USE OF GENERALIZED GAMMA DISTRIBUTION IN MODELING LIFETIME DATA

by

HONGBO YU

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN MATHEMATICS

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2017

Acknowledgements

First of all, I would especially like to thank my thesis supervisor, Dr. Suvra Pal, for his guidance and encouragement throughout the preparation of this thesis. I greatly appreciate his time, effort and advice towards the completion of the thesis, which has eventually helped me to earn my Master of Science degree.

I am always grateful to my parents, who have provided me love, emotional support and inspiration all through my life. I must express my very profound gratitude to my wife and two daughters for providing me with unfailing support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without them.

I am also grateful to my colleagues in Operations Research Group at BNSF Railway. I appreciate their support and encouragement during my internship period.

Last, but not the least, I would like to thank the thesis committee members for their time and effort in reviewing this work.

April 20th, 2017

Abstract

USE OF GENERALIZED GAMMA DISTRIBUTION IN MODELING LIFETIME DATA

HONGBO YU, MS

The University of Texas at Arlington, 2017

Supervising Professor: Suvra Pal

In this study, we have considered analysis of lifetime or survival data with right censoring, which is the most common form of censoring encountered in practice. Assuming a fully parametric setup, the main objective is to consider a wider family of distributions for the lifetime and then find the maximum likelihood estimates of the model parameters using some optimization technique available in R statistical software.

In this work, the generalized gamma distribution is considered as the distribution for the lifetime which is flexible in the sense that it contains some of the commonly used lifetime distributions, such as Weibull, gamma, and lognormal, as its special case. This flexibility allows us to carry out a formal test of hypothesis to determine a particular distribution within this family that provides an adequate fit to the data.

Another objective is to carry out an extensive Monte Carlo simulation study to demonstrate the performance of the estimation method and the flexibility of the generalized gamma family. To demonstrate the flexibility of the generalized gamma family, we carried out a model discrimination using the likelihood ratio test and information-based criteria.

Finally, we illustrate the estimation method and the flexibility of the generalized gamma family using a real data.

Table of Contents

List of Illustrations

List of Tables

Chapter 1 Introduction

1.1 Introduction to survival analysis

Survival analysis is a branch of statistics that deals with the analysis of lifetime data, defined as the time until the occurrence of an event of interest. Let *T* denote the lifetime variable. This event may be death, occurrence of a tumor, development of some disease, recurrence of a disease, equipment breakdown, cessation of breast feeding, and so forth. [1] For the purpose of survival analysis, three functions of time are usually defined.

1.2 Functions of time

To analyze lifetime data, three functions of time are usually important, which are the survival function, the cumulative distribution function and the hazard function. We define them as follows:

1.2.1 Survival function

The primary quantity of interest is the survival function, which is a function describing the proportion of individuals surviving beyond a given time point. The random variable *T* is non-negative, continuous and is defined in the interval [0, ∞). Its survival function is defined as

$$S(t) = P(T > t) = \int_t^{\infty} f(u)du \,, \tag{1.1}$$

where $f(\cdot)$ is the density function of *T*. The survival function is also called the reliability function in engineering field.

The graph of *t* against $S(t)$ is called the survival curve. Figure 1-1 is an example of survival curves, showing the percentage of patients surviving with respect to days since they were discharged. Obviously, every survival function $S(t)$ is monotonically decreasing

