p$, $i = 1,\dots,n$.

Observed information matrix

COM-Poisson cure rate model: The score functions, for a fixed value of ϕ, are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_l} &= -\sum_{i\in\Delta*}x_{il}\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}}{z_{2,i}z_{01,i}}x_{il} + \sum_{i\in\Delta_0}\frac{z_{2,i}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{z_{01,i}(1+z_{1i})},\\
\frac{\partial l}{\partial\gamma_h} &= \sum_{i\in\Delta_1}\frac{\partial \log f(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_1}\left(\frac{z_{21,i}}{z_{2,i}} - 1\right)\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}.
\end{aligned}
$$

Hence, the components of the observed information matrix, for a fixed value of ϕ, are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{-\sum_{i\in\Delta*}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}})^2}\\
&+\sum_{i\in\Delta_1}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{2,i}z_{01,i})^2}[z_{2,i}(z_{21,i}z_{01,i}+z_{31,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}})-z_{21,i}[z_{21,i}+\frac{z_{2,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}'_i\boldsymbol{\beta}}]\\
&+\sum_{i\in\Delta_0}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{1,i}+1)^2z^2_{01,i}}[(\frac{z_{21,i}}{z_{01,i}}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}+z_{2,i})(z_{1,i}+1)z_{01,i}-z_{2,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}[\frac{z_{02,i}}{z_{01,i}}(z_{1,i}+1)+z_{2,i}]]\},\\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_h} &= -\{\sum_{i\in\Delta_1}x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}\frac{z_{31,i}z_{2,i}-z^2_{21,i}}{z^2_{2,i}z_{01,i}}\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}\\
&+\sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{z_{01,i}(1+z_{1,i})}(z_{21,i}-\frac{z^2_{2,i}}{1+z_{1,i}})\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}\},\\
-\frac{\partial^2 l}{\partial\gamma_h\partial\gamma_{h'}} &= -\{\sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}\\
&+\sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}\\
&+\sum_{i\in\Delta_1}(\frac{z_{31,i}z_{2,i}-z^2_{21,i}}{z^2_{2,i}})\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\\
&+\sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}+\sum_{i\in\Delta_0}\frac{z_{21,i}z_{1,i}+z_{21,i}-z^2_{2,i}}{(1+z_{1,i})^2}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\},
\end{aligned}
$$

for $l, l' = 0, \dots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \dots, 2p$, $i = 1, \dots, n$.

Bernoulli cure rate model: The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_l} &= \sum_{i\in\Delta_1}x_{il}-\sum_{i\in\Delta*}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}+\sum_{i\in\Delta_0}w_ix_{il},\\
\frac{\partial l}{\partial\gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial\log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial\gamma_j}+\sum_{i\in\Delta_0}w_i\frac{\partial\log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},
\end{aligned}
$$

where $w_i = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{\gamma})}$.

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{-\sum_{i\in\Delta*}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2} + \sum_{i\in\Delta_0} x_{il}x_{il'}w_i(1-w_i)\},\\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} &= -\{\sum_{i\in\Delta_0} x_{il}w_i(1-w_i)\frac{\partial \text{log}S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\},\\
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} &= -\{\sum_{i\in\Delta_1}\frac{\partial\text{log}f(t_i;\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} + \sum_{i\in\Delta_0} w_i\frac{\partial\text{log}S^2(t_i;\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} + \sum_{i\in\Delta_0} w_i(1-w_i)\frac{\partial\text{log}S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\frac{\partial\text{log}S(t_i;\boldsymbol{\gamma})}{\partial\gamma_{j'}}\},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

Poisson cure rate model: The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_l} &= \sum_{i\in\Delta_1}\frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta*}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma})-1)}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}},\\
\frac{\partial l}{\partial\gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial\text{log}f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j} + \sum_{i\in\Delta*}\frac{\partial S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}).
\end{aligned}
$$

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{\sum_{i\in\Delta_1}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})-e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})]^2} + \sum_{i\in\Delta*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}(S(t_i;\boldsymbol{\gamma})-1)\},\\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} &= -\{\sum_{i\in\Delta*}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\},\\
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} &= -\{\sum_{i\in\Delta_1}\frac{\partial\text{log}f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} + \sum_{i\in\Delta*}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})\},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

Geometric cure rate model: The score functions are

$$\frac{\partial l}{\partial \beta_l} = \sum_{i\in\Delta_1} x_{il} - 2\sum_{i\in\Delta_1} \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l} - \sum_{i\in\Delta_0} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0} \frac{S(t_i;\boldsymbol{\gamma}) e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} (x_{il} - \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l}),$$
$$\frac{\partial l}{\partial \gamma_j} = \sum_{i\in\Delta_1} (\frac{\partial \log f(t_i, \boldsymbol{x}_{ic}, \boldsymbol{\gamma})}{\partial \gamma_j} - \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \gamma_j}) - \sum_{i\in\Delta*} \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \gamma_j}.$$

Hence, the components of the observed information matrix are

$$\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_l \partial \beta_{l'}} &= -\{-2\sum_{i\in\Delta_1} \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \beta_{l'}} + \sum_{i\in\Delta_0} \frac{x_{il} x_{il'} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2} (S(t_i;\boldsymbol{\gamma}) - 1) \\
&\quad - \sum_{i\in\Delta_0} \frac{S(t_i;\boldsymbol{\gamma}) e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \beta_{l'}} - \sum_{i\in\Delta_0} \frac{x_{il'} e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{\gamma})}{(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2} \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l}\}, \\
-\frac{\partial^2 l}{\partial \beta_l \partial \gamma_j} &= -\{-\sum_{i\in\Delta_1} \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j} - \sum_{i\in\Delta*} \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j}\}, \\
-\frac{\partial^2 l}{\partial \gamma_j \partial \gamma_{j'}} &= -\{\sum_{i\in\Delta_1} (\frac{\partial \log f(t_i, \boldsymbol{x}_{ic}, \boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} - \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}}) - \sum_{i\in\Delta*} \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}}\},
\end{aligned}$$

where $R_G(t_i, \boldsymbol{\theta}) = 1 - e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma}) - 1)$, and

$$\begin{aligned}
\frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \gamma_j} &= \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i, \boldsymbol{\theta})} \frac{\partial S(t_i, \boldsymbol{\theta})}{\partial \gamma_j}, \quad \frac{\partial \log R_G(t_i, \boldsymbol{\theta})}{\partial \beta_l} = \frac{x_{il} e^{\boldsymbol{x}'\boldsymbol{\beta}}(1 - S(t_i, \boldsymbol{\theta}))}{R_G(t_i, \boldsymbol{\theta})}, \\
\frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}} &= \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i, \boldsymbol{\theta})} \frac{\partial S^2(t_i, \boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}} - \frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i, \boldsymbol{\theta})^2} \frac{\partial S(t_i, \boldsymbol{\theta})}{\partial \gamma_j} \frac{\partial S(t_i, \boldsymbol{\theta})}{\partial \gamma_{j'}}, \\
\frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \beta_l} &= \frac{x_{il} x_{il'} e^{\boldsymbol{x}'\boldsymbol{\beta}}(1 - S(t_i, \boldsymbol{\theta}))}{R_G(t_i, \boldsymbol{\theta})^2}, \quad \frac{\partial \log R_G^2(t_i, \boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j} = -\frac{x_{il} e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i, \boldsymbol{\theta})^2} \frac{\partial S(t_i, \boldsymbol{\theta})}{\partial \gamma_j},
\end{aligned}$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

2.3 Estimation of parameters

We propose an Expectation-Maximization (EM) algorithm for obtaining the MLE of $\boldsymbol{\theta}$, and a profile likelihood approach for the estimation of the dispersion patameter ϕ. It is well-known that EM is an effective technique for finding the MLEs of unknown parameters of a model involving unobserved variables [for further details, refer to McLachlan and Krishnan (2007)]. In our model, the random variable I_i's are observed for i in the set Δ_1 and unobserved for i in the set Δ_0, where $I_i = 1$ if the individual is susceptible and $I_i = 0$ if the individual is cured. Let us denote the set of complete data by $(\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{I})$=$\{(t_1, \delta_1, \boldsymbol{x}_1, I_1), \ldots, (t_n, \delta_n, \boldsymbol{x}_n, I_n)\}$. The complete data likelihood function is then

$$L_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I}) \propto \prod_{i\in\Delta_1} f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) \prod_{i\in\Delta_0} p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)^{1-I_i}[(1-p_0(\boldsymbol{\beta}, \boldsymbol{x}_i))S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta})]^{I_i}, \tag{2.26}$$

where $\mathbf{I} = (I_1, \ldots, I_n)'$, $\boldsymbol{x}_{ic} = (x_{i1}, \ldots, x_{ip})'$ and $\boldsymbol{x}_i = (1, \boldsymbol{x}'_{ic})'$. The corresponding complete log-likelihood function

$$\begin{aligned} l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I}) = \text{constant} &+ \sum_{i\in\Delta_1} \text{log} f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) + \sum_{i\in\Delta_0}(1-I_i)\text{log} p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) \\ &+ \sum_{i\in\Delta_0} I_i \text{log}[1-p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)] + \sum_{i\in\Delta_0} I_i \text{log} S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}). \end{aligned} \tag{2.27}$$

2.3.1 E-step

The expectation step is achieved by calculating the expected value of the complete data log-likelihood function with respect to the conditional distribution of the unobserved I_i's

($i \in \Delta_0$), given the observed data $\boldsymbol{O} = \{(t_i, \delta_i, \boldsymbol{x}_i), i \in \Delta_1\}$ and the current estimates of the parameters $\boldsymbol{\theta}^{(k)} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ for a fixed value of ϕ. Let us denote the function as

$$Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) = E(l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta}, \boldsymbol{I})|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}), \tag{2.28}$$

at the k-th iteration step. In our model, I_i's are Bernoulli random variables and we can easily find the conditional expectation if the ith individual being susceptible is

$$\boldsymbol{\pi}_i^{(k)} = \mathrm{E}(I_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}) = P(I_i = 1|T > t) = \frac{(1 - p_0(\boldsymbol{\beta}^{(k)}, \boldsymbol{x}_i))S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}^{(k)})}{S_p(t_i, \boldsymbol{x}_i; \boldsymbol{\theta}^{(k)})}|_{\theta=\theta^{(k)}}. \tag{2.29}$$

Now, for a fixed value of ϕ, the Q function is given by

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) &= \sum_{I \in \Delta_1} \log f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) + \sum_{i \in \Delta_0} (1 - \boldsymbol{\pi}_i^{(k)}) \log p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) \\ &+ \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log[1 - p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)] + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}) \end{aligned} \tag{2.30}$$

which can be further simplified as

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) &= -\sum_{i \in \Delta *} \log(1 + e^{\boldsymbol{x}_i' \boldsymbol{\beta}}) + \sum_{i \in \Delta_1} \log f(t_i, \boldsymbol{\gamma}) - \sum_{i \in \Delta_1} \log S(t_i, \boldsymbol{\gamma}) \\ &+ \sum_{i \in \Delta_1} \log z_{2,i} + \sum_{i \in \Delta_0} \boldsymbol{\pi}_i^{(k)} \log z_{1,i}, \end{aligned} \tag{2.31}$$

where

$$z_{1,i} = z_1(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{\{\eta_i S(t_i; \boldsymbol{\gamma})\}^j}{(j!)^{\phi}}, \tag{2.32}$$

$$z_{2,i} = z_2(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{j\{\eta_i S(t_i; \boldsymbol{\gamma})\}^j}{(j!)^{\phi}}. \tag{2.33}$$

2.3.2 M-step

The M-step is achieved by maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function in (2.31) in order to obtain the improved estimate of $\boldsymbol{\theta}$, i.e.,

$$\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}). \tag{2.34}$$

The MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not have explicit expressions. We therefore adopt the numerical maximization carried out by Newton-Raphson method.

For a fixed value of ϕ, the E-step and M-step are alternated until the parameter estimate converges to a desired level of accuracy. The parameter ϕ is determined by using the profile likelihood technique. We consider a range of ϕ with small increment, and then for each value of ϕ, the MLEs of other parameters are found, and the estimate with the largest likelihood is chosen as the final estimate. The following subsections present explicit forms of the first- and second-order derivatives of the Q function as well as update function for the general case of COM-Poisson distribution.

2.3.3 Results for the COM-Poisson cure rate model

The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows:

$$
\begin{aligned}
\frac{\partial Q}{\partial \beta_l} &= -\sum_{i\in\Delta*} x_{il}\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}}{z_{2,i}z_{01,i}}x_{il} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{2,i}}{z_{1,i}z_{01,i}}x_{il},\\
\frac{\partial Q}{\partial \gamma_h} &= \sum_{i\in\Delta_1}\frac{\partial \log f(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{2,i}}{z_{1,i}}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h},\\
\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} &= -\sum_{i\in\Delta*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}\\
&\quad + \sum_{i\in\Delta_1}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(z_{2,i}z_{01,i})^2}[z_{2,i}(z_{21,i}z_{01,i}+z_{31,i}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) - z_{21,i}[z_{21,i}+\frac{z_{2,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]\\
&\quad + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(z_{1,i}z_{01,i})^2}[z_{1,i}(z_{21,i}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}+z_{2,i}z_{01,i}) - z_{2,i}[z_{2,i}+\frac{z_{1,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}_i'\boldsymbol{\beta}}],\\
\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_h} &= \sum_{i\in\Delta_1} x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2 z_{01,i}}\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}\\
&\quad + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}z_{1,i}-z_{2,i}^2}{z_{1,i}^2 z_{01,i}}\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h},\\
\frac{\partial^2 Q}{\partial\gamma_h\partial\gamma_{h'}} &= \sum_{i\in\Delta_1}\frac{\partial \log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} + \sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} +\\
&\quad + \sum_{i\in\Delta_1}(\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2})\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\\
&\quad + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{2,i}}{z_{1,i}}\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{21,i}z_{1,i}-z_{2,i}^2}{z_{1,i}^2}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$, where

$$
\begin{aligned}
z_{21,i} &= z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^2\{\eta_i S(t_i;\boldsymbol{\gamma})\}^j}{(j!)^\phi}, \quad z_{31,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^3\{\eta_i S(t_i;\boldsymbol{\gamma})\}^j}{(j!)^\phi},\\
z_{01,i} &= z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j\eta_i^j}{(j!)^\phi}, \quad z_{02,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^2\eta_i^j}{(j!)^\phi},
\end{aligned}
$$

with $z_{1,i}$ and $z_{2,i}$ being as in (2.32) and (2.33), respectively.

2.3.4 Standard errors and asymptotic confidence intervals

We may approximate the asymptotic variance-covariance matrix of the MLEs $(\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}')'$ by inverting the observed Fisher information matrix of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for a fixed value of ϕ. The components of the observed Fisher information matrix can be calculated from the negative of the second-order derivatives of the complete data likelihood function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ (for detailed infotmation, refer to Louis (1982)). Thus, we can obtain the standard errors of the estimates and then construct corresponding asymptotic confidence intervals for the parameters.

2.3.5 Estimation of the cure rate and its standard error

Suppose $\hat{\boldsymbol{\beta}}$ is the MLE of the regression coefficient $\boldsymbol{\beta}$. The estimated cure rate for group i is then $\hat{p}_{0i} = (1 + e^{\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}})^{-1}$, for $i = 1, ..., \tau$. The standard error of $\hat{p}_{0i}$ can be found through delta method as

$$\mathrm{sd}(\hat{p_{0i}}) = \sqrt{\left(\frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_0}, ..., \frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_p}\right) \mathrm{var}(\hat{\boldsymbol{\beta}}) \left(\frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_0}, ..., \frac{\partial \hat{p_{0i}}}{\partial \hat{\beta}_p}\right)'}. \tag{2.35}$$

2.3.6 Equivalent models

Proposition: The population survival function under Geometric and Bernoulli cure rate models are equivalent through re-parametrization if baseline odds follow log-logistic distribution.

Proof: Cure rates p_0 under Geometric and Bernoulli cure rate models are $1 - \eta_1$ and $\frac{1}{1+\eta_3}$, respectively. If we equate these cure rates, we obtain a relationship between η_1 and η_3 as

$$\eta_3 = \frac{\eta_1}{1-\eta_1}. \tag{2.36}$$

Suppose $S_1 = \frac{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}}}{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} + t^{\gamma_1}}$ and $S_3 = \frac{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}}}{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} + t^{\gamma_1}}$ are the survival functions of the susceptible group under Geometric and Bernoulli cure rate models, respectively. Then, the population survival function under the Bernoulli case is

$$S_p = \frac{1+\eta_3 S_3}{1+\eta_3}. \tag{2.37}$$

If we fix the relationship between γ_{30}, γ_{10}, γ_{32} and γ_{12} as

$$\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} = \gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} p_0 = \gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} (1-\eta_1), \tag{2.38}$$

then we obtain from (2.37) that

$$S_p = 1-\eta_1+\eta_1 S_3 = 1-\eta_1 \frac{t^{\gamma_1}}{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} + t^{\gamma_1}} = 1 - \frac{t^{\gamma_1}\eta_1}{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}}(1-\eta_1) + t^{\gamma_1}} = \frac{1-\eta_1}{1-\eta_1 S_1}, \tag{2.39}$$

which is the population survival function for the Geometric cure rate model under proportional odds assumption. Thus, these two models in this case are re-parametrizations of each other.

2.4 Simulation study

An extensive Monte Carlo simulation study is carried out in this paper to evaluate the performance of the proposed methodology by varying the sample size, cure rate, censoring proportion, the parameters of distributions, and underling lifetime distribution. We try to mimic the cutaneous melanoma data, and consider 4 possible categories for the individuals, namely, $x = 0, 1, 2, 3$. Three different sample sizes are considered in the study: $n = 200$ $(50, 42, 53, 55)$, $n = 400$ $(95, 102, 97, 106)$, $n = 800$ $(200, 168, 212, 220)$ to reflect small, medium and large sample sizes. Moreover, if we assume that $\boldsymbol{\beta} = (\beta_0, \beta_1)$ has two parameters, fixing the cure rates for the first and fourth categories would be enough to cover all cases as the cure rates for the second and third categories can then be obtained from $\boldsymbol{\beta}$. Here, we take $(p_{00}, p_{03}) = (0.4, 0.2)$ and $(p_{00}, p_{03}) = (0.6, 0.25)$ with respect to categories one and four for low and high cure rates, respectively. Also, the cure rate would be in a decreasing order in this way. The $\boldsymbol{\beta}$'s are then

$$\beta_0 = \log(1/p_{00} - 1) \ \ , \ \ \beta_1 = (\log(1/p_{03} - 1) - \beta_0)/3. \tag{2.40}$$

We thus obtain the true value of $\boldsymbol{\beta}$ as (0.405,0.321) and (-0.405,0.501), respectively. In addition, we consider light and heavy censored data in the simulation. The light and heavy censoring rates are $(0.52, 0.45, 0.37, 0.3)$ and $(0.65, 0.49, 0.4, 0.35)$ for the low cure rates, $(0.7, 0.57, 0.45, 0.34)$ and $(0.8, 0.64, 0.5, 0.38)$ for the high cure rates, respectively. It is natural to assume that the probability of censored population for the susceptible group equal to the difference between the probability of getting censored and cured; i.e.,

$$P(Y \geq C_x \cap M \geq 1 | X = x) = c_x - p_{0x}. \tag{2.41}$$

If we assume the censoring time C_x follows an exponential distribution with rate λ_x on $x = 0, 1, 2, 3$, Eq. (2.41) can be re-written as

$$\lambda_x \int_0^\infty S_p(C_x) e^{-\lambda_x c_x} dc_x = c_x. \tag{2.42}$$

The choice of (γ_0, γ_1) in the underlying distribution of the proportional odds survival model are $(0.571, 0.307)$ and $(1.75, 3.25)$ for Weibull and log-logistic distributions, respectively. The odds parameter is specified by $\gamma_2 = -0.75$ to ensure a decreasing lifetime for the four nodule categories. We consider an inverse transform sampling method to simulate the actual survival lifetime Y_i for each individual under different competing risks, i.e., $w_i = \frac{1}{\gamma_1}[\log(1 + (\frac{1}{u} - 1)e^{x_i\gamma_2})]^{\gamma_0}$ and $w_i = \gamma_0(\frac{u}{1-u}e^{x_i\gamma_2})^{1/\gamma_1}, i = 1, \ldots, n$, under the proportional odds model with Weibull and log-logistic baseline distribution, respectively, where u follows an uniform distribution over 0 to 1.

Under the above setting, the procedure to generate the data from different cure rate models is as follows.

Geometric cure rate model: For each individual, we simulate the number of competing risk M_i from Geometric distributions with probability $P(M_i = 0) = p_{0x}$; and we simulate the censoring time C_x from exponential distribution with rate λ_x, with the parameter λ_x found from (2.41). If M_i does not equal zero, we simulate M_i number of actual lifetimes $\{Y_{i1}, .., Y_{iM_i}\}$ from proportional odds survival model, the actual lifetimes Y_i is defined as $\min\{Y_{i1}, .., Y_{iM_i}\}$ and the observed lifetime T_i is taken as the minimum of all the actual lifetimes and the censoring time, i.e., $T_i = \min\{Y_{i1}, .., Y_{iM_i}, C_i\}$. If $Y_i > C_i$, we make the censoring indicator $\delta_i = 0$, otherwise $\delta_i = 1$. On the other hand, $M_i = 0$ means the individual is censored, we assign C_i to the actual lifetime T_i, and the censoring indicator is taken to be $\delta_i = 1$.

Poisson cure rate model: In this case, the procedure is the same as the Geometric cure rate model except that M_i is simulated from Poisson distribution with parameter $-\log(p_{0x})$.

Bernoulli cure rate model: There are two ways to do the data generation in this case. One is the same as Geometric cure rate model except that M_i is simulated from Bernoulli distribution with probability of success as $1 - p_{0x}$. Another way is a little bit simpler since M_i can only be taken as 0 or 1 in this case. For each individual, we simulate the censoring time C_x from exponential distribution with rate λ_x. Then, we simulate an uniform random variable U_i and if $U_i \leq p_{0x}$, the observed lifetime T_i is set to C_x; otherwise, we generate the observed lifetime T_i from the proportional odds survival model.

In our simulation study, 1000 Monte Carlo runs were considered in each scenario. The estimates were calculated through EM method. We stopped our estimation if the absolute difference between two consecutive estimates was less than 10^{-5}. We calculated the empirical Bias, standard errors(SE), root Mean Square Error (RMSE), and 95% coverage probabilities (CPs) for the estimates of the parameters. In addition, we computed the cure rate, SE and 95% CPs. Here, the initial values of the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ were taken from a grid of parameters with a range from 80% to 120% of the true value, and those values having the maximum likelihood were chosen as the initial value.

Tables 2.1-2.6 present the bias, SE, RMSE, and coverage probabilities for the three special cases. We can see that the estimates are quite accurate under different cure rate models. The Bias, standard error along with RMSE get reduced as the sample size increases. The same follows when the censoring is light or the cure rate is high. The standard errors and RMSE of β_0 are always larger than those for the other parameters. The coverage probabilities of the confidence intervals based on the asymptotic normality of the MLEs are quite close to the nominal level in most of the cases. To summarize, a larger sample size, and

Table 2.1: Bias, SE, RMSE, and CP for the estimates of the parameters of the Geometric cure rate model under proportional odds with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.003	0.208	0.212	94.7	-0.405	-0.005	0.205	0.208	94.5
β_1	0.327	0.003	0.116	0.121	94.5	0.501	0.007	0.111	0.111	94
γ_0	1.75	0.01	0.147	0.151	93.4	1.75	-0.001	0.154	0.155	93.8
γ_1	3.25	0.021	0.186	0.195	93.6	3.25	0.033	0.205	0.202	95.6
γ_2	-0.75	-0.003	0.149	0.153	94.7	-0.75	-0.001	0.156	0.154	95.2
n=400, HC										
β_0	0.405	0.012	0.25	0.248	96.1	-0.405	-0.002	0.257	0.265	95.1
β_1	0.327	0.002	0.134	0.133	95.6	0.501	0.004	0.131	0.133	95.3
γ_0	1.75	0.019	0.18	0.182	95.3	1.75	0.014	0.198	0.197	95.3
γ_1	3.25	0.026	0.199	0.195	96	3.25	0.045	0.225	0.236	94.6
γ_2	-0.75	-0.011	0.17	0.165	95.8	-0.75	-0.015	0.18	0.178	95.3
n=800, LC										
β_0	0.405	0.004	0.146	0.147	94.3	-0.405	-0.002	0.145	0.146	94.4
β_1	0.327	0.001	0.08	0.078	95.2	0.501	0.001	0.077	0.076	95.7
γ_0	1.75	0.01	0.104	0.106	95.1	1.75	0	0.11	0.111	95.4
γ_1	3.25	0.005	0.13	0.132	95	3.25	0.019	0.144	0.144	94.9
γ_2	-0.75	-0.006	0.103	0.102	95.6	-0.75	-0.003	0.109	0.111	93.8
n=800, HC										
β_0	0.405	0.007	0.177	0.175	95.6	-0.405	-0.003	0.183	0.184	95.7
β_1	0.327	-0.002	0.092	0.091	95.3	0.501	0.003	0.091	0.093	95.5
γ_0	1.75	0.012	0.127	0.126	95.5	1.75	0	0.14	0.144	94.4
γ_1	3.25	0.011	0.14	0.139	94.7	3.25	0.017	0.157	0.155	95.1
γ_2	-0.75	-0.007	0.118	0.115	95.5	-0.75	-0.001	0.126	0.128	95.2

smaller censoring proportion would result in more accurate estimates.

2.5 Model discrimination

The COM-Poisson distribution contains many commonly used discrete distributions for different selection of ϕ. It would therefore be of interest to select a suitable ϕ and make full use of the COM-Poisson distribution to get the best fit for the data. So, we focus here on a model discrimination among the three special cases of the COM-Poisson distribution.

Table 2.2: Bias, SE, RMSE, and CP for the estimates of the parameters of the Poisson cure rate model under proportional odds model with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.004	0.207	0.202	95.8	-0.405	-0.007	0.205	0.204	95.2
β_1	0.327	0.003	0.115	0.113	96.1	0.501	0.004	0.111	0.108	95.4
γ_0	1.75	0.001	0.12	0.115	95	1.75	-0.005	0.14	0.141	94.8
γ_1	3.25	0.031	0.189	0.195	95.5	3.25	0.039	0.208	0.208	96
γ_2	-0.75	-0.004	0.12	0.119	95.1	-0.75	-0.003	0.134	0.137	94.9
n=400, HC										
β_0	0.405	0.009	0.247	0.248	94	-0.405	-0.004	0.256	0.258	95.7
β_1	0.327	0	0.131	0.133	93.6	0.501	0.007	0.13	0.129	95.5
γ_0	1.75	0.004	0.141	0.136	95.7	1.75	0.006	0.173	0.168	95.2
γ_1	3.25	0.048	0.203	0.213	94.4	3.25	0.041	0.227	0.24	94.2
γ_2	-0.75	-0.013	0.133	0.132	95.3	-0.75	-0.009	0.152	0.159	94
n=800, LC										
β_0	0.405	0.001	0.146	0.141	95.2	-0.405	-0.006	0.145	0.15	94.2
β_1	0.327	0	0.079	0.077	95.6	0.501	0.001	0.077	0.078	95.3
γ_0	1.75	0	0.086	0.084	96	1.75	0.002	0.101	0.101	95.3
γ_1	3.25	0.014	0.132	0.131	94.8	3.25	0.007	0.145	0.138	95.8
γ_2	-0.75	-0.001	0.084	0.082	95.4	-0.75	-0.001	0.094	0.094	94.8
n=800, HC										
β_0	0.405	0.008	0.175	0.176	94.7	-0.405	-0.007	0.182	0.182	95.3
β_1	0.327	-0.002	0.091	0.094	94	0.501	0.003	0.09	0.09	95.1
γ_0	1.75	0.006	0.102	0.1	95.3	1.75	0.001	0.124	0.125	95
γ_1	3.25	0.018	0.142	0.148	93.7	3.25	0.019	0.159	0.163	94.7
γ_2	-0.75	-0.008	0.093	0.096	94.6	-0.75	-0.001	0.107	0.111	93.7

Table 2.3: Bias, SE, RMSE, and CP for the estimates of the parameters of the Bernoulli cure rate model under proportional odds model with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.002	0.208	0.214	94.6	-0.405	-0.02	0.205	0.207	94.8
β_1	0.327	0.002	0.116	0.117	95.7	0.501	0.011	0.111	0.114	94.2
γ_0	1.75	0.003	0.112	0.114	94.9	1.75	0.002	0.136	0.137	94.1
γ_1	3.25	0.025	0.186	0.193	95.3	3.25	0.032	0.206	0.204	95.2
γ_2	-0.75	-0.004	0.11	0.111	94.6	-0.75	-0.006	0.126	0.127	94.9
n=400, HC										
β_0	0.405	-0.009	0.25	0.257	94	-0.405	0.004	0.256	0.244	96.3
β_1	0.327	0.009	0.134	0.138	94.9	0.501	-0.002	0.13	0.127	96.2
γ_0	1.75	0.011	0.128	0.126	95.6	1.75	0.01	0.162	0.164	94.4
γ_1	3.25	0.028	0.2	0.212	94	3.25	0.037	0.225	0.236	94.2
γ_2	-0.75	-0.012	0.12	0.114	96.4	-0.75	-0.013	0.14	0.146	93.8
n=800, LC										
β_0	0.405	0	0.146	0.146	94.5	-0.405	-0.005	0.145	0.146	95.1
β_1	0.327	0.002	0.08	0.079	95.5	0.501	0.005	0.077	0.078	94.6
γ_0	1.75	0.001	0.079	0.076	95.6	1.75	0.001	0.097	0.097	94.9
γ_1	3.25	0.015	0.131	0.135	94.6	3.25	0.018	0.144	0.148	94.7
γ_2	-0.75	-0.002	0.077	0.076	94.8	-0.75	-0.002	0.088	0.087	95.2
n=800, HC										
β_0	0.405	-0.005	0.177	0.181	94	-0.405	-0.001	0.182	0.17	96.6
β_1	0.327	0.004	0.092	0.092	95.4	0.501	0	0.09	0.085	96.7
γ_0	1.75	0.006	0.091	0.089	94.8	1.75	0.003	0.116	0.117	94.9
γ_1	3.25	0.011	0.14	0.136	95.9	3.25	0.016	0.157	0.156	94.7
γ_2	-0.75	-0.005	0.084	0.083	95.3	-0.75	-0.004	0.098	0.1	94.1

Table 2.4: Bias, SE, RMSE, and CP for the estimates of the parameters of the Geometric cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	-0.002	0.208	0.209	94.589	-0.405	-0.011	0.197	0.199	94.6
β_1	0.327	0.004	0.117	0.119	94.6	0.501	0.002	0.107	0.108	95.6
γ_0	0.571	-0.001	0.033	0.034	94.399	0.571	0.011	0.037	0.044	95.76
γ_1	0.307	0.005	0.037	0.035	95.884	0.307	0.007	0.037	0.04	94.333
γ_2	-0.75	-0.003	0.153	0.155	93.265	-0.75	0.004	0.157	0.15	94.425
n=400, HC										
β_0	0.405	0.007	0.274	0.282	94.874	-0.405	-0.028	0.26	0.279	94.668
β_1	0.327	0.001	0.143	0.144	95.3	0.501	0.005	0.129	0.133	94.684
γ_0	0.571	0	0.036	0.037	95.19	0.571	0.012	0.046	0.046	93.89
γ_1	0.307	0.005	0.052	0.054	94.874	0.307	0.015	0.057	0.058	93.443
γ_2	-0.75	-0.002	0.183	0.184	95.33	-0.75	0.005	0.189	0.181	93.552
n=800, LC										
β_0	0.405	-0.004	0.146	0.148	94.45	-0.405	-0.015	0.139	0.144	93.3
β_1	0.327	0.002	0.081	0.081	94.289	0.501	0.005	0.075	0.074	95.2
γ_0	0.571	0.001	0.023	0.025	92.893	0.571	0.009	0.026	0.035	93.318
γ_1	0.307	0.002	0.026	0.026	94.864	0.307	0.007	0.026	0.029	93.251
γ_2	-0.75	0	0.106	0.108	94.914	-0.75	0.001	0.109	0.106	94.49
n=800, HC										
β_0	0.405	0.002	0.191	0.191	94.271	-0.405	-0.014	0.183	0.196	93.493
β_1	0.327	0.002	0.098	0.096	95.395	0.501	0.001	0.09	0.091	95.495
γ_0	0.571	0.002	0.026	0.027	93.594	0.571	0.01	0.031	0.037	88.934
γ_1	0.307	0.002	0.037	0.037	94.456	0.307	0.009	0.039	0.041	91.247
γ_2	-0.75	0	0.127	0.125	94.586	-0.75	-0.005	0.131	0.125	94.153

Table 2.5: Bias, SE, RMSE, and CP for the estimates of the parameters of the Poisson cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.005	0.21	0.207	96	-0.405	0.001	0.207	0.212	94.3
β_1	0.327	0.002	0.118	0.115	95.7	0.501	0.006	0.113	0.118	94.4
γ_0	0.571	-0.003	0.034	0.035	94.1	0.571	-0.005	0.037	0.038	94.5
γ_1	0.307	0	0.028	0.028	95.4	0.307	0.002	0.031	0.032	93.8
γ_2	-0.75	-0.009	0.12	0.119	95.1	-0.75	-0.009	0.132	0.132	95.3
n=400, HC										
β_0	0.405	0.007	0.262	0.257	95.2	-0.405	0.011	0.273	0.267	95.7
β_1	0.327	0.003	0.139	0.141	94.6	0.501	-0.001	0.136	0.134	95.9
γ_0	0.571	-0.003	0.036	0.036	94.3	0.571	-0.005	0.041	0.041	93.9
γ_1	0.307	0.004	0.036	0.037	94.2	0.307	0.005	0.044	0.046	94.5
γ_2	-0.75	-0.005	0.136	0.14	93.7	-0.75	-0.011	0.153	0.157	95.4
n=800, LC										
β_0	0.405	0.003	0.148	0.145	95.7	-0.405	0	0.147	0.146	95.3
β_1	0.327	0.001	0.081	0.08	95.4	0.501	0.001	0.078	0.073	96.2
γ_0	0.571	-0.001	0.024	0.025	93.3	0.571	-0.002	0.027	0.027	94.8
γ_1	0.307	-0.001	0.02	0.02	95.5	0.307	0.001	0.022	0.023	95.2
γ_2	-0.75	-0.008	0.083	0.082	95.3	-0.75	-0.004	0.092	0.093	94.8
n=800, HC										
β_0	0.405	0.007	0.185	0.187	94.8	-0.405	-0.004	0.193	0.197	95
β_1	0.327	-0.003	0.096	0.098	95.3	0.501	0.001	0.095	0.095	94.7
γ_0	0.571	-0.001	0.026	0.025	95.4	0.571	-0.003	0.029	0.028	94.7
γ_1	0.307	0.003	0.026	0.026	95	0.307	0.003	0.032	0.033	94
γ_2	-0.75	-0.001	0.095	0.097	95	-0.75	-0.006	0.107	0.11	95

Table 2.6: Bias, SE, RMSE, and CP for the estimates of the parameters of the Bernoulli cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.011	0.209	0.207	96.2	-0.405	-0.008	0.207	0.206	96.2
β_1	0.327	0.003	0.118	0.118	95.6	0.501	0.009	0.113	0.112	95.8
γ_0	0.571	-0.001	0.034	0.034	95	0.571	-0.004	0.038	0.038	94.1
γ_1	0.307	0.002	0.023	0.023	94.9	0.307	0.003	0.028	0.028	95.6
γ_2	-0.75	-0.004	0.107	0.103	95.5	-0.75	-0.006	0.12	0.125	94.2
n=400, HC										
β_0	0.405	-0.001	0.255	0.266	93.6	-0.405	-0.004	0.266	0.266	95.9
β_1	0.327	0.007	0.138	0.142	93.9	0.501	0.003	0.135	0.135	94.7
γ_0	0.571	-0.003	0.036	0.036	94.7	0.571	-0.005	0.041	0.04	94.7
γ_1	0.307	0.004	0.028	0.029	94.5	0.307	0.003	0.036	0.038	94.2
γ_2	-0.75	-0.002	0.116	0.121	94.5	-0.75	-0.011	0.135	0.14	94.4
n=800, LC										
β_0	0.405	0.005	0.147	0.152	94	-0.405	0.002	0.146	0.149	94.9
β_1	0.327	0.002	0.081	0.081	95.9	0.501	0.001	0.078	0.08	94.6
γ_0	0.571	0	0.024	0.024	95	0.571	-0.002	0.027	0.027	95
γ_1	0.307	0.001	0.016	0.017	93.4	0.307	0.001	0.02	0.02	94.6
γ_2	-0.75	0	0.074	0.073	95.7	-0.75	-0.004	0.084	0.086	94.2
n=800, HC										
β_0	0.405	-0.001	0.181	0.187	93.8	-0.405	-0.009	0.189	0.187	95.9
β_1	0.327	0.004	0.095	0.094	95.1	0.501	0.006	0.094	0.095	96.1
γ_0	0.571	-0.002	0.026	0.025	95.1	0.571	-0.002	0.029	0.028	94.6
γ_1	0.307	0.001	0.02	0.02	95.1	0.307	0.003	0.026	0.027	95.1
γ_2	-0.75	-0.003	0.081	0.083	94.2	-0.75	-0.002	0.095	0.095	94.5

Table 2.7: Bias, SE, RMSE, and CP for the cure rates of the Geometric cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0 (0.002)	0.049 (0.049)	0.051 (0.05)	74 (76.7)
		p_{02}	0.325	-0.001 (0)	0.03 (0.03)	0.03 (0.029)	93.5 (94.4)
		p_{03}	0.257	-0.001 (0)	0.028 (0.029)	0.027 (0.027)	94.6 (95.3)
		p_{04}	0.2	0.001 (0.001)	0.036 (0.037)	0.036 (0.036)	88 (88.8)
	heavy	p_{01}	0.4	-0.001 (0)	0.059 (0.064)	0.058 (0.066)	71.4 (70.5)
		p_{02}	0.325	-0.002 (0)	0.035 (0.038)	0.035 (0.04)	91.7 (91.2)
		p_{03}	0.257	-0.002 (0)	0.03 (0.032)	0.03 (0.032)	94.8 (94.5)
		p_{04}	0.2	0 (0.001)	0.038 (0.041)	0.039 (0.04)	87.9 (89.4)
800	light	p_{01}	0.4	0 (0.001)	0.035 (0.035)	0.035 (0.035)	73.9 (75.3)
		p_{02}	0.325	-0.001 (0.001)	0.021 (0.022)	0.022 (0.022)	92.1 (92.8)
		p_{03}	0.257	0 (0.001)	0.019 (0.02)	0.019 (0.02)	94.9 (94.3)
		p_{04}	0.2	0 (0.001)	0.025 (0.026)	0.024 (0.025)	89.5 (87.4)
	heavy	p_{01}	0.4	-0.001 (0)	0.042 (0.045)	0.042 (0.045)	71 (70.7)
		p_{02}	0.325	-0.001 (0)	0.025 (0.027)	0.025 (0.027)	92.1 (90.2)
		p_{03}	0.257	0 (-0.001)	0.021 (0.022)	0.02 (0.022)	95.9 (96.1)
		p_{04}	0.2	0.001 (0)	0.026 (0.028)	0.026 (0.027)	89.9 (90.8)
400	light	p_{01}	0.6	0 (0.002)	0.049 (0.047)	0.049 (0.047)	78.9 (79.7)
		p_{02}	0.476	0 (0.003)	0.033 (0.032)	0.034 (0.032)	92 (93.5)
		p_{03}	0.355	-0.001 (0.002)	0.03 (0.031)	0.031 (0.029)	94.4 (95.6)
		p_{04}	0.25	-0.001 (0.003)	0.038 (0.039)	0.039 (0.037)	87.6 (89.2)
	heavy	p_{01}	0.6	-0.001 (0.005)	0.061 (0.061)	0.063 (0.066)	73.4 (75.3)
		p_{02}	0.476	0 (0.006)	0.04 (0.042)	0.041 (0.044)	88.6 (88.8)
		p_{03}	0.355	-0.001 (0.005)	0.033 (0.035)	0.033 (0.033)	95.2 (95.1)
		p_{04}	0.25	0 (0.005)	0.041 (0.042)	0.041 (0.04)	89.9 (92.2)
800	light	p_{01}	0.6	0 (0.003)	0.035 (0.033)	0.035 (0.034)	77.5 (77.4)
		p_{02}	0.476	0 (0.003)	0.024 (0.023)	0.024 (0.024)	92 (91.6)
		p_{03}	0.355	0 (0.002)	0.021 (0.022)	0.021 (0.022)	94.4 (93.8)
		p_{04}	0.25	0.001 (0.001)	0.027 (0.027)	0.026 (0.026)	89.4 (88.1)
	heavy	p_{01}	0.6	0 (0.002)	0.044 (0.044)	0.044 (0.046)	74 (74.7)
		p_{02}	0.476	0 (0.003)	0.029 (0.03)	0.029 (0.032)	90.1 (87.4)
		p_{03}	0.355	0 (0.003)	0.023 (0.024)	0.024 (0.024)	95.3 (94.7)
		p_{04}	0.25	0 (0.003)	0.028 (0.029)	0.03 (0.028)	89.5 (91.9)

Table 2.8: Bias, SE, RMSE, and CP for the cure rates of the Poisson cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0 (0)	0.049 (0.05)	0.048 (0.049)	75.9 (75)
		p_{02}	0.325	-0.001 (-0.001)	0.03 (0.03)	0.029 (0.03)	93.9 (93.5)
		p_{03}	0.257	-0.001 (-0.001)	0.028 (0.028)	0.027 (0.028)	95.7 (95.3)
		p_{04}	0.2	0 (0.001)	0.035 (0.036)	0.035 (0.036)	88.2 (88.9)
	heavy	p_{01}	0.4	-0.001 (0)	0.058 (0.062)	0.059 (0.06)	73.2 (73.6)
		p_{02}	0.325	-0.001 (-0.001)	0.034 (0.037)	0.034 (0.036)	91.8 (91.8)
		p_{03}	0.257	-0.001 (-0.001)	0.029 (0.031)	0.029 (0.032)	95 (94.3)
		p_{04}	0.2	0.001 (0.001)	0.038 (0.039)	0.038 (0.041)	88.8 (87.2)
800	light	p_{01}	0.4	0 (0)	0.035 (0.035)	0.034 (0.035)	77 (75.3)
		p_{02}	0.325	0 (-0.001)	0.021 (0.022)	0.021 (0.021)	93.3 (94.7)
		p_{03}	0.257	0 (0)	0.019 (0.02)	0.02 (0.02)	94.4 (95.3)
		p_{04}	0.2	0.001 (0)	0.025 (0.025)	0.025 (0.025)	89.3 (87.1)
	heavy	p_{01}	0.4	-0.001 (-0.001)	0.042 (0.044)	0.042 (0.044)	71 (72.2)
		p_{02}	0.325	-0.001 (0)	0.025 (0.026)	0.025 (0.026)	91.2 (91.4)
		p_{03}	0.257	0 (0.001)	0.021 (0.022)	0.021 (0.022)	96.1 (94.8)
		p_{04}	0.2	0.001 (0.002)	0.026 (0.027)	0.027 (0.028)	88 (88.2)
400	light	p_{01}	0.6	0.001 (-0.001)	0.049 (0.049)	0.048 (0.051)	79 (77.8)
		p_{02}	0.476	0.001 (-0.002)	0.033 (0.033)	0.034 (0.033)	91 (92.2)
		p_{03}	0.355	0.001 (-0.002)	0.03 (0.031)	0.031 (0.031)	94.1 (95.1)
		p_{04}	0.25	0.001 (-0.001)	0.038 (0.039)	0.038 (0.04)	87.5 (86.8)
	heavy	p_{01}	0.6	0 (-0.004)	0.06 (0.065)	0.061 (0.063)	73.5 (76.3)
		p_{02}	0.476	0 (-0.002)	0.04 (0.043)	0.041 (0.042)	90.5 (90.3)
		p_{03}	0.355	-0.002 (-0.001)	0.033 (0.034)	0.033 (0.034)	94.7 (95.2)
		p_{04}	0.25	-0.001 (0.001)	0.041 (0.042)	0.04 (0.042)	90.3 (91.7)
800	light	p_{01}	0.6	0.001 (0)	0.035 (0.035)	0.036 (0.035)	77.6 (76.4)
		p_{02}	0.476	0.001 (0)	0.024 (0.024)	0.024 (0.024)	91.9 (90.5)
		p_{03}	0.355	0.001 (0)	0.021 (0.022)	0.021 (0.021)	95.2 (94.3)
		p_{04}	0.25	0.001 (0)	0.026 (0.027)	0.026 (0.025)	88.5 (90.7)
	heavy	p_{01}	0.6	0.001 (0)	0.043 (0.046)	0.043 (0.047)	75 (74.2)
		p_{02}	0.476	0.001 (0.001)	0.029 (0.031)	0.029 (0.031)	90 (89.5)
		p_{03}	0.355	0.001 (0.001)	0.023 (0.024)	0.023 (0.024)	94.6 (95.3)
		p_{04}	0.25	0.001 (0.002)	0.028 (0.029)	0.029 (0.029)	90.1 (91)

Table 2.9: Bias, SE, RMSE, and CP for the cure rates of the Bernoulli cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0.001 (-0.002)	0.049 (0.05)	0.051 (0.049)	74.4 (75.1)
		p_{02}	0.325	0 (-0.002)	0.03 (0.03)	0.031 (0.03)	92.2 (92.1)
		p_{03}	0.257	0 (-0.002)	0.028 (0.028)	0.028 (0.029)	94.8 (93)
		p_{04}	0.2	0.001 (-0.001)	0.036 (0.036)	0.035 (0.037)	88 (86.4)
	heavy	p_{01}	0.4	0.004 (0.002)	0.059 (0.06)	0.061 (0.063)	71.3 (71.9)
		p_{02}	0.325	0.001 (0)	0.035 (0.036)	0.035 (0.037)	91.6 (90.9)
		p_{03}	0.257	-0.001 (-0.001)	0.03 (0.031)	0.029 (0.031)	95.6 (95.5)
		p_{04}	0.2	0 (0)	0.038 (0.039)	0.038 (0.04)	89.3 (88.3)
800	light	p_{01}	0.4	0 (-0.001)	0.035 (0.035)	0.035 (0.036)	76 (72.3)
		p_{02}	0.325	0 (-0.001)	0.021 (0.021)	0.021 (0.022)	92 (92.3)
		p_{03}	0.257	0 (-0.001)	0.019 (0.02)	0.019 (0.02)	95.5 (93.7)
		p_{04}	0.2	0 (-0.001)	0.025 (0.025)	0.024 (0.025)	88.5 (87.2)
	heavy	p_{01}	0.4	0.002 (0.001)	0.042 (0.043)	0.043 (0.045)	72.6 (71.2)
		p_{02}	0.325	0.001 (0)	0.025 (0.026)	0.025 (0.027)	90.5 (90)
		p_{03}	0.257	0 (0)	0.021 (0.021)	0.02 (0.022)	96.5 (94.5)
		p_{04}	0.2	0 (0)	0.026 (0.027)	0.025 (0.027)	89.4 (89.1)
400	light	p_{01}	0.6	0.004 (0.001)	0.049 (0.049)	0.049 (0.049)	78.4 (76.9)
		p_{02}	0.476	0.002 (0)	0.033 (0.033)	0.033 (0.034)	92.8 (92.1)
		p_{03}	0.355	0 (-0.002)	0.03 (0.031)	0.031 (0.032)	94.6 (94.4)
		p_{04}	0.25	0 (-0.001)	0.038 (0.039)	0.039 (0.039)	88.3 (87)
	heavy	p_{01}	0.6	-0.002 (-0.001)	0.061 (0.063)	0.058 (0.063)	76 (75.3)
		p_{02}	0.476	0 (0)	0.04 (0.042)	0.038 (0.042)	92.5 (90.8)
		p_{03}	0.355	0.001 (0)	0.033 (0.034)	0.032 (0.034)	95.7 (95.5)
		p_{04}	0.25	0.003 (0.001)	0.041 (0.043)	0.041 (0.042)	90 (88.6)
800	light	p_{01}	0.6	0.001 (-0.001)	0.035 (0.035)	0.035 (0.035)	77.1 (76.3)
		p_{02}	0.476	0 (-0.001)	0.024 (0.024)	0.024 (0.024)	92 (90.9)
		p_{03}	0.355	-0.001 (-0.001)	0.021 (0.022)	0.022 (0.023)	93.7 (93.1)
		p_{04}	0.25	-0.001 (0)	0.026 (0.027)	0.027 (0.028)	88.8 (86.3)
	heavy	p_{01}	0.6	0 (0.001)	0.043 (0.045)	0.041 (0.045)	75.8 (73.9)
		p_{02}	0.476	0 (0.001)	0.029 (0.03)	0.028 (0.029)	90.8 (90.7)
		p_{03}	0.355	0.001 (0)	0.023 (0.024)	0.023 (0.024)	94.7 (95.4)
		p_{04}	0.25	0.001 (0)	0.028 (0.029)	0.028 (0.03)	90 (89.7)

We simulated 1000 random samples from the following five choice of ϕ from COM-Poisson distributions: $\phi = 0$ (Geometric), $\phi = 0.5$, $\phi = 1$ (Poisson), $\phi = 2$, $\phi \to \infty$ (Bernoulli). Two different sample sizes were considered: $n = 400$ (95, 102, 97, 106) and $n = 800$ (200, 168, 212, 220) for small and large sample sizes. The light and heavy censoring rates considered were (0.52, 0.45, 0.37, 0.3) and (0.65,0.49,0.4,0.35) with cure rates (0.4,0.2) if the lifetime follows proportional odds model under Weibull baseline with parameter $\gamma_0 = 0.571$, $\gamma_1 = 0.307$, $\gamma_2 = -0.75$. The light and heavy censoring rates were taken as (0.7, 0.57,0.45,0.34) and (0.8, 0.64, 0.5, 0.38) with cure rate (0.6, 0.25) if the lifetime follows proportional odds model under log-logistic baseline with parameter $\gamma_0 = 01.75$, $\gamma_1 = 3.25$, $\gamma_2 = -0.75$. Here, we carry out the model discrimination by two methods, namely, Likelihood-based method and information-based method.