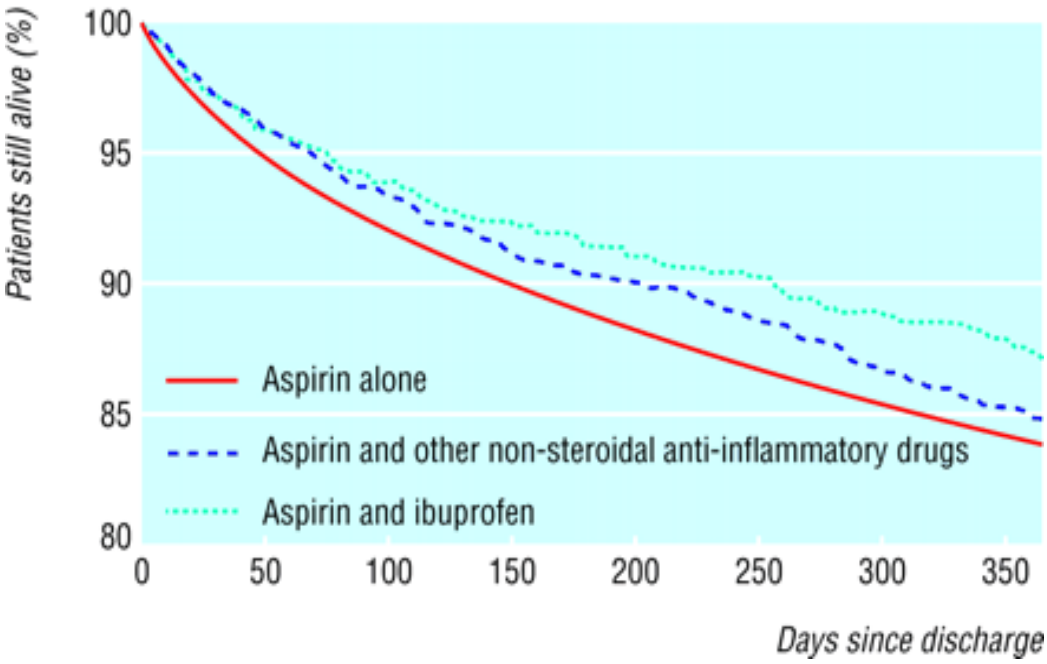

Figure 1-1 An example of survival curves: elderly US survivors of myocardial infarction prescribed aspirin alone, aspirin and ibuprofen, or aspirin and a different non-steroidal anti-inflammatory drug on discharge from hospital. [2]

with respect to time and we have $0 \leq S(u) \leq S(t) \leq 1$ for all $u > t$. The time $t = 0$ represents some origin, typically the beginning of a study or the start of operation of a system. $S(0)$ is commonly unity but can be less to represent the probability that the system fails immediately upon operation. Furthermore, under standard scenario, we have $S(\infty)$ to be zero.

1.2.2 Cumulative distribution function

The cumulative distribution function, conventionally denoted by F, is defined as the complement of the survival function, i. e.,

$$F(t) = P(T \leq t) = 1 - S(t). \quad (1.2)$$

If *F* is differentiable, then the derivative of *F*(*t*) with respect to *t* is called the probability density function and is given by

$$f(t) = F'(t) = \frac{dF(t)}{dt}. \quad (1.3)$$

The probability density function *f* is sometimes called the event density, which is the rate of event occurrence per unit time.

1.2.3 Hazard function

An alternative characterization of the distribution of *T* is given by the hazard function, which is also known as the instantaneous failure rate or hazard rate, and is defined as

$$H(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}. \quad (1.4)$$

The numerator of this expression is the conditional probability that the event will occur in the interval [*t, t*+Δ*t*) given that it has not occurred before. Thus, the limit of the ratio (hazard function) indicates the instantaneous rate of event occurrence at time *t* given that it did not occur before. [3] It can also be expressed as the ratio of the probability density function *f*(*t*) to the survival function *S*(*t*), i. e.,

$$H(t) = \frac{f(t)}{S(t)} = -[\frac{d}{dt}\{S(t)\}]/S(t) = -\frac{d}{dt}\log S(t). \quad (1.5)$$

Like the survival function, the hazard function must also be non-negative. While it is different to the survival function with respect to the decreasing property, the hazard function can be increasing, decreasing or constant.

1.3 Censoring

Observations are called censored when information about them is incomplete. A sample contains censored observations if the only information available about some of the observations is that they are below or above a specified value. Censoring is an important issue in survival analysis, representing a particular type of missing data.

Usually, there are three main types of censoring in survival analysis: left censoring, right censoring and interval censoring. Left censoring occurs when the event of interest had already occurred even before we started to make observation on a unit, which is very rarely encountered. Right censoring occurs when the actual event time is not observed but is known to exceed a certain value. It is the most common form of censoring encountered in practice. It often arises because the data collection had to be stopped at a certain time for practical considerations. Interval censoring occurs if the observation is not made continuously over time but at specific time point only. In this case, when a failed unit is observed, its exact failure time is not known, but is known to have failed between the present and previous inspection times.

1.4 Modeling survival time

Unlike ordinary regression models, survival methods correctly incorporate information from both censored and uncensored observations in estimating important model parameters. In this section, we review some commonly used models in survival analysis. A number of models are available to analyze the relationship of a set of predictor variables with the survival time. Methods include parametric, nonparametric and semi-parametric approaches. Each model has its own advantages and disadvantages.

The range of survival analysis models vary from the fully non-parametric, like the Kaplan-Meier method, to semi-parametric models and to fully parametric models where we specify the distribution of the underlying hazard.

1.4.1 Nonparametric modeling

Nonparametric methods do not require the knowledge of the underlying distribution of the failure time *T*. Hence, it provides a flexible way to deal with the data in many practical situations. A nonparametric estimator of the survival function, the Kaplan Meier method, is widely used to estimate and graph survival probabilities as a function of time. It can be used to obtain univariate descriptive statistics for survival data, including the median survival time, and compare the survival experience for two or more groups of subjects.

1.4.2 Semi-parametric modeling

In statistics, a semi-parametric model is a statistical model that has parametric and nonparametric components. The Cox proportional hazards regression model is an example of semiparametric model, which is popular for the analysis of survival data. It has more assumptions than those nonparametric methods described above, while in contrast to the parametric models, it makes no assumptions about the shape of the so-called baseline hazard function.

Parametric regression models for analyzing lifetime data assume that a particular parametric distribution, for instance, the Weibull distribution, is suitable to model the lifetime. Although one can check for the distributional assumption and there might be extensive past experience that suggest the suitability of a particular parametric distribution, in biomedical science, the situation may be different. With humans as the experimental units, every population has its own characteristics and may be different from other. Hence, previous experience may not be suitable to guide what might happen in the particular

population under study. Thus, it is important to analyze lifetime data without making any particular distributional assumption for the lifetime.

1.4.3 Parametric modeling

Parametric methods assume that the underlying distribution of the survival times follow certain known probability distributions. Popular ones include the exponential, Weibull, lognormal, and gamma distributions. The description of the distribution of the survival times and the change in their distribution as a function of predictors is of interest. Model parameters in these settings are usually estimated using an appropriate modification of maximum likelihood.

In this study, we have considered a full parametric setup and assumed right censoring as the form of censoring. The main objective is to consider the wider class of generalized gamma distribution for the lifetime and then find the maximum likelihood estimates of the model parameters using some optimization technique available in R statistical software.

The other objective is to carry out an extensive Monte Carlo simulation study to demonstrate the performance of the estimation method and also to demonstrate the flexibility of the generalized gamma family. To demonstrate the flexibility of the generalized gamma family, we carried out a model discrimination using both the likelihood ratio test and information-based criteria.

The remaining part of this thesis is organized as follows:

In chapter 2, we introduce the generalized gamma distribution and discuss some of its special cases, such as Weibull, lognormal, and gamma distributions. The form of the survival data is also introduced in this chapter. For the purpose of simulation study, an algorithm to generate data is discussed.

In chapter 3, we discuss the parameter estimation by maximum likelihood method for the generalized gamma family. An extensive Monte Carlo simulation study is carried out to determine the performance of the estimation method and results obtained are discussed in detail.

In chapter 4, the flexibility of generalized gamma family is utilized to carry out model discrimination and model selection within this family by using the likelihood-based method and information based criteria, namely the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

In chapter 5, we illustrate the method of estimation and the flexibility of the generalized gamma family using a real data on lung cancer.

Finally, in chapter 6, we make some concluding remarks.

Chapter 2 Model setup with generalized gamma distribution

2.1 Generalized gamma distribution

The early generalization of gamma distribution can be traced back to Amoroso who discussed the generalized gamma distribution and applied it to fit income rates. [4] Johnson et al. [5] proposed a four parameter generalized gamma distribution which reduces to the generalized gamma distribution with three parameters defined by Stacy [6] when the location parameter is set to zero. Agarwal and Al-Saleh used generalized gamma distribution to study hazard rates. [7] Balakrishnan and Peng used this distribution in the context of frailty model. [8] Cordeiro et al. [9] derived another generalization of Stacy's generalized gamma distribution using exponentiated method, and applied it to lifetime and survival analysis.

The generalized gamma distribution, which was introduced by Stacy, [10] presents a flexible family with varying shapes and hazard functions often suitable for modeling survival data. It is a three-parameter distribution with its probability density function given by

$$f(t;\,\boldsymbol{\theta}) = \begin{cases} \dfrac{q\left(q^{-2}\right)^{q^{-2}}(\lambda t)^{q^{-2}\left(\frac{q}{\sigma}\right)}\exp\left[-q^{-2}(\lambda t)^{\left(\frac{q}{\sigma}\right)}\right]}{[\Gamma(q^{-2})\sigma t]}, & q > 0, \\ (\sqrt{2\pi}\,\sigma t)^{-1}\exp\left\{-\dfrac{[\log(\lambda t)]^2}{2\sigma^2}\right\}, & q = 0, \end{cases} \tag{2.1}$$

where $t > 0$ is the lifetime, $\boldsymbol{\theta} = (q, \sigma, \lambda)'$ with $q \geq 0$, $\sigma > 0$, and $\lambda > 0$. Here, σ and q are the shape parameters and λ is a scale parameter. $\Gamma(\cdot)$ represents the complete gamma function.