2.5.1 Likelihood-based method

We consider a likelihood ratio test for the null hypothesis H_0 that the competing risk follows one of the three special cases of COM-Poisson distribution, namely, Geometric $\phi = 0$, Poisson $\phi = 1$, and Bernoulli $\phi \longrightarrow \infty$ versus the alternative hypothesis H_a that the competing risk follows the general COM-Poisson distribution. The test statistic is $\Lambda = -2(\hat{l}_0 - \hat{l})$, where $\hat{l}_0$ and $\hat{l}$ are the values of the maximized log-likelihood function under the null and alternative hypotheses, respectively. The asymptotic distribution of the test statistic Λ, under $H_0 : \phi = 1$ (Poisson) follows a χ^2 distribution with one degree of freedom. However, the boundary distribution of the test statistic Λ when $\phi = 0$ (Geometric) and $\phi \longrightarrow \infty$ (Bernoulli) has a mixture distribution of χ_0^2 and χ_1^2 distributions such that $P(\Lambda \le \lambda) = \frac{1}{2} + \frac{1}{2}P(\Lambda_1 \le \lambda_1)$, where $\Lambda_1 \sim \chi_1^2$, χ_0^2 is chi-square distribution with 0 degrees of freedom and χ_1^2 is the chi-square distribution with one degree of freedom.

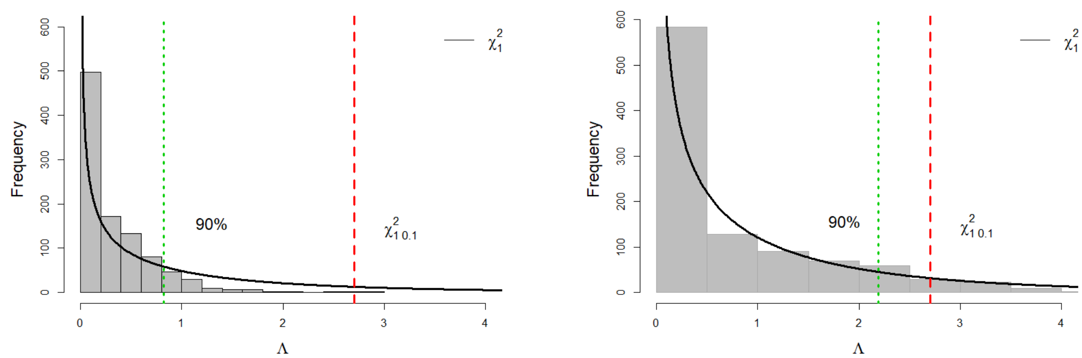

Figure 2.1: Histogram of Λ for the Poisson cure rate model under proportional odds assumption with log-logistic baseline, n=400 (left), n=4000 (right).

The values of ϕ used in the profile likelihood approach for the COM-Poisson distribution are $\{0,0.25,0.5,2/3,1,1.5,2,4,\infty\}$. Figure 2.1 provides the histograms of the test statistics Λ on the Poisson cure rate model with proportional odds assumption under log-logistic baseline when sample size equals to 400 and 4000 over 1000 generated datasets. These plots also display the probability density of chi-square distribution with one degree of freedom, and 90% quantiles. The histogram of Λ is not close to the asymptotic distribution when sample size is small while they become close as the sample size increases. This suggests a parametric bootstrap method would be better when the sample size is small, and we will describe this method in detail later in the illustrative example section. Incidentally, the same was observed under the proportional odds assumption with Weibull baseline.

2.5.2 Information-based method

We use Akaike information criterion (AIC) and Bayesian information criterion (BIC) for model selection among Geometric, Poisson, and Bernoulli distribution cure rate models.

The AIC and BIC are given by

$$\text{AIC} = 2k - 2\hat{l} \text{ and BIC} = k\log(n) - 2\hat{l}, \tag{2.43}$$

where k is the number of model parameters to be estimated, $\hat{l}$ is the maximized likelihood value, and n is the sample size. We select the model with the smallest value of AIC or BIC. AIC and BIC would give us the same selection rate since $k = 5$, $n = 400$ or 800 are the same in each of the scenarios among Geometric, Poisson, and Bernoulli distributions, i.e., the model with the largest $\hat{l}$ is the model that fits the data best. We examine the total relative bias (TRB) and total root mean square error (TRMSE) due to misspecification of the cure rate model. TRB is the sum of the absolute bias of the estimated cure rates to that of the true cure rates for each of the four groups. Similarly, TRMSE is the sum of the absolute MSE of the estimated cure rates. TRB and TRMSE due to misspecification is defined as the difference of TRB and TRMSE between the true model and the wrongly specified model (see Tables 2.10 and 2.11).

Table 2.12 shows that the selection rates under AIC or BIC for the correct models increase as the sample size increases, while they decrease as censoring rate increases, and the selection rates for the correct models are always the highest among all the cases. If the true model is under Weibull baseline, the rate to select log-logistic as the true baseline is low, and the rate becomes even lower for large sample size, or under light censoring.

2.6 Illustration with melanoma data

Now, let us consider the melanoma data described earlier in Section 1.9. In this case, we have four different cases ($x = 0, 1, 2, 3$) and the corresponding sample sizes to be $n_1 = 111$,

Table 2.10: TRB (in %) in the estimation of cure proportions when fitting different models for a given true model under proportional odds due to misspecification.

			Weibull Baseline			Loglogistic Baseline		
n	censoring	Fitted Model	Geometric	Poisson	Bernoulli	Geometric	Poisson	Bernoulli
400	light	Geometric	-	1.4	-1.5	-	0.3	-0.5
		Poisson	48	-	-0.4	0.2	-	0.8
		Bernoulli	80.8	1.4	-	0.4	0.3	-
	heavy	Geometric	-	0.9	1.5	-	2.6	1.1
		Poisson	45.9	-	1.1	0.4	-	-1.4
		Bernoulli	99.8	0	-	1.1	1.3	-
800	light	Geometric	-	1.4	-0.3	-	1.2	1.4
		Poisson	47.1	-	1.7	1	-	1.5
		Bernoulli	82.7	2	-	-0.1	-1.1	-
	heavy	Geometric	-	-0.1	1.5	-	2.5	1.5
		Poisson	48.3	-	2.5	0.9	-	1.4
		Bernoulli	100.6	-0.5	-	0.8	0.3	-

Table 2.11: TRMSE in the estimation of cure proportions when fitting different models for a given true model under proportional odds due to misspecification.

			Weibull Baseline			Loglogistic Baseline		
n	censoring	Fitted Model	Geometric	Poisson	Bernoulli	Geometric	Poisson	Bernoulli
400	light	Geometric	-	0.999	0.961	-	0.3	-0.5
		Poisson	2.081	-	1.407	0.2	-	0.8
		Bernoulli	3.354	1.434	-	0.4	0.3	-
	heavy	Geometric	-	1.013	1.023	-	2.6	1.1
		Poisson	1.637	-	1.266	0.4	-	-1.4
		Bernoulli	3.73	0.991	-	1.1	1.3	-
800	light	Geometric	-	1.026	1	-	1.2	1.4
		Poisson	3.692	-	2.071	1	-	1.5
		Bernoulli	6.693	1.539	-	-0.1	-1.1	-
	heavy	Geometric	-	0.967	0.97	-	2.5	1.5
		Poisson	3.26	-	1.38	0.9	-	1.4
		Bernoulli	7.868	0.996	-	0.8	0.3	-

Table 2.12: Selection rates based on Akaike information criterion under different settings. The data were simulated from proportional odds with Weibull baseline (γ_0= 0.571 γ_1= 0.307 γ_2= -0.75).

n	censoring	Fitted Model		True COM-Poisson Model $\phi = 0$	$\phi = 0.5$	$\phi = 1$	$\phi = 2$	$\phi \longrightarrow \infty$
400	light	Weibull	Geometric	**0.47**	0.0551	0.0828	0.0262	0.004
			Poisson	0.235	0.2906	**0.4919**	0.2553	0.2913
			Bernoulli	0.19	0.3567	0.3758	0.5146	**0.6996**
		log-logistic	Geometric	0.039	0.0802	0.0141	0.0585	0
			Poisson	0.037	0.1283	0.0222	0.0898	0.003
			Bernoulli	0.029	0.0892	0.0131	0.0555	0.002
	heavy	Weibull	Geometric	**0.344**	0.0452	0.1047	0.0241	0.011
			Poisson	0.197	0.2199	0.3797	0.204	0.279
			Bernoulli	0.229	0.3624	0.3867	0.5065	**0.677**
		log-logistic	Geometric	0.065	0.1014	0.0393	0.0894	0.01
			Poisson	0.103	0.1627	0.0584	0.0985	0.012
			Bernoulli	0.062	0.1084	0.0312	0.0774	0.011
800	light	Weibull	Geometric	**0.587**	0.0512	0.0849	0.0271	0.002
			Poisson	0.266	0.4137	**0.5905**	0.3601	0.2402
			Bernoulli	0.105	0.2801	0.3104	0.4935	**0.7578**
		log-logistic	Geometric	0.015	0.0753	0.0051	0.0411	0
			Poisson	0.012	0.1265	0.0081	0.0431	0
			Bernoulli	0.015	0.0532	0.001	0.0351	0
	heavy	Weibull	Geometric	**0.433**	0.051	0.1085	0.0301	0.002
			Poisson	0.245	0.301	**0.5136**	0.2417	0.283
			Bernoulli	0.154	0.311	0.3126	0.5035	**0.709**
		log-logistic	Geometric	0.055	0.088	0.0111	0.0662	0.003
			Poisson	0.074	0.159	0.0352	0.0832	0.001
			Bernoulli	0.039	0.09	0.0191	0.0752	0.002

$n_2 = 137$, $n_3 = 87$, $n_4 = 82$. The percentage of censored observations for the four groups were 67.57%, 61.31%, 52.87%, 32.93%. See Figure 1.1 for a plot of the lifetimes of susceptible individuals from the four nodule groups.

The initial values for the EM algorithm were chosen in the following way. We consider the censored rate as the over-estimated cured rate of groups one and four, and then calculated β_0 and β_1. The initial guess for $\boldsymbol{\gamma}$ were estimated from the linear relationship between Nelson-Aalen estimates of log odds and log t, i.e.,

$$\log O(t;\boldsymbol{\gamma}) = -\gamma_1 \log(t) + \gamma_1 \log(\gamma_0) + x\gamma_2, \tag{2.44}$$

$$\log O(t;\boldsymbol{\gamma}) = x\gamma_2 - \log(e^{(\gamma_1 t)^{1/\gamma_0}} - 1) \approx -\frac{1}{\gamma_0}\log(t) - \frac{1}{\gamma_0}\log(\gamma_1) + x\gamma_2, \tag{2.45}$$

for log-logistic and Weibull baseline distributions, respectively. We then fitted these data by Geometric ($\phi = 0$), COM-Poisson ($\phi = 0.5$), Poisson ($\phi = 1$), COM-Poisson ($\phi = 2$) and Bernoulli ($\phi \approx \infty$) cure rate models. These models, along with cure rate models proposed by Balakrishnan and Pal (2016), and Balakrishnan *et al.* (2017) are compared on the basis of AIC and BIC. From Table 2.13, we observe that $\hat{\phi} \approx \infty$ and $\hat{\phi} = 0$ provide the maximized log-likelihood values, which means that the Bernoulli and Geometric cure rate models provide a good fit for the data under log-logistic odds and Weibull odds models, respectively. Moreover, the maximized log-likelihood value increases and decreases as ϕ increases under the log-logistic odds and Weibull odds, respectively. The proposed model based on Weibull odds provides smaller values of $\hat{l}$, AIC, and most of BIC values than the model based on Weibull lifetimes and the proportional hazards lifetimes with a Weibull baseline. The proposed model based on log-logistic odds provides the smallest $\hat{l}$, AIC and most of BIC values among all the models. In conclusion, $\hat{\phi} = 0$ under proportional odds model with a log-logistic distribution as the baseline gives the best fit for the data.

Table 2.13: AIC, BIC and $\hat{l}$ under different cure rate models

	COM-Poisson	Geometric $\phi = 0$	$\phi = 0.5$	Poisson $\phi = 1$	$\phi = 2$	Bernoulli $\phi = \infty$
Proportional	AIC	1022.863	1022.845	1022.821	1022.768	**1022.683**
Odds	BIC	1043.029	1043.011	1042.986	1042.933	**1042.849**
log-logistic	$\hat{l}$	-506.432	-506.423	-506.41	-506.384	**-506.342**
Proportional	AIC	**1025.644**	1026.014	1026.374	1026.862	1027.414
Odds	BIC	**1045.809**	1046.179	1046.539	1047.027	1047.579
Weibull	$\hat{l}$	**-507.822**	-508.007	-508.187	-508.431	-508.707
Proportional	AIC	**1028.6766**	1032.4676	1034.1606	1036.0434	1038.948
Hazard	BIC	**1048.842**	1052.633	1054.326	1056.2088	1059.114
Weibull*	$\hat{l}$	**-509.3383**	-511.2338	-512.0803	-513.0217	-514.474
	AIC	**1026.838**	1032.388	1034.788	1037.792	1043.182
Weibull**	BIC	**1042.97**	1048.52	1050.92	1053.924	1059.314
	$\hat{l}$	**-509.419**	-512.194	-513.394	-514.896	-517.591

$*$ is taken from Balakrishnan and Pal (2016), while $**$ is taken from Balakrishnan *et al.* (2017).

Table 2.14: Estimates, standard errors and 95% CI for the cure rates

Cure rate model, Baseline	Bernoulli, log-logistic			Geometric, Weibull		
X	$\hat{p_0}$	SE	95% C.I.	$\hat{p_0}$	SE	95% C.I.
1	0.602	0.053	(0.497 , 0.706)	0.63	0.047	(0.538 , 0.721)
2	0.508	0.038	(0.434 , 0.583)	0.537	0.034	(0.471 , 0.603)
3	0.415	0.037	(0.343 , 0.486)	0.442	0.033	(0.377 , 0.507)
4	0.327	0.049	(0.231 , 0.422)	0.351	0.046	(0.261 , 0.44)

From Table 2.13, we can see that the maximum difference between $\hat{l}$ is only 0.09, which means that the difference between the maximized likelihood value is $e^{0.09} = 1.09$ among different ϕ's. Table 2.14 presents the estimates, standard errors and 95% CI of the cure rates stratified by nodule category. The confidence interval for all the models do not overlap for the first and forth nodule categories. Table 2.15 presents the MLES and their standard errors for different cure rate models under proportional odds assumption. It is to be noted that $\hat{l}$ is very close to each other under proportional odds model with log-logistic baseline even though the estimates are quite different. The same behaviour is also seen in the model discrimination section.

Table 2.15: MLEs and SEs of the model parameters

Cure rate model,PO Baseline	Bernoulli, log-logistic		Geometric, Weibull	
Param	MLEs	SE	MLEs	SE
β_0	-0.413	0.2226	-0.53	0.2003
β_1	0.379	0.1151	0.382	0.1049
γ_0	2.461	0.2995	0.488	0.0361
γ_1	2.266	0.1877	0.298	0.0338
γ_2	-0.473	0.1254	-0.293	0.1385

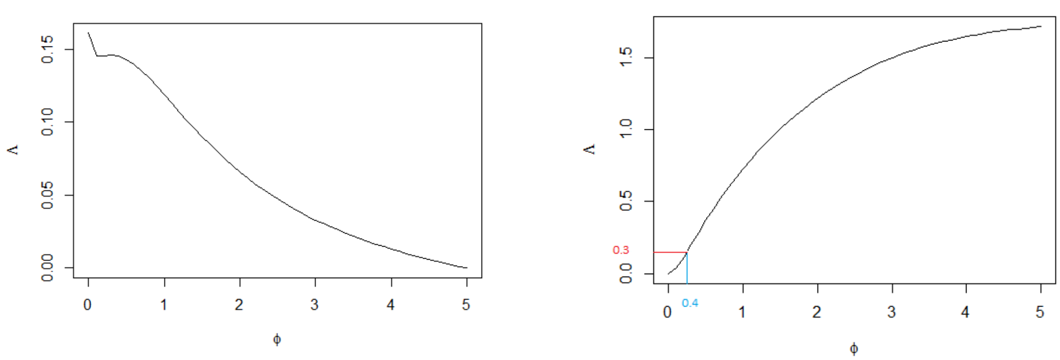

Figure 2.2: The plot of $\Lambda = -2(\hat{l_0} - \hat{l_1})$ vs ϕ under log-logistic baseline (left) and Weibull baseline (right) for the cutaneous melanoma data.

In order to further investigate the effect of ϕ under the COM-Poisson distribution, we fix ϕ from 0 to 5 with an increment of 0.1 and evaluate the maximum log-likelihood value for each ϕ through likelihood approach. Then, we test the null hypothesis $H_0 : \phi = \infty$ vs. $H_1 : 0 \leq \phi < \infty$ using the likelihood ratio test for the log-logistic odds baseline, and also $H_0 : \phi = 0$ vs. $H_1 : \phi > 0$ under the Weibull odds baseline. The likelihood ratio test statistic is given by $\Lambda = -2(\hat{l}_0 - \hat{l})$. Figure 2.2 shows that the likelihood ratio test statistic decreases and increases as ϕ increases for log-logistic and Weibull odds, respectively, which suggest that the maximized likelihoods increase and decrease as ϕ increases. As we mentioned earlier in the model discrimination part, the asymptotic distribution is not suitable when the sample size is small. So, we use a bootstrap method to obtain the distribution of the likelihood ratio test statistic Λ. We generated 1000 samples from Geometric, Poisson and Bernoulli cure rate models under proportional odds model with log-logistic and Weibull distributions, respectively. For each of the dataset, we fit the true cure rate model as well as the COM-Poisson cure rate model, and then we calculate the values of Λ. The histograms of Λ are given in Figure 2.3. The p-value is the proportion of times Λ is greater than the corresponding value determined from the data. We obtained p-values of 0.142, 0.132 and 0.681 if we test for Geometric, Poisson and Bernoulli cure rate models with log-logistic odds. Also, we obtained p-values of 0.599, 0.001 and 0.000 if we test for Geometric, Poisson and Bernoulli cure rate models under Weibull odds. Moreover, it would be of interest to get an acceptable range of ϕ if we are using the Weibull baseline for the proportional odds model. Figure 2.2 presents the values of Λ against ϕ with $\phi \in [0, 5]$. We may reject the null hypothesis $H_0 : \phi = 0$ with 10% level of significance if Λ is greater than 0.3. This implies that $\phi \in [0, 0.4)$, and the Geometric model under Weibull odds adequately fits the data.

We also set up a test on the effect of the proportional odds parameter, namely, H_0 : $\gamma_2 = 0$ as the null hypothesis vs. $H_1 : \gamma_2 \neq 0$ as alternative hypothesis for Geometric, Poisson, Bernoulli cure rate models with $\phi = 0, 1, \infty$ under Weibull (log-logistic) baseline. Note that the covariate or the nodule categories would not affect the analysis if $\gamma_2 = 0$, and the lifetime would just follow a Weibull (log-logistic) distribution. The test statistic values turned out to be 3.194 (14.95), 10.414 (19.564), 17.767 (25.99) with the corresponding p-values as 0.0739 (0.00011), 0.00125 (9.73×10^{-6}), 0.000025 (3.4×10^{-7}). Most of the p-values were less than 0.05, which shows that the proportional odds model does provide a better fit than a constant lifetime model over the four nodule categories.

Deviance residual is examined to check the error, which is defined as

$$D_i = sign(I_i + \log\hat{S}_p(t_i))\sqrt{-2(I_i + \log\hat{S}_p(t_i) + I_i\log[-(\log\hat{S}_p(t_i))])}. \tag{2.46}$$

Figures 2.6, 2.7 and 2.4, 2.5 present the deviance residuals as well as QQ plots for various fitted cure rate models. It can be seen that the deviance residuals are distributed around 0, and do satisfy the normality assumption.

2.6.1 Model diagnosis

In this section, we check if the proportional odds model with log-logistic baseline assumption on the lifetime is suitable for the cutaneous melanoma data that we have discussed in the previous section. From (2.45), we can see that $\log\ \hat{O}(t_{(i)})$ and $\log\ t_{(i)}$ should have a

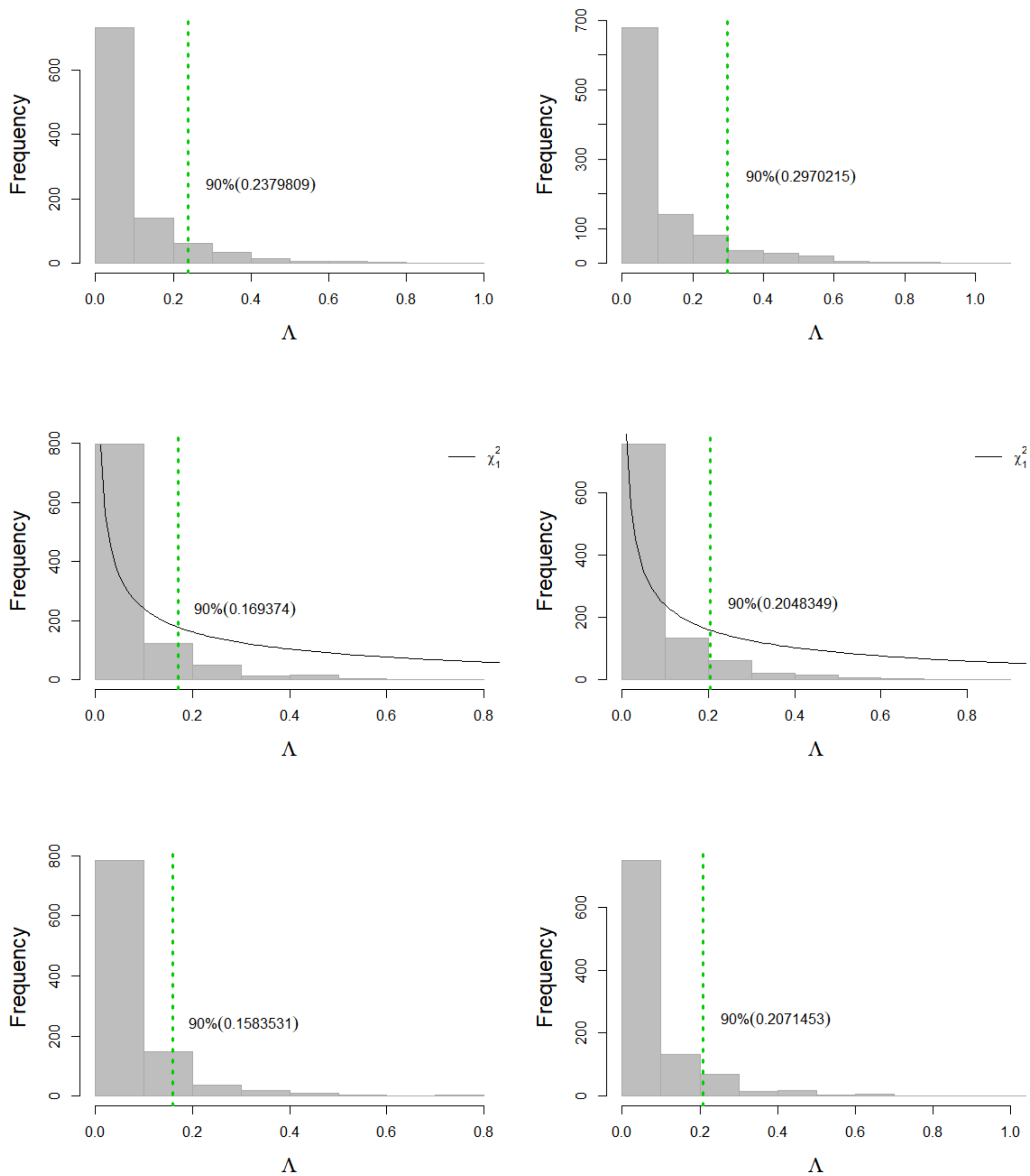

Figure 2.3: The histogram of $\Lambda = -2(\hat{l}_0 - \hat{l})$ from 1000 generated datasets with respect to MLEs on Geometric (top), Poisson (middle), Bernoulli (bottom) cure rate model with the lifetime distribution as a proportional odds model with log-logistic (left) and Weibull (right) baseline.

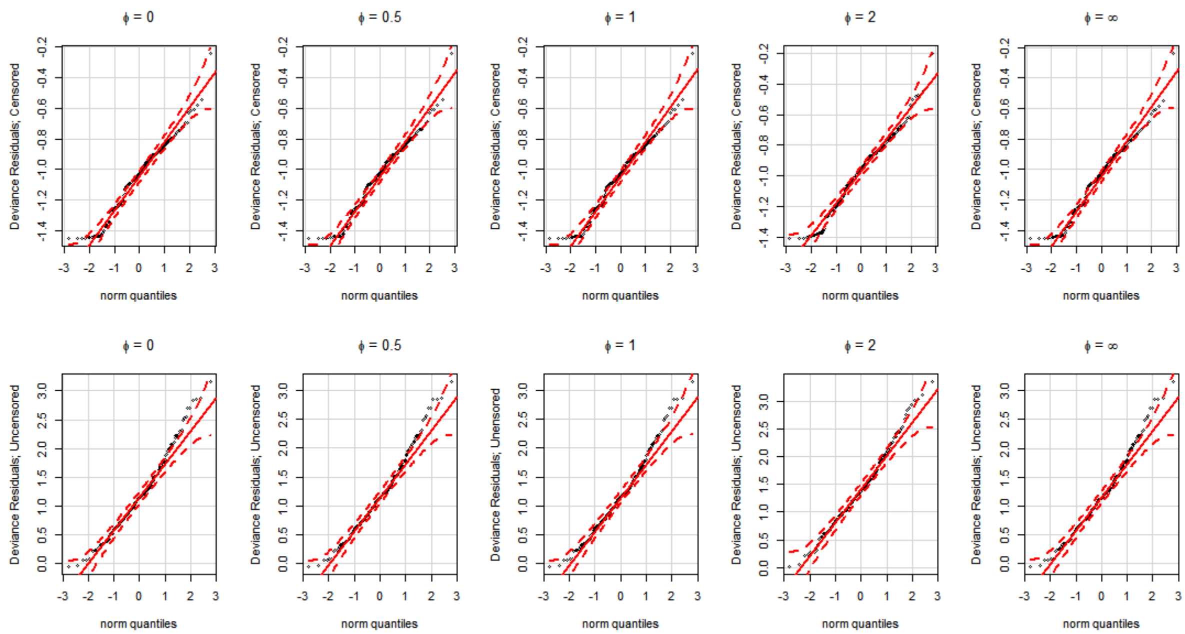

Figure 2.4: QQ plot for deviance residuals of proportional odds model with log-logistic baseline for the cutaneous melanoma data.

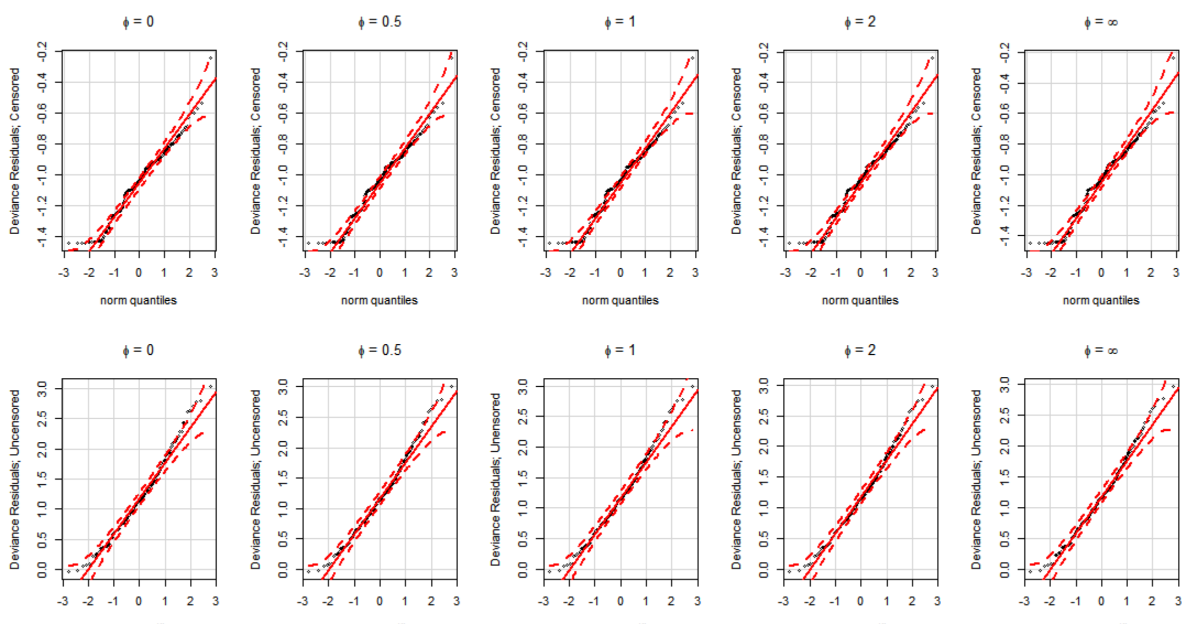

Figure 2.5: QQ plot for deviance residuals of proportional odds model with Weibull baseline for the cutaneous melanoma data.

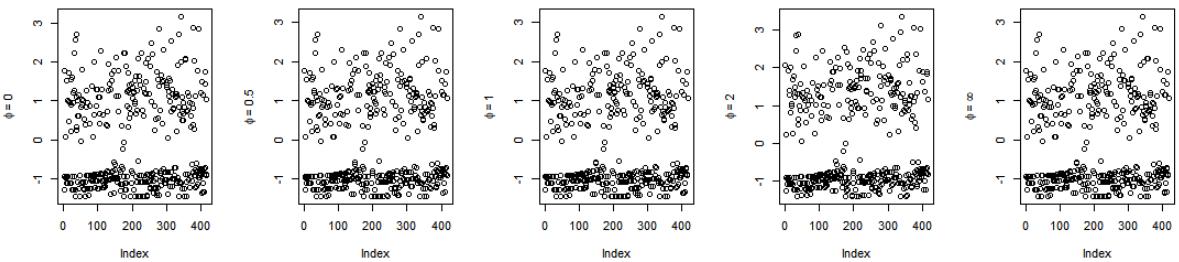

Figure 2.6: Deviance residuals of proportional odds model with log-logistic baseline for the cutaneous melanoma data.

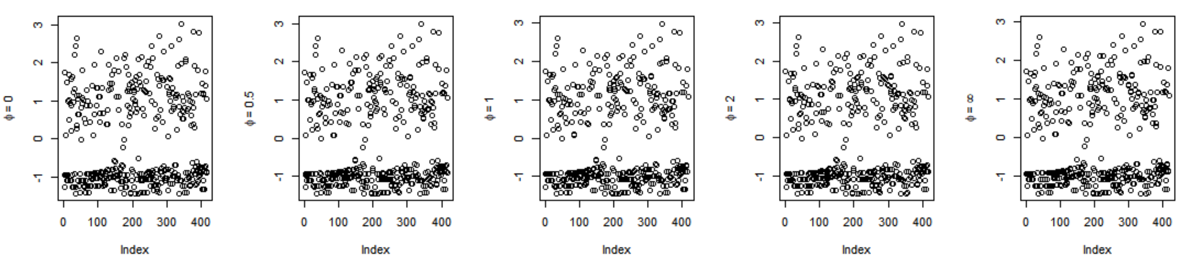

Figure 2.7: Deviance residuals of proportional odds model with Weibull baseline for the cutaneous melanoma data.

linear relationship between them. For the cutaneous melanoma data, we calculated the cumulative hazard function $\hat{H}(t)$ as the observed hazard through the non-parametric Nelson-Aalen estimator, and then get the log-odds function accordingly as

$$\hat{H}(t) = \sum_{t_i \le t} \frac{d_i}{n_i};\ \hat{S}(t) = e^{-\hat{H}(t)},\quad \log(\hat{O}) = \log(\hat{S}(t)) - \log(1 - \hat{S}(t)), \tag{2.47}$$

where d_i and n_i are the number of events and total individuals at risk at time t_i, respectively. Figure 2.8 presents the scatter plot of $\log\ \hat{O}(t_{(i)})$ vs. $\log\ t_{(i)}$ for each category. The plot shows almost a linear relationship among the four groups. It is to be noticed that there is an intersection among $x = 2$ and $x = 3$ for short survival times which violates our proportional odds assumption. Figure 2.9 shows the difference of log-odds between nodule categories 1, 2, 3 and baseline 0. In this figure, the log odds for each of the nodule categories are calculated by using the linear interpolation within the range of discrete points from 0.4 to 2. From our proportional odds assumption, $\log\ O - \log\ O_0 = \boldsymbol{x\gamma}_2$, we know that the difference should be a linear horizontal line and does not depend on the time. However, the lines in Figure 2.9 look parallel, but do show give a little curvature. Since our data include the cured individuals and are not independent, linear regression test may not be good for model diagnosis.

We use the parametric bootstrap and Monte Carlo methods to develop a goodness-of-fit test to check whether the Bernoulli cure rate model with proportional odds assumption under log-logistic baseline is sufficient for modeling these data. The critical region for this test will be to the left. We simulated 1000 datasets based on Bernoulli cure rate model with proportional odds survival assumption under log-logistic baseline. The parameters are $\beta_0 = -0.413$, $\beta_1 = 0.379$, $\gamma_0 = 2.461$, $\gamma_1 = 2.266$, $\gamma_2 = -0.473$. The censored proportion for the nodule categories are $(0.676, 0.613, 0.529, 0.329)$, respectively. We calculated the

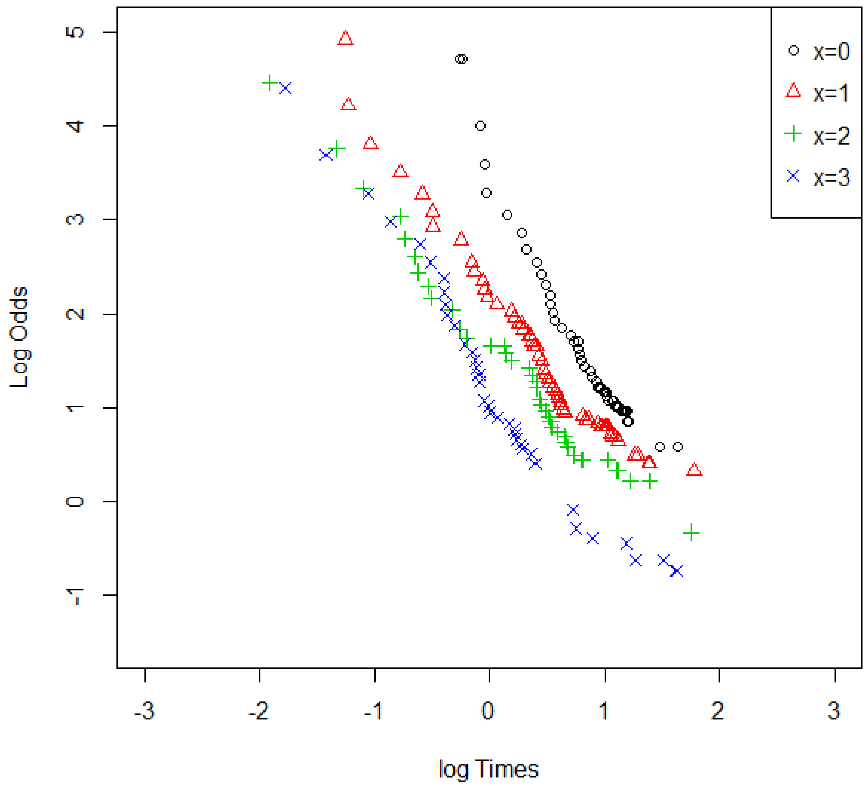

Figure 2.8: Log-odds against log t for the cutaneous melanoma data based on Nelson-Aalen estimator on patients who died.

values of maximum likelihood $\hat{l}_1, ..., \hat{l}_{1000}$ for each of these generated datasets, and order them $(\hat{l})_1, ..., \hat{l}_{(1000)}$. Then, we determine the proportion of times $\hat{l}$ is smaller than the maximum likelihood we obtained from the data as -506.342. Figure 2.10 presents the histogram of the log-likelihood over 1000 simulated datasets. We obtained the p-value for this test as 0.895, which indicates that our model is quite suitable for these data.

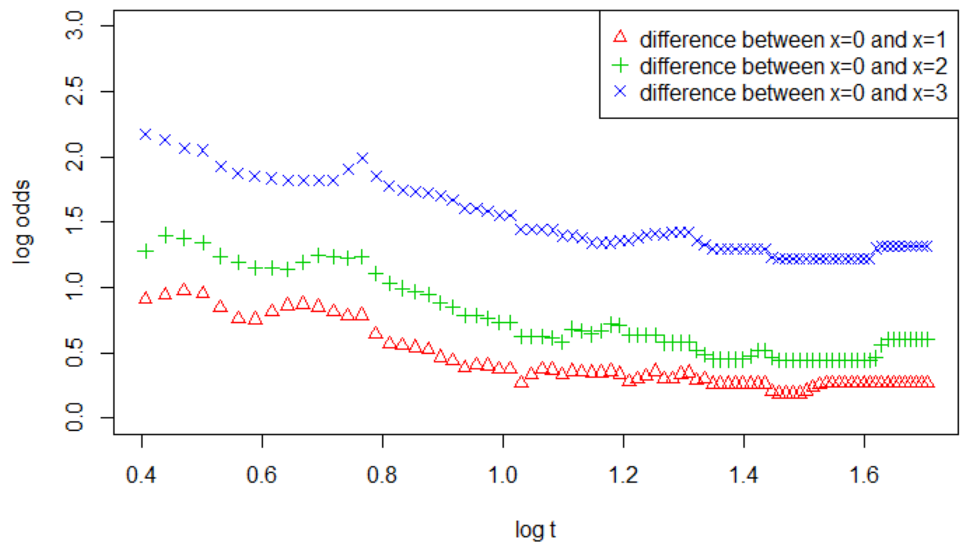

Figure 2.9: Difference of log-odds between nodule categories against log t for the cutaneous melanoma data based on Nelson-Aalen estimator.

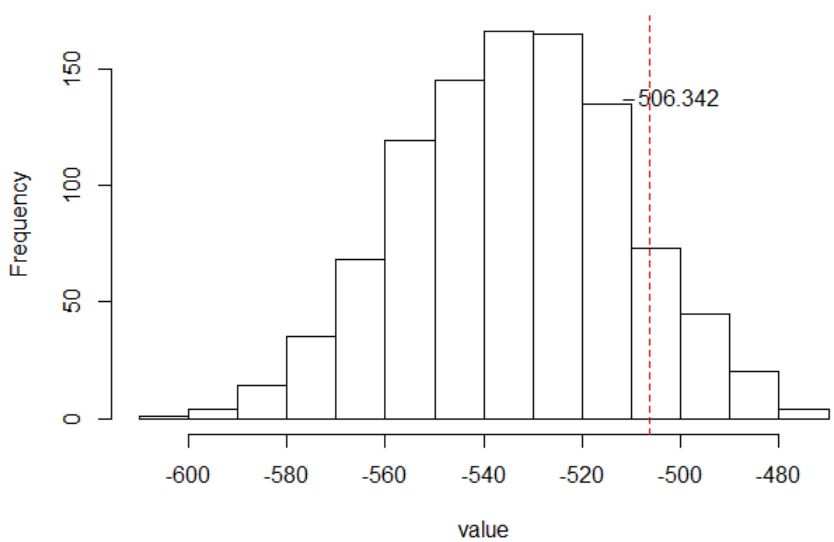

Figure 2.10: Histogram of $\hat{l}$

2.6.2 Estimation of conditional cure rate

Suppose an individual is in nodule category j, if the information is available for the individual survived up to time t, then he or she would be interested in finding the conditional cure rate based on the individual having survived up to time t, i.e. $P(I = 0|T > t)$ for $x = j$. For $x = j$, the estimate of this probability is then given by,

$$\hat{q}_{0j} = \hat{P}(I = 0|T > t) = \frac{\hat{P}(I = 0\ ,\ T > t)}{\hat{P}(T > t)}, \tag{2.48}$$

which can be estimated through Bayesian theorem,

$$\begin{aligned}\hat{q}_{0j} = \hat{P}(I = 0|T > t) &= \frac{\hat{P}(I = 0\ ,\ T > t)}{\hat{P}(T > t)} \\ &= \frac{\hat{P}(T > t|I = 0)\hat{P}(I = 0)}{\hat{P}(T > t)} \\ &= \frac{\hat{P}(I = 0)}{\hat{P}(T > t)} \\ &= \frac{\hat{p}_{0j}}{\hat{p}_{0j} + (1 - \hat{p}_{0j})\hat{S}_{sj}(t)}.\end{aligned} \tag{2.49}$$

Based on the melanoma data described earlier in Section 1.9, the corresponding plot on the conditional cure rate for the Bernoulli cure rate model with proportional odds model on log-logistic baseline are presented in Figure 2.11.

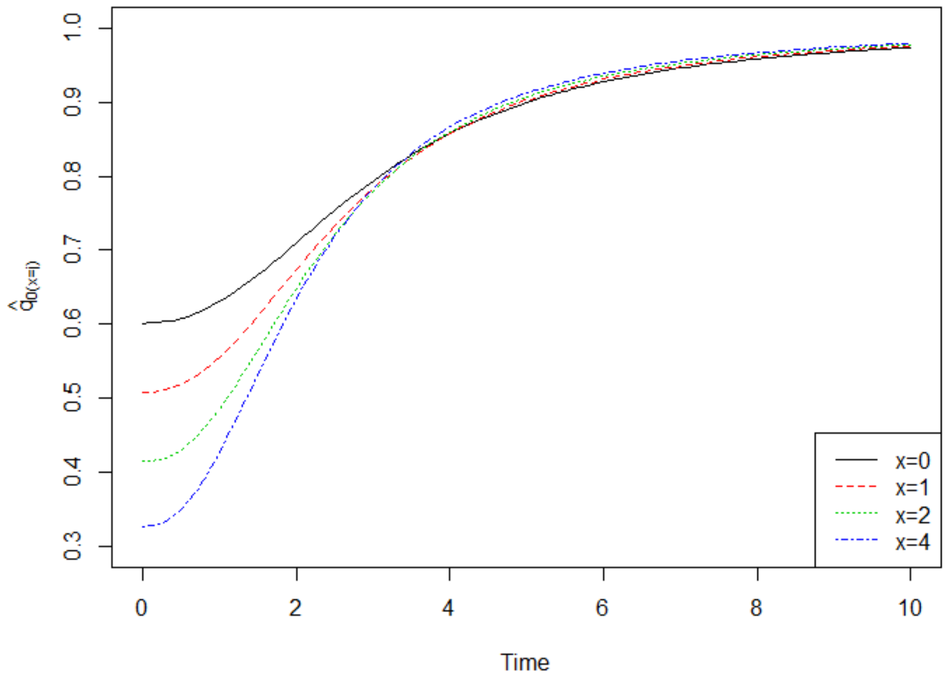

Figure 2.11: Cure rate given an individual has survived up to a specific time t over covariate groups on Bernoulli cure rate model with proportional odds model on log-logistic baseline.

2.6.3 Sample size determination

Sample size and power are key elements of study design. It would be of interest to develop a sample size and power analysis study later on for future study based on the estimated parameters and information matrix. Consider the hypothesis testing of the key factor that may affect the cure rate, i.e. for one covariate case,

$$H_0 : \hat{\beta} = 0 \;\; v.s. \;\; H_a : \hat{\beta} = \beta^* \tag{2.50}$$

The probability of at least $1 - \beta$ to reject H_0 when H_a is true is given by

$$P\left(\frac{\hat{\beta} - 0}{\sqrt{var(\hat{\beta})}} > z_\alpha \middle| H_0\right) = \alpha,$$

where z_α is the upper percentage point of the standard normal distribution.

$$P\left(\frac{\hat{\beta} - 0}{\sqrt{var(\hat{\beta})}} > z_\alpha \middle| H_a\right) \geq 1 - \gamma$$

Then the sample size can be determined from the following inequality.

$$\begin{aligned} P\left(\frac{\hat{\beta} - 0}{\sqrt{var(\hat{\beta})}} - \frac{\beta^*}{\sqrt{var(\hat{\beta})}} > z_\alpha - \frac{\beta^*}{\sqrt{var(\hat{\beta})}}\right) &\geq 1 - \gamma \\ z_\alpha - \frac{\beta^*}{\sqrt{var(\hat{\beta})}} &\leq z_{1-\gamma} \\ \frac{\beta^*}{\sqrt{var(\hat{\beta})}} &\geq z_\alpha - z_{1-\gamma}, \text{ where } var(\hat{\beta}) = \frac{\sigma^2}{n} \\ n &\geq \left(\frac{z_\alpha - z_{1-\gamma}}{\frac{\beta^*}{\sigma}}\right)^2. \end{aligned}$$

Here, σ can be determined from the illustrative data example.

Chapter 3

Destructive Cure Rate Model under Proportional Odds Lifetime

3.1 Introduction

Assume M follows a weighted Poisson distribution undergoing a damaging process as discussed earlier in Section 1.4. The probability mass functions of M and D, cure rate, population survival function and population density function for destructive length-biased Poisson (DLBP), destructive negative binomial (DNB) and destructive exponentially weighted Poisson (DEWP) cure rate models are all summarized in Table 3.1.

Table 3.1: The model, cure rate, population survival function and population density function for three special cases of weighted Poisson competing variable.

Model	$P[M=m;\eta,\phi]$	$P[D=d;\eta,\phi,p]$	p_0	$S_p(y)$	$f_p(y)$
DNB	$\frac{\Gamma(m+\frac{1}{\phi})}{\Gamma(\frac{1}{\phi})m!}(\frac{\eta\phi}{1+\eta\phi})^m(\frac{1}{1+\eta\phi})^{\frac{1}{\phi}}$	$\frac{\Gamma(d+\frac{1}{\phi})}{\Gamma(\frac{1}{\phi})d!}(\frac{\eta\phi p}{1+\eta\phi p})^d(\frac{1}{1+\eta\phi p})^{\frac{1}{\phi}}$	$(\frac{1}{1+\eta\phi p})^{\frac{1}{\phi}}$	$(1+\eta\phi pF(y))^{-\frac{1}{\phi}}$	$\frac{\eta pf(y)S_p(y)}{1+\eta\phi pF(y)}$
DEWP	$e^{-\eta e^{\phi}}\frac{(\eta e^{\phi})^m}{m!}$	$e^{-\eta pe^{\phi}}\frac{(\eta pe^{\phi})^d}{d!}$	$e^{-\eta pe^{\phi}}$	$e^{-\eta pe^{\phi}F(y)}$	$S_p(y)f(y)\eta pe^{\phi}$
DLBP	$e^{-\eta}\frac{\eta^{m-1}}{m!}$	$\frac{e^{-\eta p}(\eta p)^d}{d!}(1-p+\frac{d}{\eta})$	$(1-p)e^{-\eta p}$	$e^{-\eta pF(y)}(1-pF(y))$	$p\{\eta+\frac{1}{1-pF(y)}\}S_p(y)f(y)$

with $m = 0, 1, 2, 3 \ldots, d = 0, 1, 2, 3 \ldots, d < m$.

In this Chapter, we assume a proportional odds model for the distribution of W_j, with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w) = \theta O_0(w), \tag{3.1}$$

where $O(w) = S(w)/F(w)$ is the odds ratio (that is, the ratio of survival to failure by w), the term θ is linked to covariates in the form $e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2+\boldsymbol{z}'\boldsymbol{\gamma}_3}$ with $\boldsymbol{x}_c = (x_1, \ldots, x_p)'$ being a vector of p covariates, $\boldsymbol{z} = (z_1, \ldots, z_g)'$ being a vector of g covariates, $\boldsymbol{\gamma}_2 = (\gamma_1, \ldots, \gamma_p)'$, and $\boldsymbol{\gamma}_3 = (\gamma_{p+1}, \ldots, \gamma_{p+g})'$ are the proportional odds regression coefficients. We can further get the survival function and the corresponding probability density function (p.d.f.) for W_j as follows:

$$S(w) = [1 + e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2-\boldsymbol{z}'\boldsymbol{\gamma}_3}(S_0(w)^{-1} - 1)]^{-1},\ w > 0, \tag{3.2}$$

$$f(w) = f_0(w)e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2-\boldsymbol{z}'\boldsymbol{\gamma}_3}[(1 - S_0(w))e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2-\boldsymbol{z}'\boldsymbol{\gamma}_3} + S_0(w)]^{-2},\ w > 0. \tag{3.3}$$

The rest of this chapter proceeds as follows. Section 3.2 describes the data and the likelihood function. The estimation of the cure rate and other model parameters and their properties are discussed in Section 3.3. In Section 3.4, an extensive Monte Carlo simulation study is carried out. In Section 3.5, we conduct a model discrimination between DNB and DEWP models using information- and likelihood-based methods. In addition, a malignant melanoma data example is analyzed using the models and inferential methods developed here in Section 3.6 for illustrative purpose.

3.2 Data and the likelihood

Censoring is a common occurrence in lifetime data analysis primarily due to the life expectancy exceeding the duration of trial follow-up. Suppose the time-to-event of event of interest (say, death) is not completely observed and is subject to non-informative right censoring, which means that the lifetime above a certain value is not completely observed. Then, the observation time T_i for the ith subject would be the minimum of the censoring time C_i and the actual lifetime Y_i, i.e.,

$$T_i = \min\{Y_i, C_i\},\ i = 1, \ldots, n. \tag{3.4}$$

We define an indicator function $\delta_i = I(Y_i \leq C_i)$ for the i-th subject such that $\delta_i = 1$ if the lifetime is observed while $\delta_i = 0$ if the lifetime is right censored, sets Δ_0 and Δ_1 correspond to all the i's being equal to 0 or 1, and set Δ^* contains all the i's. It is to be noted that the cure rate p_{0i} is purely a function of η_i and p_i. We, therefore, relate the parameter p_i to covariate $\boldsymbol{x}_i = (1, x_{ic})$ by the logistic link function, and η_i to covariate $\boldsymbol{z}_i$ by the log-linear link function for all the destructive cure rate models. More specifically, we set

$$p_i = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}_1}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}_1}}, \;\; \eta_i = e^{\boldsymbol{z}_i'\boldsymbol{\beta}_2}, \; i = 1, \ldots, n, \tag{3.5}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are covariates, and $\boldsymbol{\beta}_1$, and $\boldsymbol{\beta}_2$ are the vectors of the regression coefficients. It is to be noted to be that $\boldsymbol{\beta}_2$ should not include an intercept term to retain identifiability.

For *n* pairs of observations $(\boldsymbol{t}, \boldsymbol{\delta})$=$\{(t_1, \delta_1), \ldots, (t_n, \delta_n)\}$ corresponding to n subjects, the observed data likelihood function under non-informative censoring is then given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \{f_p(t_i; \boldsymbol{\theta})\}^{\delta_i} \{S_p(t_i; \boldsymbol{\theta})\}^{1-\delta_i}, \tag{3.6}$$

where $\boldsymbol{\theta}$ is the set of parameters $(\phi, \boldsymbol{\beta}', \boldsymbol{\gamma}')$, which is equivalent to

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i \in \Delta_1} f_p(t_i; \boldsymbol{\theta}) \prod_{i \in \Delta_0} \{p_0 + (1-p_0) S_s(t_i; \boldsymbol{\theta})\}. \tag{3.7}$$

In this work, we consider two baseline distributions of the proportional odds survival model for the time-to-event random variable, namely, Weibull and Log-logistic distributions. It should be noted here that Log-logistic distribution actually processes the proportional odds property, while Weibull distribution does not. The survival function and the p.d.f. of W under a Weibull baseline are as follows:

$$S(w, \boldsymbol{\gamma}) = [1 + e^{-\boldsymbol{x}_c'\boldsymbol{\gamma}_2 - \boldsymbol{z}'\boldsymbol{\gamma}_3}(e^{(\gamma_1 w)^{1/\gamma_0}} - 1)]^{-1}, \tag{3.8}$$

$$f(w, \boldsymbol{\gamma}) = \frac{(\gamma_1 w)^{1/\gamma_0} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3 - (\gamma_1 w)^{1/\gamma_0}}}{[e^{-(\gamma_1 w)^{1/\gamma_0}}(e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3} - 1) + 1]^2 \gamma_0 w}, \tag{3.9}$$

where $w > 0$, $\gamma_0 > 0$ and $\gamma_1 > 0$ are the shape and scale parameters of the baseline

Weibull distribution, respectively. Instead, if we assume that the baseline distribution is a Log-logistic distribution with $\gamma_0 > 0$ and $\gamma_1 > 0$ as the scale and shape parameters, respectively, the corresponding odds function of W_i is seen to be

$$O(w; \boldsymbol{x}_c, \boldsymbol{z}, \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1}}{w^{\gamma_1}} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3} = O_0(w; \boldsymbol{x}_c, \boldsymbol{z}, \boldsymbol{\gamma}) e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3}. \tag{3.10}$$

Thus, W_i still follows a two-parameter log-logistic distribution ($\gamma_0, \gamma_1 > 0$) with shape parameter γ_1 and scale parameter $\gamma_0 e^{-(\boldsymbol{x}'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3)/\gamma_1}$. The corresponding survival and density functions are

$$S(w, \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3}}{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3} + w^{\gamma_1}}, \tag{3.11}$$

$$f(w, \boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3} \gamma_1 t^{\gamma_1 - 1}}{(\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2 + \boldsymbol{z}'\boldsymbol{\gamma}_3} + w^{\gamma_1})^2}, w > 0. \tag{3.12}$$

Note that the mean does not exist if $\gamma_1 < 1$ and the variance does not exist if $\gamma_1 < 2$.

3.3 Estimation of model parameters

We develop here an Expectation-Maximization algorithm to obtain the MLE of $\boldsymbol{\theta}$, and a profile likelihood approach for the estimation of the parameter ϕ. It is well-known that EM is an effective technique for finding the maximum likelihood estimates of unknown parameters of a model involving unobserved variables [for further discussion, see McLachlan and Krishnan (2007)]. In our model, the random variable I_i's are observed for i in the set Δ_1 while they are unobserved for i in the set Δ_0. Let us denote the set of complete data by

$$(\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{I}) = \{(t_1, \delta_1, \boldsymbol{x}_1, \boldsymbol{z}_1, I_1), \ldots, (t_n, \delta_n, \boldsymbol{x}_n, \boldsymbol{z}_n, I_n)\}.$$

Then, the complete data likelihood function is given by,

$$L_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\delta}, \boldsymbol{I}) \propto \prod_{i\in\Delta_1} f_p(t_i, \boldsymbol{x}_i, \boldsymbol{z}_i, \boldsymbol{\theta}) \prod_{i\in\Delta_0} p_0(\boldsymbol{\beta}, \boldsymbol{x}_i, \boldsymbol{z}_i)^{1-I_i}[(1-p_0(\boldsymbol{\beta}, \boldsymbol{x}_i, \boldsymbol{z}_i))S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta})]^{I_i}, \tag{3.13}$$

where $\mathbf{I} = (I_1, \ldots, I_n)'$, $\boldsymbol{x}_{ic} = (x_{i1}, \ldots, x_{ip})'$, $\boldsymbol{x}_i = (1, \boldsymbol{x}'_{ic})'$, and $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{ig})'$. The corresponding complete log-likelihood function (without the constant term) is as follows:

$$\begin{aligned} l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\delta}, \boldsymbol{I}) = & \sum_{i\in\Delta_1} \log f_p(t_i, \boldsymbol{x}_i, \boldsymbol{z}_i, \boldsymbol{\theta}) + \sum_{i\in\Delta_0} (1-I_i)\log p_0(\boldsymbol{\beta}, \boldsymbol{x}_i, \boldsymbol{z}_i) \\ & + \sum_{i\in\Delta_0} I_i\{\log S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}) - p_0(\boldsymbol{\beta}, \boldsymbol{x}_i, \boldsymbol{z}_i)\}. \end{aligned} \tag{3.14}$$

The complete log-likelihood (without the constant term) for the special cases of DLBP, DNB, and DEWP are as follows:

$$l_c = \sum_{i\in\Delta_1} \log p_i\left[\eta_i + \frac{1}{1-p_iF(t_i)}\right] f(t_i) + \sum_{i\in\Delta^*}\left[\log(1-p_iF(t_i)) - \eta_i p_i F(t_i)\right], \tag{3.15}$$

$$\begin{aligned} l_c = & \sum_{i\in\Delta_1}\left(\log\{\eta_i p_i f(y)\} - (1+\frac{1}{\phi}) + \log[1+\eta_i\phi p_i F(y)]\right) \\ & - \frac{1}{\phi}\sum_{i\in\Delta_0} \log\left[1+\eta_i\phi p_i F(y)\right], \end{aligned} \tag{3.16}$$

$$l_c = \sum_{i\in\Delta_1}\left(\log f(t_i) + \log\eta_i + \log p_i + \phi\right) - \sum_{i\in\Delta^*} \eta_i p_i e^{\phi} F(t_i), \tag{3.17}$$

where $F(t_i)$ follows the proportional odds assumption in (3.1).

3.3.1 E-step

The expectation step is achieved by calculating the expected value of the complete data log-likelihood function with respect to the conditional distribution of the unobserved I_i's ($i \in \Delta_0$), given the observed data $\boldsymbol{O} = \{(t_i, \delta_i, \boldsymbol{x}_i, \boldsymbol{z}_i), i \in \Delta_1\}$ under the current estimates of the parameters $\boldsymbol{\theta}^{(k)} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ for a fixed value of ϕ. Let us define the Q function as

$$Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) = E(l_c(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\delta}, \boldsymbol{I})|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}), \tag{3.18}$$

at the k-th iteration step. In our model, I_i's are Bernoulli random variables and we can therefore easily find the conditional expectation of the ith observation that is censored in the following way:

$$\begin{aligned}
\pi_i^{(k)} &= \mathrm{E}(I_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}) \\
&= P(I_i = 1|T > t) + P(I_i = 0|T > t)0 \\
&= \frac{P(T > t|I_i = 1)P(I_i = 1)}{P(T > t)} + 0 \\
&= \frac{(1 - p_0(\boldsymbol{\beta}^{(k)}, \boldsymbol{x}_i\boldsymbol{z}_i))S_s(t_i, \boldsymbol{x}_{ic}, \boldsymbol{z}_i; \boldsymbol{\theta}^{(k)})}{S_p(t_i, \boldsymbol{x}_i, \boldsymbol{z}_i; \boldsymbol{\theta}^{(k)})}|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}}.
\end{aligned} \tag{3.19}$$

For the special cases of DLBP, DNB, and DEWP, $\pi_i^{(k)}$ are as follows:

$$\text{DLBP} : \pi_i^{(k)} = 1 - \frac{e^{-\eta_i^{(k)} p_i^{(k)} S(t_i)^{(k)}}(1 - p_i^{(k)})}{1 - p_i^{(k)} F(t_i)^{(k)}}, \tag{3.20}$$

$$\text{DNP} : \pi_i^{(k)} = 1 - \left\{\frac{1 + \eta_i^{(k)} \phi_i p_i^{(k)} F(t_i)^{(k)}}{1 + \eta_i^{(k)} \phi_i^{(k)} p_i^{(k)}}\right\}^{\frac{1}{\phi}}, \tag{3.21}$$

$$\text{DEWP} : \pi_i^{(k)} = 1 - exp\left[- \eta_i^{(k)} p_i^{(k)} e^{\phi} S(t_i)^{(k)}\right]. \tag{3.22}$$

Now, for a fixed value of ϕ, the Q function is given by

$$\begin{aligned} Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)}) = &\sum_{i\in\Delta_1} \log f_p(t_i,\boldsymbol{x}_i,\boldsymbol{z}_i,\boldsymbol{\theta}) + \sum_{i\in\Delta_0}\{(1-\pi_i^{(k)})\log p_0(\boldsymbol{\beta},\boldsymbol{x}_i,\boldsymbol{z}_i) \\ &+ \pi_i^{(k)}\log[S_p(t_i,\boldsymbol{x}_i,\boldsymbol{z}_i;\boldsymbol{\theta}) - p_0(\boldsymbol{\beta},\boldsymbol{x}_i,\boldsymbol{z}_i)]\}. \end{aligned} \tag{3.23}$$

For the special cases of DLBP, DNB, and DEWP, the Q function takes on the following forms:

$$\begin{aligned} Q = &\sum_{i\in\Delta_1}\{\log p_i[\eta_i(1-p_iF(t_i))+1]f(t_i) - \eta_i p_i F(t_i)\} + \sum_{i\in\Delta_0}(1-\pi_i^{(k)})[\log q_i - \eta_i p_i] \\ &+ \sum_{i\in\Delta_0}\pi_i^{(k)}\log[e^{-\eta_i p_i F(t_i)}(1-p_iF(t_i)) - q_i e^{-\eta_i p_i}], \end{aligned} \tag{3.24}$$

$$\begin{aligned} Q = &\sum_{i\in\Delta_1}[\log(\eta_i p_i f(t_i)) - (1+\frac{1}{\phi})\log(1+\eta_i\phi p_i F(t_i))] - \sum_{i\in\Delta_0}\frac{1}{\phi}(1-\pi_i^{(k)})\log(1+\eta_i\phi p_i) \\ &+ \sum_{i\in\Delta_0}\pi_i^{(k)}\log[(1+\eta_i\phi p_i F(t_i))^{-\frac{1}{\phi}} - (1+\eta_i\phi p_i)^{-\frac{1}{\phi}}], \end{aligned} \tag{3.25}$$

$$\begin{aligned} Q = &\sum_{i\in\Delta_1}(\log[f(t_i)\eta_i p_i] + \phi - \eta_i p_i e^{\phi}F(t_i)) + \sum_{i\in\Delta_0}\{(1-\pi_i^{(k)})\eta_i p_i e^{\phi} \\ &+ \pi_i^{(k)}\log[e^{-\eta_i p_i e^{\phi}F(t_i)} - e^{-\eta_i p_i e^{\phi}}]\}, \end{aligned} \tag{3.26}$$

where $F(t_i)$ and $S(t_i)$ follow the proportional odds assumption in (3.1).

3.3.2 M-step

The M-step is achieved by maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function in order to obtain the improved estimate of $\boldsymbol{\theta}$ at the $(k+1)-$th iteration step, i.e.,

$$\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}). \tag{3.27}$$

As the MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not have explicit expressions, a numerical maximization is carried out by using one-step Newton-Raphson method. The explicit forms for the first- and second-order derivatives of the $Q(\boldsymbol{\theta}^{*}, \boldsymbol{\pi}^{(k)})$ function that are required for the numerical maximization procedure are presented in the Appendix.

For a fixed value of ϕ, the E-step and the M-step are alternated until the parameter estimates have converged to a desired level of accuracy. The parameter ϕ is determined by using the profile likelihood technique. We consider a range of ϕ with small increment for the profile likelihood estimation, and the accuracy of the estimates would be determined by the value of the range and the increment. For each value of ϕ, the MLEs of all the other parameters are found, and then the value of ϕ with the largest likelihood value is chosen as the final estimate of the parameter.

3.3.3 Standard errors and asymptotic confidence intervals

We can approximate the asymptotic variance-covariance matrix of the MLEs $(\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}')'$ by inverting the observed Fisher information matrix of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, under a fixed value of ϕ. Then, the components of the observed Fisher information matrix can be calculated from the negative second-order derivatives of the complete data likelihood function from the EM algorithm $l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\delta})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. Explicit forms of the components of the score function [for detailed information, refer to Louis (1982)] and the observed information matrix is given in the Appendix. In this way, we can obtain the standard errors as square roots of the variances and then construct the corresponding asymptotic confidence interval for each parameter.

3.3.4 Estimate of the cure rate and its standard error

The cure rate p_{0i} for group i is a function of $\boldsymbol{\beta}$. Suppose $\hat{\boldsymbol{\beta}}$ is the MLE of the regression coefficient $\boldsymbol{\beta}$. Then, the MLE of p_{0i} can be readily found and then the standard error of $\hat{p}_{0i}$ can be found through delta method as,

$$\mathrm{sd}(\hat{p_{0i}}) = \sqrt{\left(\frac{\partial\hat{p_{0i}}}{\partial\hat{\beta}_0}, ..., \frac{\partial\hat{p_{0i}}}{\partial\hat{\beta}_l}\right)\mathrm{var}(\hat{\boldsymbol{\beta}})\left(\frac{\partial\hat{p_{0i}}}{\partial\hat{\beta}_0}, ..., \frac{\partial\hat{p_{0i}}}{\partial\hat{\beta}_l}\right)'}. \tag{3.28}$$

3.3.5 Estimate of the destructive probability and its standard error

Suppose $\hat{\boldsymbol{\beta}}$ is the MLE of the regression coefficient $\boldsymbol{\beta}$. Then, the estimate of the destructive probability for group i is calculated as

$$\hat{p}_i = \frac{e^{\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_1}}{1 + e^{\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_1}},\ i = 1, ..., n. \tag{3.29}$$

The standard error of $\hat{p}_i$ can be found once again through delta method as

$$\mathrm{sd}(\hat{p}_i) = \sqrt{\left(\frac{\partial\hat{p}_i}{\partial\hat{\beta}_0}...\frac{\partial\hat{p}_i}{\partial\hat{\beta}_l}\right)\mathrm{var}(\hat{\boldsymbol{\beta}})\left(\frac{\partial\hat{p}_i}{\partial\hat{\beta}_0}, ...\frac{\partial\hat{p}_i}{\partial\hat{\beta}_l}\right)'}. \tag{3.30}$$

3.4 Simulation study

In this section, we perform an extensive Monte Carlo simulation study to demonstrate the EM algorithm developed in the preceding section. For illustrative purpose, we mimic a classical melanoma survival data described in Chapter 1, and consider various scenarios by varying the sample size, destructive probability, and censoring rate. From a medical point of view, the thickness of the tumour cell and whether the ulcer is present or not are thought to be important prognostic variables to determine the chance of survival from malignant melanoma. Therefore, tumour thickness $\boldsymbol{x}$ and ulceration state $\boldsymbol{z}$ are chosen as two covariates. A preliminary analysis of the data shows that 44% of the patients had ulceration status as present. So, we first generated a random number from Uniform(0,1), and if it was greater than 0.44, we set ulceration status to be 0 which means the ulceration is absent. The real data showed the mean of tumour thickness to be 1.81 mm for this group. Hence, we generated the tumour thickness from an exponential distribution with rate parameter as 1/1.81=0.552. Otherwise, we set the ulceration status to be 1 which means the ulceration is present and generated the tumour thickness from a Weibull distribution. The real data suggests the mean and standard deviation of tumour thickness are 4.34 mm and 3.22 mm for this group. So, from the mean and the standard deviation, we find the shape and scale parameters α_1 and α_2 through the following equations:

$$\left(\frac{3.22}{4.34}\right)^2 + 1 - \frac{\Gamma(1+\frac{2}{\alpha_1})}{\Gamma(1+\frac{1}{\alpha_1})^2} = 0, \quad \frac{4.34}{\Gamma(1+\frac{1}{\alpha_1})} - \alpha_2 = 0. \tag{3.31}$$

These values of α_1 and α_2 were then used to simulate the tumor thickness from the corresponding Weibull distribution.

The censored times were simulated from an exponential distribution with rate 0.05 and

0.1 corresponding to small and large censoring, respectively. We considered the sample sizes to be 300 and 600 corresponding to small and large sample sizes, respectively. In order to generate the damage distribution, we fixed minimum and maximum values of p corresponding to the minimum and maximum values in the covariate $\boldsymbol{x}$. The regression coefficients were then found through the equations

$$\min(\boldsymbol{p}) = \frac{e^{\beta_0+\beta_1 \min(\boldsymbol{x})}}{1+e^{\beta_0+\beta_1 \min(\boldsymbol{x})}}, \quad \max(\boldsymbol{p}) = \frac{e^{\beta_0+\beta_1 \max(\boldsymbol{x})}}{1+e^{\beta_0+\beta_1 \max(\boldsymbol{x})}}. \tag{3.32}$$

We fixed the minimum and maximum values of p as (0.2,0.6) or (0.3,0.9) for small or large destructive probability. Note that $\boldsymbol{x}$ are random for each simulation, and β_0 and β_1 were therefore not fixed throughout the simulation. The value of η was chosen to be 3.

For each individual in the Destructive Negative Binomial cure rate model, we generated the censoring time C_i from exponential distribution with rate 0.05 and 0.1 corresponding to small and large censoring. To simulate the number of competing risks, we generated a random variable M_i from a Negative Binomial distribution with parameter $\frac{1}{\phi}$ and probability of success $\frac{1}{1+\phi\eta_i}$. If $M_i = 0$, then we set $D_i = 0$. If $M_i > 0$, we generated D_i from binomial distribution with size M_i and probability of success p. If $D_i = 0$, the observed lifetime was taken to be the censoring time. Otherwise, we generated D_i random variables $\{Y_1, \ldots, Y_{D_i}\}$ from a proportional odds model with a parametric baseline. Here, we used (0.34,0.13,-0.21,1.9) for the Weibull baseline, and (8.15,2.88,-0.24,2.55) for the log-logistic baseline for the parameter vector $(\gamma_0, \gamma_1, \gamma_2, \gamma_3)$. The observed lifetimes were then taken to be the minimum of the censoring time and the actual lifetime, i.e., $T_i = \min\{Y_1, \ldots, Y_{D_i}, C_i\}$. The simulation procedure for the Destructive exponentially weighted Poisson cure rate model and Destructive Length-biased Poisson model were the same as that of the destructive negative binomial cure rate model except for the change in the corresponding competing causes random variable M_i.

We considered 1000 simulations for each specific model within each case. The code was written in R-software. We examine the performance of the proposed estimation method through bias, root mean square error, and coverage probability of confidence intervals based on the asymptotic normality of the MLEs. Tables 3.2 - 3.7 present the simulation results for all the models. From these tables, we observe that the estimates are quite accurate under different cure rate models. The Bias, standard error along with RMSE get reduced as the sample size increase. That is also the case when the censoring is light or the destructive probability is high. The coverage probabilities of the confidence intervals based on the asymptotic normality of the MlEs are quite close to the nominal level in most of the cases. In conclusion, good estimates are achieved with large sample, light censoring, and high destructive probability.

Table 3.2: Estimates, bias, standard error, RMSE and CP for **DLBP** under proportional odds model with Weibull baseline.

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	β_0	-1.387	-1.397	-0.009	0.2	0.2	0.207	91.1	95.2	-1.399	-0.013	0.139	0.139	0.145	91	96.4
		β_1	0.111	0.119	0.008	0.062	0.06	0.059	90.7	94.9	0.106	0.003	0.044	0.04	0.04	89.9	95
		β_2	1.099	1.085	-0.014	0.335	0.335	0.321	90.4	95.9	1.104	0.005	0.226	0.226	0.222	90.5	95.5
		γ_0	0.34	0.337	-0.003	0.028	0.028	0.028	88.6	94.8	0.338	-0.002	0.02	0.02	0.02	89.2	94.5
		γ_1	0.13	0.13	0	0.007	0.007	0.007	89.3	94.2	0.13	0	0.005	0.005	0.005	89.8	94.9
		γ_2	-0.21	-0.213	-0.003	0.06	0.06	0.061	90.3	95.5	-0.213	-0.003	0.045	0.046	0.043	87.8	93.9
		γ_3	1.9	1.916	0.016	0.388	0.388	0.393	91.2	95.5	1.916	0.016	0.272	0.272	0.277	89.9	95.5
	high	β_0	-0.849	-0.868	-0.019	0.223	0.224	0.218	89.8	94.8	-0.854	-0.006	0.153	0.153	0.151	90.5	95.8
		β_1	0.189	0.203	0.014	0.089	0.085	0.083	88.6	95	0.179	0.004	0.062	0.058	0.056	87.8	93.4
		β_2	1.099	1.098	0	0.304	0.304	0.3	90	95.8	1.111	0.012	0.204	0.205	0.207	89.7	95.1
		γ_0	0.34	0.338	-0.002	0.024	0.024	0.024	91	94.8	0.338	-0.002	0.017	0.017	0.017	89.2	93.8
		γ_1	0.13	0.13	0	0.006	0.006	0.006	91.2	95.3	0.13	0	0.005	0.005	0.004	89.4	94.9
		γ_2	-0.21	-0.216	-0.006	0.054	0.054	0.053	89.5	94.6	-0.212	-0.002	0.038	0.038	0.037	89.2	93.8
		γ_3	1.9	1.939	0.039	0.341	0.343	0.347	90.9	94.7	1.918	0.018	0.253	0.253	0.244	88.8	94.4
heavy	low	β_0	-1.387	-1.412	-0.025	0.25	0.251	0.253	90.4	95.4	-1.41	-0.023	0.172	0.173	0.176	91	95.1
		β_1	0.111	0.125	0.013	0.078	0.078	0.074	88.5	94.5	0.109	0.007	0.052	0.049	0.049	89.3	94.7
		β_2	1.099	1.07	-0.029	0.42	0.421	0.407	90.6	96	1.1	0.002	0.284	0.284	0.276	89.7	95.7
		γ_0	0.34	0.335	-0.005	0.032	0.032	0.032	88.4	93.9	0.338	-0.002	0.023	0.023	0.023	90.2	93.9
		γ_1	0.13	0.131	0.001	0.009	0.009	0.009	88.6	93.6	0.131	0.001	0.006	0.006	0.006	90.2	95.2
		γ_2	-0.21	-0.216	-0.006	0.073	0.073	0.072	89.2	94.7	-0.212	-0.002	0.051	0.051	0.05	89.9	94.3
		γ_3	1.9	1.94	0.04	0.488	0.49	0.476	89.5	95.3	1.921	0.021	0.341	0.341	0.334	89.1	94.7
	high	β_0	-0.849	-0.872	-0.023	0.275	0.276	0.266	89.3	94.1	-0.853	-0.005	0.183	0.182	0.184	89.3	95.2
		β_1	0.189	0.208	0.019	0.11	0.108	0.102	88.9	93.6	0.181	0.006	0.074	0.071	0.067	89	94
		β_2	1.099	1.098	-0.001	0.375	0.375	0.383	92.2	96.4	1.11	0.012	0.263	0.263	0.26	90.2	96
		γ_0	0.34	0.336	-0.004	0.027	0.027	0.027	89.6	93.6	0.337	-0.003	0.019	0.02	0.019	89.3	94.6
		γ_1	0.13	0.131	0.001	0.008	0.008	0.008	89.7	95.9	0.13	0	0.006	0.006	0.005	89.3	94.3
		γ_2	-0.21	-0.219	-0.009	0.062	0.063	0.06	89.4	94.2	-0.214	-0.004	0.042	0.043	0.043	88.5	95.2
		γ_3	1.9	1.958	0.058	0.416	0.42	0.422	90.8	95.3	1.929	0.029	0.303	0.305	0.295	89	93.8

Table 3.3: Estimates, bias, standard error, RMSE and CP for **DLBP** under proportional odds model with Log-logistic baseline.

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	β_0	-1.387	-1.403	-0.015	0.235	0.235	0.232	89.2	94.5	-1.402	-0.015	0.157	0.158	0.162	91.1	94.9
		β_1	0.111	0.123	0.011	0.072	0.071	0.069	90.1	94.4	0.107	0.004	0.049	0.046	0.046	89.6	95.2
		β_2	1.099	1.072	-0.026	0.422	0.423	0.405	90.9	95.8	1.099	0	0.278	0.277	0.273	90.3	95.5
		γ_0	8.15	8.137	-0.013	0.907	0.906	0.902	87.5	92.8	8.141	-0.009	0.615	0.615	0.635	89.9	94.8
		γ_1	2.88	2.946	0.066	0.281	0.289	0.28	90.3	95.5	2.917	0.037	0.197	0.201	0.197	90	94.8
		γ_2	-0.24	-0.245	-0.005	0.074	0.074	0.07	88.5	93.5	-0.241	-0.001	0.051	0.051	0.049	88.3	93.6
		γ_3	2.55	2.614	0.064	0.508	0.512	0.494	89.3	94.6	2.58	0.03	0.355	0.356	0.347	90.1	93.9
	high	β_0	-0.849	-0.875	-0.026	0.248	0.249	0.246	90.7	94.9	-0.851	-0.003	0.17	0.17	0.169	89.9	95.1
		β_1	0.189	0.208	0.019	0.101	0.099	0.095	88.9	94.1	0.177	0.003	0.067	0.063	0.062	88.9	94.4
		β_2	1.099	1.105	0.006	0.392	0.392	0.39	91.6	96.1	1.116	0.018	0.269	0.269	0.262	90.1	95.4
		γ_0	8.15	8.233	0.083	0.778	0.782	0.772	89.2	94.5	8.184	0.034	0.532	0.533	0.539	89.8	95.2
		γ_1	2.88	2.93	0.05	0.242	0.247	0.238	88.6	94.1	2.905	0.025	0.166	0.168	0.167	91	95.5
		γ_2	-0.24	-0.247	-0.007	0.059	0.06	0.058	90.6	94.8	-0.245	-0.005	0.041	0.041	0.041	88.9	95.8
		γ_3	2.55	2.607	0.057	0.464	0.467	0.44	88.2	93.6	2.587	0.037	0.309	0.311	0.306	89.7	95.1
heavy	low	β_0	-1.387	-1.428	-0.041	0.323	0.325	0.314	89.5	94.1	-1.409	-0.022	0.21	0.211	0.218	90.5	96
		β_1	0.111	0.129	0.017	0.1	0.1	0.094	87.7	93	0.11	0.007	0.067	0.065	0.063	88.8	94.1
		β_2	1.099	1.088	-0.01	0.687	0.686	0.709	93.7	96.7	1.106	0.008	0.44	0.44	0.422	90.2	96.1
		γ_0	8.15	8.119	-0.031	1.271	1.271	1.234	85.3	90.8	8.164	0.014	0.812	0.811	0.852	91.1	94.6
		γ_1	2.88	2.986	0.106	0.359	0.375	0.355	89.9	94.6	2.932	0.052	0.243	0.249	0.247	90.2	95.4
		γ_2	-0.24	-0.249	-0.009	0.095	0.095	0.087	87.2	93.2	-0.245	-0.005	0.062	0.063	0.061	89.2	93.7
		γ_3	2.55	2.696	0.146	0.729	0.743	0.7	89.1	93.9	2.62	0.07	0.493	0.498	0.471	88.1	93.3
	high	β_0	-0.849	-0.879	-0.03	0.35	0.351	0.334	88.3	94.1	-0.848	0	0.224	0.224	0.228	90.2	95.1
		β_1	0.189	0.218	0.029	0.141	0.14	0.13	88.4	93.9	0.181	0.007	0.088	0.084	0.082	87.8	94
		β_2	1.099	1.102	0.004	0.922	0.921	1.17	96.9	98.3	1.135	0.037	0.427	0.429	0.418	92	97.4
		γ_0	8.15	8.291	0.141	1.08	1.089	1.057	88.7	93.6	8.235	0.085	0.713	0.717	0.723	90.8	95.1
		γ_1	2.88	2.959	0.079	0.312	0.321	0.297	89.7	94.6	2.909	0.029	0.208	0.21	0.207	89.9	94.8
		γ_2	-0.24	-0.252	-0.012	0.072	0.073	0.069	88.8	93.7	-0.246	-0.006	0.049	0.049	0.05	88.9	95.9
		γ_3	2.55	2.677	0.127	0.708	0.719	1.077	89.7	94.5	2.613	0.063	0.437	0.441	0.433	89.9	95.2

Table 3.4: Estimates, bias, standard error, RMSE and CP for **DNB** under proportional odds model with Weibull baseline.

The range of ϕ are $(0.10, 0.12, 0.14, \ldots, 0.94, 0.96, 0.98, 1)$

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	ϕ	0.5	0.487	-0.013	0.428	0.428	NaN	NaN	NaN	0.333	-0.007	0.036	0.037	0.035	87	92.3
		β_0	-1.387	-1.455	-0.068	0.337	0.344	0.333	91.3	95.6	0.131	0.001	0.011	0.011	0.01	87.8	93.2
		β_1	0.111	0.151	0.039	0.204	0.207	0.143	90.5	96.1	-0.227	-0.017	0.087	0.089	0.078	85.9	91.8
		β_2	1.099	1.092	-0.006	0.33	0.33	0.32	87.9	93.7	1.956	0.056	0.584	0.586	0.517	86.1	92
		γ_0	0.34	0.333	-0.007	0.037	0.038	0.036	87.2	92.2	0.506	0.006	0.426	0.426	NaN	NaN	NaN
		γ_1	0.13	0.131	0.001	0.01	0.01	0.01	90.7	94.3	-1.427	-0.04	0.242	0.245	0.226	87.2	92.8
		γ_2	-0.21	-0.218	-0.008	0.077	0.078	0.077	90.2	94.9	0.12	0.017	0.077	0.077	0.071	89.4	94.5
		γ_3	1.9	1.941	0.041	0.53	0.531	0.504	88.1	93.2	1.103	0.005	0.249	0.249	0.228	86	91.7
	high	ϕ	0.5	0.467	-0.033	0.408	0.409	NaN	NaN	NaN	0.337	-0.003	0.027	0.027	0.026	87.3	92.2
		β_0	-0.849	-0.935	-0.086	0.385	0.394	0.374	89.3	94.7	0.13	0	0.007	0.007	0.007	88.7	94.3
		β_1	0.189	0.26	0.071	0.338	0.345	0.258	87.2	90.4	-0.213	-0.003	0.057	0.057	0.054	88.4	94
		β_2	1.099	1.121	0.022	0.337	0.337	0.279	83.5	90.5	1.925	0.025	0.38	0.381	0.355	88.2	93.6
		γ_0	0.34	0.334	-0.006	0.031	0.032	0.03	87.2	93.2	0.488	-0.012	0.402	0.402	NaN	NaN	NaN
		γ_1	0.13	0.13	0	0.009	0.009	0.008	87.5	93.9	-0.9	-0.051	0.26	0.265	0.252	88.6	94.5
		γ_2	-0.21	-0.224	-0.014	0.071	0.072	0.067	87.9	93.7	0.209	0.035	0.148	0.149	0.126	84.3	90.5
		γ_3	1.9	1.94	0.04	0.483	0.484	0.43	85.8	92	1.103	0.004	0.254	0.254	0.202	81.1	88.8
heavy	low	ϕ	0.5	0.497	-0.003	0.429	0.429	NaN	NaN	NaN	0.337	-0.003	0.022	0.023	0.022	87.3	93.1
		β_0	-1.387	-1.478	-0.091	0.435	0.445	0.41	90	94.3	0.13	0	0.006	0.006	0.006	88.9	94
		β_1	0.111	0.169	0.058	0.257	0.262	0.183	89.9	95.8	-0.214	-0.004	0.047	0.047	0.047	90.2	95.2
		β_2	1.099	1.095	-0.004	0.416	0.416	0.386	87.5	92.5	1.934	0.034	0.336	0.338	0.304	86.6	92.8
		γ_0	0.34	0.33	-0.01	0.042	0.043	0.041	86	90.6	0.517	0.017	0.43	0.43	NaN	NaN	NaN
		γ_1	0.13	0.131	0.001	0.012	0.012	0.012	90.5	94.8	-1.437	-0.05	0.286	0.29	0.274	87.3	94.2
		γ_2	-0.21	-0.223	-0.013	0.096	0.097	0.091	87.9	94.1	0.124	0.021	0.096	0.097	0.091	90.3	95.6
		γ_3	1.9	1.971	0.071	0.655	0.659	0.608	87.8	93.7	1.115	0.017	0.296	0.297	0.277	87.1	94
	high	ϕ	0.5	0.453	-0.047	0.408	0.41	NaN	NaN	NaN	0.336	-0.004	0.031	0.032	0.03	86.7	92.3
		β_0	-0.849	-0.987	-0.138	0.567	0.584	0.49	89	94.3	0.131	0.001	0.009	0.009	0.008	89	94.4
		β_1	0.189	0.362	0.173	1.714	1.72	0.584	85.2	90.5	-0.215	-0.005	0.068	0.068	0.064	86.8	93.6
		β_2	1.099	1.128	0.029	0.408	0.408	0.336	83.2	90	1.946	0.046	0.445	0.448	0.427	88.6	93.4
		γ_1	0.13	0.131	0.001	0.011	0.011	0.01	87.8	93.2	0.13	0	0.008	0.008	0.007	87.7	93.2
		γ_2	-0.21	-0.227	-0.017	0.087	0.089	0.078	85.9	91.8	-0.217	-0.007	0.055	0.056	0.055	89.9	93.9
		γ_3	1.9	1.956	0.056	0.584	0.586	0.517	86.1	92	1.934	0.034	0.413	0.415	0.364	85.3	91

Table 3.5: Estimates, bias, standard error, RMSE and CP for **DNB** under proportional odds model with Log-logistic baseline.

The range of ϕ are $(0.10, 0.12, 0.14, \ldots, 0.98, 1)$

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	ϕ	0.5	0.457	-0.043	0.423	0.424	NaN	NaN	NaN	0.486	-0.014	0.418	0.418	NaN	NaN	NaN
		β_0	-1.387	-1.48	-0.093	0.39	0.401	0.369	88.9	94.9	-1.44	-0.054	0.264	0.269	0.25	87.3	93.3
		β_1	0.111	0.158	0.047	0.249	0.253	0.167	88.5	95.2	0.123	0.02	0.086	0.086	0.081	89.3	95.1
		β_2	1.099	1.099	0	0.398	0.398	0.377	87.9	93.2	1.1	0.002	0.294	0.294	0.27	87	92.5
		γ_0	8.15	8.151	0.001	1.264	1.263	1.278	88.3	93.5	8.106	-0.044	0.89	0.89	0.88	89	93.2
		γ_1	2.88	3.002	0.122	0.374	0.394	0.363	89.2	94.7	2.943	0.063	0.25	0.257	0.252	90.7	95.5
		γ_2	-0.24	-0.245	-0.005	0.097	0.097	0.089	86.6	92.9	-0.239	0.001	0.068	0.068	0.063	88.1	93
		γ_3	2.55	2.643	0.093	0.672	0.678	0.634	86.7	93.2	2.598	0.048	0.473	0.475	0.444	88.2	93.4
	high	ϕ	0.5	0.449	-0.051	0.4	0.404	NaN	NaN	NaN	0.481	-0.019	0.404	0.405	NaN	NaN	NaN
		β_0	-0.849	-0.961	-0.112	0.45	0.463	0.417	89.1	93.9	-0.902	-0.053	0.295	0.3	0.279	88.2	93.1
		β_1	0.189	0.278	0.088	0.446	0.453	0.34	86.3	90.8	0.21	0.035	0.154	0.156	0.138	85	90.8
		β_2	1.099	1.124	0.026	0.391	0.392	0.335	84.3	91	1.106	0.007	0.286	0.286	0.24	83.6	91.2
		γ_0	8.15	8.161	0.011	1.074	1.073	1.066	87.9	91.7	8.156	0.006	0.765	0.764	0.747	88.3	93.2
		γ_1	2.88	2.979	0.099	0.3	0.315	0.303	90.6	95.2	2.936	0.056	0.218	0.225	0.212	89.5	94.4
		γ_2	-0.24	-0.254	-0.014	0.076	0.077	0.075	90	95	-0.249	-0.009	0.053	0.054	0.053	89.3	94.1
		γ_3	2.55	2.649	0.099	0.582	0.59	0.535	87	93.8	2.608	0.058	0.424	0.428	0.376	85.8	92.5
heavy	low	ϕ	0.5	0.45	-0.05	0.42	0.423	NaN	NaN	NaN	0.478	-0.022	0.422	0.422	NaN	NaN	NaN
		β_0	-1.387	-1.471	-0.084	0.574	0.58	0.519	89.1	94.6	-1.465	-0.078	0.349	0.357	0.33	88.1	93.6
		β_1	0.111	0.16	0.049	0.296	0.299	0.313	88.3	93.6	0.124	0.021	0.113	0.114	0.106	89.3	94.7
		β_2	1.099	1.277	0.178	1.463	1.473	4.487	87.2	92	1.143	0.044	0.493	0.495	0.671	85.8	92.3
		γ_0	8.15	8.32	0.17	2.352	2.357	2.228	86.5	90.8	8.062	-0.088	1.184	1.186	1.188	87.7	92.8
		γ_1	2.88	3.072	0.192	0.502	0.537	0.468	88.1	93.4	2.987	0.107	0.321	0.338	0.323	90.7	95
		γ_2	-0.24	-0.262	-0.022	0.143	0.145	0.119	84.4	90.2	-0.246	-0.006	0.089	0.09	0.08	85.9	92
		γ_3	2.55	2.892	0.342	1.657	1.691	4.534	85.2	91.6	2.678	0.128	0.687	0.699	0.804	87.9	93
	high	ϕ	0.5	0.441	-0.059	0.409	0.413	NaN	NaN	NaN	0.439	-0.061	0.399	0.404	NaN	NaN	NaN
		β_0	-0.849	-0.987	-0.137	1.165	1.173	1.644	88.689	93.594	-0.913	-0.065	0.395	0.4	0.36	87.4	92.2
		β_1	0.189	0.432	0.243	2.231	2.241	3.489	85.185	90.39	0.215	0.041	0.254	0.254	0.195	84.4	89.6
		β_2	1.099	1.319	0.22	1.406	1.422	0.585	87.588	92.693	1.13	0.031	0.439	0.44	0.379	86.8	92.4
		γ_0	8.15	8.248	0.098	1.555	1.557	1.498	86.887	91.892	8.156	0.006	1.006	1.005	0.998	88.9	93.6
		γ_1	2.88	3.019	0.139	0.397	0.421	0.382	89.489	94.895	2.96	0.08	0.279	0.29	0.265	87.7	93.7
		γ_2	-0.24	-0.263	-0.023	0.104	0.107	0.093	87.487	93.193	-0.256	-0.016	0.068	0.069	0.065	88.7	94.2
		γ_3	2.55	2.875	0.325	1.539	1.572	0.786	87.087	93.594	2.656	0.106	0.578	0.587	0.525	87.8	93.8

Table 3.6: Estimates, bias, standard error, RMSE and CP for **DEWP** under proportional odds model with Weibull baseline.

The range of $(0, .0050.01, \ldots, 0.19, 0.195, 0.2)$.