The family of generalized gamma distribution include several commonly used lifetime distributions as its particular case. For instance, (2.1) reduces to

1) the Weibull distribution when $q = 1$,
2) the lognormal distribution when $q = 0$, and
3) the gamma distribution when $q/\sigma = 1$.

Thus, it has considerable flexibility to capture the properties of a distribution that may not be possible when using its special case. This motivates us to use the generalized gamma distribution as the distribution of the lifetime as it would enable us to carry out a model discrimination within this family to select a particular parametric lifetime distribution that provides the best fit to the data. The corresponding survival function for the expression (2.1) is given by

$$S(t;\lambda,\sigma,q) = \begin{cases} \frac{\Gamma\left(q^{-2},q^{-2}(\lambda t)^{\frac{q}{\sigma}}\right)}{\Gamma(q^{-2})}, & q > 0, \\ 1 - \Phi\left(\frac{\log(\lambda t)}{\sigma}\right), & q = 0, \end{cases} \tag{2.2}$$

where

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

is the upper incomplete gamma function and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. [10]

We next discuss the form of the density function and survival function for the special cases of generalized gamma distribution.

1. Weibull distribution:

If $q = 1$ in (2.1), the density and survival functions are reduced to

$$f_w(t) = (\lambda t)^{1/\sigma} \exp[-(\lambda t)^{1/\sigma}]/\sigma t, \text{ and} \tag{2.3}$$

$$S_w(t) = \exp[-(\lambda t)^{1/\alpha}]. \quad (2.4)$$

2. Lognormal distribution:

If $q = 0$ in (2.1), the density and survival functions are reduced to

$$f_l(t) = (\sqrt{2\pi}\sigma t)^{-1}\exp\{-\frac{[log(\lambda t)]^2}{2\sigma^2}\}, \text{ and} \quad (2.5)$$

$$S_l(t) = 1 - \Phi\left[\frac{\log(\lambda t)}{\sigma}\right]. \quad (2.6)$$

3. Gamma distribution:

If $q/\sigma = 1$ in (2.1), the density and survival functions are reduced to

$$f_g(t) = \sigma(\frac{\lambda t}{\sigma^2})^{\sigma^{-2}}\exp[-(\lambda t)/\sigma^2]/\Gamma(t/\sigma), \text{ and} \quad (2.7)$$

$$S_g(t) = \Gamma[1/\sigma^2,\ (\lambda t)/\sigma^2]/\Gamma(1/\sigma^2). \quad (2.8)$$

Maximum-likelihood estimation of the parameters and quasi maximum likelihood estimators for its subfamily (two-parameter gamma distribution) can be found in. [11-14] Hwang, T. et al. [15] introduced a new moment estimation of parameters of the generalized gamma distribution using its characterization.

2.2 Simulation study: Model setup and data generation

2.2.1 Model setup for simulation study

In this study, we consider a situation where the lifetime T is not completely observed and is subject to right censoring, which is the most common situation encountered in practice. For the *i-th* subject, let Y_i denote the actual failure time and C_i denote the censoring time. In a sample of size n, the observed time is then given by $T_i = \min\{Y_i, C_i\}$ for $i = 1, 2, \ldots, n$. Let $\delta_i = I(Y_i \le C_i)$ denote the right censoring indicator. It is obvious that if $Y_i \le C_i$, then $T_i = \min\{Y_i, C_i\} = Y_i$, and $\delta_i = 1$. While if $Y_i > C_i$, then $T_i = \min$

$\{Y_i, C_i\} = C_i$, and $\delta_i = 0$. Thus, the observed data based on n subjects can be represented by $(t_1, \delta_1), (t_2, \delta_2), \ldots, (t_n, \delta_n)$.

In designing an experiment, subjects are usually divided into two groups: treatment and control. Subjects assigned to the treatment group receive the therapy, whereas those in the control group receive the standard therapy. In our study, we introduce a covariate to indicate which group a subject belongs. Thus, X is a binary covariate and can be defined as follows:

$$X = \begin{cases} 1, & if\ the\ subject\ belongs\ to\ treatment\ group \\ 0, & if\ the\ subject\ belongs\ to\ control\ group. \end{cases}$$

Table 2-1 summaries an example of a data set and variables of interest in this study. Next, we describe the model setup using Weibull as the distribution of the lifetime T. For other distributions, the technique remains the same.