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	ϕ	0.1	0.118	0.018	0.097	0.099	NaN	NaN	NaN	0.113	0.013	0.098	0.099	NaN	NaN	NaN
		β_0	-1.387	-1.425	-0.037	0.31	0.312	0.278	85.7	92.3	-1.425	-0.038	0.229	0.232	0.193	82.9	91.1
		β_1	0.111	0.127	0.016	0.088	0.087	0.086	91.4	95.9	0.111	0.008	0.058	0.056	0.053	90.2	94.8
		β_2	1.099	1.088	-0.01	0.272	0.272	0.268	89.6	94	1.1	0.002	0.193	0.193	0.191	89.1	94.3
		γ_0	0.34	0.335	-0.005	0.034	0.034	0.034	87.2	92.4	0.337	-0.003	0.024	0.024	0.024	88.7	93.6
		γ_1	0.13	0.131	0.001	0.009	0.009	0.009	88.8	94.9	0.13	0	0.006	0.006	0.006	88.9	95.2
		γ_2	-0.21	-0.218	-0.008	0.07	0.071	0.068	89	95	-0.213	-0.003	0.048	0.048	0.048	90.8	94.2
		γ_3	1.9	1.941	0.041	0.459	0.461	0.456	88.8	94.6	1.925	0.025	0.311	0.312	0.323	90.4	95.5
	high	ϕ	0.1	0.128	0.028	0.094	0.098	NaN	NaN	NaN	0.123	0.023	0.094	0.096	NaN	NaN	NaN
		β_0	-0.849	-0.918	-0.069	0.316	0.323	0.297	86	93.5	-0.905	-0.057	0.232	0.239	0.203	83.3	89.3
		β_1	0.189	0.216	0.027	0.156	0.157	0.139	87	91.1	0.191	0.016	0.093	0.091	0.085	91	94.4
		β_2	1.099	1.12	0.021	0.244	0.245	0.225	86.3	91.6	1.105	0.006	0.167	0.167	0.162	89.3	93.9
		γ_0	0.34	0.336	-0.004	0.029	0.029	0.028	87.5	93.5	0.338	-0.002	0.021	0.021	0.02	87.8	93.2
		γ_1	0.13	0.13	0	0.008	0.008	0.008	90	94.7	0.13	0	0.006	0.006	0.005	90.2	94.2
		γ_2	-0.21	-0.218	-0.008	0.062	0.063	0.059	89	94.4	-0.211	-0.001	0.043	0.043	0.042	89.3	94.1
		γ_3	1.9	1.948	0.048	0.379	0.382	0.383	91.3	95.2	1.923	0.023	0.28	0.281	0.27	89	94
heavy	low	ϕ	0.1	0.117	0.017	0.098	0.099	NaN	NaN	NaN	0.115	0.015	0.098	0.099	NaN	NaN	NaN
		β_0	-1.387	-1.447	-0.059	0.372	0.377	0.342	87	92.8	-1.434	-0.047	0.265	0.269	0.233	84.7	90.8
		β_1	0.111	0.138	0.027	0.122	0.123	0.116	91	95.1	0.113	0.01	0.074	0.073	0.066	89.7	94.6
		β_2	1.099	1.09	-0.009	0.34	0.34	0.326	87.6	92.9	1.108	0.009	0.239	0.24	0.231	89.8	94.5
		γ_0	0.34	0.333	-0.007	0.038	0.039	0.038	87.5	92.8	0.336	-0.004	0.027	0.028	0.027	89.2	94.2
		γ_1	0.13	0.131	0.001	0.011	0.011	0.011	89.8	94.4	0.131	0.001	0.008	0.008	0.008	90.2	95
		γ_2	-0.21	-0.219	-0.009	0.086	0.086	0.082	88.4	94.1	-0.215	-0.005	0.059	0.059	0.057	88.9	93.7
		γ_3	1.9	1.957	0.057	0.573	0.576	0.55	88.2	94.6	1.951	0.051	0.387	0.39	0.386	89.1	94.4
	high	ϕ	0.1	0.132	0.032	0.092	0.098	NaN	NaN	NaN	0.123	0.023	0.094	0.097	NaN	NaN	NaN
		β_0	-0.849	-0.942	-0.093	0.383	0.394	0.368	86.2	93.1	-0.916	-0.068	0.261	0.27	0.248	85.8	91.4
		β_1	0.189	0.238	0.049	0.243	0.246	0.198	86.6	91.4	0.197	0.022	0.111	0.11	0.105	89.4	94.3
		β_2	1.099	1.135	0.037	0.302	0.304	0.275	86.9	92.4	1.111	0.013	0.198	0.199	0.196	89.2	94.9
		γ_0	0.34	0.335	-0.005	0.031	0.032	0.032	89.1	94	0.337	-0.003	0.024	0.024	0.023	88.7	93.1
		γ_1	0.13	0.131	0.001	0.01	0.01	0.009	90	95.4	0.131	0.001	0.007	0.007	0.007	88.9	94.2
		γ_2	-0.21	-0.22	-0.01	0.071	0.072	0.069	89.2	94.5	-0.213	-0.003	0.049	0.049	0.048	90.6	94.6
		γ_3	1.9	1.983	0.083	0.464	0.472	0.458	89.3	95.2	1.945	0.045	0.334	0.337	0.322	87.7	93.3

Table 3.7: Estimates, bias, standard error, RMSE and CP for **DEWP** under proportional odds model with Log-logistic baseline.

The range of ϕ are $(0, .005, 0.01, \ldots, 0.19, 0.195, 0.2)$

				$n = 300$							$n = 600$						
C	p	Par	True	Estimate	Bias	se	RMSE	S.E.	CP		Estimate	Bias	se	RMSE	S.E.	CP	
									90%	95%						90%	95%
light	low	ϕ	0.1	0.135	0.035	0.093	0.099	NaN	NaN	NaN	0.121	0.021	0.096	0.099	NaN	NaN	NaN
		β_0	-1.387	-1.474	-0.087	0.323	0.335	0.291	85	91.7	-1.439	-0.053	0.227	0.233	0.201	84.4	91
		β_1	0.112	0.176	0.064	0.228	0.236	0.176	89.2	94.4	0.136	0.033	0.14	0.142	0.109	87.9	93.9
		β_2	1.099	1.053	-0.045	0.428	0.43	0.398	84.7	89.7	1.076	-0.023	0.314	0.315	0.294	87.2	91.7
		γ_0	8.15	8.202	0.052	1.199	1.2	1.167	87.4	92.8	8.147	-0.003	0.828	0.827	0.811	88.8	93.8
		γ_1	2.88	2.994	0.114	0.362	0.379	0.345	89.1	94	2.932	0.052	0.251	0.256	0.239	89.6	95
		γ_2	-0.24	-0.25	-0.01	0.09	0.09	0.086	88.9	94	-0.244	-0.004	0.061	0.061	0.061	89.8	94.7
		γ_3	2.55	2.645	0.095	0.675	0.682	0.667	89.6	95	2.603	0.053	0.468	0.47	0.471	91.5	96.4
	high	ϕ	0.1	0.146	0.046	0.086	0.098	NaN	NaN	NaN	0.133	0.033	0.091	0.097	NaN	NaN	NaN
		β_0	-0.848	-0.962	-0.114	0.321	0.34	0.294	84.2	90.5	-0.929	-0.082	0.223	0.238	0.202	82.8	90.1
		β_1	0.19	0.292	0.102	0.345	0.357	0.286	84.6	89.3	0.231	0.056	0.204	0.211	0.165	85.2	91.4
		β_2	1.099	1.12	0.021	0.38	0.38	0.338	82.9	89.2	1.098	-0.001	0.277	0.277	0.251	83.9	89.8
		γ_0	8.15	8.193	0.043	0.965	0.966	0.957	89.2	94.1	8.214	0.064	0.673	0.676	0.685	90.6	95.7
		γ_1	2.88	2.958	0.078	0.298	0.308	0.288	88.6	94.8	2.902	0.022	0.207	0.208	0.201	88.2	94.7
		γ_2	-0.24	-0.259	-0.019	0.075	0.078	0.074	88.8	94.2	-0.25	-0.01	0.054	0.055	0.052	90.3	94.7
		γ_3	2.55	2.711	0.161	0.582	0.603	0.556	89.1	94	2.612	0.062	0.41	0.414	0.389	87.6	93
heavy	low	ϕ	0.1	0.144	0.044	0.088	0.099	NaN	NaN	NaN	0.131	0.031	0.094	0.099	NaN	NaN	NaN
		β_0	-1.387	-1.522	-0.135	0.452	0.472	0.402	86.533	92.663	-1.473	-0.087	0.286	0.299	0.268	86.7	91.9
		β_1	0.112	0.2	0.088	0.373	0.382	0.268	88.442	94.07	0.157	0.054	0.217	0.223	0.16	87.6	92.8
		β_2	1.099	1.139	0.041	0.714	0.715	0.67	85.628	89.95	1.096	-0.003	0.47	0.469	0.433	84.8	89.7
		γ_0	8.15	8.235	0.085	1.828	1.829	1.704	85.729	90.854	8.142	-0.008	1.065	1.065	1.1	89.6	93.9
		γ_1	2.88	3.061	0.181	0.469	0.503	0.451	88.241	94.07	2.957	0.077	0.323	0.332	0.307	89	95
		γ_2	-0.24	-0.268	-0.028	0.126	0.129	0.114	86.533	92.563	-0.25	-0.01	0.081	0.081	0.079	87	94
		γ_3	2.55	2.82	0.27	1.019	1.053	0.972	89.749	94.171	2.668	0.118	0.642	0.652	0.642	89.6	95.1
	high	ϕ	0.1	0.151	0.051	0.083	0.097	NaN	NaN	NaN	0.142	0.042	0.088	0.098	NaN	NaN	NaN
		β_0	-0.848	-0.997	-0.149	0.429	0.454	0.4	84.523	91.156	-0.963	-0.115	0.277	0.3	0.268	85.4	91.2
		β_1	0.19	0.351	0.161	0.598	0.617	0.523	83.92	89.246	0.255	0.08	0.28	0.291	0.217	83.7	89.2
		β_2	1.099	1.223	0.124	0.648	0.66	0.644	85.829	91.156	1.128	0.03	0.391	0.392	0.363	85.5	91.1
		γ_0	8.15	8.218	0.068	1.327	1.328	1.327	88.04	92.563	8.214	0.064	0.923	0.925	0.922	89.2	93.7
		γ_1	2.88	3.001	0.121	0.388	0.406	0.37	88.141	93.467	2.929	0.049	0.265	0.269	0.255	89.6	94
		γ_2	-0.24	-0.274	-0.034	0.099	0.105	0.094	87.236	92.965	-0.257	-0.017	0.067	0.069	0.064	88.3	94
		γ_3	2.55	2.864	0.314	0.856	0.912	0.858	87.739	93.266	2.682	0.132	0.541	0.557	0.525	89.7	94.8

3.5 Model discrimination

Among the three models discussed in the preceding sections, DLBP contains one less parameter than the other two models. So, DLBP may get selected often based on AIC or BIC if all three models have a similar log-likelihood value (cf. Occam's razor principle). For this reason, it will be interesting to do model discrimination between DNB and DEWP cure rate models. We use AIC and BIC as criterion to select the best model, which are

$$AIC = 2p - 2\hat{l}, BIC = p\log(n) - 2\hat{l}, \tag{3.33}$$

where $\hat{l}$ is the maximized log-likelihood value corresponding to the model, p is the number of parameters estimated, and n is the sample size. The model to be chosen as the best model should have the highest likelihood value, or the smallest AIC/BIC value.

Tables 3.8 and 3.9 present the selection rates based on AIC, BIC and maximized log-likelihood value for DNB and DEWP models under proportional odds lifetimes under various cases. Because DNB and DEWP contain the same number of parameters, same selection rate would result based on AIC/BIC/$\hat{l}$. From the tables, we see that the rate to select DNB as a better model increases from 51.1% to 61.4% for Weibull baseline and from 50.6% to 60.2% for Log-logistic baseline when ϕ increases from 0.2 to 3. In addition, the rate to select DEWP as a better model increases from 57.6% to 89.1% for Weibull baseline and from 57.7% to 79.2% for Log-logistic baseline when ϕ increases from 0.1 to 1. Figure 3.1 displays the histograms of $\hat{l}$ by fitting the simulated data (ϕ=2) by DNB and DEWP. The two plots are almost identical which means DEWP and DNB are almost the same in this case. Notice that had we simulated data from DEWP with $\phi = 2$ from Weibull baseline, DEWP would almost always to be selected as a better model compared to DNB cure rate model.

Table 3.8: Selection rates based on AIC, BIC and maximized log-likelihood values for DNB, DEWP under proportional odds ($n = 600$, light censoring, low destruction).

True Models			Fitted Models					
PO baseline		ϕ	AIC		BIC		log-lik	
			DNB	DEWP	DNB	DEWP	DNB	DEWP
Weibull	DNB	0.2	0.511	0.489	0.511	0.489	0.511	0.489
		0.5	0.525	0.475	0.525	0.475	0.525	0.475
		2	0.586	0.414	0.586	0.414	0.586	0.414
		2(n=2000)	0.633	0.367	0.633	0.367	0.633	0.367
		3	0.614	0.386	0.614	0.386	0.614	0.386
	DEWP	0.1	0.424	0.576	0.424	0.576	0.424	0.576
		1	0.109	0.891	0.109	0.891	0.109	0.891
		2	0	1	0	1	0	1
log-logistic	DNB	0.2	0.506	0.494	0.506	0.494	0.506	0.494
		0.5	0.513	0.487	0.513	0.487	0.513	0.487
		2	0.529	0.471	0.529	0.471	0.529	0.471
		3	0.602	0.398	0.602	0.398	0.602	0.398
	DEWP	0.1	0.423	0.577	0.423	0.577	0.423	0.577
		1	0.208	0.792	0.208	0.792	0.208	0.792
		2	0.036	0.964	0.036	0.964	0.036	0.964

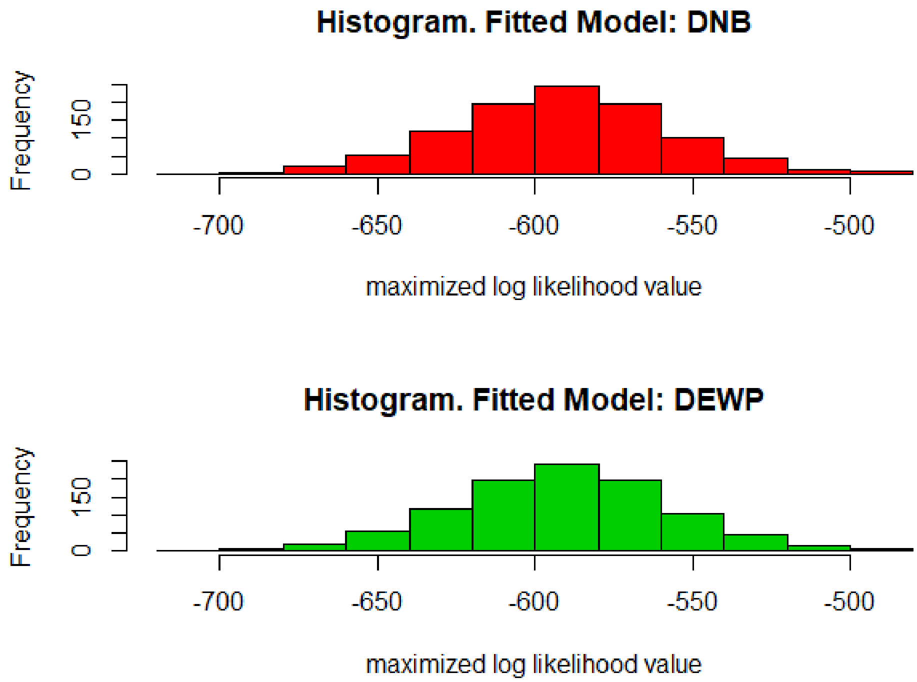

Figure 3.1: Histograms of $\hat{l}$ based on 1000 simulated data from DNB cure rate model (with ϕ=0.2). The value of ϕ to be selected for the profile likelihood method are $(0.1, 0.12, 0.14, \ldots, 0.58, 0.6)$.

Table 3.9: Selection rates based on AIC, BIC and maximized log-likelihood values for DNB, DEWP under proportional odds lifetimes (n=600).

True Models			Fitted Models					
	C	p	AIC		BIC		log-lik	
	DNB ($\phi = 2$)		DNB	DEWP	DNB	DEWP	DNB	DEWP
Weibull	light	high	0.616	0.384	0.616	0.384	0.616	0.384
	heavy	low	0.546	0.454	0.546	0.454	0.546	0.454
		high	0.598	0.402	0.598	0.402	0.598	0.402
Log-logistic	light	high	0.586	0.414	0.586	0.414	0.586	0.414
	heavy	low	0.537	0.463	0.537	0.463	0.537	0.463
		high	0.581	0.419	0.581	0.419	0.581	0.419
	DEWP ($\phi = 2$)		DNB	DEWP	DNB	DEWP	DNB	DEWP
Weibull	light	high	0.028	0.972	0.028	0.972	0.028	0.972
	heavy	low	0.005	0.995	0.005	0.995	0.005	0.995
		high	0.068	0.932	0.068	0.932	0.068	0.932
Log-logistic	light	high	0.029	0.971	0.029	0.971	0.029	0.971
	heavy	low	0.034	0.966	0.034	0.966	0.034	0.966
		high	0.091	0.909	0.091	0.909	0.091	0.909

Figures 3.2-3.5 plot the mean, relative bias, and relative RMSE of cure rate corresponding to x with different z by fitting DNB and DEWP under proportional odds lifetimes with Weibull and Log-logistic baseline distributions. When we simulated data from DEWP, the mean estimates of the cure rate from both model are very close to the true value. The relative bias is very close to zero, especially for $z = 0$ with the range being from -0.015 to 0.005. However, DEWP tends to have a smaller relative bias and relative RMSE compared to DNB. When we simulated data from DNB, the mean estimates of the cure rate from both

model are very close to the true value. The relative bias is very close to zero, especially for $z = 0$ with the range being from -0.01 to 0.005. However, DNB tends to have a smaller relative bias and relative RMSE compared to DNB. In conclusion, the rate of selecting the true model as a better model is larger if we have a larger ϕ, or larger sample size, or light censoring, or low destruction probability.

3.6 Illustrative data analysis

Let us now consider the melanoma data described earlier in Section 1.9. In this case, we select tumor thickness (in mm) and ulceration status (absent: $n = 115$; present: $n = 90$) as two covariates. We linked the destructive probability p to tumor thickness and the parameter η to ulceration status, respectively. The regression coefficients were chosen to be $\boldsymbol{\beta}_1 = (\beta_0, \beta_1)$ and $\boldsymbol{\beta}_2 = \beta_2$. The parameter $\boldsymbol{\beta}_2$ does not include an intercept term in order to avoid non-identifiability problem as mentioned earlier. A preliminary analysis of Kaplan-Meier estimates against time by ulceration status (Figure 3.6) indicates there exists cure rates for the group with and without ulceration. We also plot in Figure 3.7 the log-odds based on Nelson-Aalen estimates over time stratified by ulceration status to validate the proportional odds assumption. The two curves are almost parallel and do not intersect, which suggests that the proportional odds assumption on the lifetimes is quite reasonable.

We fitted destructive length-biased Poisson, destructive exponentially weighted Poisson and destructive negative binomial cure rate models under proportional odds lifetimes with Weibull and log-logistic baseline to these data. In order to employ the profile likelihood method on the estimation of ϕ for each model, we chose a range of ϕ to be (-2,2) and (0,10) with an increment of 0.1 in DEWP and DNB models, respectively. These resulted in 41 and 101 distinct values of ϕ. We found the estimates for each value of ϕ, and chose the one

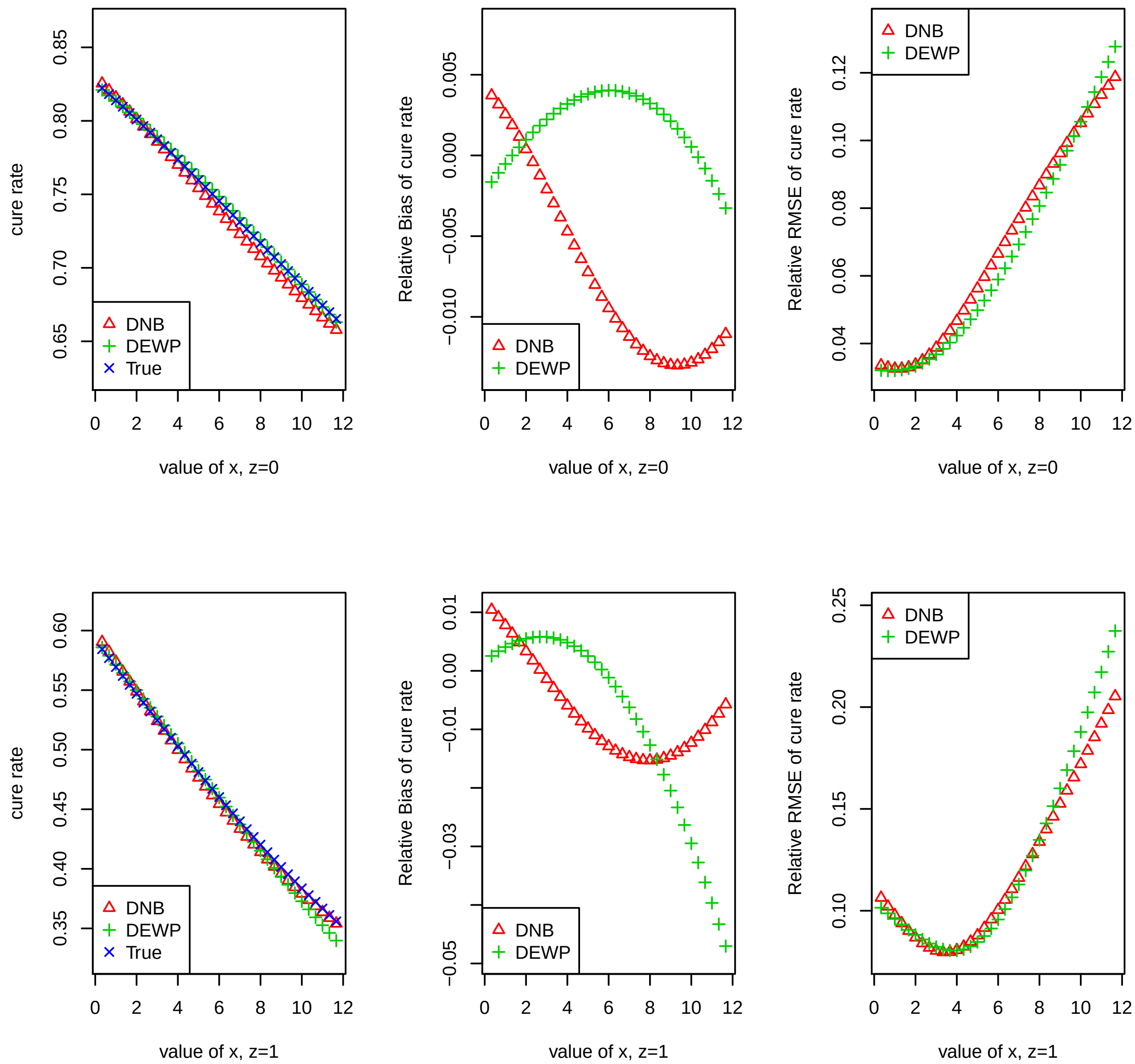

Figure 3.2: Estimates, relative bias and relative RMSE of cure rate by fitting DNB and DEWP under proportional odds lifetimes with Weibull baseline. These figures are based on 1000 simulated data from DNB($\phi = 0.2$) with $n = 600$, on light censoring and low destruction.

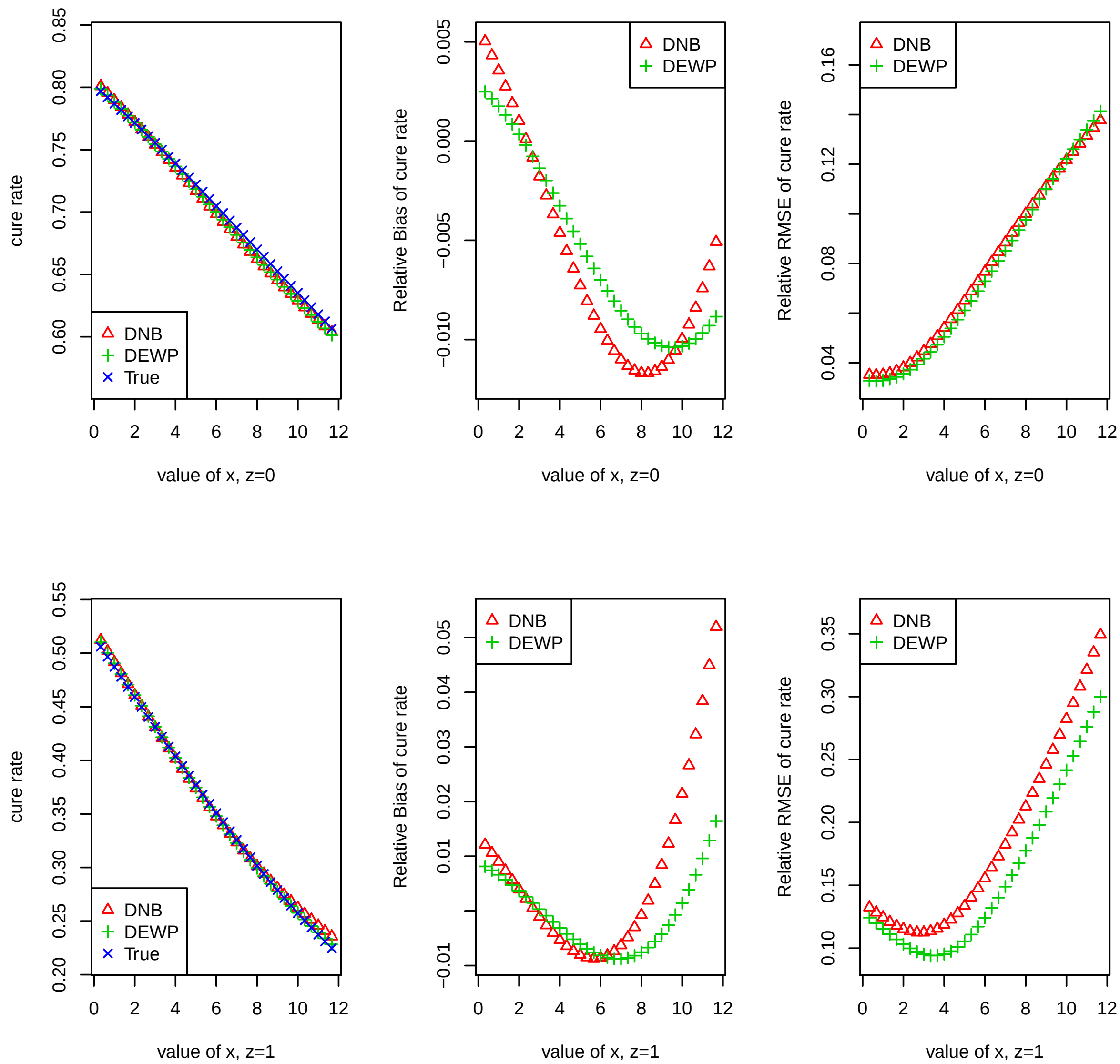

Figure 3.3: Estimates, relative bias and relative RMSE of cure rate by fitting DNB and DEWP under proportional odds lifetimes with Weibull baseline. These figures are based on 1000 simulated dataset from DEWP($\phi = 0.1$) with $n = 600$, on light censoring and low destruction.

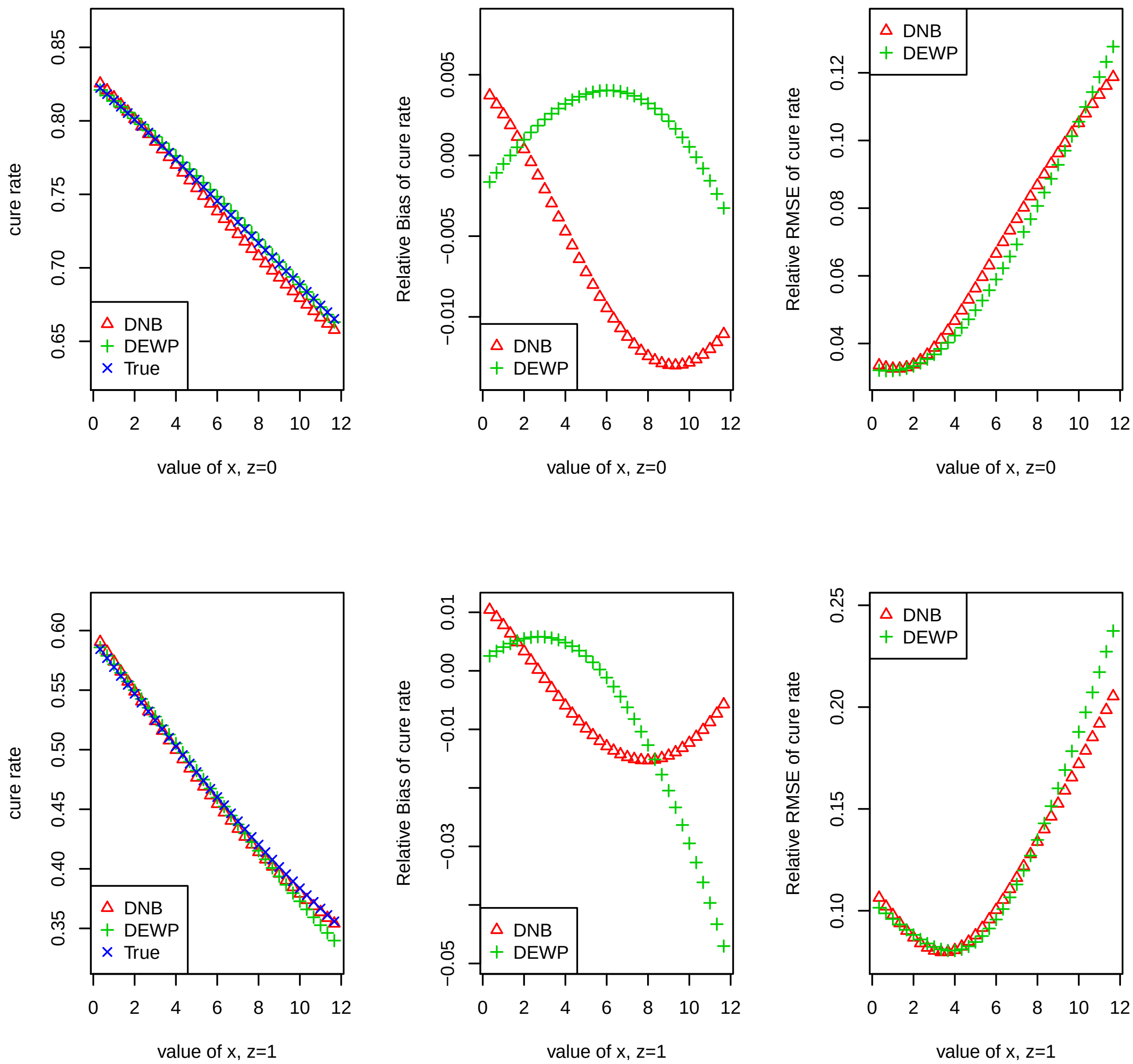

Figure 3.4: Estimates, relative bias and relative RMSE of cure rate by fitting DNB and DEWP under proportional odds lifetimes with log-logistic baseline. These figures are based on 1000 simulated data from DNB($\phi = 0.2$) with $n = 600$, on light censoring and low destruction.

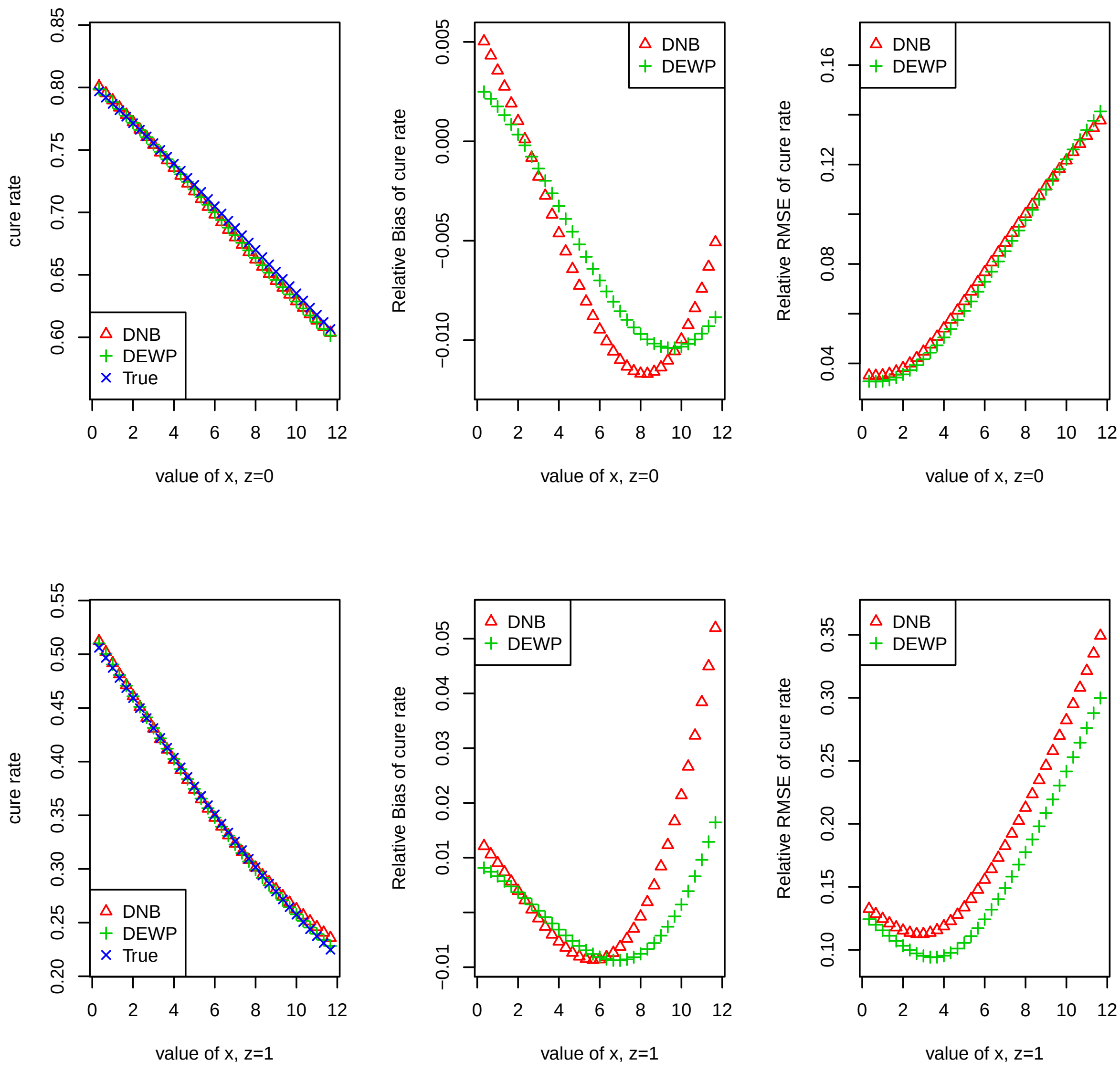

Figure 3.5: Estimates, relative bias and relative RMSE of cure rate by fitting DNB and DEWP under proportional odds lifetimes with log-logistic baseline. These figures are based on 1000 simulated dataset from DEWP($\phi = 0.1$) with $n = 600$, on light censoring and low destruction.

corresponding to the maximized log-likelihood value as the MLE.

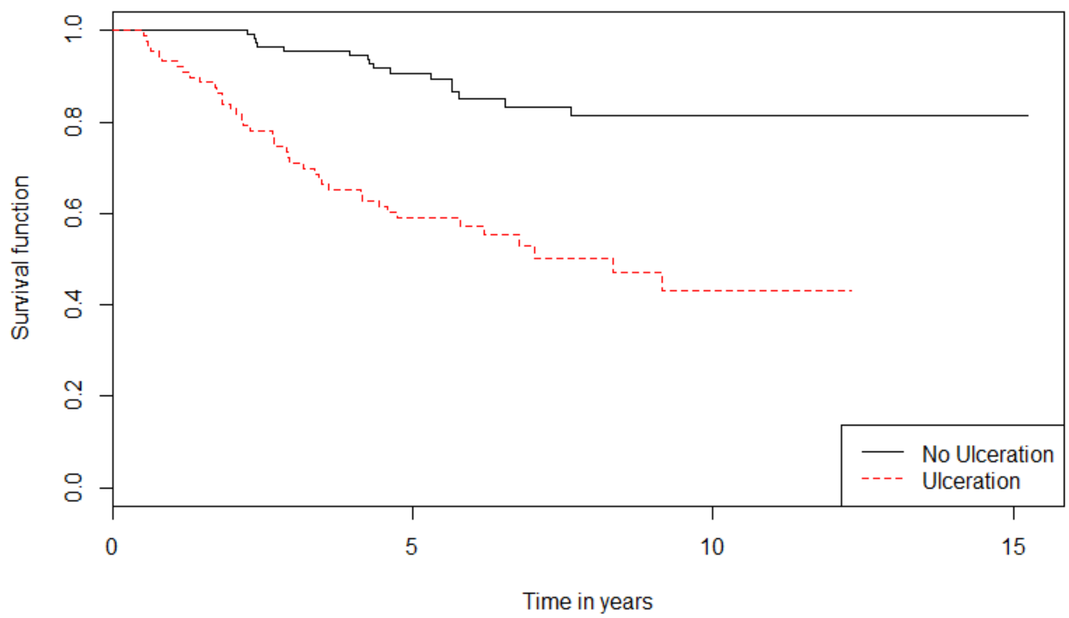

Figure 3.6: Kaplan-Meier curves for cutaneous melanoma data by ulceration status.

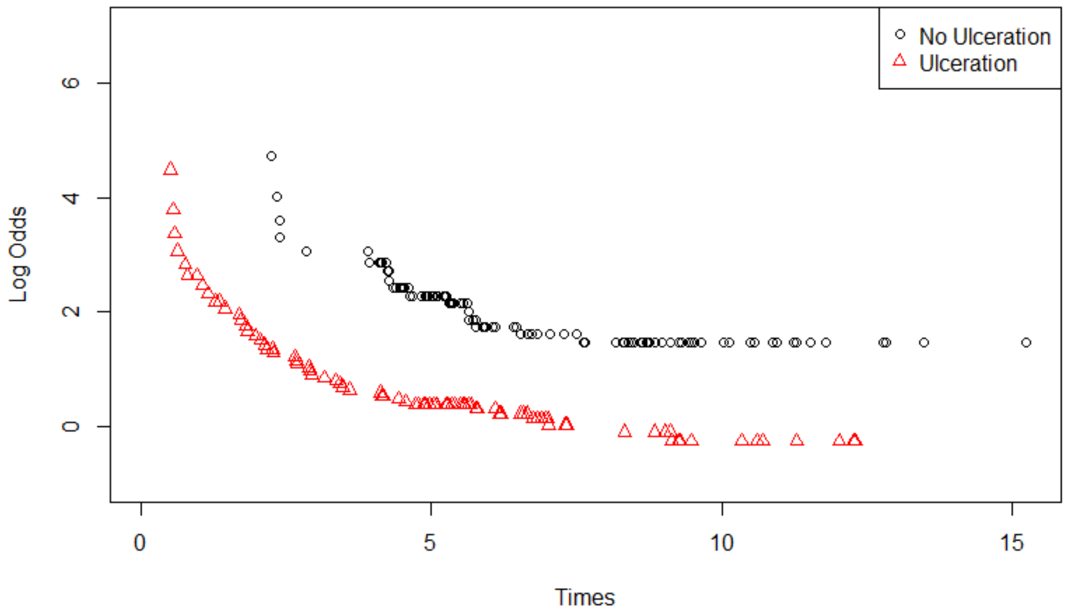

Figure 3.7: log-odds based on Nelson-Aalen estimates, based on ulceration status.

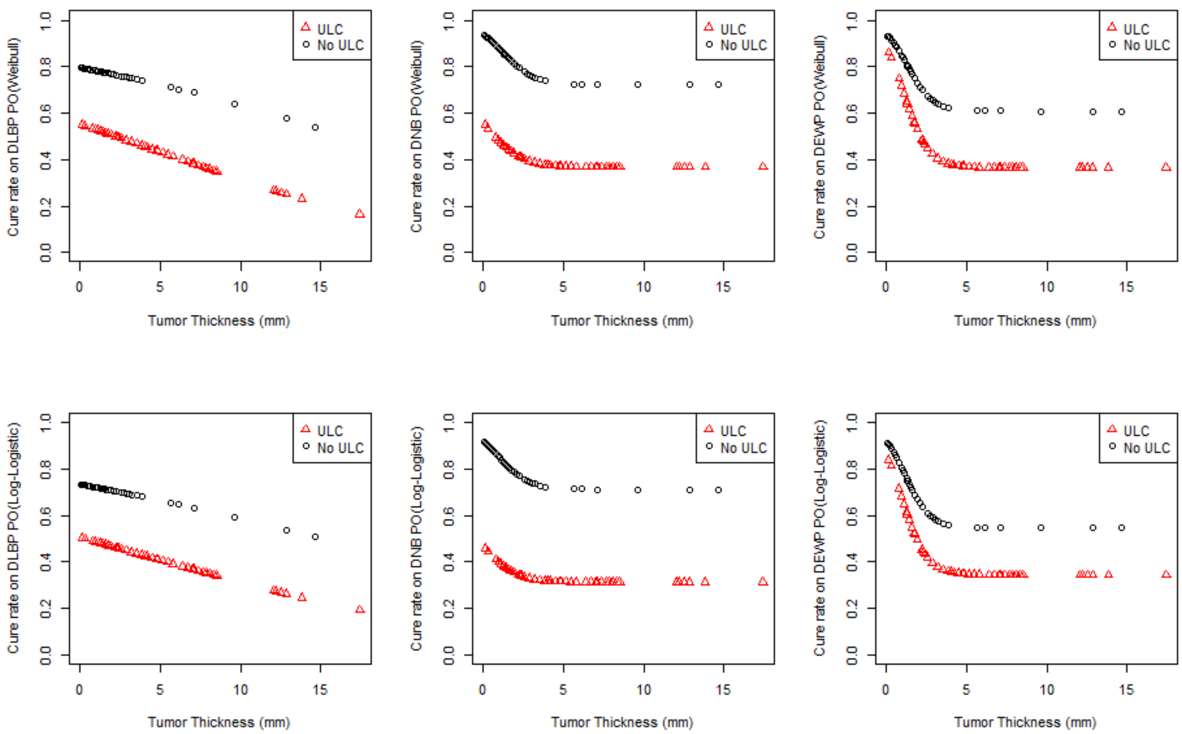

Figure 3.8: Cure rate against tumour thickness in mm stratified by ulceration (ULC) status.

Figure 3.8 displays the estimated cure rate against tumour thickness with respect to the six different cure rate models. The graph shows that the cure rate is always higher when ulceration status is absent as compared to when it is present, and the cure rate decreases with higher tumour thickness value as expected. It implies that if the tumour thickness is small and there is no ulcer, it is highly likely that the patient gets cured compared.

Note that $\phi = 1$ or $\phi = 0$ in DNB or DEWP leads to a destructive geometric model or destructive Poisson model. To test $\phi = 1$ for the full DNB model, we found the likelihood ratio test statistic to be 2.12 with corresponding p-value as 0.145. To test $\phi = 0$ for the full DEWP model, we found the likelihood ratio test statistic to be 0.16 with corresponding p-value as 0.69. The test statistic is taken as $\Lambda = -2(\hat{l}_0 - \hat{l})$, where $\hat{l}_0$ and $\hat{l}$ are the values of the maximized log-likelihood function under the null and alternative hypotheses, respectively. The asymptotic distribution of the test statistic Λ, under H_0 :, follows a χ^2 distribution with one degree of freedom. To validate the heterogeneity among the lifetimes of the susceptibles, we set up the null hypothesis as $\gamma_2 = \gamma_3 = 0$ for the DNB model, and found the likelihood-ratio test statistic to be 4.296 with corresponding p-value 0.12. The test statistic is taken as $\Lambda = -2(\hat{l}_0 - \hat{l})$, where $\hat{l}_0$ and $\hat{l}$ are the values of the maximized log-likelihood function under the null and alternative hypotheses, respectively. The asymptotic distribution of the test statistic Λ, under H_0, follows a χ^2 distribution with two degree of freedom.

We also considered the case when $p = 1$, i.e., the model becomes a standard cure rate model and the destructive mechanism is absent in the malignant cells. In case of DNB, we can get geometric and negative binomial cure rate models upon setting $\phi = 1, p = 1$ and $p = 1$, respectively. In the case of DEWP, we can get exponentially weighted Poisson and Poisson cure rate models by setting $\phi = 0, p = 1$ and $p = 1$, respectively. When $p = 1$, we can link both x and z to η by a log-linear link function of the form $\eta = exp\{\beta_0 + \beta_1 x + \beta_2 z\}$ to incorporate the two covariates and retain the number of parameters. To test the presence of destruction in DNB, we set the null hypothesis as $H_0 : p_i = 1$ vs. alternative hypothesis $H_a : p_i < 1$ for each individual i, and use a Wald type test statistic $\frac{(\hat{p}_i - 1)^2}{var(\hat{p}_i)}$. It is to be noted that the value of destructive probability p_i is bounded between zero and one, and also we cannot use the standard chi-square distribution as asymptotic distribution for the test statistic. Instead, we use a parametric bootstrap approach to determine the distribution of the test statistic. The critical value at 5% level turned out to be 31.3 from 10,000 bootstrap iterations. Figure 3.9 displays the value of tumour thickness versus the test statistic at 5% level of significance. The test statistic value decreases with increase in tumour thickness, as expected. This implies that when the tumour is not thick, for example, greater than 1.29 mm, we do not find enough evidence towards the presence of destruction.

In addition, we also linked η to tumour thickness and p to ulceration status of the form

$$p = \frac{e^{\beta_0 + z\beta_1}}{1 + e^{\beta_0 + z\beta_1}}, \quad \eta = e^{x\beta_2}. \tag{3.34}$$

Again, the coefficient corresponding to x does not include an intercept term to avoid non-identifiability problem. The resulting maximized log-likelihood value was always lower. Table 3.10 presents the estimates of ϕ as well as the maximized log-likelihood value for various models. The destructive negative binomial cure rate model with proportional odds under Weibull lifetimes gives the maximized log-likelihood value as -199.76 and with minimum AIC of 415.51 and BIC of 442.1 with $\phi = 6$. Table 3.11 provides the estimates, standard error and 95% confidence interval corresponding to this model.

Table 3.10: The maximized log-likelihood value for various destructive cure rate models with other link functions.

			Weibull PO				Log-logistic PO			
Link function	Model	k	$\hat{\phi}$	$\hat{l}$	AIC	BIC	$\hat{\phi}$	$\hat{l}$	AIC	BIC
$p_i = \frac{1}{1+e^{-\beta_0 - x_i\beta_1}}$	DLBP	7	NA	-200.85	415.69	438.95	NA	-203.66	421.31	449.34
$\eta_i = e^{z\beta_2}$	DNB	8	6	**-199.76**	**415.51**	442.1	5.5	-200.29	416.57	448.61
	DEWP	8	-0.7	-203.57	423.14	449.73	-0.5	-201.23	418.45	450.48
	DG	7	NA	-200.82	415.65	438.91	NA	-200.99	415.98	444
	DP	7	NA	-203.45	420.91	444.17	NA	-201.9	417.81	445.84
$\eta_i = e^{\beta_0+\beta_1 x_i+\beta_2 z_i}$	NB	8	10	-200.3	416.59	443.18	8.6	-200.94	417.88	449.92
$p_i = 1$	G	7	NA	-203.44	420.88	444.14	NA	-206.35	426.71	454.74
	EWP	8	-1.40	-203.65	423.31	449.89	-1.1	-203.71	423.43	455.46
	P	7	NA	-203.65	421.31	444.57	NA	-203.71	421.43	449.45
$p_i = \frac{1}{1+e^{-\beta_0 - z_i\beta_1}}$	DLBP	7	NA	-203.77	421.53	444.79	NA	-203.77	421.53	449.56
$\eta_i = e^{x\beta_2}$	DNB	8	9.3	-202.00	420.01	446.59	8.2	-202	420.01	452.04
	DEWP	8	-0.3	-203.65	423.31	449.89	-0.3	-203.65	423.31	455.34

Note that G: Geometric, P: Poisson, NB: Negative Binomial, D: Destructive,

k is the number of parameter.

Table 3.11: The MLEs and standard errors of the parameters of the destructive negative binomial cure rate model with proportional odds under Weibull lifetimes for the melanoma data.

Parameter	Estimates	S.E.	95% CI
β_0	-2.51	1.00	(-4.46, -0.55)
β_1	1.11	0.58	(-0.03, 2.25)
β_2	4.20	1.21	(1.83 , 6.57)
γ_1	0.34	0.04	(0.26 , 0.42)
γ_2	0.13	0.03	(0.08 , 0.18)
γ_3	-0.21	0.10	(-0.40, -0.02)
γ_4	1.99	1.26	(-0.48, 4.45)

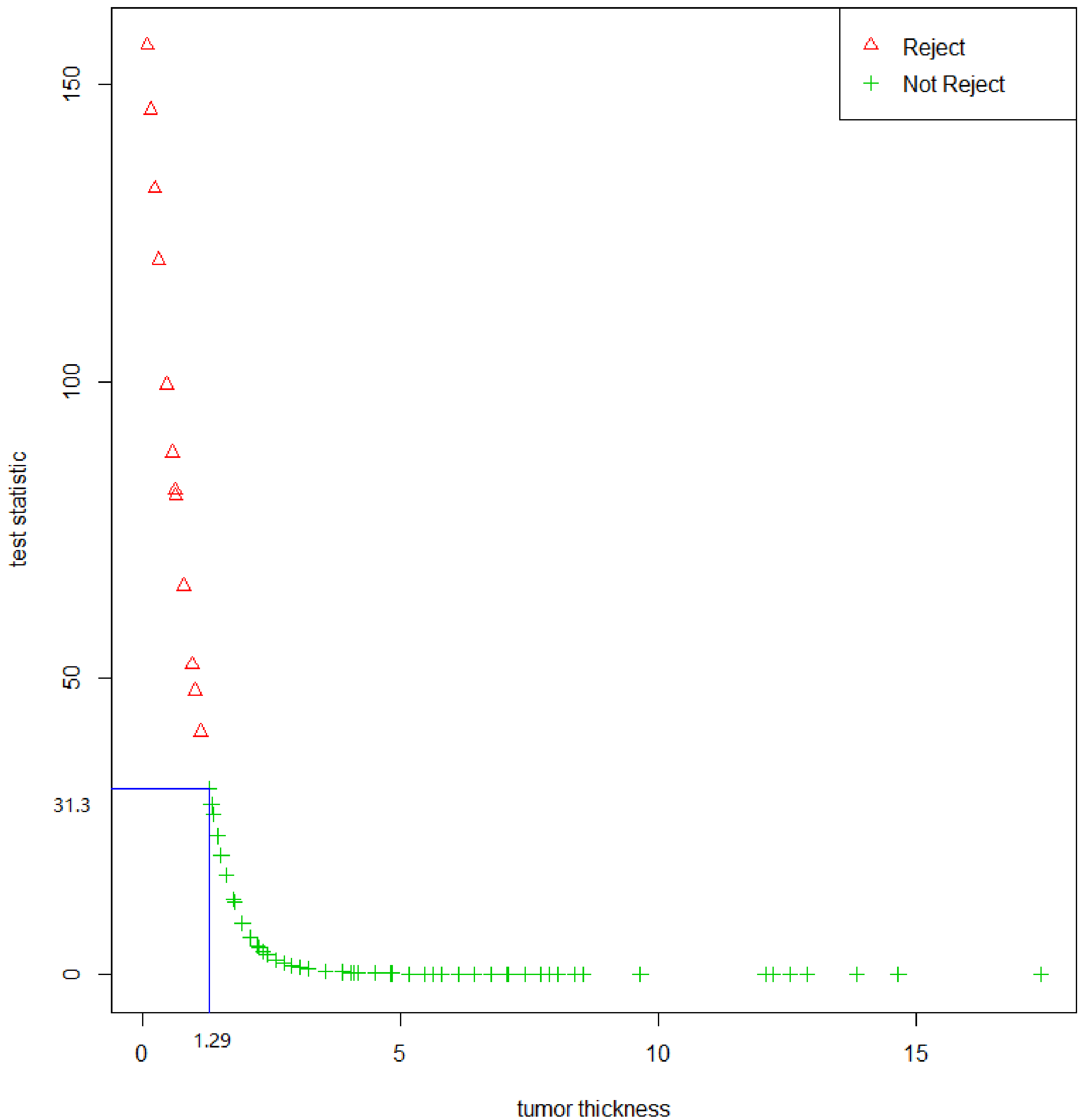

Figure 3.9: $\hat{p}_i = \frac{e^{\hat{\beta}_0+\hat{\beta}_1 x_i}}{1+e^{\hat{\beta}_0+\hat{\beta}_1 x_i}}$. Test statistic=$\frac{(\hat{p}_i-1)^2}{var(\hat{p}_i)}$ for testing $H_0 : p_i = 1$ v.s. $H_a : p_i < 1$ at 5% level of significance.

Chapter 4

Gamma Frailty Cure Rate Model under Proportional Odds Lifetime

4.1 Introduction

In this Chapter, we assume a proportional odds model with frailty term for the distribution of W_j $(j = 1, \dots)$, with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w, \boldsymbol{x}|r) = r\theta O_0(w), \tag{4.1}$$

where $O(w) = S(w)/F(w)$ is the odds of survival up to time w, the proportionality term of θ is linked to covariates as $e^{\boldsymbol{\alpha}'\boldsymbol{x}_c}$ with $\boldsymbol{x}_c = (x_1, \dots, x_p)'$ is a vector of p covariates, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)'$ is the proportional odds regression coefficients, $O_0(w)$ is the baseline odds function, and r is the frailty term following a gamma distribution with shape $k > 0$ and scale $\xi > 0$. The mean and variance of r are k/ξ and k/ξ^2, as mentioned in Section

1.6.

The rest of this Chapter proceeds as follows. Section 4.2 describes the form of the data and the likelihood, while the estimation of the model parameters and associated inferential issues are discussed in Section 4.3. In Section 4.4, an extensive Monte Carlo simulation study is carried out. In Section 4.5, we discuss model discrimination using information- and likelihood-based methods. A data on cutaneous melanoma as well as a simulated data are then analyzed in Section 4.6 for illustrative purpose.

4.2 Data and the likelihood

Here again, we assume the data are subject to non-informative right censoring. Hence, the observation time T_i would be the minimum of the censoring time C_i and the actual lifetime Y_i for the ith subject, i.e.,

$$T_i = \min\{Y_i, C_i\},\ i = 1, \dots, n. \tag{4.2}$$

We define an indicator variable $\delta_i = I(Y_i \leq C_i)$ for the i-th subject such that $\delta_i = 1$ if the lifetime is observed while $\delta_i = 0$ if the lifetime is right censored, Δ_0 and Δ_1 are sets with all the i's equal to 0 and 1, respectively, and set Δ^* contains all the i's. It is to be noted that the cure rate $p_0 = Z(\eta, \phi)^{-1}$ is purely a function of η for a fixed value of ϕ. The range of $1/p_0$ is from 1 to ∞ and it is monotone in η. Therefore, it is natural to use a logistic link function of the form $H_\phi(\eta) = 1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$ to link the covariates $\boldsymbol{x}$ to the cured proportion p_{0i}, i.e.,

$$p_{0i} = p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) = Z(\eta, \phi)^{-1} = H_\phi(\eta)^{-1} = (1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}, \tag{4.3}$$

where p_{0i} is the cured proportion for the ith category, $\boldsymbol{x}_i = (1, \boldsymbol{x}'_{ic})' = (1, x_{i1}, \ldots, x_{ip})'$ is a vector of $p+1$ covariates, and $\boldsymbol{\beta}$ is the vector of regression coefficients. Under this link function, η can be calculated from $H_\phi^{-1}(.)$ analytically for the Geometric, Poisson and Bernoulli distributions, and by using numerical method for the general COM-Poisson distribution.

For *n* pairs of observations ($\boldsymbol{t}, \boldsymbol{\delta}$)=$\{(t_1, \delta_1)$,...,$(t_n, \delta_n)\}$ corresponding to n individuals, the observed data likelihood function under non-informative censoring is then given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \{f_p(t_i; \boldsymbol{\theta})\}^{\delta_i} \{S_p(t_i; \boldsymbol{\theta})\}^{1-\delta_i} \prod_{i \in \Delta^*} f_r(r_i), \tag{4.4}$$

where $\boldsymbol{\theta}$ is the set of parameters $(\phi, \boldsymbol{\beta}', \boldsymbol{z}', \boldsymbol{\gamma}')$, which is equivalent to

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i \in \Delta_1} f_p(t_i; \boldsymbol{\theta}) \prod_{i \in \Delta_0} \{p_0 + (1-p_0) S_s(t_i; \boldsymbol{\theta})\} \prod_{i \in \Delta^*} f_r(r_i). \tag{4.5}$$

Here, we consider two baseline distributions for the proportional odds survival model corresponding to the time-to-event random variable, namely, Weibull and log-logistic distributions. It should also be noted that log-logistic distribution in fact processes the proportional odds property, while the Weibull distribution does not. The survival function and p.d.f. of W under a Weibull baseline, for example, are

$$S(w, \gamma_0, \gamma_1) = [1 + e^{-\boldsymbol{x}'_c \boldsymbol{\alpha}} / r(e^{(\gamma_1 w)^{1/\gamma_0}} - 1)]^{-1}, w > 0, \tag{4.6}$$

$$f(w, , \gamma_0, \gamma_1) = (\gamma_1 w)^{1/\gamma_0} r e^{\boldsymbol{x}'_c \boldsymbol{\alpha} - (\gamma_1 w)^{1/\gamma_0}} [e^{-(\gamma_1 w)^{1/\gamma_0}} (r e^{\boldsymbol{x}'_c \boldsymbol{\alpha}} - 1) + 1]^{-2} / (\gamma_0 w), w > 0, \tag{4.7}$$

where $\gamma_0 > 0$ and $\gamma_1 > 0$ are the shape and scale parameters of the baseline Weibull distribution, respectively. Instead, if we assume the baseline distribution to be a log-logistic distribution with $\gamma_0 > 0$ and $\gamma_1 > 0$ as the scale and shape parameters, respectively, then the corresponding odds function of W_i is given by

$$O(w; \boldsymbol{x}'_c, \boldsymbol{\alpha}|r) = \frac{\gamma_0^{\gamma_1}}{w^{\gamma_1}} r e^{\boldsymbol{x}'_c \boldsymbol{\alpha}} = O_0(w, ; \gamma_0, \gamma_1) r e^{\boldsymbol{x}'_c \boldsymbol{\alpha}}. \tag{4.8}$$

We observe that W_i still follows a two-parameter log-logistic distribution $(\gamma_0, \gamma_1 > 0)$ with shape parameter γ_1 and scale parameter $\gamma_0 e^{-\boldsymbol{x}'_c \boldsymbol{z}/\gamma_1}$, and with corresponding survival function

$$S(w; \boldsymbol{x}'_c \boldsymbol{\alpha}|r) = \frac{r\gamma_0^{\gamma_1} e^{\boldsymbol{x}'_c \boldsymbol{\alpha}}}{r\gamma_0^{\gamma_1} e^{\boldsymbol{x}'_c \boldsymbol{\alpha}} + w^{\gamma_1}}, w > 0. \tag{4.9}$$

Note that the mean does not exist if $\gamma_1 < 1$ and the variance does not exist if $\gamma_1 < 2$.

4.3 Estimation of parameters

In this section, we propose an Expectation-Maximization (EM) algorithm for obtaining the MLEs of $\boldsymbol{\theta}$, and a profile likelihood approach for the estimation of the dispersion patameter ϕ. In our model, the random variable I_i's are observed for i in the set Δ_1 and unobserved for i in the set Δ_0, where $I_i = 1$ if the individual is susceptible and $I_i = 0$ if the individual is cured. Let us denote the set of complete data by

$$(\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{x}, \boldsymbol{I}) = \{(t_1, \delta_1, \boldsymbol{x}_1, I_1), \ldots, (t_n, \delta_n, \boldsymbol{x}_n, I_n)\}.$$

The complete data likelihood is then given by

$$L_c(\boldsymbol{t},\boldsymbol{\delta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y}) = \prod_{i\in\Delta_1}\{(1-p_{0i})f(t_i|r_i)\}\prod_{i\in\Delta_0}\{p_{0i}^{1-I_i}[(1-p_{0i})S(t_i|r_i)]^{I_i}\}\prod_{i\in\Delta^*}f_r(r_i), \tag{4.10}$$

where $\mathbf{I}=(I_1,\ldots,I_n)'$, $\boldsymbol{x}_{ic}=(x_{i1},\ldots,x_{ip})'$ and $\boldsymbol{x}_i=(1,\boldsymbol{x}_{ic}')'$. The corresponding complete log-likelihood function

$$\begin{aligned} l_c(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{x},\boldsymbol{\delta},\boldsymbol{I}) = \text{constant} &+ \sum_{i\in\Delta_1}\text{log}f_p(t_i,\boldsymbol{x}_i,\boldsymbol{\theta}) + \sum_{i\in\Delta_0}(1-I_i)\text{log}p_0(\boldsymbol{\beta},\boldsymbol{x}_i) \\ &+ \sum_{i\in\Delta_0}I_i\text{log}[1-p_0(\boldsymbol{\beta},\boldsymbol{x}_i)] + \sum_{i\in\Delta_0}I_i\text{log}S_s(t_i,\boldsymbol{x}_{ic};\boldsymbol{\theta}) + \sum_{i\in\Delta^*}f_r(r_i). \end{aligned} \tag{4.11}$$

4.3.1 E-step

The expectation step is achieved by calculating the expected value of the complete data log-likelihood function with respect to the conditional distribution of the unobserved I_i's ($i\in\Delta_0$), given the observed data $\boldsymbol{O}=\{(t_i,\delta_i,\boldsymbol{x}_i), i\in\Delta_1\}$ and the current estimates of the parameters $\boldsymbol{\theta}^{(k)}=(\boldsymbol{\beta}',\boldsymbol{\gamma}')'$, for a fixed value of ϕ. Let us denote the function as

$$Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)}) = E(l_c(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{x},\boldsymbol{\delta},\boldsymbol{I})|\boldsymbol{O},\boldsymbol{\theta}^{(k)}) \tag{4.12}$$

at the k-th iteration step. In our model, I_i's are Bernoulli random variables and we can easily find the conditional expectation for the ith individual who is susceptible to be

$$E(K_i(t_i|r_i)|\boldsymbol{O},\boldsymbol{\theta}^{(k)}) = \frac{E(K_i(t_i|r_i)f_p(t_i|Y)^{\delta_i}S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}{E(f_p(t_i|Y)^{\delta_i}S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}, \tag{4.13}$$

$$E(I_iK_i(t_i|r_i)|\boldsymbol{O},\boldsymbol{\theta}^{(k)}) = \frac{(1-p_{0i}(\boldsymbol{x}_i,\boldsymbol{\beta}^{(k)}))E(S_s(t_i|Y)|\boldsymbol{\theta}^{(k)})}{(1-p_{0i}(\boldsymbol{x}_i,\boldsymbol{\beta}^{(k)}))E(K_i(t_i|r_i)S_s(t_i|Y)|\boldsymbol{\theta}^{(k)})+p_{0i}(\boldsymbol{x}_i,\boldsymbol{\beta}^{(k)})}, \tag{4.14}$$

where $K(.)$ is a function of t_i conditional on r_i. Now, for a fixed value of ϕ, the Q function is given by

$$Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)}) = Q_1(\boldsymbol{\beta},\gamma_0,\gamma_1,\boldsymbol{\alpha}) + Q_2(\xi), \tag{4.15}$$

with

$$Q_1(\boldsymbol{\beta},\gamma_0,\gamma_1,\boldsymbol{\alpha}) = \sum_{i\in\Delta_1}E_{3,i} - \sum_{i\in\Delta_1}E_{4,i} + \sum_{i\in\Delta_1}E_{5,i} + \sum_{i\in\Delta_0}E_{6,i} - \sum_{i\in\Delta^*}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{z}_i}), \tag{4.16}$$

$$Q_2(\xi) = \sum_{i\in\Delta_0}n\xi log(\xi) - n\log\Gamma(\xi) + \sum_{i\in\Delta^*}((\xi-1)E_{2i} - \xi E_{1i}), \tag{4.17}$$

where

$$E_{1i} = E(r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),\ \ E_{2i} = E(\log r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),\ \ E_{3i} = E(\log f_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

$$E_{4i} = E(\log S_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),\ \ E_{5i} = E(\log z_{2,i}|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),\ \ E_{6i} = E(I_i\log z_{1,i}|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

and

$$z_{1,i} = z_1(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{\{\eta_i S(t_i|r_i)\}^j}{(j!)^{\phi}}, \quad z_{2,i} = z_2(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{j\{\eta_i S(t_i|r_i)\}^j}{(j!)^{\phi}}.$$

4.3.2 M-step

The M-step is achieved by maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})-$function in (4.15) in order to obtain the improved estimate of $\boldsymbol{\theta}$, i.e.,

$$\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}). \tag{4.18}$$

The MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not have explicit expressions, and so numerical maximization is carried out by using the Newton-Raphson method, for example.

For a fixed value of ϕ, the E-step and M-step are alternated until the parameter estimates converge to a desired level of accuracy. The parameter ϕ is determined by using the profile likelihood technique. We consider a range of ϕ with small increment, and then for each value of ϕ, the MLEs of other parameters are found, and the value of ϕ with the largest likelihood is chosen as the final estimate. The following subsections present explicit forms of the first- and second-order derivatives of the Q function as well as the update function for the case of COM-Poisson distribution, which are necessary for the numerical computation process.

4.3.3 Results for the COM-Poisson cure rate model with gamma frailty

The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^{*}, \boldsymbol{\pi}^{(k)})$ function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows:

$$\frac{\partial Q_1}{\partial \alpha_l} = \sum_{i\in\Delta_1} E_{3,i(\alpha_l)} - \sum_{i\in\Delta_1} E_{4,i(\alpha_l)} + \sum_{i\in\Delta_1} E_{5,i(\alpha_l)} + \sum_{i\in\Delta_0} E_{6,i(\alpha_l)},$$
$$\frac{\partial Q_1}{\partial \gamma_k} = \sum_{i\in\Delta_1} E_{3,i(\gamma_k)} - \sum_{i\in\Delta_1} E_{4,i(\gamma_k)} + \sum_{i\in\Delta_1} E_{5,i(\gamma_k)} + \sum_{i\in\Delta_0} E_{6,i(\gamma_k)},$$
$$\frac{\partial Q_1}{\partial \beta_h} = \sum_{i\in\Delta_1} E_{5,i(\beta_h)} + \sum_{i\in\Delta_0} E_{6,i(\beta_h)} - \sum_{i\in\Delta^*} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}},$$
$$\frac{\partial Q}{\partial \xi} = n\log(\xi) + n - n\psi_0(\xi) + \sum_{i\in\Delta^*}(E_{3i} - E_{2i}),$$
$$\frac{\partial^2 Q_1}{\partial \alpha_l \partial \alpha_{l'}} = \sum_{i\in\Delta_1} E_{3,i(\alpha_l\alpha_{l'})} - \sum_{i\in\Delta_1} E_{4,i(\alpha_l\alpha_{l'})} + \sum_{i\in\Delta_1} E_{5,i(\alpha_l\alpha_{l'})} + \sum_{i\in\Delta_0} E_{6,i(\alpha_l\alpha_{l'})},$$
$$\frac{\partial^2 Q_1}{\partial \alpha_l \gamma_k} = \sum_{i\in\Delta_1} E_{3,i(\alpha_l\gamma_k)} - \sum_{i\in\Delta_1} E_{4,i(\alpha_l\gamma_k)} + \sum_{i\in\Delta_1} E_{5,i(\alpha_l\gamma_k)} + \sum_{i\in\Delta_0} E_{6,i(\alpha_l\gamma_k)},$$
$$\frac{\partial^2 Q_1}{\partial \gamma_k \gamma_{k'}} = \sum_{i\in\Delta_1} E_{3,i(\gamma_k\gamma_{k'})} - \sum_{i\in\Delta_1} E_{4,i(\gamma_k\gamma_{k'})} + \sum_{i\in\Delta_1} E_{5,i(\gamma_k\gamma_{k'})} + \sum_{i\in\Delta_0} E_{6,i(\gamma_k\gamma_{k'})},$$
$$\frac{\partial^2 Q_1}{\partial \alpha_l \partial \beta_h} = \sum_{i\in\Delta_1} E_{5,i(\alpha_l\beta_h)} + \sum_{i\in\Delta_0} E_{6,i(\alpha_l\beta_h))}, \quad \frac{\partial^2 Q_1}{\partial \gamma_k \partial \beta_h} = \sum_{i\in\Delta_1} E_{5,i(\gamma_k\beta_h)} + \sum_{i\in\Delta_0} E_{6,i(\gamma_k\beta_h))},$$
$$\frac{\partial^2 Q_1}{\partial \beta_h \partial \beta_{h'}} = \sum_{i\in\Delta_1} E_{5,i(\beta_h\beta_{h'})} + \sum_{i\in\Delta_0} E_{6,i(\beta_h\beta_{h'})} - \sum_{i\in\Delta^*} \frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2}, \frac{\partial^2 Q}{\partial \xi \partial \xi} = \frac{n}{\xi} - n\psi_1(\xi).$$

The first- and second-order derivatives of the E function with respect to different parameters are as follows:

$$E_{3,i(.)} = M_{1,1}((\log f_i).), \;\; E_{3,i(..)} = M_{1,2}((\log f_i)..), \;\; E_{4,i(.)} = M_{1,1}((\log S_i).),$$
$$E_{4,i(..)} = M_{1,2}((\log S_i)..), \;\; E_{5,i(.)} = M_{1,1}((\log z_{2i}).), \;\; E_{5,i(..)} = M_{1,2}((\log z_{2i})..)$$
$$E_{6,i(.)} = M_{2,1}((\log z_{2i}).), \;\; E_{6,i(..)} = M_{2,2}((\log z_{2i})..).$$

The $M_{..}(.)$ functions are given by

$$M_{1,1}(K(t_i|r_i).) = \frac{E(\frac{\partial K(t_i|r_i)}{\partial .} f_p(t_i|Y)^{\delta_i} S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}{E(f_p(t_i|Y)^{\delta_i} S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}, \tag{4.19}$$

$$M_{1,2}(K(t_i|r_i)..) = \frac{E(\frac{\partial^2 K(t_i|r_i)}{\partial . \partial .} f_p(t_i|Y)^{\delta_i} S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}{E(f_p(t_i|Y)^{\delta_i} S_p(t_i|Y)^{1-\delta_i}|\boldsymbol{\theta}^{(k)})}, \tag{4.20}$$

$$M_{2,1}(K(t_i|r_i).) = \frac{(1-p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)}))E(\frac{\partial \log z_{2i}}{\partial .} S_s(t_i|Y)|\boldsymbol{\theta}^{(k)})}{(1-p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)}))E(S_s(t_i|Y)|\boldsymbol{\theta}^{(k)}) + p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)})}, \tag{4.21}$$

$$M_{2,2}(K(t_i|r_i)..) = \frac{(1-p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)}))E(\frac{\partial^2 \log z_{2i}}{\partial . \partial .} S_s(t_i|Y)|\boldsymbol{\theta}^{(k)})}{(1-p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)}))E(S_s(t_i|Y)|\boldsymbol{\theta}^{(k)}) + p_{0i}(\boldsymbol{x}_i, \boldsymbol{\beta}^{(k)})}, \tag{4.22}$$

where . can be γ_k, α_l or β_h. The first- and second-order derivatives of the $\log f_i$ and $\log S_i$ with respect to γ_k, α_l, β_h under proportional odds model with log-logistic as well as Weibull baseline are presented in the Appendix. The first- and second-order derivatives of the $\log z_{1,i}$ with respect to γ_k, α_l, β_h are as follows:

$$\begin{aligned}
\frac{\partial \log z_{1,i}}{\partial \alpha_l} &= \frac{z_{2,i}}{z_{1,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l}, \quad \frac{\partial \log z_{1,i}}{\partial \beta_h} = \frac{z_{2,i}}{z_{01,i} z_{1,i}} x_{ih} e^{\boldsymbol{\beta}'_i \boldsymbol{x}_i}, \\
\frac{\partial^2 \log z_{1,i}}{\partial \alpha_l \partial \beta_h} &= \frac{x_{ih} e^{\boldsymbol{\beta}' \boldsymbol{x}_i}}{z_{1,i}^2 z_{01,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l} \{z_{21,i} z_{1,i} - z_{2,i}^2\}, \\
\frac{\partial^2 \log z_{1,i}}{\partial \alpha_l \partial \alpha_{l'}} &= \frac{z_{21,i} z_{1,i} - z_{2,i}^2}{z_{1,i}^2} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_{l'}} + \frac{z_{2,i}}{z_{1,i}} \frac{\partial^2 \log S(t_i|r_i)}{\partial \alpha_l \partial \alpha_{l'}}, \\
\frac{\partial^2 \log z_{1,i}}{\partial \beta_h \partial \beta_{h'}} &= \frac{x_{ih} x_{ih'} e^{\boldsymbol{\beta}'_i \boldsymbol{x}_i}}{z_{01,i}^2 z_{1,i}} \left[z_{2,i} z_{01,i} + \left(z_{21,i} - z_{2,i} \frac{z_{02,i}}{z_{01,i}} - \frac{z_{2,i}^2}{z_{1,i}} \right) e^{\boldsymbol{\beta}'_i \boldsymbol{x}_i} \right].
\end{aligned}$$

The first- and second-order derivatives of $\log z_{2,i}$ with respect to γ_k, α_l, β_h are as follows:

$$
\begin{aligned}
&\frac{\partial \log z_{2,i}}{\partial \alpha_l} = \frac{z_{21,i}}{z_{2,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l}, \quad \frac{\partial \log z_{2,i}}{\partial \beta_h} = \frac{z_{21,i} x_{ih} e^{\boldsymbol{\beta}_i' \boldsymbol{x}_i}}{z_{01,i} z_{2,i}}, \quad \frac{\partial \log z_{2,i}}{\partial \gamma_k} = \frac{z_{21,i}}{z_{2,i}} \frac{\partial \log S(t_i|r_i)}{\partial \gamma_k},\\
&\frac{\partial^2 \log z_{2,i}}{\partial \alpha_l \partial \beta_h} = \frac{(z_{31,i} z_{2,i} - z_{21,i}^2) x_{ih} e^{\boldsymbol{\beta}' \boldsymbol{x}_i}}{z_{2,i}^2 z_{01,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l},\\
&\frac{\partial^2 \log z_{2,i}}{\partial \beta_h \partial \beta_{h'}} = \frac{x_{ih} x_{ih'} e^{\boldsymbol{\beta}_i' \boldsymbol{x}_i}}{z_{01,i}^2 z_{2,i}} \left[z_{21,i} z_{01,i} + \left(z_{31,i} - z_{21,i} \frac{z_{02,i}}{z_{01,i}} - \frac{z_{21,i}^2}{z_{2,i}} \right) e^{\boldsymbol{\beta}_i' \boldsymbol{x}_i} \right],\\
&\frac{\partial^2 \log z_{2,i}}{\partial \alpha_l \partial \alpha_{l'}} = \frac{z_{21,i}}{z_{2,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l \alpha_{l'}} + \left(\frac{z_{31,i}}{z_{2,i}} - \frac{z_{21,i}^2}{z_{2,i}^2} \right) \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_{l'}},\\
&\frac{\partial^2 \log z_{2,i}}{\partial \gamma_k \partial \gamma_{k'}} = \frac{z_{21,i}}{z_{2,i}} \frac{\partial \log S(t_i|r_i)}{\partial \gamma_k \gamma_{k'}} + \left(\frac{z_{31,i}}{z_{2,i}} - \frac{z_{21,i}^2}{z_{2,i}^2} \right) \frac{\partial \log S(t_i|r_i)}{\partial \gamma_k} \frac{\partial \log S(t_i|r_i)}{\partial \gamma_{k'}},\\
&\frac{\partial^2 \log z_{2,i}}{\partial \alpha_l \gamma_k} = \frac{z_{21,i}}{z_{2,i}} \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l \partial \gamma_k} + \left(\frac{z_{31,i}}{z_{2,i}} - \frac{z_{21,i}^2}{z_{2,i}^2} \right) \frac{\partial \log S(t_i|r_i)}{\partial \alpha_l} \frac{\partial \log S(t_i|r_i)}{\partial \gamma_k},
\end{aligned}
$$

where

$$
\begin{aligned}
z_{1,i} &= z_1(t_i) = \sum_{=1}^{\infty} \frac{\{\eta_i S(t_i|r_i)\}^j}{(j!)^\phi}, \quad z_{2,i} = z_2(t_i) = \sum_{j=1}^{\infty} \frac{j\{\eta_i S(t_i|r_i)\}^j}{(j!)^\phi},\\
z_{01,i} &= z_{01}(t_i) = \sum_{j=1}^{\infty} \frac{j\eta_i^j}{(j!)^\phi}, \quad z_{21,i} = z_{21}(t_i) = \sum_{j=1}^{\infty} \frac{j^2\{\eta_i S(t_i|r_i)\}^j}{(j!)^\phi},\\
z_{31,i} &= z_{31}(t_i) = \sum_{j=1}^{\infty} \frac{j^3\{\eta_i S(t_i|r_i)\}^j}{(j!)^\phi}, \quad z_{02,i} = z_{02}(t_i) = \sum_{j=1}^{\infty} \frac{j^2\eta_i^j}{(j!)^\phi}.
\end{aligned}
$$

4.4 Simulation study

In this section, an extensive Monte Carlo simulation study is carried out for the special cases to illustrate the performance of the proposed model and the method of inference. We vary the sample size, censoring proportion, and underling baseline distribution to consider different scenarios. We mimic the cutaneous melanoma data, and consider 4 possible categories for the individuals, namely, $x = 0, 1, 2, 3$. Two different sample sizes were considered in the study: $n = 800$ $(200, 168, 212, 220)$ and $n = 2000$ $(500, 420, 530, 550)$ to reflect medium and large sample sizes. Moreover, if we assume that $\boldsymbol{\beta} = (\beta_0, \beta_1)$ has two parameters, fixing the cure rates for the first and fourth categories would be enough to cover all cases as the cure rates for the second and third categories can then be obtained from $\boldsymbol{\beta}$. Here, we took $(p_{00}, p_{03}) = (0.4, 0.2)$ with respect to categories one and four for the cure rate. Also, the cure rate would be in a decreasing order in this way. The $\boldsymbol{\beta}$'s turn out to be

$$\beta_0 = \ln(1/p_{00} - 1) \ , \ \ \beta_1 = (\ln(1/p_{03} - 1) - \beta_0)/3. \tag{4.23}$$

We thus obtain the true value of $\boldsymbol{\beta}$ as (0.405,0.321). In addition, we consider light and heavy censoring cases in the simulation. The light and heavy censoring rates are (0.52, 0.45, 0.37, 0.3) and (0.65, 0.49, 0.4, 0.35) for the low cure rates, (0.7, 0.57, 0.45, 0.34) and (0.8, 0.64, 0.5, 0.38) for the high cure rates, respectively. Suppose the probability of getting censored and cured for group x are c_x and p_{0x}, respectively. It is natural to consider the proportion of censored individuals for the susceptible group to be equal to the difference between the probability of getting censored and cured; i.e.,

$$P(Y \geq C_x \cap M \geq 1 | X = x) = c_x - p_{0x}, \tag{4.24}$$

where the censoring time C_x follows an exponential distribution with rate λ_x on $x = 0, 1, 2, 3$. The choice of (ξ, γ_0) in the underlying distribution of the proportional odds survival model were taken to be $(0.571, 0.307)$ and $(1.75, 3.25)$ for Weibull and log-logistic distributions, respectively. The odds parameter was specified as $\gamma_1 = -0.75$ to ensure a decreasing lifetime for the four nodule categories. We consider an inverse transform sampling method to simulate the actual survival lifetime w_i for each individual under different competing risks, i.e.,

$$w_i = \frac{1}{\gamma_1}\left[\log\left(1+\left(\frac{1}{u}-1\right)r_i e^{x_i\alpha}\right)\right]^{\gamma_0}, \tag{4.25}$$

$$w_i = \gamma_0\left(\frac{u}{1-u}r_i e^{x_i\alpha}\right)^{1/\gamma_1}, i = 1,\ldots,n, \tag{4.26}$$

under the proportional odds model with Weibull and log-logistic baseline distributions, respectively, where u follows an uniform distribution over 0 to 1.

Under the above setting, the procedure to generate the data from different cure rate models is as follows.

Geometric cure rate model: For each individual, we simulate the number of competing risks M_i from Geometric distributions with probability $P(M_i = 0) = p_{0x}$; and we simulate the censoring time C_x from exponential distribution with rate λ_x. If M_i does not equal zero, we simulate M_i number of actual lifetimes $\{Y_{i1},..,Y_{iM_i}\}$ from proportional odds survival model, the actual lifetime is defined as $Y_i = \min\{Y_{i1},..,Y_{iM_i}\}$ and the observed lifetime T_i is taken as the minimum of all the actual lifetimes and the censoring time, i.e., $T_i = \min\{Y_i, C_i\}$. If $Y_i > C_i$, we make the censoring indicator $\delta_i = 0$, otherwise $\delta_i = 1$. On the other hand, $M_i = 0$ means the individual is cured, and so we assign C_i to be the lifetime, and the censoring indicator is taken to be $\delta_i = 0$.

Poisson cure rate model: In this case, the procedure is the same as the Geometric cure rate model except that M_i is simulated from Poisson distribution with parameter $-\log(p_{0x})$.

Bernoulli cure rate model: There are two ways to do the data generation in this case. One is the same as Geometric cure rate model except that M_i is simulated from Bernoulli distribution with probability of success as $1 - p_{0x}$. Another way is a little bit simpler since M_i can only take 0 or 1 in this case. For each individual, we simulate the censoring time C_x from exponential distribution with rate λ_x. Then, we simulate an uniform random variable U_i and if $U_i \le p_{0x}$, the observed lifetime T_i is set to C_x; otherwise, we generate the observed lifetime T_i from the proportional odds survival model.

In our simulation study, 1000 Monte Carlo runs were considered in each scenario. The estimates were calculated through the EM method. We stopped the iteration in estimation when the difference in the log-likelihood values between two consecutive estimates was less than 10^{-7}. We calculated the empirical Bias, standard error (SE), root Mean Square Error (RMSE), and 95% coverage probabilities (CPs) for all the parameters. Here, the initial values of the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ were taken from a grid of parameter values, and those values having the maximum likelihood were then chosen as the initial value.

Tables 4.1 - 4.9 present the bias, SE, RMSE, and CPs for the three special cases. We can see that the estimates are quite accurate under different cure rate models with gamma frailty. The bias, SE, along with RMSE get reduced for low censoring. The coverage probabilities of the confidence intervals based on the asymptotic normality of the MLEs are quite close to the nominal level in most of the cases except for $1/\xi$. This becomes better when the sample size n increases.

Table 4.1: True values of parameters, Bias, SE, RMSE and CP for the Geometric cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.006	0.108	0.104	95.689	-0.75	0.009	0.122	0.12	95.39
	γ_0	1.25	-0.003	0.065	0.065	93.653	1.25	-0.005	0.079	0.08	93.388
	γ_1	3.75	-0.022	0.189	0.148	96.527	3.75	-0.023	0.202	0.158	97.045
	β_1	0.405	0.003	0.146	0.145	95.329	0.405	0.005	0.175	0.17	95.868
	β_2	0.327	0.004	0.08	0.081	94.371	0.327	0.002	0.092	0.088	96.222
	$1/\xi$	0.152	-0.061	0.085	0.082	54.371	0.152	-0.06	0.086	0.082	55.083

Table 4.2: True values of parameters, Bias, SE, RMSE and CP for the Poisson cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.006	0.085	0.086	94.97	-0.75	0.01	0.094	0.099	94.097
	γ_0	1.25	-0.005	0.054	0.054	94.611	1.25	-0.006	0.064	0.065	93.388
	γ_1	3.75	-0.025	0.151	0.156	92.575	3.75	-0.021	0.161	0.164	93.388
	β_1	0.405	0.005	0.145	0.143	95.329	0.405	-0.001	0.174	0.167	95.632
	β_2	0.327	0.001	0.079	0.079	94.85	0.327	0.004	0.091	0.09	95.396
	$1/\xi$	0.152	-0.06	0.053	0.08	53.653	0.152	-0.052	0.056	0.075	61.039

Table 4.3: True values of parameters, Bias, SE, RMSE and CP for the Bernoulli cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.004	0.079	0.08	94.85	-0.75	0.007	0.085	0.088	94.097
	γ_0	1.25	-0.006	0.052	0.052	94.97	1.25	-0.01	0.058	0.06	94.215
	γ_1	3.75	-0.028	0.154	0.147	95.569	3.75	-0.015	0.165	0.163	95.277
	β_1	0.405	-0.005	0.146	0.146	94.731	0.405	0.009	0.175	0.184	94.097
	β_2	0.327	0.004	0.08	0.08	95.808	0.327	-0.001	0.092	0.097	93.743
	ξ	6.6	9.696	11.836	16.138	100	6.6	9.81	12.23	17.045	100
	$1/\xi$	0.152	-0.054	0.063	0.075	58.922	0.152	-0.051	0.065	0.074	60.803

Table 4.4: True values of parameters, Bias, SE, RMSE and CP for the Geometric cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.004	0.069	0.068	94.897	-0.75	0.003	0.079	0.075	97.054
	γ_0	1.25	-0.002	0.041	0.043	93.317	1.25	-0.003	0.05	0.049	94.993
	γ_1	3.75	-0.019	0.124	0.1	96.719	3.75	-0.002	0.143	0.103	98.38
	β_1	0.405	0.002	0.092	0.092	95.018	0.405	-0.002	0.11	0.111	94.845
	β_2	0.327	0.001	0.05	0.051	95.261	0.327	0.002	0.058	0.057	95.582
	$1/\xi$	0.152	-0.04	0.058	0.061	66.1	0.152	-0.019	0.069	0.04	83.652

Table 4.5: True values of parameters, Bias, SE, RMSE and CP for the Poisson cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.005	0.054	0.056	93.816	-0.75	0.002	0.06	0.056	96.159
	γ_0	1.25	-0.004	0.034	0.036	94.079	1.25	0	0.041	0.039	96.302
	γ_1	3.75	-0.02	0.096	0.103	92.763	3.75	-0.015	0.102	0.104	94.168
	β_1	0.405	0	0.092	0.095	94.605	0.405	0.005	0.109	0.104	96.302
	β_2	0.327	0.003	0.05	0.049	95.789	0.327	-0.001	0.057	0.055	95.875
	$1/\xi$	0.152	-0.035	0.039	0.055	69.868	0.152	-0.022	0.042	0.041	82.077

Table 4.6: True values of parameters, Bias, SE, RMSE and CP for the Bernoulli cure rate model with gamma frailty under proportional odds with log-logistic baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.005	0.05	0.05	94.4	-0.75	0.005	0.054	0.055	95.395
	γ_0	1.25	-0.006	0.033	0.034	92.7	1.25	-0.006	0.037	0.038	94.194
	γ_1	3.75	-0.029	0.097	0.101	93.4	3.75	-0.02	0.105	0.102	94.394
	β_1	0.405	0	0.092	0.093	95	0.405	-0.004	0.11	0.107	95.596
	β_2	0.327	0.003	0.05	0.05	95.4	0.327	0.001	0.058	0.057	94.795
	$1/\xi$	0.152	-0.044	0.041	0.063	60.8	0.152	-0.034	0.045	0.056	70.47

Table 4.7: True values of parameters, Bias, SE, RMSE and CP for the Geometric cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.006	0.119	0.109	96.134	-0.75	0.008	0.142	0.13	96.482
	γ_0	0.571	0.003	0.035	0.025	97.597	0.571	0.003	0.039	0.026	98.593
	γ_1	0.307	0.003	0.027	0.027	96.029	0.307	0.005	0.04	0.04	94.684
	β_1	0.405	-0.003	0.149	0.143	95.716	0.405	0.002	0.2	0.2	95.888
	β_2	0.327	0.005	0.083	0.081	95.82	0.327	0.004	0.102	0.1	95.884
	$1/\xi$	0.152	-0.039	0.107	0.062	71.578	0.152	-0.035	0.114	0.058	75.879

Table 4.8: True values of parameters, Bias, SE, RMSE and CP for the Poisson cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.004	0.085	0.082	95.929	-0.75	0	0.098	0.097	95.391
	γ_0	0.571	0.005	0.024	0.024	94.781	0.571	0.003	0.026	0.027	94.589
	γ_1	0.307	0.003	0.021	0.021	94.572	0.307	0.003	0.028	0.029	94.289
	β_1	0.405	-0.001	0.149	0.147	95.929	0.405	0.002	0.19	0.195	94.689
	β_2	0.327	0	0.082	0.083	94.05	0.327	0.003	0.098	0.098	95.09
	$1/\xi$	0.152	-0.042	0.058	0.062	63.779	0.152	-0.036	0.062	0.056	71.643

Table 4.9: True values of parameters, Bias, SE, RMSE and CP for the Bernoulli cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
800	α	-0.75	0.005	0.076	0.075	95.687	-0.75	0	0.084	0.084	94.4
	γ_0	0.571	0.005	0.026	0.025	95.587	0.571	0.003	0.028	0.027	95.2
	γ_1	0.307	0.003	0.017	0.017	95.186	0.307	0.003	0.021	0.022	95
	β_1	0.405	0.006	0.148	0.155	94.283	0.405	0.005	0.185	0.183	95.6
	β_2	0.327	0.002	0.082	0.083	93.882	0.327	0.001	0.097	0.099	94.4
	$1/\xi$	0.152	-0.042	0.069	0.064	69.007	0.152	-0.036	0.072	0.059	74.6

Table 4.10: True values of parameters, Bias, SE, RMSE and CP for the Geometric cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.003	0.115	0.092	97.355	-0.75	0.001	0.135	0.1	92.026
	γ_0	0.571	0.001	0.027	0.02	98.105	0.571	0.002	0.031	0.02	96.229
	γ_1	0.307	0.003	0.028	0.022	98.083	0.307	0.002	0.042	0.03	93.827
	β_1	0.405	-0.004	0.15	0.122	97.679	0.405	0.01	0.198	0.151	91.918
	β_2	0.327	0.004	0.083	0.07	97.376	0.327	-0.002	0.099	0.077	91.866
	$1/\xi$	0.152	-0.026	0.049	0.05	78.312	0.152	-0.02	0.055	0.042	83.162

Table 4.11: True values of parameters, Bias, SE, RMSE and CP for the Poisson cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.001	0.054	0.054	94.751	-0.75	-0.001	0.062	0.062	94.954
	γ_0	0.571	0.001	0.015	0.015	95.304	0.571	0	0.016	0.017	93.493
	γ_1	0.307	0.001	0.013	0.013	94.475	0.307	0	0.017	0.017	95.485
	β_1	0.405	-0.001	0.094	0.094	95.028	0.405	0.003	0.119	0.117	94.29
	β_2	0.327	0.001	0.052	0.052	94.061	0.327	-0.001	0.061	0.062	93.094
	$1/\xi$	0.152	-0.008	0.034	0.022	92.403	0.152	-0.007	0.034	0.022	92.961

Table 4.12: True values of parameters, Bias, SE, RMSE and CP for the Bernoulli cure rate model with gamma frailty under proportional odds with Weibull baseline with light censoring (LC) and heavy censoring (HC).

		LC					HC				
n	Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
2000	α	-0.75	0.001	0.048	0.048	95.286	-0.75	0	0.053	0.052	95
	γ_0	0.571	0.003	0.016	0.016	95.286	0.571	0.001	0.017	0.016	96.2
	γ_1	0.307	0.001	0.011	0.011	93.882	0.307	0.002	0.013	0.013	94.3
	β_1	0.405	0.002	0.093	0.092	94.885	0.405	0.007	0.116	0.115	94.8
	β_2	0.327	0.001	0.052	0.05	95.085	0.327	-0.002	0.061	0.061	94.8
	$1/\xi$	0.152	-0.024	0.036	0.044	77.332	0.152	-0.02	0.037	0.039	81.8

4.5 Data illustration

In this section, a clinical data as well as a simulated data are used for illustrating the model and the inferential methods developed in the preceding sections.

4.5.1 Cutaneous melanoma data

Let us now consider the cutaneous melanoma data described earlier in Section 1.9 and also analyzed in Chapter 2. In this case, the subjects were divided into four different categories ($x = 0, 1, 2, 3$), with corresponding sample sizes $n_1 = 111$, $n_2 = 137$, $n_3 = 87$, $n_4 = 82$. The percentage of censored observations for the groups are 67.57%, 61.31%, 52.87%, 32.93%. See Figure 1.2 for a plot of the lifetimes of susceptibles. We fitted the data by Geometric ($\phi = 0$), COM-Poisson ($\phi = 0.5$), Poisson ($\phi = 1$), COM-Poisson ($\phi = 2$) and Bernoulli ($\phi \approx \infty$) cure rate model under proportional odds models with gamma frailty with Weibull and log-logistic baseline distributions.

Tables 4.13 and 4.14 represent the MLEs and SEs of the parameters under different cure rate models with gamma frailty. Here, we chose the convergence criterion to be the absolute difference between two consecutive likelihood values to be less than 10^{-7}. Also 40,000 sample points were simulated for the Monte Carlo integration in the expectation step. In order to check the accuracy of the above used criteria, we generated 100,000 sample points for Monte Carlo integration for Bernoulli cure rate model with gamma frailty under proportional odds model for Weibull baseline, and set convergence criteria to be the absolute difference between two consecutive likelihood values to be less than 10^{-8}. The estimates obtained were -0.5221, 0.4663, 0.2847, -0.9557, 0.3975, 0.1166 and their corresponding standard errors were 0.1224, 0.0475, 0.0383, 0.2921, 0.1063, 2.2382, for the parameters $\alpha, \gamma_0, \gamma_1, \beta_0, \beta_1, 1/\xi$, which are quite close to the values reported in Table 4.13 and Table 4.14.

Table 4.13: MLEs of the model parameters for different PO cure models with gamma frailty

Par	PO frailty Weibull baseline					PO frailty Log-lositic baseline				
ϕ	0	0.5	1	2	∞	0	0.5	1	2	∞
α	-0.3271	-0.3897	-0.4224	-0.4622	-0.5250	-0.2877	-0.3591	-0.3903	-0.4268	-0.4882
γ_0	0.4542	0.4616	0.4631	0.4637	0.4636	3.5800	3.4719	3.3720	3.2505	3.1431
γ_1	0.2591	0.2660	0.2713	0.2778	0.2835	2.3258	2.3033	2.3058	2.3155	2.3284
β_0	-0.8749	-0.8990	-0.9111	-0.9273	-0.9543	-0.7693	-0.7682	-0.7707	-0.7759	-0.7851
β_1	0.3717	0.3793	0.3829	0.3879	0.3972	0.3713	0.3699	0.3709	0.3732	0.3767
$1/\xi$	0.1483	0.1393	0.1390	0.1400	0.1405	0.1509	0.1487	0.1484	0.1479	0.1479

Table 4.14: SEs of the model parameters for different PO cure models with gamma frailty

Par	PO frailty Weibull baseline					PO frailty Log-lositic baseline				
ϕ	0	0.5	1	2	∞	0	0.5	1	2	∞
α	0.1433	0.1326	0.1293	0.1267	0.1237	0.1642	0.1482	0.1345	0.1399	0.1642
γ_0	0.0481	0.0482	0.0484	0.0487	0.0486	0.7984	0.7252	0.5891	0.6371	0.7984
γ_1	0.0430	0.0412	0.0404	0.0396	0.0386	0.2426	0.2412	0.2404	0.2407	0.2426
β_0	0.2996	0.2978	0.2964	0.2948	0.2925	0.3212	0.3204	0.3227	0.3217	0.3212
β_1	0.1056	0.1061	0.1061	0.1062	0.1064	0.1149	0.1141	0.1154	0.1148	0.1149
$1/\xi$	0.2960	0.2771	0.2757	0.2769	0.2772	0.3187	0.3115	0.3139	0.3111	0.3064

Table 4.15 lists the AIC, BIC and $\hat{l}$ for proportional odds cure rate model under Weibull and log-logistic baseline distributions with and without frailty term. The proportional odds model under Weibull baseline with frailty term has a higher $\hat{l}$. But, the increase in log-likelihood is not significant enough for AIC or BIC to select this model with frailty over the model without frailty.

Table 4.15: AIC, BIC and $\hat{l}$ for different models for the melanoma data

		Geometric		Poisson		Bernoulli
	ϕ	0	0.5	1	2	∞
PO	AIC	**1026.8062**	1026.8642	1027.0028	1027.1876	1027.3077
frailty CRM	BIC	**1051.1469**	1051.2049	1051.3435	1051.5283	1051.6484
Weibull	$\hat{l}$	**-507.4031**	-507.4321	-507.5014	-507.5938	-507.6538
PO	AIC	1025.0739	1025.0517	1025.0276	1024.9765	**1024.8932**
frailty CRM	BIC	1049.4146	1049.3924	1049.3683	1049.3172	**1049.2339**
log-logistic	$\hat{l}$	-506.5369	-506.5259	-506.5138	-506.4882	**-506.4466**
PO	AIC	**1025.644**	1026.014	1026.374	1026.862	1027.414
CRM	BIC	**1045.809**	1046.179	1046.539	1047.027	1047.579
Weibull	$\hat{l}$	**-507.822**	-508.007	-508.187	-508.431	-508.707
PO	AIC	1022.863	1022.845	1022.821	1022.768	**1022.683**
CRM	BIC	1043.029	1043.011	1042.986	1042.933	**1042.849**
log-logistic	$\hat{l}$	-506.432	-506.423	-506.41	-506.384	**-506.342**

The proportional odds model under log-logistic baseline with frailty term has a lower $\hat{l}$. The MLEs and SEs of the cure rate model with gamma frailty under proportional odds model for log-logistic baseline is presented in Table 4.16. The frailty parameter ξ becomes very large, and the proportional odds frailty cure rate model tends to proportional odds assumptions with ordinary cure rate model, and so the likelihood for the model with and without frailty term turn out to be the same; see Table 4.17 for a list of $\hat{l}$, AIC and BICs values.

Table 4.16: MLEs and SEs of the parameters in CRM with gamma frailty under PO with log-logistic baseline (convergence criteria: two consecutive likelihood values to be less than $< 10^{-10}$)

	MLEs					SEs				
ϕ	0	0.5	1	2	∞	0	0.5	1	2	∞
α	-0.271	-0.343	-0.374	-0.411	-0.473	0.059478	0.138998	0.134526	0.130778	0.135997
γ_0	3.460	3.355	3.256	3.137	3.033	0.121427	0.651613	0.609339	0.562915	0.301182
γ_1	2.262	2.240	2.242	2.253	2.266	0.026793	0.187205	0.187253	0.187609	0.186913
β_0	-0.775	-0.774	-0.776	-0.782	-0.791	0.223368	0.31806	0.318435	0.319483	0.236902
β_1	0.373	0.372	0.373	0.375	0.379	0.091576	0.113683	0.113849	0.114416	0.090989
$1/\xi$	0.0003	0.0003	0.0003	0.0003	0.0003	5.87E-16	8.73E-22	7.85E-22	6.96E-22	5.06E-22

Table 4.17: AIC, BIC and $\hat{l}$ for CRM with gamma frailty under PO with log-logistic baseline (convergence criteria: two consecutive likelihood values to be less than $< 10^{-10}$)

		Geometric	Poisson			Bernoulli
	ϕ	0	0.5	1	2	∞
PO	AIC	1024.8640	1024.8456	1024.8212	1024.7684	1024.6836
frailty CRM	BIC	1049.2047	1049.1863	1049.1619	1049.1091	1049.0243
log-logistic	l	-506.432	-506.423	-506.411	-506.384	**-506.342**
PO	AIC	1022.863	1022.845	1022.821	1022.768	**1022.683**
CRM	BIC	1043.029	1043.011	1042.986	1042.933	**1042.849**
log-logistic	$\hat{l}$	-506.432	-506.423	-506.41	-506.384	**-506.342**

4.5.2 Simulated dataset and analysis

Based on the setting of the cutaneous melanoma data, we simulated from proportional odds cure rate model with gamma frailty having log-logistic and Weibull baseline distributions, respectively. Two simulated datasets were obtained by specifying ϕ as ∞ (Bernoulli) for Weibull baseline and 1 (Poisson) for log-logistic baseline. Here, we chose the cure rates for the first and fourth groups to be $p_{01} = 0.4$, $p_{04} = 0.2$; the censoring proportions for the four groups to be $(0.520.450.370.30)$, $\xi = 0.5$, $\alpha = -0.75$. $n = 800$ $(200, 168, 212, 220)$, and (γ_0, γ_1) to be $(0.571, 0.307)$ and $(1.25, 3.75)$ for Weibull and log-logistic baseline, receptively. The simulated datasets for censored and uncensored individuals are plotted in Figures 4.1 and 4.2. The corresponding datasets can be downloaded through the links https://www.dropbox.com/s/wg47izgk9qt15pr/simudata1.csv?dl=0 and https://www.dropbox.com/s/ohefxa4l4jh9vtj/simudata2.csv?dl=0 , respectively.

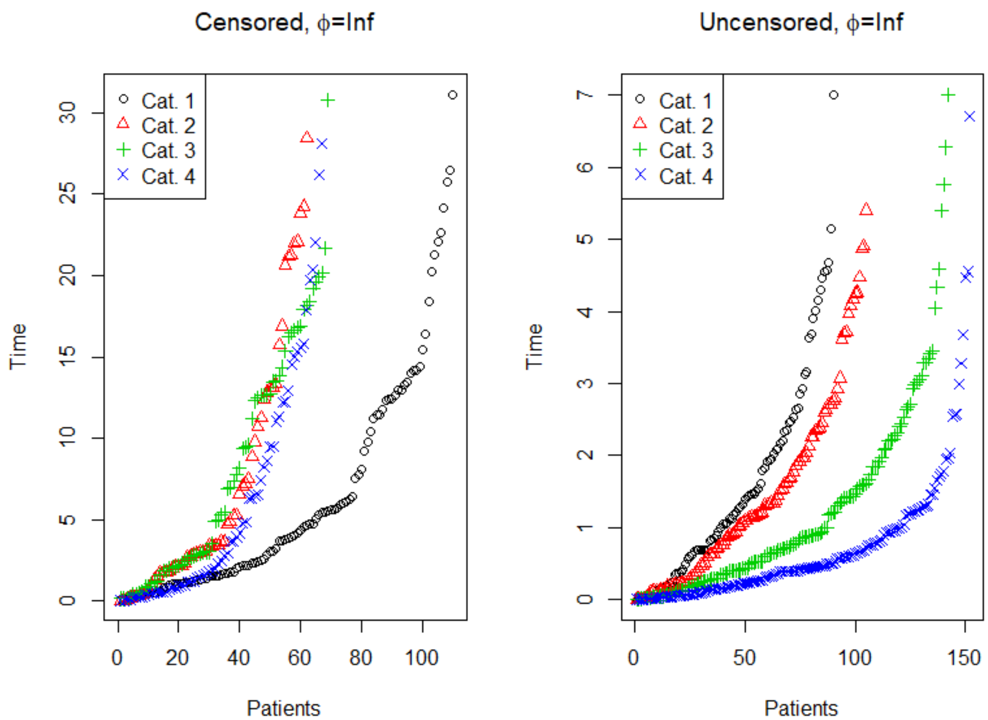

Figure 4.1: Simulated datasets based on FCRM under PO with Weibull baseline

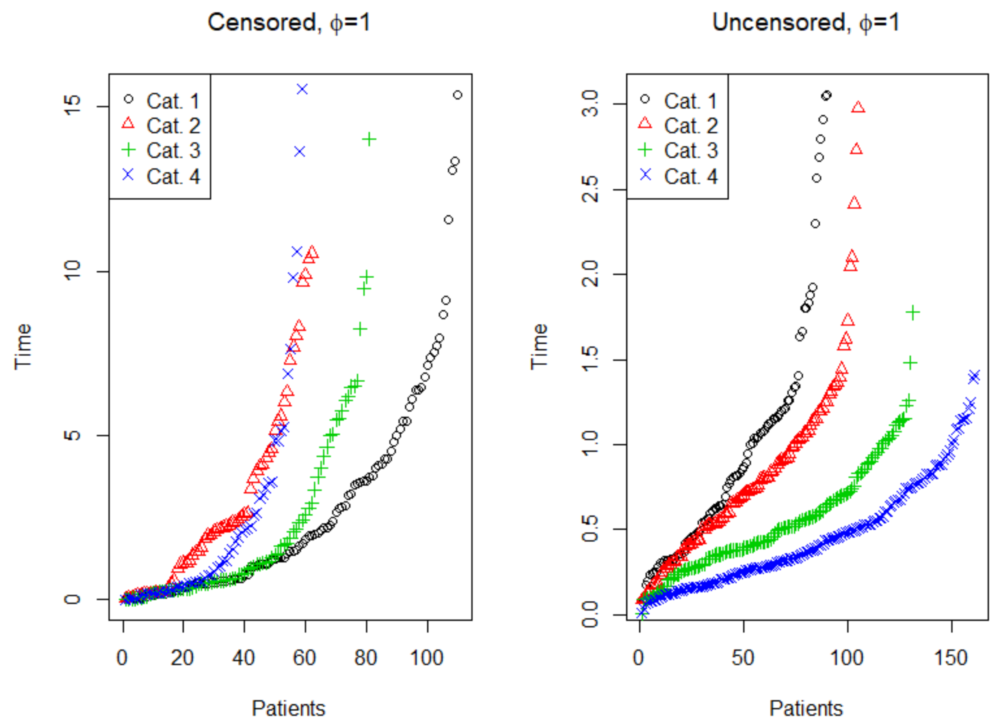

Figure 4.2: Simulated datasets based on FCRM under PO with log-logistic baseline

To compare the results between cure rate models with and without frailty term, we then fitted the simulated data by cure rate models under proportional odds model with and without frailty terms.

Table 4.18 presents the MLEs and SEs of the cure rate model parameters. The estimates are quite close to the true parameter values. Table 4.19 compares the AIC, BIC and maximized log-likelihood value $\hat{l}$ for different models. These results show that the cure rate model with frailty term always has larger $\hat{l}$ and smaller AIC and BIC, which suggests that the frailty cure rate model provides a better fit than the cure rate model without frailty when the true model contains frailty term.

Table 4.18: MLEs and SEs of the frailty cure rate model parameters

With frailty	PO Weibull baseline		PO log-logistic baseline			
	MLEs	SE	MLES	SE	MLES	SE
ϕ	∞	-	0.5	-	1	-
α	-0.7033	0.2601	-0.8026	0.1539	-0.8434	0.1519
γ_0	0.578	0.1645	1.4274	0.1041	1.3634	0.0986
γ_1	0.3248	0.0623	3.6396	0.3806	3.6542	0.3837
β_1	0.4347	0.1562	0.463	0.1546	0.4629	0.1545
β_2	0.3016	0.0849	0.3864	0.0866	0.3864	0.0866
$1/\xi$	1.9816	0.9981	1.9444	0.407	1.9543	0.4103

Table 4.19: $\hat{l}$, AIC, and BIC for different models

	ϕ	With frailty			Without frailty		
		$\hat{l}$	AIC	BIC	$\hat{l}$	AIC	BIC
PO Weibull baseline	0	-973.041	1958.081	1986.189	-1104.6	2219.192	2242.615
	0.5	-972.3	1956.599	1984.707	-985.998	1981.997	2005.42
	1	-972.437	1956.874	1984.982	-984.349	1978.697	2002.12
	2	-972.701	1957.403	1985.511	-982.614	1975.227	1998.65
	∞	**-972.213**	1956.425	1984.533	-980.446	1970.891	1994.314
PO log-logistic baseline	0	-578.338	1168.677	1196.784	-585.12	1180.241	1203.664
	0.5	**-578.325**	1168.65	1196.758	-584.727	1179.453	1202.876
	1	-578.328	1168.655	1196.763	-584.765	1179.529	1202.952
	2	-578.352	1168.705	1196.813	-585.001	1180.002	1203.425
	∞	-578.337	1168.674	1196.782	-585.229	1180.457	1203.88

One important indicator is the probability an individual is cured conditional on that individual having survived up to a specific time t, i.e., $P(I = 0|T > t)$. The estimate of this probability is given by

$$\hat{P}(I=0|T>t) = \frac{p_0(1-\int_0^\infty S(t|r)f(r)dr)}{(1-p_0)\int_0^\infty S(t|r)f(r)dr) + p_0}, \tag{4.27}$$

where $S(t|y)$ is the survival function under proportional odds model which is

$$S(t|r) = \frac{rS_0(t)e^{\boldsymbol{\alpha}'\boldsymbol{x}}}{rS_0(t)e^{\boldsymbol{\alpha}'\boldsymbol{x}} + F_0(t)}. \tag{4.28}$$

The corresponding plots are presented in Figures 4.3 and 4.4.

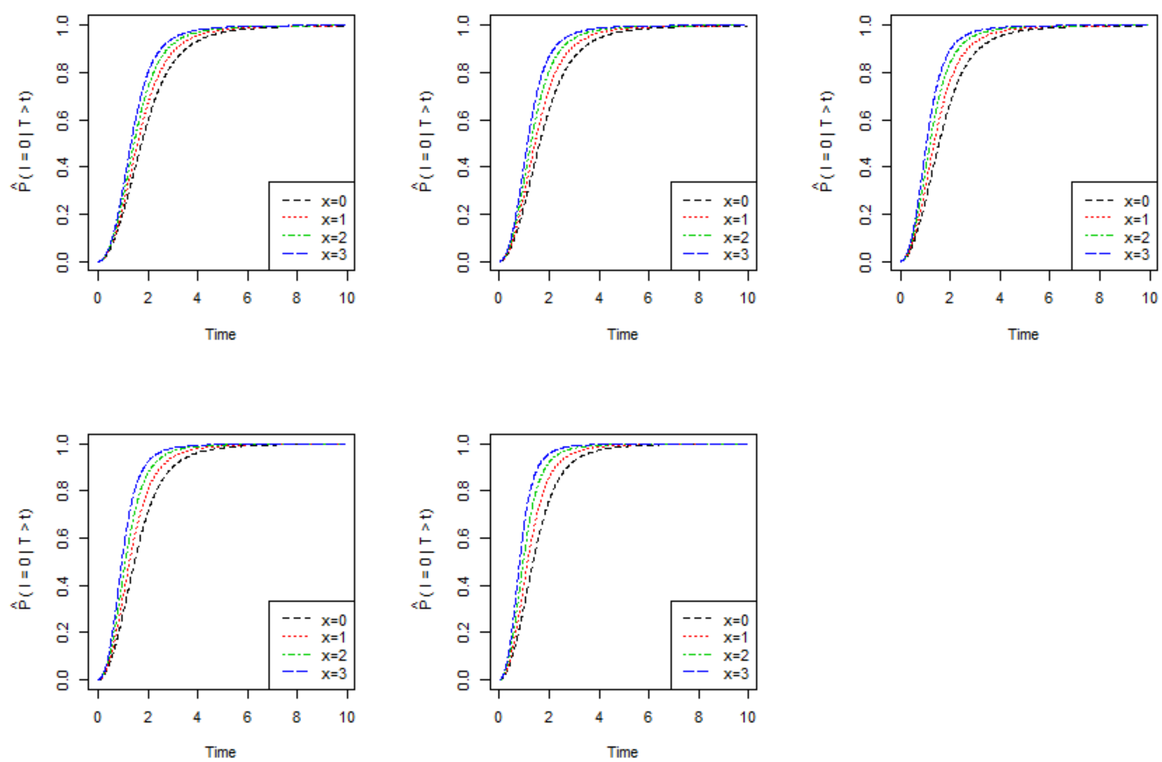

Figure 4.3: Cure rate given an individual has survived up to a specific time t over four covariate groups (From the left to right, the models are proportional odds cure rate model with log-logistic baseline $\phi = (0, 0.5, 1, 2, \infty)$).

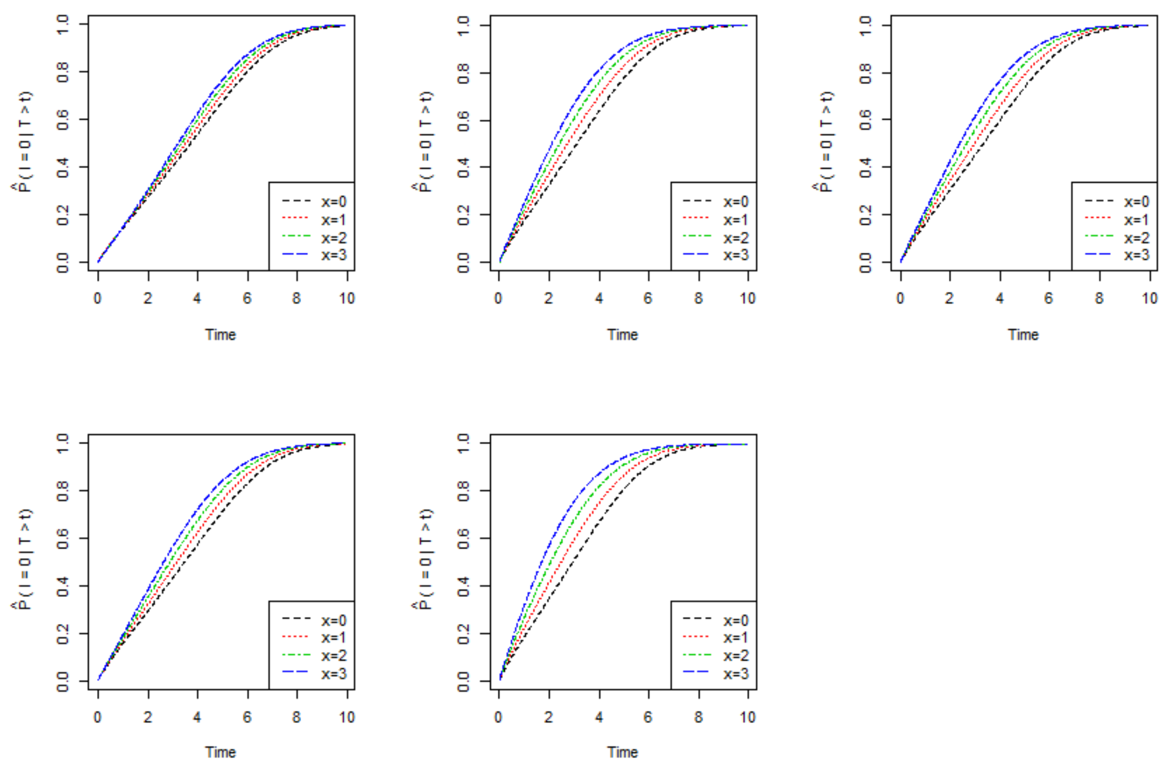

Figure 4.4: Cure rate given an individual has survived up to a specific time t over four covariate groups (From the left to right, the models are proportional odds cure rate model with Weibull baseline $\phi = (0, 0.5, 1, 2, \infty)$).

Chapter 5

Summary and Conclusions

Due to the significant improvement in medical science and effective treatments for various diseases in the past several decades, it has become possible that some patients, with certain types of cancer, can be cured meaning they do not experience the disease again for a long period of time following a good prognosis and efficient treatment. Hence, the cure rate becomes an important indicator for evaluating the effectiveness of a treatment. Cure and destructive cure rate models have become important tools in survival analysis literature, particularly in the study involving a surviving fraction and time to relapse. For this reason, generalizing these models in different ways (e.g., proportional odds lifetimes) and by accommodating more realistic assumptions are interpretations is highly desirable.

5.1 Summary of research

In this thesis, cure rate and destructive cure rate models under proportional odds lifetime for the susceptibles are studied in detail.

In Chapter 2, we have developed a flexible COM-Poisson cure rate model under a proportional odds assumption for the lifetime distribution of susceptibles with the baseline function being that of a Weibull distribution or a log-logistic distribution. An EM algorithm has been developed for the maximum likelihood estimation of the model parameters of the proposed cure rate model. We have performed an extensive Monte Carlo simulation study by varying sample sizes, censoring proportion, cure rates, and the parameters in the distributions to evaluate the performance of the proposed model and the fitting methodology. Overall, the developed methodology provides accurate estimates of the model parameters as well as of the cure rates. Moreover, a real data on cutaneous melanoma has been analyzed and model diagnosis has been performed for illustrative purpose.

In Chapter 3, we have developed the EM algorithm for the destructive cure rate model by assuming the competing risks to follow negative binomial, exponentially weighted Poisson and length-biased Poisson models with proportional odds lifetimes with baseline odds to have either Weibull or log-logistic distributions, based on right censored data. We have carried out an extensive Monte Carlo simulation study and have also discussed a model discrimination between the models. We have used a malignant melanoma data to examine the performance of the proposed models and the methods of inference developed in this Chapter.

In Chapter 4, we have developed a flexible proportional odds COM-Poisson cure rate model with gamma frailty with proportional odds lifetime distribution for the susceptibles having baseline function to be either a Weibull distribution or a log-logistic distribution. An EM algorithm has been developed for the maximum likelihood estimation of the parameters of the proposed cure rate model. An extensive Monte Carlo simulation study has been performed by varying sample sizes, censoring proportions, cure rates, and the parameter values in the distributions to evaluate the performance of the proposed model and fitting methodology. Overall, the proposed method provides accurate estimates of the model parameters as well as of the cure rates. Moreover, a real data on cutaneous melanoma and a simulated dataset have been analyzed and model discrimination has been performed for illustrative purpose.

5.2 Future work

There are many possible extensions to this work. First, the whole thesis is under frequentist framework, a natural extension would be to consider the Bayesian approach by assigning appropriate prior assumptions on the parameters of interest and then develop the corresponding inference. Second, the underlying distribution can be varied for more general scenarios. For example, we consider log-logistic and Weibull as baseline distributions for proportional odds model, but one can investigate the use of other general baseline distributions. One may instead use non-parametric specification of the baseline distribution in the proportional odds assumptions for the lifetime of susceptibles. This may resolve the subjectivity invovled in the choice of the baseline distribution. Finally, one may consider the use of a much more generalized model, called transformation survival model, which includes proportional odds, proportional hazards, as well as many other commonly used survival models as special cases for the underlying distribution of the lifetimes of susceptibles.

Appendix A

Appendix corresponding to Chapter 2

As mentioned earlier, the COM-Poisson distribution includes the Bernoulli, Poisson and Geometric distributions as special cases. Here, we detail the steps of the EM algorithm for these three special cure models.

A.1 Bernoulli cure rate model

Let the competing cause random variable M follow a Bernoulli distribution with probability of success $\eta/(1+\eta)$. The probability density function for the whole population can then be expressed as

$$f_p(t_i, \boldsymbol{\theta}) = \frac{\eta}{1+\eta} f(t_i; \boldsymbol{\gamma}). \tag{A.1}$$

The survival function for the susceptible group is just the survival function for the time to event W, i.e., $S_s(t_i; \boldsymbol{\theta}) = S(t_i; \boldsymbol{\gamma})$. The inverse of the cure rate under this setting is $1/p_0 = 1 + \eta$. We, therefore, have $H_\phi(\eta) = 1 + \eta$ under the logistic link function with a fixed value of ϕ, which implies $\eta = e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$. The $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function is then given by

$$\sum_{i\in\Delta_1} \boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{i\in\Delta_1} \log f(t_i; \boldsymbol{x}_i, \boldsymbol{\gamma}) - \sum_{i\in\Delta*} \log(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma}). \tag{A.2}$$

It is readily seen that some of the terms in the Q function are only corresponding to $\boldsymbol{\beta}$ while the others are only corresponding to $\boldsymbol{\gamma}$. So, it can be split into two parts as follows:

$$Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) = Q_1(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) + Q_2(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}), \tag{A.3}$$

$$Q_1(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}) = \sum_{I\in\Delta_1} \log f(t_i; \boldsymbol{x}_i, \boldsymbol{\gamma}) + \sum_{I_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i; \boldsymbol{x}_{ic}, \boldsymbol{\gamma}), \tag{A.4}$$

$$Q_2(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) = \sum_{i\in\Delta_1} \boldsymbol{x}_i'\boldsymbol{\beta} - \sum_{i\in\Delta*} \log(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i'\boldsymbol{\beta}, \tag{A.5}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i; \boldsymbol{\gamma}^{(k)})}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i; \boldsymbol{\gamma}^{(k)})} \tag{A.6}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$
\begin{aligned}
\frac{\partial Q}{\partial \gamma_j} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j},\\
\frac{\partial Q}{\partial \beta_l} &= \sum_{i\in\Delta_1} x_{il} - \sum_{i\in\Delta *} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} x_{il},\\
\frac{\partial^2 Q}{\partial \beta_l \partial \beta_{l'}} &= -\sum_{i\in\Delta *} \frac{x_{il} x_{il'} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2},\\
\frac{\partial^2 Q}{\partial \gamma_j \partial \gamma_{j'}} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}},
\end{aligned}
$$

for $l, l' = 0, \ldots, p,\ j, j' = 1, 2, h = 21, \ldots, 2p, i = 1, \ldots, n.$

A.2 Poisson cure rate model

Let the competing cause random variable M follow a Poisson distribution. The probability density function for the whole population in this case can be expressed as

$$f_p(t_i;\boldsymbol{\theta}) = [\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] f(t_i;\boldsymbol{\gamma})(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{(S(t_i;\boldsymbol{\gamma})-1)}, \tag{A.7}$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = [(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{\gamma})} - 1] e^{-\boldsymbol{x}_i\boldsymbol{\beta}}. \tag{A.8}$$

The cure rate is $p_0 = e^{-\eta}$. We would then have $H_\phi(\eta) = e^\eta$ under the logistic link with a fixed value of ϕ, which implies that $\eta = \ln(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$. The $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function is then given by

$$Q = \sum_{I\in\Delta_1} \log[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] + \sum_{I\in\Delta_1} \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma}) + \sum_{I\in\Delta_1} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$$
$$- \sum_{i\in\Delta*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)}\log((1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}-1) \tag{A.9}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = 1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}})^{-S(t_i;\boldsymbol{\gamma}^{(k)})} \tag{A.10}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\begin{aligned}
\frac{\partial Q}{\partial \beta_l} &= \sum_{i\in\Delta_1} \frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} - \sum_{i\in\Delta*}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} \\
&\quad + \sum_{i\in\Delta_0}\frac{\boldsymbol{\pi}^{(k)}S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}, \\
\frac{\partial Q}{\partial \gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial \log f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j} + \sum_{i\in\Delta_1}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) \\
&\quad + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} &= \sum_{i\in\Delta_1}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}\{\frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}[1-\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}]+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\}-\sum_{i\in\Delta*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}\\
&+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}[1-\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}-1}],\\
\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\\
&+\sum_{i\in\Delta_0}\frac{\boldsymbol{\pi}^{(k)}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\frac{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}(1+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}))}{[1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},\\
\frac{\partial^2 Q}{\partial\gamma_j\partial\gamma_{j'}} &= \sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_1}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})\\
&+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\\
&-\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}[\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})]^2}{[1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_{j'}},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $j, j' = 1, 2$, $h = 21, \ldots, 2p$, $i = 1, \ldots, n$.

A.3 Geometric cure rate model

Let the competing cause random variable M follow a Geometric distribution. The probability density function for the whole population in this case can be expressed as

$$
f_p(t_i;\boldsymbol{\theta}) = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}f(t_i,\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})^2}, \tag{A.11}
$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = \frac{S(t_i;\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})}, \tag{A.12}$$

where $R_G(t_i,\boldsymbol{\theta}) = 1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}} - e^{\boldsymbol{x}_i'\boldsymbol{\beta}}S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})$. The cure rate under this setting is $p_0 = 1-\eta$, and under the logistic link with a fixed value of ϕ, we would have $H_\phi(\eta) = (1-\eta)^{-1}$, which implies that $\eta = e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}$. The $Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)})$ function is then given by

$$\begin{aligned} Q = &\sum_{I\in\Delta_1}\boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{I\in\Delta_1}\log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma}) - \sum_{I\in\Delta_1}2\log R_G(t_i,\boldsymbol{\theta}) - \sum_{i\in\Delta_0}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\boldsymbol{x}_i'\boldsymbol{\beta} \\ &+ \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\log S(t_i;\boldsymbol{\gamma}) - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\log R_G(t_i,\boldsymbol{\theta}), \end{aligned} \tag{A.13}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{S(t_i;\boldsymbol{\gamma}^{(k)})e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}} \tag{A.14}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\begin{aligned} \frac{\partial Q}{\partial \beta_l} &= \sum_{i\in\Delta_1}x_{il} - 2\sum_{i\in\Delta_1}\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} - \sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}x_{il} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}, \\ \frac{\partial Q}{\partial \gamma_j} &= \sum_{i\in\Delta_1}\left(\frac{\partial\log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j} - 2\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\right) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\left(\frac{\partial\log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j} - \frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\right), \end{aligned}$$

$$\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} = -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}} - \sum_{i\in\Delta_0}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}},$$
$$\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_j} = -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j},$$
$$\frac{\partial Q^2}{\partial\gamma_j\partial\gamma_{j'}} = \sum_{i\in\Delta_1}(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} - 2\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}(\frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}),$$

$$\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}, \quad \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} = \frac{x_{ih}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})},$$
$$\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_{j'}},$$
$$\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_l} = \frac{x_{il}x_{il'}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}(1-S(t_i,\boldsymbol{\theta})), \quad \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} = -\frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j},$$

for $l, l' = 0,\ldots,p$, $j, j' = 1, 2$, $h = 21,\ldots,2p$, $i = 1,\ldots,n$.

A.4 Observed information matrix

COM-Poisson cure rate model: The score functions, for a fixed value of ϕ, are

$$\frac{\partial l}{\partial\beta_l} = -\sum_{i\in\Delta*}x_{il}\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}}{z_{2,i}z_{01,i}}x_{il} + \sum_{i\in\Delta_0}\frac{z_{2,i}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{z_{01,i}(1+z_{1i})},$$
$$\frac{\partial l}{\partial\gamma_h} = \sum_{i\in\Delta_1}\frac{\partial \log f(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}} - 1)\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}.$$

Hence, the components of the observed information matrix, for a fixed value of ϕ, are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} =& -\Big\{ -\sum_{i\in\Delta*}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}})^2}+ \\
&+\sum_{i\in\Delta_1}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{2,i}z_{01,i})^2}\Big[z_{2,i}(z_{21,i}z_{01,i}+z_{31,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}})-z_{21,i}[z_{21,i}+\frac{z_{2,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}'_i\boldsymbol{\beta}}\Big] \\
&+\sum_{i\in\Delta_0}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{1,i}+1)^2z_{01,i}^2}\Big[(\frac{z_{21,i}}{z_{01,i}}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}+z_{2,i})(z_{1,i}+1)z_{01,i} \\
&-z_{2,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}[\frac{z_{02,i}}{z_{01,i}}(z_{1,i}+1)+z_{2,i}]\Big]\Big\}, \\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_h} =& -\Big\{\sum_{i\in\Delta_1}x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2z_{01,i}}\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h} \\
&+\sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{z_{01,i}(1+z_{1,i})}(z_{21,i}-\frac{z_{2,i}^2}{1+z_{1,i}})\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}\Big\}, \\
-\frac{\partial^2 l}{\partial\gamma_h\partial\gamma_{h'}} =& -\Big\{\sum_{i\in\Delta_1}\frac{\partial \log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}+\sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} \\
&+\sum_{i\in\Delta_1}(\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2})\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}} \\
&+\sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}+ \\
&+\sum_{i\in\Delta_0}\frac{z_{21,i}z_{1,i}+z_{21,i}-z_{2,i}^2}{(1+z_{1,i})^2}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\Big\},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

A.4.1 Bernoulli cure rate model

The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_l} &= \sum_{i\in\Delta_1} x_{il} - \sum_{i\in\Delta *} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0} w_i x_{il},\\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_0} w_i \frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j},
\end{aligned}
$$

where $w_i = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{\gamma})}$.

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_l \partial \beta_{l'}} &= -\Big\{ -\sum_{i\in\Delta *} \frac{x_{il}x_{il'} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2} + \sum_{i\in\Delta_0} x_{il}x_{il'} w_i(1-w_i) \Big\},\\
-\frac{\partial^2 l}{\partial \beta_l \partial \gamma_j} &= -\Big\{ \sum_{i\in\Delta_0} x_{il} w_i(1-w_i) \frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial \gamma_j} \Big\},\\
-\frac{\partial^2 l}{\partial \gamma_j \partial \gamma_{j'}} &= -\Big\{ \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}}\\
&\quad + \sum_{i\in\Delta_0} w_i \frac{\partial \log S^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} + \sum_{i\in\Delta_0} w_i(1-w_i) \frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial \gamma_j} \frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial \gamma_{j'}} \Big\},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

A.4.2 Poisson cure rate model

The score functions are

$$\begin{aligned}
\frac{\partial l}{\partial \beta_l} &= \sum_{i\in\Delta_1} \frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta*} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma})-1)}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}},\\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta*} \frac{\partial S(t_i;\boldsymbol{\gamma})}{\partial \gamma_j}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}).
\end{aligned}$$

Hence, the components of the observed information matrix are

$$\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_l \partial \beta_{l'}} &= -\Bigg\{ \sum_{i\in\Delta_1} \frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})-e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})]^2}\\
&\qquad + \sum_{i\in\Delta*} \frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}(S(t_i;\boldsymbol{\gamma})-1)\Bigg\},\\
-\frac{\partial^2 l}{\partial \beta_l \partial \gamma_j} &= -\Bigg\{ \sum_{i\in\Delta*} \frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\Bigg\},\\
-\frac{\partial^2 l}{\partial \gamma_j \partial \gamma_{j'}} &= -\Bigg\{ \sum_{i\in\Delta_1} \frac{\partial \log f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j\partial \gamma_{j'}} + \sum_{i\in\Delta*} \frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})\Bigg\},
\end{aligned}$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

A.4.3 Geometric cure rate model

The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_l} &= \sum_{i\in\Delta_1} x_{il} - 2\sum_{i\in\Delta_1}\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} - \sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} \\
&\quad + \sum_{i\in\Delta_0}\frac{S(t_i;\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\left(x_{il} - \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}\right), \\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i\in\Delta_1}\left(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j} - \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\right) - \sum_{i\in\Delta*}\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j}.
\end{aligned}
$$

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\Bigg\{-2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}} + \sum_{i\in\Delta_0}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}(S(t_i;\boldsymbol{\gamma})-1) \\
&\quad - \sum_{i\in\Delta_0}\frac{S(t_i;\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}} - \sum_{i\in\Delta_0}\frac{x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}S(t_i;\boldsymbol{\gamma})}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}\Bigg\}, \\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} &= -\Bigg\{-\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} - \sum_{i\in\Delta*}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j}\Bigg\}, \\
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} &= -\Bigg\{\sum_{i\in\Delta_1}\Big(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}\Big) - \sum_{i\in\Delta*}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}\Bigg\},
\end{aligned}
$$

where $R_G(t_i,\boldsymbol{\theta}) = 1 - e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma})-1)$;

$$
\begin{aligned}
&\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}, \quad \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} = \frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})}, \\
&\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} - \frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_{j'}}, \\
&\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_l} = \frac{x_{il}x_{il'}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})^2}, \quad \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} = -\frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j},
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

Appendix B

Appendix corresponding to Chapter 3

B.1 Destructive length-biased Poisson cure rate model

The required first- and second-order derivatives of the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows;

$$
\begin{aligned}
\frac{\partial Q}{\partial \beta_{1l}} &= \sum_{i \in \Delta_1} x_{il} q_i \{1 - p_i F(t_i) A_i G_i\} - \sum_{i \in \Delta_0} x_{il} p_i \{(1 - \pi_i^{(k)}) D_i + \pi_i^{(k)} [q_i F(t_i) S_p(t_i) A_i - p_{0i} D_i] J_i\}, \\
\frac{\partial Q}{\partial \beta_{2k}} &= \sum_{i \in \Delta_1} z_{ik} \eta_i \{\frac{1}{A_i} - p_i F(t_i)\} - \sum_{i \in \Delta_0} z_{ik} \eta_i p_i (1 - \pi_i^{(k)} + \pi_i^{(k)} B_i), \\
\frac{\partial Q}{\partial \gamma_r} &= \sum_{i \in \Delta_1} \left\{ \frac{\partial \log f(t_i)}{\partial \gamma_r} - p_i \eta_i \frac{\partial F(t_i)}{\partial \gamma_r} K_i \right\} - \sum_{i \in \Delta_0} \pi_i^{(k)} l_{sp} J_i.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\beta_{il}\partial\beta_{il'}} &= -\sum_{i\in\Delta_1} x_{il}x_{il'}p_iq_i\Bigg[1+F(t_i)\Bigg\{G_i(\frac{E_i}{H_i}+r_iA_i)-\frac{2\eta_iE_i}{C_i^2}\Bigg\}\Bigg] \\
&- \sum_{i\in\Delta_0} x_{il}x_{il'}(1-\pi^{(k)})p_iq_i(D_i-\eta_ip_i) \\
&- \sum_{i\in\Delta_0} x_{il}x_{il'}\pi^{(k)}J_i[E_iS_p(t_i)\{E_i+r_iC_i-E_iC_i^2\} \\
&- p_ip_{0i}\{D_i(q_i-p_iD_i)-p_i\eta_iq_i\}+(C_iE_iS_p(t_i)-p_ip_{0i}D_i)^2J_i], \\
\frac{\partial^2 Q}{\partial\beta_{il}\partial\beta_{2k}} &= -\sum_{i\in\Delta_1} x_{il}z_{ik}\eta_ip_iq_iF(t_i)\{1+\frac{1}{C_i^2}\}-\sum_{i\in\Delta_0} x_{il}z_{ik}\eta_ip_i\{(1-\pi_i^{(k)})q_i \\
&+ \pi_i^{(k)}p_ip_{0i}S(t_i)S_p(t_i)J_i^2[D_i-q_iF(t_i)A_i]+\pi_i^{(k)}q_iB_i\}, \\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} &= \sum_{i\in\Delta_1} z_{ik}z_{ik'}\eta_i\{\frac{1}{A_i}-p_iF(t_i)-\frac{\eta}{A_i^2}\} \\
&- \sum_{i\in\Delta_0} z_{ik}z_{ik'}p_i\eta_i\{1-\pi_i^{(k)}+\pi_i^{(k)}\eta_ip_i(B_i^2-F_i)+\pi_i^{(k)}B_i\}, \\
\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_r} &= -\sum_{i\in\Delta_1} x_{il}q_ip_i\frac{\partial F(t_i)}{\partial\gamma_r}\Bigg\{A_iG_i+p_iF(t_i)\Bigg[\frac{G_i}{H_i^2}-\frac{2A_i\eta_i}{C_i^3}\Bigg]\Bigg\} \\
&- \sum_{i\in\Delta_0} x_{il}p_i\pi_i^{(k)}J_iS_p(t_i)A_i\frac{\partial F(t_i)}{\partial\gamma_r}\Bigg\{q_i\Bigg[1-F(t_i)A_ip_i+\frac{F(t_i)p_i}{C_iH_i}\Bigg] \\
&+ J_ip_i(q_iF(t_i)S_p(t_i)A_i-p_{0i}D_0)\Bigg\} \\
\frac{\partial^2 Q}{\partial\beta_{2k}\partial\gamma_r} &= \sum_{i\in\Delta_1} z_{ik}p_i\eta_i\{\frac{1}{C_i}(\frac{\eta_i}{A_i}-1)-1\}\frac{\partial F(t_i)}{\partial\gamma_r} \\
&- \sum_{i\in\Delta_0} z_{ik}\pi^{(k)}\eta_ip_iJ_i\frac{\partial F(t_i)}{\partial\gamma_r}\{S_p(t_i)-p_{0i}I_iS(t_i)J_i\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial\gamma_r\gamma_{r'}} = & \sum_{i\in\Delta_1}\left\{\frac{\partial^2 \log f(t_i)}{\partial\gamma_r\gamma_{r'}} - p_i\eta_i\left[K_i\frac{\partial^2 F(t_i)}{\partial\gamma_r\gamma_{r'}} + \frac{\partial F(t_i)}{\partial\gamma_r}\frac{\partial F(t_i)}{\partial\gamma_{r'}}\frac{\eta_i p_i}{C_i^2}\right]\right\} \\
& - \sum_{i\in\Delta_0}\pi_i^{(k)}p_iJ_i\left\{A_iS_p(t_i)\frac{\partial F^2(t_i)}{\partial\gamma_r\partial\gamma_{r'}}\right. \\
& \left. + \frac{\partial F(t_i)}{\partial\gamma_r}\frac{\partial F(t_i)}{\partial\gamma_{r'}}\left(\frac{S_p(t_i)p_i}{H_i^2} - A_iI_i + A_iS_p(t_i)I_iJ_i\right)\right\},
\end{aligned}
$$

with

$$
\begin{aligned}
A_i &= \eta_i + \frac{1}{H_i}, \\
B_i &= (S_p(t_i)F(t_i) - p_{0i})J_i, \\
C_i &= \eta_i H_i + 1, \\
D_i &= \eta_i q_i + 1, \\
E_i &= \frac{p_i q_i F(t_i)}{H_i}, \\
F_i &= (S_p(t_i)F(t_i)^2 - p_{0i})J_i, G_i = 1 - \frac{1}{C_i^2}, \\
J_i &= \frac{1}{S_p(t_i) - p_{0i}}, \quad K_i = 1 + \frac{1}{C_i} \\
H_i &= 1 - p_iF(t_i), \\
I_i &= p_iA_iS_p(t_i), \\
l_{sp} &= I_i\frac{\partial F(t_i)}{\partial\gamma_r}, \\
r_i &= 1 - 2p_i.
\end{aligned}
$$

B.2 Destructive negative binomial cure rate model

The required first- and second-order derivatives of the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows:

$$
\begin{aligned}
\frac{\partial Q}{\partial \beta_{1l}} &= \sum_{i\in\Delta_1} x_{il} q_i C_i - \sum_{i\in\Delta_0} (1-\pi_i^{(k)}) x_{il}\eta_i p_i q_i A_i - \sum_{i\in\Delta_0} \pi_i^{(k)} x_{il}\eta_i p_i q_i D_i E_i,\\
\frac{\partial Q}{\partial \beta_{2k}} &= \sum_{i\in\Delta_1} z_{ik} C_i - \sum_{i\in\Delta_0} (1-\pi_i^{(k)}) z_{ik}\eta_i p_i A_i - \sum_{i\in\Delta_0} \pi_i^{(k)} z_{ik}\eta_i p_i D_i E_i,\\
\frac{\partial Q}{\partial \gamma_r} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i)}{\partial \gamma_r} - \sum_{i\in\Delta_1} (1+\phi) B_i \eta_i p_i \frac{\partial F(t_i)}{\partial \gamma_r} - \sum_{i\in\Delta_0} \pi_i^{(k)} \eta_i p_i E_i B_i S_p(t_i) \frac{\partial F(t_i)}{\partial \gamma_r},\\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \beta_{1l'}} &= -\sum_{i\in\Delta_1} x_{il} x_{il'} p_i q_i B_i (1 + \eta_i r_i F(t_i) + \phi \eta_i q_i F(t_i) C_i)\\
&\quad - \sum_{i\in\Delta_0} x_{il} x_{il'} (1-\pi_i^{(k)}) \eta_i p_i q_i A_i F_i\\
&\quad - \sum_{i\in\Delta_0} x_{il} x_{il'} \pi_i^{(k)} \eta_i p_i q_i E_i \{\eta_i p_i q_i p_{0i} S_p(t_i)(F(t_i)B_i - A_i)^2 E_i\\
&\quad - p_{0i} A_i F_i + F(t_i) S_p(t_i) B_i J_i\},\\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \beta_{2k}} &= -(1+\phi) \sum_{i\in\Delta_1} x_{il} z_{ik} \eta_i p_i q_i F(t_i) B_i^2 - \sum_{i\in\Delta_0} x_{il} z_{ik} (1-\pi^{(k)}) \eta_i p_i q_i A_i^2\\
&\quad - \sum_{i\in\Delta_0} x_{il} z_{ik} q_i \pi^{(k)} K_i,\\
\frac{\partial^2 Q}{\partial \beta_{2k}\partial \beta_{2k'}} &= -(1+\phi) \sum_{i\in\Delta_1} z_{ik} z_{ik'} \eta_i p_i F(t_i) B_i^2\\
&\quad - \sum_{i\in\Delta_0} z_{ik} z_{ik'} (1-\pi^{(k)}) \eta_i p_i A_i^2 - \sum_{i\in\Delta_0} z_{ik} z_{ik'} \pi^{(k)} K_i,\\
\frac{\partial^2 Q}{\partial \beta_{1l}\partial \gamma_r} &= -\sum_{i\in\Delta_1} x_{il} q_i \eta_i p_i B_i \frac{\partial F(t_i)}{\partial \gamma_r} (1 + \phi C_i) - \sum_{i\in\Delta_0} \pi_i^{(k)} x_{il} \eta_i p_i q_i G_i,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q}{\partial \beta_{2k}\partial \gamma_r} &= -\sum_{i\in\Delta_1} z_{ik}\eta_i p_i B_i \frac{\partial F(t_i)}{\partial \gamma_r}(1+\phi C_i) - \sum_{i\in\Delta_0} \pi_i^{(k)} z_{ik}\eta_i p_i G_i,\\
\frac{\partial^2 Q}{\partial \gamma_r\partial \gamma_{r'}} &= \sum_{i\in\Delta_1} \frac{\partial^2 \log f(t_i)}{\partial \gamma_r\partial \gamma_{r'}} + \sum_{i\in\Delta_1}(1+\phi)B_i\eta_i p_i\left[\eta_i\phi p_i B_i \frac{\partial F(t_i)}{\partial \gamma_r}\frac{\partial F(t_i)}{\partial \gamma_{r'}} - \frac{\partial^2 F(t_i)}{\partial \gamma_r \gamma_{r'}}\right]\\
&\quad - \sum_{i\in\Delta_0} \pi_i^{(k)}\eta_i p_i B_i E_i S_p(t_i)\left\{\frac{\partial^2 F(t_i)}{\partial \gamma_r\partial \gamma_{r'}} + B_i\eta_i p_i \frac{\partial F(t_i)}{\partial \gamma_r}\frac{\partial F(t_i)}{\partial \gamma_{r'}}\left[E_i p_{0i} - \phi\right]\right\}.
\end{aligned}
$$

with

$$
\begin{aligned}
&A_i = \frac{1}{1+\eta_i p_i \phi}, B_i = \frac{1}{1+\eta_i p_i\phi F(t_i)}, q_i = \frac{1}{1+e^{\boldsymbol{x}_i\boldsymbol{\beta}_1}},\\
&r_i = \frac{1-e^{\boldsymbol{x}_i\boldsymbol{\beta}_1}}{1+e^{\boldsymbol{x}_i\boldsymbol{\beta}_1}}, C_i = \frac{1-\eta_i p_i F(t_i)}{1+\eta_i p_i \phi F(t_i)},\\
&D_i = S_p(t_i)F(t_i)B_i - p_{0i}A_i,\ \ E_i = \frac{1}{S_p(t_i)-p_{0i}},\\
&G_i = B_i S_p(t_i)E_i\{1 - F(t_i)B_i\eta_i p_i(1+\phi) + \eta_i p_i D_i E_i\}\frac{\partial F(t_i)}{\partial \gamma_r},\\
&J_i = 1 - 2p_i - \phi\eta_i p_i q_i B_i F(t_i)),\\
&F_i = 1 - 2p_i - \phi\eta_i p_i q_i A_i,\\
&K_i = \eta_i p_i E_i\{F(t_i)S_p(t_i)B_i^2 - p_{0i}A_i^2 + \eta_i p_i p_{0i} S_p(t_i)(F(t_i)B_i - A_i)^2 E_i\}.
\end{aligned}
$$

B.3 Destructive exponentially weighted Poisson cure rate model

The required first- and second-order derivatives of the $Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)})$ function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows;

$$\frac{\partial Q}{\partial\beta_{1l}} = \sum_{i\in\Delta_1} x_{il}q_i(1-l_{qi}F(t_i)) - \sum_{i\in\Delta_0}(1-\pi_i^{(k)})x_{il}q_il_{qi} - \sum_{i\in\Delta_0}\pi_i^{(k)}A_iB_ix_{il}q_il_{qi},$$
$$\frac{\partial Q}{\partial\beta_{2k}} = \sum_{i\in\Delta_1} z_{ik}(1-l_{qi}F(t_i)) - \sum_{i\in\Delta_0} z_{ik}(1-\pi_i^{(k)})l_{qi} - \sum_{i\in\Delta_0}\pi_i^{(k)}A_iB_iz_{ik}l_{qi},$$
$$\frac{\partial Q}{\partial\gamma_r} = \sum_{i\in\Delta_1}[\frac{\partial \log f(t_i)}{\partial\gamma_r} - \frac{\partial F(t_i)}{\partial\gamma_r}l_{qi}] - \sum_{i\in\Delta_0}\pi_i^{(k)}l_{qi}S_p(t_i)B_i\frac{\partial F(t_i)}{\partial\gamma_r},$$

$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{1l'}} = -\sum_{i\in\Delta_1} x_{il}x_{il'}p_iq_i\{1+\eta_ir_ie^{\phi}F(t_i)\}$$
$$- \sum_{i\in\Delta_0} x_{il}x_{il'}q_il_{qi}\{(1-\pi_i^{(k)})r_i + \pi_i^{(k)}B_i(r_iA_i+q_iC_i)\},$$
$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\beta_{2k}} = -\sum_{i\in\Delta_1} x_{il}z_{ik}q_il_{qi}F(t_i) - \sum_{i\in\Delta_0} x_{il}z_{ik}q_il_{qi}(1-\pi_i^{(k)}+\pi_i^{(k)}B_i(A_i+C_i)),$$
$$\frac{\partial^2 Q}{\partial\beta_{2k}\partial\beta_{2k'}} = -\sum_{i\in\Delta_1} z_{ik}z_{ik'}l_{qi}F(t_i) - \sum_{i\in\Delta_0} z_{il}z_{ik'}l_{qi}(1-\pi_i^{(k)}+\pi_i^{(k)}B_i(A_i+C_i)),$$
$$\frac{\partial^2 Q}{\partial\beta_{1l}\partial\gamma_r} = -\sum_{i\in\Delta_1} x_{il}q_il_{qi}\frac{\partial F(t_i)}{\partial\gamma_r} - \sum_{i\in\Delta_0}\pi_i^{(k)}x_{il}q_iE_i,$$
$$\frac{\partial^2 Q}{\partial\beta_{2l}\partial\gamma_r} = -\sum_{i\in\Delta_1} z_{ik}l_{qi}\frac{\partial F(t_i)}{\partial\gamma_r} - \sum_{i\in\Delta_0}\pi_i^{(k)}z_{ik}E_i,$$
$$\frac{\partial^2 Q}{\partial\gamma_r\partial\gamma_{r'}} = \sum_{i\in\Delta_1}\left[\frac{\partial^2 \log f(t_i)}{\partial\gamma_r\partial\gamma_{r'}} - \frac{\partial^2 F(t_i)}{\partial\gamma_r\partial\gamma_{r'}}l_{qi}\right]$$
$$- \sum_{i\in\Delta_0}\pi_i^{(k)}l_{qi}S_p(t_i)B_i\left[\frac{\partial F(t_i)}{\partial\gamma_r}\frac{\partial F(t_i)}{\partial\gamma_{r'}}l_{qi}p_{0i}B_i + \frac{\partial^2 F(t_i)}{\partial\gamma_r\partial\gamma_{r'}}\right],$$

with

$$
\begin{aligned}
&A_i = S_p(t_i)F(t_i) - p_{0i}, \;\; B_i = \frac{1}{S_p - p_{0i}}, \;\; r_i = 1 - 2p_i, ,\\
&C_i = \eta_i p_i p_{0i} e^{\phi} S_p(t_i) S(t_i)^2 B_i,\\
&l_{qi} = \eta_i p_i e^{\phi},\\
&E_i = [1 - p_{0i} l_{qi} S(t_i) B_i] l_{qi} S_p(t_i) \frac{\partial F(t_i)}{\partial \gamma_r} B_i.
\end{aligned}
$$

B.4 Observed information matrix

B.4.1 Destructive length-biased Poisson cure rate model

The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_{1l}} &= \sum_{i \in \Delta_1} x_{il} q_i \left\{ 1 + \frac{p_i F(t_i)}{C_i H_i} \right\} - \sum_{i \in \Delta^*} x_{il} \eta_i p_i q_i F(t_i) - \sum_{i \in \Delta^*} x_{il} E_i,\\
\frac{\partial l}{\partial \beta_{2k}} &= \sum_{i \in \Delta_1} \frac{z_{ik} \eta_i}{A_i} - \sum_{i \in \Delta^*} z_{il} p_i \eta_i F(t_i),\\
\frac{\partial l}{\partial \gamma_r} &= \sum_{i \in \Delta_1} \frac{p_i}{C_i H_i} \frac{\partial F(t_i)}{\partial \gamma_r} + \sum_{i \in \Delta_1} \frac{\partial \log f(t_i)}{\partial \gamma_r} - \sum_{i \in \Delta^*} \frac{p_i C_i}{H_i} \frac{\partial F(t_i)}{\partial \gamma_r}.
\end{aligned}
$$

Hence, the components of the negative observed information matrix are

$$
\begin{aligned}
\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{1l'}} &= -\sum_{i\in\Delta_1} x_{il}x_{il'}p_iq_i\left[1+F(t_i)\left\{G_i\left(r_iA_i+\frac{E_i}{H_i}\right)\right\}-\frac{2\eta_iE_i}{C_i^2}\right] \\
&\quad -\sum_{ii\in\Delta_0} x_{il}x_{il'}E_i\{E_i+r_iC_i\}, \\
\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{2k}} &= -\sum_{i\in\Delta_1}\frac{x_{il}z_{ik}\eta_ip_iq_iF(t_i)}{C_i^2}-\sum_{i\in\Delta^*}x_{il}z_{ik}\eta_ip_iq_iF(t_i), \\
\frac{\partial^2 l}{\partial\beta_{2k}\partial\beta_{2k'}} &= \sum_{i\in\Delta_1}\frac{z_{ik}z_{ik'}\eta_i}{A_iC_i}-\sum_{i\in\Delta^*}z_{ik}z_{ik'}\eta_ip_iF(t_i), \\
\frac{\partial^2 l}{\partial\beta_{1l}\partial\gamma_r} &= \sum_{i\in\Delta_1}\frac{x_{il}p_i}{C_iH_i}\left\{q_i+\frac{E_i(1+2\eta_iH_i)}{C_i}\right\}\frac{\partial F(t_i)}{\partial\gamma_r}-\sum_{i\in\Delta^*}x_{il}p_iq_i\frac{\partial F(t_i)}{\partial\gamma_r}\left(\eta_i+\frac{1}{H_i^2}\right), \\
\frac{\partial^2 l}{\partial\beta_{2k}\partial\gamma_r} &= -\sum_{i\in\Delta_1}\frac{z_{ik}\eta_ip_i}{C_i^2}\frac{\partial F(t_i)}{\partial\gamma_r}-\sum_{i\in\Delta^*}z_{il}p_i\eta_i\frac{\partial F(t_i)}{\partial\gamma_r}, \\
\frac{\partial^2 l}{\partial\gamma_r\partial\gamma_{r'}} &= \sum_{i\in\Delta_1}\frac{p_i}{C_iH_i}\left\{\frac{\partial F^2(t_i)}{\partial\gamma_r\partial\gamma_{r'}}+\frac{p_i(2\eta_iH_i+1)}{C_iH_i}\frac{\partial F(t_i)}{\partial\gamma_r}\frac{\partial F(t_i)}{\partial\gamma_{r'}}\right\}+\sum_{i\in\Delta_1}\frac{\partial^2\log f(t_i)}{\partial\gamma_r\partial\gamma_{r'}} \\
&\quad -\sum_{i\in\Delta^*}\frac{p_i}{H_i}\left\{C_i\frac{\partial F(t_i)}{\partial\gamma_r\partial\gamma_{r'}}+\frac{p_i}{H_i}\frac{\partial F(t_i)}{\partial\gamma_r}\frac{\partial F(t_i)}{\partial\gamma_{r'}}\right\}.
\end{aligned}
$$

B.4.2 Destructive negative binomial cure rate model

The score functions, for a fixed value of ϕ, are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_{1l}} &= \sum_{i\in\Delta_1}x_{il}q_i-(1+\phi)\sum_{i\in\Delta_1}x_{il}p_iq_i\eta_iB_i(t_i)-\sum_{i\in\Delta_0}x_{il}p_iq_i\eta_iB_i(t_i), \\
\frac{\partial l}{\partial\beta_{2k}} &= \sum_{i\in\Delta_1}z_{ik}-(1+\phi)\sum_{i\in\Delta_1}z_{ik}p_i\eta_iB_iF(t_i)-\sum_{i\in\Delta_0}z_{il}p_i\eta_iB_iF(t_i), \\
\frac{\partial l}{\partial\gamma_r} &= \sum_{i\in\Delta_1}\frac{\partial\log f(t_i)}{\partial\gamma_r}-(1+\phi)\sum_{i\in\Delta_1}\eta_ip_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}-\sum_{i\in\Delta_0}\eta_ip_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}.
\end{aligned}
$$

Hence, the components of the negative observed information matrix, for a fixed value of ϕ, are

$$\begin{aligned}
\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{1l'}} &= -\sum_{i\in\Delta_1} x_{il}x_{il'}p_iq_iB_i\{1+\eta_ir_iF(t_i)+\phi\eta_iq_iF(t_i)C_i\} \\
&\quad -\sum_{i\in\Delta_0} x_{il}x_{il'}\eta p_iq_iF(t_i)B_iJ_i, \\
\frac{\partial^2 l}{\partial\beta_{1l}\partial\beta_{2k}} &= -(1+\phi)\sum_{i\in\Delta_1} x_{il}z_{ik}\eta_ip_iq_iF(t_i)B_i^2 - \sum_{i\in\Delta_0} x_{il}z_{ik}\eta_ip_iq_iF(t_i)B_i^2, \\
\frac{\partial^2 l}{\partial\beta_{2k}\partial\beta_{2k'}} &= -(1+\phi)\sum_{i\in\Delta_1} z_{ik}z_{ik'}\eta_ip_iF(t_i)B_i^2 \\
&\quad -\sum_{i\in\Delta_0} z_{ik}z_{ik'}\eta_ip_iF(t_i)B_i^2, \\
\frac{\partial^2 l}{\partial\beta_{1l}\partial\gamma_r} &= (1+\phi)\sum_{i\in\Delta_1} x_{il}p_iq_i\eta_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}\{p_i\eta_i\phi B_iF(t_i)-1\} \\
&\quad +\sum_{i\in\Delta_0} x_{il}p_iq_i\eta_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}\{\eta_i\phi p_iB_iF(t_i)-1\}, \\
\frac{\partial^2 l}{\partial\beta_{2k}\partial\gamma_r} &= (1+\phi)\sum_{i\in\Delta_1} z_{ik}p_i\eta_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}\{p_i\eta_i\phi B_iF(t_i)-1\}+ \\
&\quad +\sum_{i\in\Delta_0} z_{ik}p_i\eta_iB_i\frac{\partial F(t_i)}{\partial\gamma_r}\{\eta_i\phi p_iB_iF(t_i)-1\}, \\
\frac{\partial^2 l}{\partial\gamma_r\partial\gamma_{r'}} &= \sum_{i\in\Delta_1}\frac{\partial \log f^2(t_i)}{\partial\gamma_r\partial\gamma_{r'}} - (1+\phi)\sum_{i\in\Delta_1}\eta_ip_iB_i\left\{\frac{\partial^2F(t_i)}{\partial\gamma_r\gamma_{r'}} - \eta_ip_i\phi B_i\frac{\partial F(t_i)}{\partial\gamma_r}\frac{F(t_i)}{\gamma_{r'}}\right\} \\
&\quad -\sum_{i\in\Delta_0}\eta_ip_iB_i\left\{\frac{\partial^2F(t_i)}{\partial\gamma_r\gamma_{r'}} - \eta_ip_i\phi B_i\frac{\partial F(t_i)}{\partial\gamma_r}\frac{F(t_i)}{\gamma_{r'}}\right\}.
\end{aligned}$$

B.4.3 Destructive exponentially weighted Poisson cure rate model

The score functions, for a fixed value of ϕ, are

$$\begin{aligned}
\frac{\partial l}{\partial \beta_{1l}} &= \sum_{i\in\Delta_1} x_{il}q_i - \sum_{i\in\Delta^*} x_{il}\eta_i p_i q_i e^{\phi}F(t_i),\\
\frac{\partial l}{\partial \beta_{2k}} &= \sum_{i\in\Delta_1} z_{ik} - \sum_{i\in\Delta^*} z_{ik}\eta_i p_i e^{\phi}F(t_i),\\
\frac{\partial l}{\partial \gamma_r} &= \sum_{i\in\Delta_1} \frac{\partial \log f(t_i)}{\partial \gamma_r} - \sum_{i\in\Delta^*} \eta_i p_i e^{\phi}\frac{\partial F(t_i)}{\partial \gamma_r}.
\end{aligned}$$

Hence, the components of the negative observed information matrix, for a fixed value of ϕ, are

$$\begin{aligned}
\frac{\partial^2 l}{\partial \beta_{1l}\partial \beta_{1l'}} &= -\sum_{i\in\Delta_1} x_{il}x_{il'}p_i q_i - \sum_{i\in\Delta^*} x_{il}x_{il'}\eta_i p_i q_i r_i e^{\phi}F(t_i),\\
\frac{\partial^2 l}{\partial \beta_{1l}\beta_{2k}} &= -\sum_{i\in\Delta^*} x_{il}z_{ik}\eta_i p_i q_i e^{\phi}F(t_i), \quad \frac{\partial^2 l}{\partial \beta_{2k}\beta_{2k}} = -\sum_{i\in\Delta^*} z_{ik}z_{ik'}\eta_i p_i e^{\phi}F(t_i),\\
\frac{\partial^2 l}{\partial \beta_{1l}\partial \gamma_r} &= -\sum_{i\in\Delta^*} x_{il}\eta_i p_i q_i e^{\phi}\frac{\partial F(t_i)}{\partial \gamma_r}, \quad \frac{\partial^2 l}{\partial \beta_{2k}\partial \gamma_r} = -\sum_{i\in\Delta^*} z_{ik}\eta_i p_i e^{\phi}\frac{\partial F(t_i)}{\partial \gamma_r},\\
\frac{\partial^2 l}{\partial \gamma_r\partial \gamma_{r'}} &= \sum_{i\in\Delta_1}^{2} \frac{\partial \log f(t_i)}{\partial \gamma_r\partial \gamma_{r'}} - \sum_{i\in\Delta^*} \eta_i p_i e^{\phi}\frac{\partial^2 F(t_i)}{\partial \gamma_r\partial \gamma_{r'},}
\end{aligned}$$

for $l, l' = 0, \ldots, s$, $x_{i0} \equiv 1$, $k, k' = 1, 2, \ldots, s'$, $i = 1, \ldots, n$

Appendix C

Appendix corresponding to Chapter 4

C.1 Results for the special cases of COM-Poisson cure rate model with gamma frailty

C.1.1 Geometric cure rate model with gamma frailty

Let the competing cause random variable M follow a Geometric distribution. Then, the susceptible survival function and population density function for susceptibles, conditional on r, are

$$S_s(t_i|r_i) = \frac{S(t_i|r_i)}{1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i)}, \tag{C.15}$$

$$f_p(t_i|r_i) = \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i} f(t_i|r_i)}{(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i))^2}, \tag{C.16}$$

respectively. The cure rate is $p_0 = 1-\eta$, and the parameter η is linked to covariates through the logistic link function $e^{\boldsymbol{\beta}'\boldsymbol{x}_i}(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-1}$. The Q-function can then be expressed as

$$
\begin{aligned}
Q &= Q_1(\boldsymbol{\beta}, \gamma_0, \gamma_1, \boldsymbol{\alpha}) + Q_2(\xi), \\
Q_1 &= \sum_{i\in\Delta_1} \{E_{4i} - 2E_{5i} + \boldsymbol{\beta}'\boldsymbol{x}_i\} - \sum_{i\in\Delta_0} \{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + \boldsymbol{\beta}'\boldsymbol{x}_i E_{1i} + E_{6i} - E_{7i}\}, \\
Q_2 &= n\xi log(\xi) - n\log\Gamma(\xi) + (\xi - 1)\sum_{i\in\Delta^*} E_{3i} - \xi \sum_{i\in\Delta^*} E_{2i},
\end{aligned}
$$

where

$$
\begin{aligned}
&E_{1i} = E(I_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \;\; E_{2i} = E(r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \;\; E_{3i} = E(\log r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \\
&E_{4i} = E(\log f_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \;\; E_{5i} = E(\log R_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \;\; E_{6i} = E(I_i \log S_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \\
&E_{7i} = E(I_i \log R_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \;\; R_i = 1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i).
\end{aligned}
$$

The required first- and second-order derivatives of $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for a fixed value of ϕ, are as follows:

$$
\begin{aligned}
\frac{\partial Q_1}{\partial \alpha_l} &= \sum_{i\in\Delta_1} E_{4i(\alpha)} - 2\sum_{i\in\Delta} E_{5i(\alpha)} + \sum_{i\in\Delta_0} E_{6i(\alpha)} - \sum_{i\in\Delta_0} E_{7i(\alpha)}, \\
\frac{\partial Q_1}{\partial \gamma_k} &= \sum_{i\in\Delta_1} E_{4i(\gamma_k)} - 2\sum_{i\in\Delta_1} E_{5i(\gamma_k)} + \sum_{i\in\Delta_0} E_{6i(\gamma_k)} - \sum_{i\in\Delta_0} E_{7i(\gamma_k)}, \\
\frac{\partial Q_1}{\partial \beta_h} &= -2\sum_{i\in\Delta_1} E_{5i(\beta_h)} + \sum_{i\in\Delta_1} x_{ik} - \sum_{i\in\Delta_0} \frac{x_{ik} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i\in\Delta_0} x_{ik} E_{1i} - \sum_{i\in\Delta_0} E_{7i(\beta_h)},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 Q_1}{\partial \alpha_l \partial \alpha_{l'}} &= \sum_{i\in\Delta_1} E_{4i(\alpha_l\alpha_{l'})} - 2\sum_{i\in\Delta} E_{5i(\alpha_l\alpha_{l'})} + \sum_{i\in\Delta_0} E_{6i(\alpha_l\alpha_{l'})} - \sum_{i\in\Delta_0} E_{7i(\alpha_l\alpha_{l'})},\\
\frac{\partial^2 Q_1}{\partial \alpha_l \partial \gamma_k} &= \sum_{i\in\Delta_1} E_{4i(\alpha_l\gamma_k)} - 2\sum_{i\in\Delta} E_{5i(\alpha_l\gamma_k)} + \sum_{i\in\Delta_0} E_{6i(\alpha_l\gamma_k)} - \sum_{i\in\Delta_0} E_{7i(\alpha_l\gamma_k)},\\
\frac{\partial^2 Q_1}{\partial \gamma_k \partial \gamma_{k'}} &= \sum_{i\in\Delta_1} E_{4i(\gamma_k\gamma_{k'})} - 2\sum_{i\in\Delta} E_{5i(\gamma_k\gamma_{k'})} + \sum_{i\in\Delta_0} E_{6i(\gamma_k\gamma_{k'})} - \sum_{i\in\Delta_0} E_{7i(\gamma_k\gamma_{k'})},\\
\frac{\partial^2 Q_1}{\partial \alpha_l \partial \beta_h} &= -2\sum_{i\in\Delta_1} E_{5i(\alpha_l\beta_h)} - \sum_{i\in\Delta_0} E_{7i(\alpha_l\beta_h)}, \quad \frac{\partial^2 Q_1}{\partial \gamma_k \partial \beta_h} = -2\sum_{i\in\Delta_1} E_{5i(\gamma_k\beta_h)} - \sum_{i\in\Delta_0} E_{7i(\gamma_k\beta_h)},\\
\frac{\partial^2 Q_1}{\partial \beta_h \partial \beta_{h'}} &= -2\sum_{i\in\Delta_1} E_{5i(\beta_h\beta_{h'})} - \sum_{i\in\Delta_0} \frac{x_{il}x_{il'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2} - \sum_{i\in\Delta_0} E_{7i(\beta_h\beta_{h'})}.
\end{aligned}
$$

The first- and second-order derivatives of the E function with respect to different parameters are as follows,

$$
\begin{aligned}
&E_{4,i(.)} = M_{1,1}((\log f_i).), \quad E_{4,i(..)} = M_{1,2}((\log f_i)..), \quad E_{5,i(.)} = M_{1,1}((\log R_i).),,\\
&E_{5,i(..)} = M_{1,2}((\log R_i)..) \quad E_{6,i(.)} = M_{2,1}((\log S_i).), \quad E_{6,i(..)} = M_{2,2}((\log S_i)..),\\
&E_{7,i(.)} = M_{2,1}((\log R_i).), \quad E_{7,i(..)} = M_{2,2}((\log R_i)..),
\end{aligned}
$$

where . can be γ_k, α_l or β_h, functions $M_{1,1}, M_{1,2}, M_{2,1}, M_{2,2}$ are as in (4.19), (4.20), (4.21), and (4.22), and

$$
\begin{aligned}
\frac{\partial \log R_i}{\partial \alpha_l} &= -\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i} S(t_i|r_i)_{\alpha_l}}{R_i}, \quad \frac{\partial \log R_i}{\partial \gamma_k} = -\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i} S(t_i|r_i)_{\gamma_k}}{R_i} \\
\frac{\partial \log R_i}{\partial \beta_h} &= \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i} - x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i} S(t_i|r_i)}{R_i}, \\
\frac{\partial^2 \log R_i}{\partial \alpha_l \partial \beta_h} &= \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i) - R_i}{R_i^2} S(t_i|r_i)_{\alpha_l} x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i} \\
\frac{\partial^2 \log R_i}{\partial \gamma_k \partial \beta_h} &= \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i) - R_i}{R_i^2} S(t_i|r_i)_{\gamma_k} x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}, \\
\frac{\partial^2 \log R_i}{\partial \beta_h \partial \beta_{h'}} &= x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i r_i) \frac{R_i - e^{\boldsymbol{\beta}'\boldsymbol{x}_i} F(t_i|r_i)}{R_i^2}, \\
\frac{\partial^2 \log R_i}{\partial \alpha_l \partial \alpha_{l'}} &= \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i} S(t_i|r_i)_{\alpha_l \alpha_{l'}} - \frac{e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i^2} S(t_i|r_i)_{\alpha_l} S(t_i|r_i)_{\alpha_{l'}}, \\
\frac{\partial^2 \log R_i}{\partial \alpha_l \partial \gamma_k} &= \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i} S(t_i|r_i)_{\alpha_l \gamma_k} - \frac{e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i^2} S(t_i|r_i)_{\alpha_l} S(t_i|r_i)_{\gamma_k}, \\
\frac{\partial^2 \log R_i}{\partial \gamma_k \partial \gamma_{k'}} &= \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i} S(t_i|r_i)_{\gamma_k \gamma_{k'}} - \frac{e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{R_i^2} S(t_i|r_i)_{\gamma_k} S(t_i|r_i)_{\gamma_{k'}}.
\end{aligned}
$$

C.1.2 Poisson cure rate model with gamma frailty

Let the competing cause random variable M follow a Poisson distribution. Then the susceptible survival function and population p.d.f., conditional on r_i, are

$$S_s(t_i|r_i) = [(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)} - 1]e^{-\boldsymbol{\beta}'\boldsymbol{x}_i}, \tag{C.17}$$

$$f_p(t_i|r_i) = (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \log(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) f(t_i|r_i), \tag{C.18}$$

respectively. The cure rate is $p_0 = e^{-\eta}$, and $\eta = \ln(1 + e^{-\boldsymbol{\beta}'\boldsymbol{z}_i})$ under the logistic link function. The Q-function can be expressed as

$$
\begin{aligned}
Q &= Q_1(\boldsymbol{\beta}, \gamma_0, \gamma_1, \boldsymbol{\alpha}) + Q_2(\xi), \qquad \text{(C.19)}\\
Q_1 &= \sum_{i\in\Delta_1}\{E_{4i}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + \log\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i}\} + \sum_{i\in\Delta_0} E_{5i} - \sum_{i\in\Delta^*}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}),\\
Q_2 &= \sum_{i\in\Delta_0} n\xi log(\xi) - n\log\Gamma(\xi) + (\xi-1)\sum_{i\in\Delta^*} E_{2i} - \xi\sum_{i\in\Delta^*} E_{1i},
\end{aligned}
$$

where

$$
\begin{aligned}
&E_{1i} = E(r_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}),\ \ E_{2i} = E(\log r_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}),\ \ E_{3i} = E(\log f_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}),\\
&E_{4i} = E(S_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}),\ \ E_{5i} = E(I_i\log B_i|\boldsymbol{O}, \boldsymbol{\theta}^{(k)}), B_i = (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)} - 1.
\end{aligned}
$$

The required first- and second-order derivatives of $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed value of ϕ, are as follows:

$$\frac{\partial Q_1}{\partial \alpha_l} = \sum_{i\in\Delta_1}\{E_{4i(\alpha_l)}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i(\alpha_l)}\} + \sum_{i\in\Delta_0} E_{5i(\alpha_l)},$$

$$\frac{\partial Q_1}{\partial \gamma_k} = \sum_{i\in\Delta_1}\{E_{4i(\gamma_k)}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i(\gamma_k)}\} + \sum_{i\in\Delta_0} E_{5i(\gamma_k)},$$

$$\frac{\partial Q_1}{\partial \beta_h} = \sum_{i\in\Delta_1} E_{4i}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i\in\Delta_0}\frac{1}{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i\in\Delta_0} E_{5i(\beta_h)} - \sum_{i\in\Delta^*}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}},$$

$$\frac{\partial^2 Q_1}{\partial \beta_h\beta_{h'}} = \sum_{i\in\Delta_1}\left(E_{4i} + \frac{1}{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})} - \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}))^2}\right)\frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2} +$$

$$+ \sum_{i\in\Delta_0} E_{5i(\beta_h\beta_{h'})} - \sum_{i\in\Delta^*}\frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2},$$

$$\frac{\partial Q_1}{\partial \alpha_l\partial\alpha_{l'}} = \sum_{i\in\Delta_1}\{E_{4i(\alpha_l\alpha_{l'})}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i(\alpha_l\alpha_{l'})}\} + \sum_{i\in\Delta_0} E_{5i(\alpha_l\alpha_{l'})},$$

$$\frac{\partial Q_1}{\partial \alpha_l\beta_h} = \sum_{i\in\Delta_1} E_{4i(\alpha_l)}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i\in\Delta_0} E_{5i(\alpha_l\beta_h)},$$

$$\frac{\partial^2 Q_1}{\partial \beta_h\gamma_k} = \sum_{i\in\Delta_1} E_{4i(\gamma_k)}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i\in\Delta_0} E_{5i(\beta_h\gamma_k)},$$

$$\frac{\partial Q_1}{\partial \alpha_l\gamma_k} = \sum_{i\in\Delta_1}\{E_{4i(\alpha_l\gamma_k)}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i(\alpha_l\gamma_k)}\} + \sum_{i\in\Delta_0} E_{5i(\alpha_l\gamma_k)},$$

$$\frac{\partial^2 Q_1}{\partial \gamma_k\gamma_{k'}} = \sum_{i\in\Delta_1}\{E_{4i(\gamma_k\gamma_{k'})}\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + E_{3i(\gamma_k\gamma_{k'})}\} + \sum_{i\in\Delta_0} E_{5i(\gamma_k\gamma_{k'})}.$$

The first- and second-order derivatives of the E function with respect to different parameters are as follows:

$$E_{3,i(.)} = M_{1,1}((\log f_i).),\;\; E_{3,i(..)} = M_{1,2}((\log f_i)..),\;\; E_{4,i(.)} = M_{1,1}((S_i).),,$$
$$E_{4,i(..)} = M_{1,2}((S_i)..),\;\; E_{5,i(.)} = M_{2,1}((\log B_i).),\;\; E_{5,i(..)} = M_{2,2}((\log B_i)..),$$

where . can be γ_k, α_l or β_h, functions $M_{1,1}, M_{1,2}, M_{2,1}, M_{2,2}$ are as in (4.19), (4.20), (4.21), and (4.22), and

$$
\begin{aligned}
&\frac{\partial \mathrm{log} B_i}{\partial \alpha_l} = B_{2i} S(t_i|r_i)_{\alpha_l}, \;\; \frac{\partial \mathrm{log} B_i}{\partial \gamma_k} = B_{2i} S(t_i|r_i)_{\gamma_k}, \;\; \frac{\partial \mathrm{log} B_i}{\partial \beta_h} = S(t_i|r_i) B_{1i} p_{0i} e^{\boldsymbol{\beta}' \boldsymbol{x}_i} x_{ih},\\
&\frac{\partial^2 \mathrm{log} B_i}{\partial \alpha_l \alpha_{l'}} = B_{2i}[S(t_i|r_i)_{\alpha_l \alpha_{l'}} + S(t_i|r_i)_{\alpha_l} S(t_i|r_i)_{\alpha_l'} B_{0i} - \frac{\partial \mathrm{log} B_i}{\partial \alpha_{l'}} S(t_i|r_i)_{\alpha_l}],\\
&\frac{\partial^2 \mathrm{log} B_i}{\partial \alpha_l \gamma_k} = B_{2i}[S(t_i|r_i)_{\alpha_l \gamma_k} + S(t_i|r_i)_{\alpha_l} S(t_i|r_i)_{\gamma_k} B_{0i} - \frac{\partial \mathrm{log} B_i}{\partial \gamma_k} S(t_i|r_i)_{\alpha_l}],\\
&\frac{\partial^2 \mathrm{log} B_i}{\partial \gamma_k \gamma_{k'}} = B_{2i}[S(t_i|r_i)_{\gamma_l \gamma_{l'}} + S(t_i|r_i)_{\gamma_l} S(t_i|r_i)_{\gamma_l'} B_{0i} - \frac{\partial \mathrm{log} B_i}{\partial \gamma_{k'}} S(t_i|r_i)_{\gamma_k}],\\
&\frac{\partial \mathrm{log} B_i}{\partial \alpha_l \partial \beta_h} = S(t_i|r_i)_{\alpha_l} x_{ih} (1 - p_0) \{B_{1i} + S(t_i|r_i) B_{2i} (1 - B_{1i})\},\\
&\frac{\partial \mathrm{log} B_i}{\partial \gamma_k \partial \beta_h} = S(t_i|r_i)_{\beta_h} x_{ih} (1 - p_0) \{B_{1i} + S(t_i|r_i) B_{2i} (1 - B_{1i})\},\\
&\frac{\partial^2 \mathrm{log} B_i}{\partial \beta_h \beta_{h'}} = \frac{\partial \mathrm{log} B_i}{\partial \beta_h} p_{0i} x_{ih'} (1 - \frac{S(t_i|r_i) e^{\boldsymbol{\beta}' \boldsymbol{x}_i}}{B_i}),
\end{aligned}
$$

where $B_{0i} = \log(1 + e^{\boldsymbol{\beta}' \boldsymbol{z}_i}), \;\; B_{1i} = \frac{(1+e^{\boldsymbol{\beta}' \boldsymbol{x}_i}) S(t_i|r_i)}{B_i}$, and $B_{2i} = B_{0i} B_{1i}$.

C.1.3 Bernoulli cure rate model with gamma frailty

Let the competing causes random variable M follow a Bernoulli distribution with probability of success $\frac{\eta}{1-\eta}$. The Q-function can be written as a sum of three parts as follows:

$$
Q = Q_1(\boldsymbol{\beta}) + Q_2(\boldsymbol{\alpha}, \gamma_0, \gamma_1) + Q_3(\xi) \tag{C.20}
$$

with

$$Q_1(\boldsymbol{\beta}) = \sum_{i\in\Delta_1} \boldsymbol{\beta}'\boldsymbol{x}_i - \sum_{i\in\Delta^*} \log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) + \sum_{i\in\Delta_0} E_{1i}\boldsymbol{\beta}'\boldsymbol{x}_i, \tag{C.21}$$

$$Q_2(\boldsymbol{\alpha},\gamma_0,\gamma_1) = \sum_{i\in\Delta_1}(\boldsymbol{\alpha}'\boldsymbol{x}_i + \log f_0(t_i) - 2E_{4i}) + \sum_{i\in\Delta_0}(E_{1i}\boldsymbol{\alpha}'\boldsymbol{x}_i + E_{1i}\log S_0(t_i) - E_{5i}),$$

$$Q_3(\xi) = n\xi log(\xi) - n\log\Gamma(\xi) + (\xi-1)\sum_{i\in\Delta^*} E_{3i} - \xi\sum_{i\in\Delta^*} E_{2i},$$

where

$$E_{1i} = E(I_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

$$E_{2i} = E(r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

$$E_{3i} = E(\log r_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

$$E_{4i} = E(\log A_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}),$$

$$E_{5i} = E(I_i\log A_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)})$$

$$A_i = r_iS_0(t_i)e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} + F_0(t_i).$$

The required first- and second-order derivatives of $Q(\boldsymbol{\theta},\boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed value of ϕ, are as follows:

$$\frac{\partial Q}{\partial \xi} = n\log(\xi) + n - n\psi_0(\xi) + \sum_{i\in\Delta^*}(E_{3i} - E_{2i}),$$
$$\frac{\partial Q}{\partial \beta_h} = \sum_{i\in\Delta_1} x_{il} + \sum_{i\in\Delta_0} E_{1i}x_{il} - \sum_{i\in\Delta^*} \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}x_{il}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}},$$
$$\frac{\partial Q}{\partial \alpha_l} = \sum_{i\in\Delta_1} x_{il} + \sum_{i\in\Delta_0} E_{1i}x_{il} - 2\sum_{i\in\Delta_1} E_{4i(\alpha_l)} - \sum_{i\in\Delta_0} E_{5i(\alpha_l)},$$
$$\frac{\partial Q}{\partial \gamma_k} = \sum_{i\in\Delta_1} \frac{\partial \log f_0(t_i)}{\partial \gamma_k} - 2\sum_{i\in\Delta_1} E_{4i(\gamma_k)} + \sum_{i\in\Delta_0} E_{1i}\frac{\partial \log S_0(t_i)}{\partial \gamma_k} - \sum_{i\in\Delta_0} E_{5i}(\gamma_k),$$
$$\frac{\partial^2 Q}{\partial \beta_h \partial \beta_{h'}} = -\sum_{i\in\Delta^*} \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}x_{il}x_{il'}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2},$$
$$\frac{\partial^2 Q}{\partial \alpha_l \gamma_k} = -2\sum_{i\in\Delta_1} E_{4i(\alpha_l\gamma_k)} - \sum_{i\in\Delta_0} E_{5i(\alpha_l\gamma_k)},$$
$$\frac{\partial^2 Q}{\partial \xi^2} = \frac{n}{\xi} - n\psi_1(\xi),$$
$$\frac{\partial^2 Q}{\partial \alpha_l \alpha_{l'}} = -2\sum_{i\in\Delta_1} E_{4i(\alpha_l\alpha_{l'})} - \sum_{i\in\Delta_0} E_{5i(\alpha_l\alpha_{l'})},$$
$$\frac{\partial^2 Q}{\partial \gamma_k \partial \gamma_{k'}} = \sum_{i\in\Delta_1}\left(\frac{\partial^2 \log f_0(t_i)}{\partial \gamma_k \partial \gamma_{k'}} - 2E_{4i(\gamma_k\gamma_{k'})}\right) + \sum_{i\in\Delta_0}\left(E_{1i}\frac{\partial^2 \log S_0(t_i)}{\partial \gamma_k \partial \gamma_{k'}} - E_{5i}(\gamma_k\gamma_{k'})\right),$$

where $\psi_0(.)$ is the digamma function given by $\psi_0(x) = \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma - \sum_{k=0}^{\infty}(\frac{1}{x+k} - \frac{1}{k+1})$, and $\gamma = \lim_{n\longrightarrow\infty}[\sum_{k=1}^{n}\frac{1}{k} - \ln n]$, and $\psi_1(x) = -\int_0^1 \frac{t^{x-1}}{1-t}(\ln t)^m dt$.

The first- and second-order derivatives of the E-function with respect to different parameters are as follows:

$$E_{4,i(.)} = M_{1,1}((\log A_i).), \;\; E_{4,i(..)} = M_{1,2}((\log A_i)..),$$
$$E_{5,i(.)} = M_{2,1}((\log A_i).), \;\; E_{5,i(..)} = M_{2,2}((\log A_i)..),$$

where . can be γ_k, α_l or β_h, functions $M_{1,1}, M_{1,2}, M_{2,1}, M_{2,2}$ as in (4.19), (4.20), (4.21), and (4.22), and

$$
\begin{aligned}
\frac{\partial \log A_i}{\partial \alpha_l} &= \frac{r_i S_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} x_{il}}{A_i}, \quad \frac{\partial \log A_i}{\partial \gamma_k} = \frac{r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} - 1}{A_i} S_{0;\gamma_k} \\
\frac{\partial^2 \log A_i}{\partial \alpha_l \partial \alpha_{l'}} &= \frac{r_i S_0(t_i) F_0(t_i) e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} x_{il} x_{il'}}{A_i^2}, \\
\frac{\partial^2 \log A_i}{\partial \alpha_l \partial \gamma_k} &= \frac{r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} x_{il}}{A_i} S_{0;\gamma_k} - \frac{\partial \log A_i}{\partial \alpha_l} \frac{\partial \log A_i}{\partial \gamma_k}, \\
\frac{\partial^2 \log A_i}{\partial \gamma_k \partial \gamma_{k'}} &= \frac{r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} - 1}{A_i} S_{0;\gamma_k \gamma_{k'}} - \frac{(r_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i} - 1)^2}{A_i^2} S_{0;\gamma_k} S_{0;\gamma_{k'}}.
\end{aligned}
$$

Please refer to Appendix D for $S_{0;\gamma_k \gamma_{k'}}$.

C.2 Observed information matrix

In this section, we present the observed information matrix for the COM-Poisson cure rate model with gamma frailty as well as for its special cases.

C.2.1 COM-Poisson cure rate model with gamma frailty

The score functions, for a fixed value of ϕ, are

$$
\begin{aligned}
\frac{\partial l}{\partial \alpha_l} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \alpha_l} z_{2,i} + f(t_i|r_i) \left[\frac{\partial z_{2,i}}{\partial \alpha_l} - \frac{\partial z_{1,i}}{\partial \alpha_l} \right] \right) f_r(r_i) dr_i \\
&\quad + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial z_{1,i}}{\partial \alpha_l} f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \gamma_k} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \gamma_k} z_{2,i} + f(t_i|r_i) \left[\frac{\partial z_{2,i}}{\partial \gamma_k} - \frac{\partial z_{1,i}}{\partial \gamma_k} \right] \right) f_r(r_i) dr_i \\
&\quad + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial z_{1,i}}{\partial \gamma_k} f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \beta_h} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial z_{2,i}}{\partial \beta_h} f_r(r_i) dr_i + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial z_{1,i}}{\partial \beta_h} f_r(r_i) dr_i - \sum_{i\in\Delta^*} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}, \\
\frac{\partial l}{\partial \xi} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{S(t_i|r_i)} z_{2,i} \frac{\partial \log f_r(r_i)}{\partial \xi} f_r(r_i) dr_i + \sum_{i\in\Delta_0} \frac{1}{B_i} \int (1+z_{1,i}) \frac{\partial \log fy(r_i)}{\partial \xi} f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \xi} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{S(t_i|r_i)} z_{2,i} \frac{\partial f_r(r_i)}{\partial \xi} dr_i + \sum_{i\in\Delta_0} \frac{1}{B_i} \int (1+z_{1,i}) \frac{\partial fy(r_i)}{\partial \xi} dr_i.
\end{aligned}
$$

Hence, the components of the observed information matrix, for a fixed value of ϕ, are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \gamma_k \partial \beta_h} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \gamma_k} \frac{\partial z_{2,i}}{\partial \beta_h} + f(t_i|r_i) \left[\frac{\partial^2 z_{2,i}}{\partial \gamma_k \beta_h} - \frac{\partial^2 z_{1,i}}{\partial \gamma_k \beta_h} \right] \right) f_r(r_i) dr_i \\
& + \sum_{i\in\Delta_1} \frac{1}{A_i^2} \left\{ \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \gamma_k} z_{2,i} + f(t_i|r_i) \left[\frac{\partial z_{2,i}}{\partial \gamma_k} - \frac{\partial z_{1,i}}{\partial \gamma_k} \right] \right) f_r(r_i) dr_i \right\} \\
& \times \left\{ \int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial z_{2,i}}{\partial \beta_h} f_r(r_i) dr_i \right\} \\
& + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial^2 z_{1,i}}{\partial \gamma_k \beta_h} f_r(r_i) dr_i - \sum_{i\in\Delta_0} \frac{1}{B_i^2} \int \frac{\partial z_{1,i}}{\partial \gamma_k} f_r(r_i) dr_i \int \frac{\partial z_{1,i}}{\partial \beta_h} f_r(r_i) dr_i \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \alpha_l \partial \beta_h} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \alpha_l} \frac{\partial z_{2,i}}{\partial \beta_h} + f(t_i|r_i) \left[\frac{\partial^2 z_{2,i}}{\partial \alpha_l \beta_h} - \frac{\partial^2 z_{1,i}}{\partial \alpha_l \beta_h} \right] \right) f_r(r_i) dr_i \\
& + \sum_{i\in\Delta_1} \frac{1}{A_i^2} \Bigg\{ \int \frac{1}{S(t_i|r_i)} \left(\frac{\partial f(t_i|r_i)}{\partial \alpha_l} z_{2,i} + f(t_i|r_i) \left[\frac{\partial z_{2,i}}{\partial \alpha_l} - \frac{\partial z_{1,i}}{\partial \alpha_l} \right] \right) f_r(r_i) dr_i \Bigg\} \\
& \times \Bigg\{ \int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial z_{2,i}}{\partial \beta_h} f_r(r_i) dr_i \Bigg\} \\
& + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial^2 z_{1,i}}{\partial \alpha_l \beta_h} f_r(r_i) dr_i - \sum_{i\in\Delta_0} \frac{1}{B_i^2} \int \frac{\partial z_{1,i}}{\partial \alpha_l} f_r(r_i) dr_i \int \frac{\partial z_{1,i}}{\partial \beta_h} f_r(r_i) dr_i \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial^2 z_{2,i}}{\partial \beta_h \partial \beta_{h'}} f_r(r_i) dr_i \\
& - \sum_{i\in\Delta_1} \frac{1}{A_i^2} \left[\int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial z_{2,i}}{\partial \beta_h} f_r(r_i) dr_i \right] \left[\int \frac{f(t_i|r_i)}{S(t_i|r_i)} \frac{\partial z_{2,i}}{\partial \beta_{h'}} f_r(r_i) dr_i \right] \\
& + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{\partial^2 z_{1,i}}{\partial \beta_h \partial \beta_{h'}} f_r(r_i) dr_i - \sum_{i\in\Delta_0} \frac{1}{B_i^2} \int \frac{\partial z_{1,i}}{\partial \beta_h} \frac{\partial z_{1,i}}{\partial \beta_{h'}} f_r(r_i) dr_i - \sum_{i\in\Delta^*} \frac{x_{ih} x_{ih'} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2} \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\gamma_k\partial\gamma_{k'}} &= -\Bigg\{\sum_{i\in\Delta_1}\frac{1}{A_i}\Bigg\{\int\frac{1}{S(t_i|r_i)}\bigg(\frac{\partial^2 f(t_i|r_i)}{\partial\gamma_k\partial\gamma_{k'}}z_{2,i} \\
&+\frac{\partial f(t_i|r_i)}{\partial\gamma_k}\frac{\partial z_{2,i}}{\partial\gamma_{k'}}+\frac{\partial f(t_i|r_i)}{\partial\gamma_{k'}}\left[\frac{\partial z_{2,i}}{\partial\gamma_k}-\frac{\partial z_{1,i}}{\partial\gamma_k}\right]+f(t_i|r_i)\left[\frac{\partial^2 z_{2,i}}{\partial\gamma_k\gamma_{k'}}-\frac{\partial^2 z_{1,i}}{\partial\gamma_k\gamma_{k'}}\right] \\
&-\frac{\partial \mathrm{log} S(t_i|r_i)}{\partial\gamma_{k'}}\left\{\frac{\partial f(t_i|r_i)}{\partial\gamma_k}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\gamma_k}-\frac{\partial z_{1,i}}{\partial\gamma_k}\right]\right\}\bigg)f_r(r_i)dr_i\Bigg\} \\
&-\sum_{i\in\Delta_1}\frac{1}{A_i^2}\int\frac{1}{S(t_i|r_i)}\left(\frac{\partial f(t_i|r_i)}{\partial\gamma_k}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\gamma_k}-\frac{\partial z_{1,i}}{\partial\gamma_{k'}}\right]\right)f_r(r_i)dr_i \\
&\int\frac{1}{S(t_i|r_i)}\left(\frac{\partial f(t_i|r_i)}{\partial\gamma_{k'}}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\gamma_{k'}}-\frac{\partial z_{1,i}}{\partial\gamma_{k'}}\right]\right)f_r(r_i)dr_i \\
&+\sum_{i\in\Delta_0}\frac{1}{B_i}\int\frac{\partial^2 z_{1,i}}{\partial\gamma_k\gamma_{k'}}f_r(r_i)dr_i-\sum_{i\in\Delta_0}\frac{1}{B_i^2}\int\frac{\partial z_{1,i}}{\gamma_k}f_r(r_i)dr_i\int\frac{\partial z_{1,i}}{\gamma_{k'}}f_r(r_i)dr_i\}\Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\alpha_l\partial\alpha_{l'}} &= -\Bigg\{\sum_{i\in\Delta_1}\frac{1}{A_i}\Bigg\{\int\frac{1}{S(t_i|r_i)}\Bigg(\frac{\partial^2 f(t_i|r_i)}{\partial\alpha_l\partial\alpha_{l'}}z_{2,i}+\frac{\partial f(t_i|r_i)}{\partial\alpha_l}\frac{\partial z_{2,i}}{\partial\alpha_{l'}}\\
&\quad+\frac{\partial f(t_i|r_i)}{\partial\alpha_{l'}}\left[\frac{\partial z_{2,i}}{\partial\alpha_l}-\frac{\partial z_{1,i}}{\partial\alpha_l}\right]+f(t_i|r_i)\left[\frac{\partial^2 z_{2,i}}{\partial\alpha_l\alpha_{l'}}-\frac{\partial^2 z_{1,i}}{\partial\alpha_l\alpha_{l'}}\right]\\
&\quad-\frac{\partial\log S(t_i|r_i)}{\partial\alpha_{l'}}\left(\frac{\partial f(t_i|r_i)}{\partial\alpha_l}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\alpha_l}-\frac{\partial z_{1,i}}{\partial\alpha_l}\right]\right)\Bigg)f_r(r_i)dr_i\Bigg\}\\
&\quad-\sum_{i\in\Delta_1}\frac{1}{A_i^2}\int\frac{1}{S(t_i|r_i)}\left(\frac{\partial f(t_i|r_i)}{\partial\alpha_l}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\alpha_l}-\frac{\partial z_{1,i}}{\partial\alpha_{l'}}\right]\right)f_r(r_i)dr_i\\
&\quad\int\frac{1}{S(t_i|r_i)}\left(\frac{\partial f(t_i|r_i)}{\partial\alpha_{l'}}z_{2,i}+f(t_i|r_i)\left[\frac{\partial z_{2,i}}{\partial\alpha_{l'}}-\frac{\partial z_{1,i}}{\partial\alpha_{l'}}\right]\right)f_r(r_i)dr_i\\
&\quad+\sum_{i\in\Delta_0}\frac{1}{B_i}\int\frac{\partial^2 z_{1,i}}{\partial\alpha_l\alpha_{l'}}f_r(r_i)dr_i-\sum_{i\in\Delta_0}\frac{1}{B_i^2}\int\frac{\partial z_{1,i}}{\alpha_l}f_r(r_i)dr_i\int\frac{\partial z_{1,i}}{\alpha_{l'}}f_r(r_i)dr_i\Bigg\},\\
-\frac{\partial^2 l}{\partial\beta_h\partial\gamma_k} &= -\Bigg\{\sum_{i\in\Delta_1}x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2z_{01,i}}\frac{\partial\log S(t_i|r_i)}{\partial\gamma_k}\\
&\quad+\sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{z_{01,i}(1+z_{1,i})}\left(z_{21,i}-\frac{z_{2,i}^2}{1+z_{1,i}}\right)\frac{\partial\log S(t_i|r_i)}{\partial\gamma_k}\Bigg\},\\
-\frac{\partial^2 l}{\partial\gamma_k\partial\gamma_{k'}} &= -\Bigg\{\sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_{k'}}+\sum_{i\in\Delta_1}\left(\frac{z_{21,i}}{z_{2,i}}-1\right)\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_{k'}}\\
&\quad+\sum_{i\in\Delta_1}\left(\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2}\right)\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_k}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{k'}}\\
&\quad+\sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_{k'}}+\sum_{i\in\Delta_0}\frac{z_{21,i}z_{1,i}+z_{21,i}-z_{2,i}^2}{(1+z_{1,i})^2}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_k}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{k'}}\Bigg\},\\
-\frac{\partial^2 l}{\partial\xi^2} &= -\left\{\frac{n}{\xi}-n\psi_1(\xi)\right\},\psi_1(x)=-\int_0^1\frac{t^{x-1}}{1-t}(\ln t)^m dt,
\end{aligned}
$$

where

$$
A_i=\int\frac{f(t_i|r_i)}{S(t_i|r_i)}z_{2,i}f_r(r_i)dr_i,\ \ B_i=\int(1+z_{1,i})f_r(r_i)dr_i
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j^* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

C.2.2 Geometric cure rate model with gamma frailty

Set $A_i = \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} f_r(r_i)dr_i$, $\; B_i = \int \frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)} f_r(r_i)dr_i$. The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_h} &= \sum_{i\in\Delta_1} x_{ih} - \sum_{i\in\Delta_1} \frac{2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \\
&\quad - \sum_{i\in\Delta_0} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} dr_i, \\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i\in\Delta_1} \frac{2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i)dr_i \\
&\quad + \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i)dr_i \\
&\quad + \sum_{i\in\Delta_0} \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i)dr_i, \\
\frac{\partial l}{\partial \xi} &= \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial f_r(r_i)}{\partial \xi} dr_i \\
&\quad + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)} \frac{\partial f_r(r_i)}{\partial \xi} dr_i.
\end{aligned}
$$

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\xi\partial\gamma_j} = -\Bigg\{ & -\sum_{i\in\Delta_1}\frac{1}{A_i^2}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f_r(r_i)}{\partial\xi}dr_i\frac{\partial A_i}{\partial\gamma_j} \\
& + 2\sum_{i\in\Delta_1}\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial f_r(r_i)}{\partial\xi}dr_i \\
& - \sum_{i\in\Delta_0}\frac{1}{B_i^2}\int\frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)}\frac{\partial f_r(r_i)}{\partial\xi}dr_i\frac{\partial B_i}{\partial\gamma_j} \\
& + \sum_{i\in\Delta_0}\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f_r(r_i)}{\partial\xi}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}dr_i\Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\xi\partial\beta_h} = -\Bigg\{ & -\sum_{i\in\Delta_1}\frac{2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{F(t_i|r_i)f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial f_r(r_i)}{\partial\xi}dr_i \\
& + 2\sum_{i\in\Delta_1}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f_r(r_i)}{\partial\xi}dr_i\int\frac{f(t_i|r_i)F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}f_r(r_i)dr_i \\
& - \sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i}\int\frac{F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f_r(r_i)}{\partial\xi}dr_i \\
& + \sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\int\frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)}\frac{\partial f_r(r_i)}{\partial\xi}dr_i\int\frac{F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}f_r(r_i)dr_i\Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \xi^2} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial^2 f_r(r_i)}{\partial \xi^2} dr_i \\
& - \sum_{i\in\Delta_1} \frac{1}{A_i^2} \left[\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial f_r(r_i)}{\partial \xi} dr_i \right]^2 \\
& + \sum_{i\in\Delta_0} \frac{1}{B_i} \int \frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)} \frac{\partial^2 f_r(r_i)}{\partial \xi^2} dr_i \\
& - \sum_{i\in\Delta_0} \frac{1}{B_i^2} \left[\int \frac{1}{1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i)} \frac{\partial f_r(r_i)}{\partial \xi} dr_i \right]^2 \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} = -\Bigg\{ & -\sum_{i\in\Delta_1} \frac{2x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \\
& + \sum_{i\in\Delta_1} \frac{6x_{ih}x_{ih'}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)F(t_i|r_i)^2 f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^4} dr_i \\
& - \sum_{i\in\Delta_1} \frac{4x_{ih}x_{ih'}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2} \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \\
& - \sum_{i\in\Delta_0} \frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} dr_i \\
& + \sum_{i\in\Delta_0} \frac{2x_{ih}x_{ih'}e^{\boldsymbol{2\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{F(t_i|r_i)^2 f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \\
& - \sum_{i\in\Delta_0} \frac{x_{ih}x_{ih'}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} dr_i \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} dr_i \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_h \partial \gamma_j} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_1} \frac{2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_1} \frac{6x_{ih}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^4} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} dr_i \\
& + \sum_{i\in\Delta_1} \frac{2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2} \int \frac{f(t_i|r_i)F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} dr_i \frac{\partial A_i}{\partial \gamma_j} \\
& + \sum_{i\in\Delta_0} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_0} \frac{2x_{ih}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} dr_i \\
& + \sum_{i\in\Delta_0} \frac{x_{ih}e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2} \int \frac{F(t_i|r_i)f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} dr_i \int \frac{f_r(r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} dr_i \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} = -\Bigg\{ & \sum_{i\in\Delta_1}\frac{2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial^2 S(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_i \\
& + \sum_{i\in\Delta_1}\frac{2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial f(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& + \sum_{i\in\Delta_1}\frac{6e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^4}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& - \sum_{i\in\Delta_1}\frac{2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\frac{\partial A_i}{\partial\gamma_{j'}} \\
& + \sum_{i\in\Delta_1}\frac{1}{A_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial^2 f(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_i \\
& + \sum_{i\in\Delta_1}\frac{2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial f(t_i|r_i)}{\partial\gamma_j}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& - \sum_{i\in\Delta_1}\frac{1}{A_i^2}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\frac{\partial A_i}{\partial\gamma_{j'}} \\
& + \sum_{i\in\Delta_0}\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial^2 S(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_i \\
& + \sum_{i\in\Delta_0}\frac{2e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& - \sum_{i\in\Delta_0}\frac{e^{2\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i \\
& \int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i\Bigg\},
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial A_i}{\partial \beta_h} &= -2x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\int \frac{f(t_i|r_i)F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3} f_r(r_i)dr_i, \\
\frac{\partial B_i}{\partial \beta_h} &= -x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\int \frac{F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2} f_r(r_i)dr_i, \\
\frac{\partial A_i}{\partial \gamma_j} &= 2e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^3}\frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i)dr_i \\
&\quad + \int \frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}F(t_i|r_i))^2}\frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i)dr_i,
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

C.2.3 Poisson cure rate model with gamma frailty

Set $A_i = \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} f(t_i|r_i) f_r(r_i) dr_i$, $B_i = \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} f_r(r_i) dr_i$. The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial \beta_h} &= \sum_{i \in \Delta_1} \frac{1}{\log(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} \\
&\quad - \sum_{i \in \Delta_1} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f(t_i|r_i) f_r(r_i) dr_i \\
&\quad - \sum_{i \in \Delta_0} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i \in \Delta_1} \frac{\log(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f(t_i|r_i) f_r(r_i) dr_i \\
&\quad + \sum_{i \in \Delta_1} \frac{1}{A_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \\
&\quad + \sum_{i \in \Delta_0} \frac{\log(1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{B_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \\
\frac{\partial l}{\partial \xi} &= \sum_{i \in \Delta_1} \frac{1}{A_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} f(t_i|r_i) \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
&\quad + \sum_{i \in \Delta_0} \frac{1}{B_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i.
\end{aligned}
$$

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\xi\partial\beta_h} = -\Bigg\{ & -\sum_{i\in\Delta_1}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i}\int\frac{F(t_i|r_i)f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)+1}}\frac{\partial f_r(r_i)}{\partial\gamma_j}dr_i \\
& +\sum_{i\in\Delta_1}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2}\int\frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}}\frac{\partial f_r(r_i)}{\partial\gamma_j}dr_i\int\frac{F(t_i|r_i)f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)+1}}f_r(r_i)dr_i \\
& -\sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i}\int\frac{F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)+1}}\frac{\partial f_r(r_i)}{\partial\gamma_j}dr_i \\
& +\sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\int\frac{1}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}}\frac{\partial f_r(r_i)}{\partial\gamma_j}dr_i\int\frac{F(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)+1}}f_r(r_i)dr_i\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \xi \partial \gamma_j} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_1} \frac{1}{A_i} \log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_1} \frac{1}{A_i^2} \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \\
& + \sum_{i\in\Delta_1} \frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i^2} \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
& \times \int \frac{f(t_i|r_i)}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{F(t_i|r_i)}} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \\
& + \sum_{i\in\Delta_0} \frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{B_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
& + \sum_{i\in\Delta_0} \frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{B_i^2} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \\
& \times \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \Bigg\}, \\
-\frac{\partial^2 l}{\partial \xi^2} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{1}{A_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} f(t_i|r_i) \frac{\partial^2 f_r(r_i)}{\partial \gamma_j \partial \gamma_j} dr_i \\
& + \sum_{i\in\Delta_0} \frac{1}{B_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial^2 f_r(r_i)}{\partial \gamma_j \partial \gamma_j} dr_i \\
& - \sum_{i\in\Delta_1} \frac{1}{A_i^2} \left[\int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} f(t_i|r_i) \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \right]^2 \\
& - \sum_{i\in\Delta_0} \frac{1}{B_i^2} \left[\int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f_r(r_i)}{\partial \gamma_j} dr_i \right]^2 \Bigg\}, \\
-\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{z_{ih} z_{ih'} e^{\boldsymbol{\beta}'\boldsymbol{x}_i} (\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) - e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2 [\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})]^2} - \sum_{i\in\Delta_1} x_{ih} x_{ih'} e^{\boldsymbol{\beta}'\boldsymbol{x}_i} \frac{D_i}{A_i} \\
& - \sum_{i\in\Delta_1} x_{ih} x_{ih'} e^{2\boldsymbol{\beta}'\boldsymbol{x}_i} \left(\frac{F_i}{A_i} + \frac{D_i^2}{A_i^2} \right) - \sum_{i\in\Delta_0} x_{ih} x_{ih'} e^{\boldsymbol{\beta}'\boldsymbol{x}_i} \frac{C_i}{B_i} - \sum_{i\in\Delta_0} x_{ih} x_{ih'} e^{2\boldsymbol{\beta}'\boldsymbol{x}_i} \left(\frac{E_i}{B_i} + \frac{C_i^2}{B_i^2} \right) \Bigg\},
\end{aligned}
$$

where

$$
\begin{aligned}
C_i &= \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f_r(r_i) dr_i,\\
D_i &= \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f(t_i|r_i) f_r(r_i) dr_i,\\
E_i &= \int (S(t_i|r_i)-2)(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-3} F(t_i|r_i) f_r(r_i) dr_i,\\
F_i &= \int (S(t_i|r_i)-2)(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-3} F(t_i|r_i) f(t_i|r_i) f_r(r_i) dr_i,
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial \beta_h \partial \gamma_j} = -\Bigg\{ & \sum_{i\in\Delta_1} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f(t_i|r_i) f_r(r_i) dr_i\\
& - \sum_{i\in\Delta_1} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i\\
& - \sum_{i\in\Delta_1} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i} \log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} F(t_i|r_i) f(t_i|r_i) f_r(r_i) dr_i\\
& + \sum_{i\in\Delta_1} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{A_i^2} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f(t_i|r_i) f_r(r_i) dr_i\\
& (\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f(t_i|r_i) f_r(r_i) dr_i\\
& + \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i)\\
& + \sum_{i\in\Delta_0} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i\\
& - \sum_{i\in\Delta_0} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}) \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i\\
& + \sum_{i\in\Delta_0} \frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2} \log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})\\
& \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)} \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i \int (1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{S(t_i|r_i)-2} F(t_i|r_i) f_r(r_i) dr_i \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} = -\Bigg\{ & \sum_{i\in\Delta_1}\frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial f(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& +\sum_{i\in\Delta_1}\frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S^2(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f(t_i|r_i)f_r(r_i)dr_i \\
& +\sum_{i\in\Delta_1}\frac{[\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})]^2}{A_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f(t_i|r_i)f_r(r_i)dr_i \\
& -\sum_{i\in\Delta_1}\frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i^2}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial f(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i \\
& \Bigg(\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f(t_i|r_i)f_r(r_i)dr_i \\
& +\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial f(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i\Bigg) \\
& +\sum_{i\in\Delta_1}\frac{1}{A_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial f^2(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_i \\
& +\sum_{i\in\Delta_1}\frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{A_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}\frac{\partial f(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i \\
& -\sum_{i\in\Delta_1}\frac{1}{A_i^2}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f(t_i|r_i)f_r(r_i)dr_i \\
& \Bigg(\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f(t_i|r_i)f_r(r_i)dr_i+\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})H_i\Bigg) \\
& +\sum_{i\in\Delta_0}\frac{\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})}{B_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S^2(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_i \\
& +\sum_{i\in\Delta_0}\frac{[\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})]^2}{B_i}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i \\
& -\sum_{i\in\Delta_0}\frac{[\log(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})]^2}{B_i^2}\int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i \\
& \int(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^{-F(t_i|r_i)}\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i)\Bigg\}
\end{aligned}
$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

C.2.4 Bernoulli cure rate model with gamma frailty

Set $A_i = \int f(t_i|r_i) f_r(r_i) dr_i$ and $B_i = \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i} S(t_i|r_i)) f_r(r_i) dr_i$. In this case, the score functions are

$$\begin{aligned}
\frac{\partial l}{\partial \beta_h} &= \sum_{i \in \Delta_1} x_{ih} - \sum_{i \in \Delta*} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i}} + \sum_{i \in \Delta_0} \frac{x_{ih} e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int S(t_i|r_i) f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \gamma_j} &= \sum_{i \in \Delta_1} \frac{1}{A_i} \int \frac{\partial f(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i + \sum_{i \in \Delta_0} \frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i} \int \frac{\partial S(t_i|r_i)}{\partial \gamma_j} f_r(r_i) dr_i, \\
\frac{\partial l}{\partial \xi} &= \sum_{i \in \Delta_1} \frac{1}{A_i} \int f(t_i|r_i) \frac{\partial f_r(r_i)}{\partial \xi} dr_i + \sum_{i \in \Delta_0} \frac{1}{B_i} \int (1 + e^{\boldsymbol{\beta}'\boldsymbol{x}_i} S(t_i|r_i)) \frac{\partial f_r(r_i)}{\partial \xi} dr_i
\end{aligned}$$

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\xi\partial\beta_h} &= -\Bigg\{\sum_{i\in\Delta_0}\frac{1}{B_i}\int x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\Bigg\},\\
-\frac{\partial^2 l}{\partial\xi\partial\gamma_j} &= -\Bigg\{\sum_{i\in\Delta_1}\frac{1}{A_i^2}\Bigg[\int\frac{\partial f(t_i|r_i)}{\partial\gamma_j}\frac{\partial f_r(r_i)}{\partial\xi}dr_iA_i-\int f(t_i|r_i)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\int\frac{\partial f(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\Bigg]\\
&\quad+\sum_{i\in\Delta_0}\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\Bigg[\int\frac{\partial S(t_i|r_i)}{\partial\gamma_j}\frac{\partial f_r(r_i)}{\partial\xi}dr_iB_i\\
&\quad-\int\Big(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\Big)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\int\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\Bigg]\Bigg\},\\
-\frac{\partial^2 l}{\partial\beta_h\partial\beta_{h'}} &= -\Bigg\{-\sum_{i\in\Delta*}\frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i})^2}\\
&\quad+\sum_{i\in\Delta_0}\frac{x_{ih}x_{ih'}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\Bigg[\int S(t_i|r_i)f_r(r_i)dr_iB_i-e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\Big(\int S(t_i|r_i)f_r(r_i)dr_i\Big)^2\Bigg]\Bigg\},\\
-\frac{\partial^2 l}{\partial\beta_h\partial\gamma_j} &= -\Bigg\{\sum_{i\in\Delta_0}\frac{x_{ih}e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\Bigg[\int\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_iB_i\\
&\quad-e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\int S(t_i|r_i)f_r(r_i)dr_i\int\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i]\Bigg]\Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} = -\Bigg\{ & \sum_{i\in\Delta_1}\frac{1}{A_i^2}\Bigg[\int\frac{\partial^2 f(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_iA_i \\
& -\int\frac{\partial f(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\int\frac{\partial f(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i\Bigg] \\
& +\sum_{i\in\Delta_0}\frac{e^{\boldsymbol{\beta}'\boldsymbol{x}_i}}{B_i^2}\Bigg[\int\frac{\partial^2 S(t_i|r_i)}{\partial\gamma_j\partial\gamma_{j'}}f_r(r_i)dr_iB_i \\
& -e^{\boldsymbol{\beta}'\boldsymbol{x}_i}\int\frac{\partial S(t_i|r_i)}{\partial\gamma_j}f_r(r_i)dr_i\int\frac{\partial S(t_i|r_i)}{\partial\gamma_{j'}}f_r(r_i)dr_i\Bigg]\Bigg\}, \\
-\frac{\partial^2 l}{\partial\xi^2} = -\Bigg\{ & \sum_{i\in\Delta_1}\frac{1}{A_i^2}\Bigg[\int f(t_i|r_i)\frac{\partial^2 f_r(r_i)}{\partial\xi^2}dr_iA_i \\
& -\int f(t_i|r_i)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\int f(t_i|r_i)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\Bigg] \\
& +\sum_{i\in\Delta_0}\frac{1}{B_i^2}\Bigg[\int\Big(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\Big)\frac{\partial^2 f_r(r_i)}{\partial\xi^2}dr_iB_i \\
& -\Big\{\int\Big(1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\Big)\frac{\partial f_r(r_i)}{\partial\xi}dr_i\Big\}^2\Bigg]\Bigg\},
\end{aligned}
$$

where

$$\int f(t_i|r_i)\frac{\partial^2 f_r(r_i)}{\partial\xi^2}dr_i = E\Bigg(f(t_i|r_i)\Bigg[\frac{\partial^2\log f_r(r_i)}{\partial\xi^2}+\frac{\partial\log f_r(r_i)}{\partial\xi}\frac{\partial\log f_r(r_i)}{\partial\xi}\Bigg]\Bigg),$$

$$\int\Big[1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\Big]\frac{\partial f_r(r_i)}{\partial\xi}dr_i = E\Bigg(\Big[1+e^{\boldsymbol{\beta}'\boldsymbol{x}_i}S(t_i|r_i)\Big]\frac{\partial\log f_r(r_i)}{\partial\xi}\Bigg),$$

$$\frac{\partial\log f_r(y)}{\partial\xi} = \log y+1+\log\xi-y-\psi_1(\xi),\quad \frac{\partial^2\log f_r(y)}{\partial\xi^2} = \frac{1}{\xi}-\psi_2(\xi)$$

for $l, l' = 0, \ldots, p$, $x_{i0} \equiv 1$, $h, h' = 0, 1, j^*$, $j* = 21, 22, \ldots, 2p$, $i = 1, \ldots, n$.

Appendix D

Derivatives with respect to proportional odds model

D.1 Weibull baseline

$S_0(t) = e^{-(\gamma_1 t)^{1/\gamma_0}}, f_0(t) = \frac{(\gamma_1 t)^{1/\gamma_0}}{\gamma_0 t} e^{-(\gamma_1 t)^{1/\gamma_0}},$

$$S_{0;\gamma_0} = \frac{(\gamma_1 t_i)^{1/\gamma_0} \log(\gamma_1 t_i)}{\gamma_0^2} S_0, \quad S_{0;\gamma_0} = -\frac{(\gamma_1 t_i)^{1/\gamma_0}}{\gamma_0 \gamma_1} S_0, \quad S_{0;\gamma_0\gamma_0} = S_{0;\gamma_0} \frac{[(\gamma_1 t_i)^{1/\gamma_0} - 1]\log(\gamma_1 t_i) - 2\gamma_0}{\gamma_0^2},$$
$$S_{0;\gamma_1\gamma_1} = S_{0;\gamma_1} \frac{[(\gamma_1 t_i)^{1/\gamma_0} - 1]\log(\gamma_1 t_i) - \gamma_0}{\gamma_0^2}, \quad S_{0;\gamma_1\gamma_0} = S_{0;\gamma_1} \frac{1 - \gamma_0 - (\gamma_1 t_i)^{1/\gamma_0}}{\gamma_0 \gamma_1};$$

$$\frac{\partial \log S_0}{\partial \gamma_0} = \frac{(\gamma_1 t_i)^{1/\gamma_0} \log(\gamma_1 t_i)}{\gamma_0^2}, \quad \frac{\partial \log S_0}{\partial \gamma_1} = -\frac{(\gamma_1 t_i)^{1/\gamma_0}}{\gamma_0 \gamma_1}, \quad \frac{\partial \log S_0^2}{\partial \gamma_0 \partial \gamma_0} = -\frac{\partial \log S_0}{\partial \gamma_0}\left(\frac{\log(\gamma_1 t_i)}{\gamma_0^2} + \frac{2}{\gamma_0}\right),$$
$$\frac{\partial \log S_0^2}{\partial \gamma_0 \partial \gamma_1} = -\frac{\partial \log S_0}{\partial \gamma_1}\left(\frac{\log(\gamma_1 t_i)}{\gamma_0^2} + \frac{1}{\gamma_0}\right), \quad \frac{\partial \log S_0^2}{\partial \gamma_1 \partial \gamma_1} = \frac{\partial \log S_0}{\partial \gamma_1}\left(\frac{1}{\gamma_0 \gamma_1} - \frac{1}{\gamma_1}\right);$$

$$\frac{\partial \text{log} f_0}{\partial \gamma_0} = \frac{\partial \text{log} S_0}{\partial \gamma_0} - \frac{1}{\gamma_0}\left(1 + \frac{\text{log}(\gamma_1 t_i)}{\gamma_0}\right), \quad \frac{\partial \text{log} f_0}{\partial \gamma_1} = \frac{1}{\gamma_0 \gamma_1} + \frac{\partial \text{log} S_0}{\partial \gamma_1},$$
$$\frac{\partial \text{log} f_0^2}{\partial \gamma_0 \partial \gamma_0} = \frac{\partial \text{log} S_0^2}{\partial \gamma_0 \gamma_0} + \frac{1}{\gamma_0^2} + \frac{2\text{log}(\gamma_1 t_i)}{\gamma_0^3}, \quad \frac{\partial \text{log} f_0^2}{\partial \gamma_0 \partial \gamma_1} = \frac{\partial \text{log} S_0^2}{\partial \gamma_0 \partial \gamma_1} - \frac{1}{\gamma_0^2 \gamma_1}, \quad \frac{\partial \text{log} f_0^2}{\partial \gamma_1 \partial \gamma_1} = \frac{\partial \text{log} S_0^2}{\partial \gamma_1 \partial \gamma_1} - \frac{1}{\gamma_0 \gamma_1^2};$$

$$\frac{\partial S}{\partial \gamma_0} = \frac{\partial S_0}{\partial \gamma_0} G_0, \quad \frac{\partial S}{\partial \gamma_1} = \frac{\partial S_0}{\partial \gamma_1} G_0, \quad \frac{\partial S}{\partial \alpha_l} = \frac{x_{il} F_0 S}{G}, \quad \frac{\partial^2 S}{\partial \gamma_0 \partial \gamma_0} = \frac{\partial^2 S_0}{\partial \gamma_0 \partial \gamma_0} G_0 - 2\frac{\partial S}{\partial \gamma_0}\frac{\partial S_0}{\partial \gamma_0} G_1$$
$$\frac{\partial^2 S}{\partial \gamma_0 \partial \gamma_1} = \frac{\partial^2 S_0}{\partial \gamma_0 \partial \gamma_1} G_0 - 2\frac{\partial S}{\partial \gamma_0}\frac{\partial S_0}{\partial \gamma_1} G_1, \quad \frac{\partial^2 S}{\partial \gamma_1 \partial \gamma_1} = \frac{\partial^2 S_0}{\partial \gamma_1 \partial \gamma_1} G_0 - 2\frac{\partial S}{\partial \gamma_1}\frac{\partial S_0}{\partial \gamma_1} G_1$$
$$\frac{\partial^2 S}{\partial \gamma_0 \partial \alpha_l} = \frac{\partial S}{\partial \gamma_0} x_{il} G_2, \quad \frac{\partial^2 S}{\partial \gamma_1 \partial \alpha_l} = \frac{\partial S}{\partial \gamma_1} x_{il} G_2, \quad \frac{\partial^2 S}{\partial \alpha_l \partial \alpha_{l'}} = \frac{\partial S}{\partial \alpha_l} x_{il'} G_2;$$

$$\frac{\partial \text{log} S}{\partial \gamma_0} = \frac{\partial \text{log} S_0}{\partial \gamma_0}\frac{1}{G}, \quad \frac{\partial \text{log} S}{\partial \gamma_1} = \frac{\partial \text{log} S_0}{\partial \gamma_1}\frac{1}{G}, \quad \frac{\partial^2 \text{log} S}{\partial \gamma_0 \partial \gamma_0} = (\frac{\partial^2 \text{log} S_0}{\partial \gamma_0 \partial \gamma_0} - \frac{\partial \text{log} S_0}{\partial \gamma_0}\frac{\partial S_0}{\partial \gamma_0} G_1)\frac{1}{G}$$
$$\frac{\partial^2 \text{log} S}{\partial \gamma_0 \partial \gamma_1} = (\frac{\partial^2 \text{log} S_0}{\partial \gamma_0 \partial \gamma_1} - \frac{\partial \text{log} S_0}{\partial \gamma_0}\frac{\partial S_0}{\partial \gamma_1} G_1)\frac{1}{G}, \quad \frac{\partial^2 \text{log} S}{\partial \gamma_1 \partial \gamma_1} = (\frac{\partial^2 \text{log} S_0}{\partial \gamma_1 \partial \gamma_1} - \frac{\partial \text{log} S_0}{\partial \gamma_1}\frac{\partial S_0}{\partial \gamma_1} G_1)\frac{1}{G}$$
$$\frac{\partial^2 \text{log} S}{\partial \gamma_0 \partial \alpha_{il}} = -\frac{\partial S}{\partial \gamma_0} x_{il}, \quad \frac{\partial^2 \text{log} S}{\partial \gamma_1 \partial \alpha_{il}} = -\frac{\partial S}{\partial \gamma_1} x_{il}, \quad \frac{\partial \text{log} S}{\partial \alpha_{il}} = \frac{x_{il} F_0}{G}, \quad \frac{\partial^2 \text{log} S}{\partial \alpha_{il} \partial \alpha_{il'}} = -\frac{\partial S}{\partial \alpha_{il}} x_{il'};$$

$$
\begin{aligned}
\frac{\partial \log f}{\partial \gamma_0} &= \frac{\partial \log f_0}{\partial \gamma_0} - 2\frac{\partial S_0}{\partial \gamma_0} G_1, \\
\frac{\partial \log f}{\partial \gamma_1} &= \frac{\partial \log f_0}{\partial \gamma_1} - 2\frac{\partial S_0}{\partial \gamma_1} G_1, \\
\frac{\partial \log f^2}{\partial \gamma_0 \partial \gamma_0} &= \frac{\partial \log f_0^2}{\partial \gamma_0 \partial \gamma_0} - 2\frac{\partial S_0^2}{\partial \gamma_0 \partial \gamma_0} G_1 + 2\left(\frac{\partial S_0}{\partial \gamma_0} G_1\right)^2, \\
\frac{\partial \log f^2}{\partial \gamma_0 \partial \gamma_1} &= \frac{\partial \log f_0^2}{\partial \gamma_0 \partial \gamma_1} - 2\frac{\partial S_0^2}{\partial \gamma_0 \partial \gamma_1} G_1 + 2\frac{\partial S_0}{\partial \gamma_0}\frac{\partial S_0}{\partial \gamma_1}(G_1)^2, \\
\frac{\partial \log f^2}{\partial \gamma_0 \partial \alpha_l} &= -2\frac{\partial S_0}{\partial \gamma_0}\frac{x_{il} y_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i}}{G^2}, \\
\frac{\partial \log f^2}{\partial \gamma_1 \partial \gamma_1} &= \frac{\partial \log f_0^2}{\partial \gamma_1 \partial \gamma_1} - 2\frac{\partial S_0^2}{\partial \gamma_1 \partial \gamma_1} G_1 + 2\left(\frac{\partial S_0}{\partial \gamma_1} G_1\right)^2 \\
\frac{\partial \log f^2}{\partial \gamma_1 \partial \alpha_l} &= -2\frac{\partial S_0}{\partial \gamma_1}\frac{x_{il} y_i e^{\boldsymbol{\alpha}'\boldsymbol{x}_i}}{G^2}, \\
\frac{\partial \log f}{\partial \alpha_l} &= x_{il}\left[2\frac{1-S_0}{G} - 1\right], \quad \frac{\partial \log f^2}{\partial \alpha_l \partial \alpha_{l'}} = -\frac{2x_{il}x_{il'}S_0F_0y_ie^{\boldsymbol{\alpha}'\boldsymbol{x}_i}}{G^2},
\end{aligned}
$$

where

$$
\begin{aligned}
G &= 1 + S_0(y_i e^{\boldsymbol{x}_i'\gamma_1} - 1), \\
G_0 &= \frac{f}{f_0}, \\
G_1 &= \frac{y_i e^{\boldsymbol{\beta}'\boldsymbol{z}_i} - 1}{G}, \\
G_2 &= 2\frac{F_0}{G} - 1, \\
G_3 &= (\gamma_1 t_i)^{1/\gamma_0} S_0 G_1 + 1.
\end{aligned}
$$

D.2 Log-logistic baseline

$$S_0 = \frac{\gamma_0^{\gamma_1}}{t_i^{\gamma_1} + \gamma_0^{\gamma_1}}, \quad S = \frac{\gamma_0^{\gamma_1} y_i e^{\boldsymbol{\alpha}'\boldsymbol{x}}}{y_i \gamma_0^{\gamma_1} e^{\boldsymbol{\alpha}'\boldsymbol{x}} + t_i^{\gamma_1}}, \quad f_0 = \frac{\gamma_0^{\gamma_1} \gamma_1 t_i^{\gamma_1 - 1}}{(t_i^{\gamma_1} + \gamma_0^{\gamma_1})^2}, \quad f = \frac{\gamma_0^{\gamma_1} \gamma_1 t_i^{\gamma_1 - 1} e^{\boldsymbol{\alpha}'\boldsymbol{x}_i}}{(t_i^{\gamma_1} + \gamma_0^{\gamma_1} e^{\boldsymbol{\alpha}'\boldsymbol{x}_i})^2}.$$

The derivatives of $S(t_i, \boldsymbol{\gamma})$ are as follows:

$$\begin{aligned}
\frac{\partial S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} &= F_0 S_0 \frac{\gamma_1}{\gamma_0}, \quad \frac{\partial S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_1} = F_0 S_0 \log\frac{\gamma_0}{t_i}, \\
\frac{\partial S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_0} &= -\frac{\partial S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} \frac{1 + \gamma_1 S_0(t_i;\boldsymbol{\gamma})}{\gamma_0} + F_0 S_0 \frac{\gamma_1}{\gamma_0} F_0 \frac{\gamma_1}{\gamma_0}, \\
\frac{\partial S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_1} &= \frac{\partial S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} \left[\frac{1}{\gamma_1} + S_0(t_i;\boldsymbol{\gamma}) \log\frac{t_i}{\gamma_0}\right] + F_0 S_0 \frac{\gamma_1}{\gamma_0} S_0 \log\frac{\gamma_0}{t_i}, \\
\frac{\partial S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_1 \partial \gamma_1} &= \frac{\partial S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_1} S_0(t_i;\boldsymbol{\gamma}) \log\frac{t_i}{\gamma_0} + F_0 \log\frac{\gamma_0}{t_i} F_0 S_0 \log\frac{\gamma_0}{t_i}.
\end{aligned}$$

The derivatives of $\log S(t_i, \boldsymbol{\gamma})$ are as follows:

$$\begin{aligned}
\frac{\partial \log S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} &= \frac{t_i^{\gamma_1}(\gamma_1/\gamma_0)}{\gamma_0^{\gamma_1} + t_i^{\gamma_1}},; \quad \frac{\partial \log S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_1} = \frac{t_i^{\gamma_1} \log(\gamma_0/t_i)}{\gamma_0^{\gamma_1} + t_i^{\gamma_1}}, \\
\frac{\partial \log S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_0} &= -\frac{\partial \log S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} \frac{1 + \gamma_1 S_0(t_i;\boldsymbol{\gamma})}{\gamma_0}, \\
\frac{\partial \log S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_1} &= \frac{\partial \log S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_0} \left[\frac{1}{\gamma_1} + S_0(t_i;\boldsymbol{\gamma}) \log\frac{t_i}{\gamma_0}\right], \\
\frac{\partial \log S_0^2(t_i;\boldsymbol{\gamma})}{\partial \gamma_1 \partial \gamma_1} &= \frac{\partial \log S_0(t_i;\boldsymbol{\gamma})}{\partial \gamma_1} S_0(t_i;\boldsymbol{\gamma}) \log\frac{t_i}{\gamma_0}.
\end{aligned}$$

The derivatives of $\log f(t_i, \boldsymbol{\gamma})$ are as follows:

$$
\begin{aligned}
\frac{\partial \log f_0(t_i,\boldsymbol{\gamma})}{\partial \gamma_0} &= \frac{\gamma_1}{\gamma_0}\frac{t_i^{\gamma_1}-\gamma_0^{\gamma_1}}{\gamma_0^{\gamma_1}+t_i^{\gamma_1}}, \quad \frac{\partial \log f_0(t_i,\boldsymbol{\gamma})}{\partial \gamma_1} = \frac{\gamma_0^{\gamma_1}-t_i^{\gamma_1}}{\gamma_0^{\gamma_1}+t_i^{\gamma_1}}\log\frac{t_i}{\gamma_0}+\frac{1}{\gamma_1},\\
\frac{\partial \log f_0^2(t_i,\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_0} &= -\frac{\partial \log f_0(t_i,\boldsymbol{\gamma})}{\partial \gamma_0}\frac{1}{\gamma_0} - \frac{2\gamma_0^{\gamma_1}t_i^{\gamma_1}}{(\gamma_0^{\gamma_1}+t_i^{\gamma_1})^2}\frac{\gamma_1^2}{\gamma_0^2},\\
\frac{\partial \log f_0^2(t_i,\boldsymbol{\gamma})}{\partial \gamma_0 \partial \gamma_1} &= \frac{\partial \log f_0(t_i,\boldsymbol{\gamma})}{\partial \gamma_0}\frac{1}{\gamma_1} + \frac{2\gamma_0^{\gamma_1}t_i^{\gamma_1}}{(\gamma_0^{\gamma_1}+t_i^{\gamma_1})^2}\frac{\gamma_1}{\gamma_0}\log\frac{t_i}{\gamma_0},\\
\frac{\partial \log f_0^2(t_i,\boldsymbol{\gamma})}{\partial \gamma_1 \partial \gamma_1} &= -\frac{2x_{il}\gamma_0^{\gamma_1}t_i^{\gamma_1}}{(\gamma_0^{\gamma_1}+t_i^{\gamma_1})^2}(\log\frac{t_i}{\gamma_0})^2 - \frac{1}{\gamma_1^2}.
\end{aligned}
$$

The derivatives of $S(t_i|y_i)$ are as follows:

$$
\begin{aligned}
\frac{\partial S(t_i)}{\partial \gamma_0} &= F(t_i)S(t_i)\frac{\gamma_1}{\gamma_0}, \quad \frac{\partial S(t_i)}{\partial \gamma_1} = -S(t_i)F(t_i)\log\frac{t_i}{\gamma_0}, \quad \frac{\partial S(t_i)}{\partial \alpha_l} = x_{il}F(t_i)S(t_i),\\
\frac{\partial S^2(t_i)}{\partial \gamma_0 \partial \gamma_0} &= \frac{\partial S(t_i)}{\partial \gamma_0}\frac{\gamma_1(F(t_i)-S(t_i))-1}{\gamma_0}, \quad \frac{\partial S^2(t_i)}{\partial \gamma_0 \partial \gamma_1} = \frac{\partial S(t_i)}{\partial \gamma_0}\left[\frac{1}{\gamma_1}-(F(t_i)-S(t_i))\log\frac{t_i}{\gamma_0}\right],\\
\frac{\partial S^2(t_i)}{\partial \gamma_0 \alpha_h} &= \frac{\partial S(t_i)}{\partial \gamma_0}(F(t_i)-S(t_i))x_{il}, \qquad \frac{\partial S^2(t_i)}{\partial \gamma_1 \partial \gamma_1} = \frac{\partial S(t_i)}{\partial \gamma_1}(S(t_i)-F(t_i))\log\frac{t_i}{\gamma_0},\\
\frac{\partial S^2(t_i)}{\partial \gamma_1 \partial \alpha_h} &= \frac{\partial S(t_i)}{\partial \gamma_1}(F(t_i)-S(t_i))x_{il}, \qquad \frac{\partial S^2(t_i)}{\partial \alpha_l \partial \alpha_{l'}} = \frac{\partial S(t_i)}{\partial \alpha_l}(F(t_i)-S(t_i))x_{il'}.
\end{aligned}
$$

The derivatives of $\log S(t_i|y_i)$ are as follows:

$$
\begin{aligned}
\frac{\partial \log S(t_i)}{\partial \gamma_0} &= F(t_i)\frac{\gamma_1}{\gamma_0}, \quad \frac{\partial \log S(t_i)}{\partial \gamma_1} = F(t_i)\log\frac{\gamma_0}{t_i}, \quad \frac{\partial \log S(t_i)}{\partial \alpha_l} = x_{il}F(t_i),\\
\frac{\partial \log S^2(t_i)}{\partial \gamma_0 \partial \gamma_0} &= -\frac{\partial \log S(t_i)}{\partial \gamma_0}\frac{1+\gamma_1 S(t_i)}{\gamma_0}, \quad \frac{\partial \log S^2(t_i)}{\partial \gamma_0 \partial \gamma_1} = \frac{\partial \log S(t_i)}{\partial \gamma_0}\left[\frac{1}{\gamma_1}+S(t_i)\log\frac{t_i}{\gamma_0}\right],\\
\frac{\partial \log S^2(t_i)}{\partial \gamma_0 \alpha_h} &= -\frac{\partial \log S(t_i)}{\partial \gamma_0}S(t_i)x_{il}, \qquad \frac{\partial \log S^2(t_i)}{\partial \gamma_1 \partial \gamma_1} = \frac{\partial \log S(t_i)}{\partial \gamma_1}S(t_i)\log\frac{t_i}{\gamma_0},\\
\frac{\partial \log S^2(t_i)}{\partial \gamma_1 \partial \alpha_l} &= -\frac{\partial \log S(t_i)}{\partial \gamma_1}S(t_i)x_{il}, \qquad \frac{\partial \log S^2(t_i)}{\partial \alpha_l \partial \alpha_{l'}} = -x_{il'}\frac{\partial S(t_i)}{\partial \alpha_l}.
\end{aligned}
$$

The derivatives of $\log f(t_i|y_i)$ are as follows:

$$
\begin{aligned}
&\frac{\partial \log f(t_i)}{\partial \gamma_0} = \frac{\gamma_1}{\gamma_0} V_i, \quad \frac{\partial \log f(t_i)}{\partial \gamma_1} = \frac{1}{\gamma_1} - V_i \log \frac{t_i}{\gamma_0}, \quad \frac{\partial \log f(t_i)}{\partial \alpha_l} = x_{il} V_i \\
&\frac{\partial \log f^2(t_i)}{\partial \gamma_0 \partial \gamma_0} = -\frac{\partial \log f(t_i)}{\partial \gamma_0}\frac{1}{\gamma_0} - W_i \frac{\gamma_1^2}{\gamma_0^2}, \quad \frac{\partial \log f^2(t_i)}{\partial \gamma_0 \partial \gamma_1} = \frac{\partial \log f(t_i)}{\partial \gamma_0}\frac{1}{\gamma_1} + W_i \frac{\gamma_1}{\gamma_0} \log \frac{t_i}{\gamma_0}, \\
&\frac{\partial \log f^2(t_i)}{\partial \gamma_0 \partial \alpha_l} = -W_i \frac{\gamma_1}{\gamma_0} x_{il}, \quad \frac{\partial \log f^2(t_i)}{\partial \gamma_1 \partial \gamma_1} = -W_i \left(\log \frac{t_i}{\gamma_0} \right)^2 - \frac{1}{\gamma_1^2}, \\
&\frac{\partial \log f^2(t_i)}{\partial \gamma_1 \partial \alpha_l} = W_i \log \frac{t_i}{\gamma_0} x_{il}, \quad \frac{\partial \log f^2(t_i)}{\partial \alpha_l \partial \alpha_{l'}} = -x_{il} x_{il'} W_i,
\end{aligned}
$$

where $V_i = F(t_i) - S(t_i),\ W_i = \frac{2\gamma_0^{\gamma_1} y_i e^{\boldsymbol{\alpha}'\boldsymbol{x}} t_i^{\gamma_1}}{(\gamma_0^{\gamma_1} y_i e^{\boldsymbol{\alpha}'\boldsymbol{x}} + t_i^{\gamma_1})^2}$.

Bibliography

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (2012). *Statistical Models Based on Counting Processes*. Springer, New York.

Balakrishnan, N., Barui, S., and Milienos F.S. (2017). Proportional hazards under ConwayMaxwell-Poisson cure rate model and associated inference. *Statistical Methods in Medical Research*, **26**(5), 2055–2077.

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. (2014). COM-Poisson cure rate models and associated likelihood-based inference with exponential and Weibull lifetimes. In *Applied Reliability Engineering and Risk Analysis Probabilistic Models and Statistical Inference*: (Eds., Frenkel, I.B., Karagrigoriou, A., and Lisnianski, A.), John Wiley & Sons, Chichester, UK, 308–348.

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.

Balakrishnan, N. and Peng, Y. (2006). Generalized gamma frailty model. *Statistics in Medicine*, **25**(16), 2797–2816.

Balakrishnan, N., Barui, S., and Milienos, F. (2017). Proportional hazards under Conway–Maxwell-Poisson cure rate model and associated inference. *Statistical Methods in Medical Research*, **26**(5), 2055–2077.

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*, **11**(1), 15–53.

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.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151.

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

Drzewiecki, K. T., Ladefoged, C, and Christensen, H. E. (1980). Biopsy and prognosis for cutaneous malignant melanomas in clinical stage I. *Scandinavian Journal of Plastic and Reconstructive Surgery*, **14**(2):141–144.

Greenwood, M. and Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, **83**(2), 255–279.

Gu, Y., Sinha, D., and Banerjee, S. (2011). Analysis of cure rate survival data under proportional odds model. *Lifetime Data Analysis*, **17**(1), 123–134.

Hoggart, C. and Griffin, J. E. (2001). A Bayesian partition model for customer attrition. In *Bayesian Methods with Applications to Science, Policy, and Official Statistics (Selected Papers from ISBA 2000), Proceedings of the Sixth World Meeting of the International Society for Bayesian Analysis*: (Ed, George, E.I.), International Society for Bayesian Analysis, Creta, Greece, 6170.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, **73**(2), 387–396.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2005). *Bayesian Survival Analysis*. Springer, New York.

Kadane, J. B., Shmueli, G., Minka, T. P., Borle, S., Boatwright, P., *et al.* (2006). Conjugate analysis of the Conway-Maxwell-Poisson distribution. *Bayesian Analysis*, **1**(2), 363–374.

Kirkwood, J. M., Ibrahim, J. G., Sondak, V. K., Richards, J., Flaherty, L. E., Ernstoff, M. S., Smith, T. J., Rao, U., Steele, M., and Blum, R. H. (2000). High-and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial e1690/s9111/c9190. *Journal of Clinical Oncology*, **18**(12), 2444–2458.

Klebanov, L. B., Rachev, S. T., and Yakovlev, A. Y. (1993). A stochastic model of radiation carcinogenesis: Latent time distributions and their properties. *Mathematical Biosciences*, **113**(1), 51–75.

Kokonendji, C. C., Kiesse, T. S., and Balakrishnan, N. (2009). Semiparametric estimation for count data through weighted distributions. *Journal of Statistical Planning and Inference*, **139**(10), 3625–3638

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

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

Liu, X., Peng, Y., Tu, D., and Liang, H. (2012). Variable selection in semiparametric cure models based on penalized likelihood, with application to breast cancer clinical trials. *Statistics in Medicine*, **31**(24), 2882–2891.

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

Lu, W. (2010). Efficient estimation for an accelerated failure time model with a cure fraction. *Statistica Sinica*, **20**, 661–674.

Mao, M. and Wang, J.-L. (2010). Semiparametric efficient estimation for a class of generalized proportional odds cure models. *Journal of the American Statistical Association*, **105**(489), 302–311.

McLachlan, G. and Krishnan, T. (2007). *The EM Algorithm and Extensions*, Second edition. John Wiley & Sons, Hoboken, New Jersey.

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. (2017). 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.

Pal, S. and Balakrishnan, N. (2018). Likelihood inference based on EM algorithm for the destructive length-biased Poisson cure rate model with Weibull lifetime. *Communications in Statistics-Simulation and Computation*, 1–17 (to appear).

Peng, Y. and Dear, K. B. (2000). A nonparametric mixture model for cure rate estimation. *Biometrics*, **56**(1), 237–243.

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.

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.

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.

Sy, J. P. and Taylor, J. M. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics*, **56**(1), 227–236.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**(3), 439–454.

Zhang, J. and Peng, Y. (2007). A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine*, **26**(16), 3157–3171.

Zhang, J. and Peng, Y. (2009). Accelerated hazards mixture cure model. *Lifetime Data Analysis*, **15**(4), 455-467.