Table 2-1 A sample of data set

Subject No.	t_i	δ_i	X_i	Observed (*Y/N*)	Group (*T/C*)
1	*3.25*	*0*	*0*	*N*	*C*
2	*2.58*	*0*	*1*	*N*	*T*
3	*4.32*	*1*	*0*	*Y*	*C*
4	*7.16*	*1*	*1*	*Y*	*T*
5	*1.57*	*0*	*1*	*N*	*T*
….	…	…	…	…	…
n	*3.42*	*1*	*1*	*Y*	*T*

We let the observed lifetime to follow the Weibull distribution with parameters λ and σ, as given in (2.3). To introduce random censoring, we assume the censoring variable C to follow an exponential distribution with a rate α. The probability of censoring is given by

$$P[Y > C] = \int_{c=0}^{\infty} P[Y > c] f_c(c) dc = \int_{c=0}^{\infty} S(c) \alpha\, e^{-\alpha c} dc,$$

$$\overset{let\ t=\alpha c}{\Longleftrightarrow} \int_{t=0}^{\infty} S(t/\alpha)\ e^{-t} dt. \tag{2.9}$$

Where $f_c(\cdot)$ is the density function of the censoring variable C and $S(\cdot)$ is the survival function of the Weibull distribution.

Note that the above probability can be expressed as

$$p = E\left[S\left(\frac{T}{\alpha}\right)\middle| T \sim \exp(1)\right], \tag{2.10}$$

and it has to be approximated using some Monte Carlo technique. Note also that for other lifetime distributions, such as lognormal, gamma, and generalized gamma, only the functional form of $S(\cdot)$ will differ.

2.2.2 Data generation for simulation

To approximate (2.10), we generate N random samples from exponential distribution with unity rate. Then, the censoring probability p can be approximated as

$$p = \frac{1}{N}\sum_{i=1}^{N} \exp\left[-\left(\frac{\lambda t_i}{\alpha}\right)^{\frac{1}{\sigma}}\right].$$

Note that for a desired value of p and for fixed values of λ and σ, the censoring rate α can be easily determined. For fixed values of p, λ, σ, and the calculated value of α, we follow the following steps to generate data:

Step 1: for the *i-th* subject, generate Y_i from the Weibull distribution with density function given by (2.3) and C_i from an exponential distribution with rate α;

Step 2: set $T_i = \min\{Y_i, C_i\}$;

Step 3: if $T_i = Y_i$, set $\delta_i = 1$, else set $\delta_i = 0$;

Step 4: for the *i-th* subject, generate the group indicator X_i from a Bernoulli distribution with probability of success 0.5.

Using the R software, the computational codes for data generation for all lifetime distributions under consideration in this study have been developed and are presented in Appendix A I.

In case of lognormal, gamma, and generalized gamma distribution, the censoring proportions can be expressed as $p = \frac{1}{N}\sum_{i=1}^{N}\left\{1 - \phi\left[\frac{\log\left(\frac{\lambda t_i}{\alpha}\right)}{\sigma}\right]\right\}$,

$p = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\Gamma\left(\sigma^{-2}, \sigma^{-2}\left(\frac{\lambda t_i}{\alpha}\right)\right)}{\Gamma\left(\sigma^{-2}\right)}\right\}$, and $p = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\Gamma\left(q^{-2}, q^{-2}\left(\frac{\lambda t_i}{\alpha}\right)^{q/\sigma}\right)}{\Gamma\left(q^{-2}\right)}\right\}$, respectively.

The data generation steps for the lognormal, gamma, and generalized gamma distributions remain the same as in the Weibull case except that in Step 1, Y_i is generated from the lifetime distribution under consideration. To generate data from generalized gamma distribution, we made use of the fact that if *W* follows a generalized gamma distribution, then $V = W^{q/\sigma}$ follows a gamma distribution with shape parameter $1/q^2$ and scale parameter $\frac{q^2}{\lambda^{q/\sigma}}$. Hence, $W = V^{\sigma/q}$ is a generalized gamma variable with probability density function given in (2.1).

Chapter 3 Parameter estimation for generalized gamma family

3.1 Estimation of parameters

In this chapter, we shall examine techniques for drawing inference about the distribution of time to some event of interest, based on a sample of right censored survival data. The data consists of a time variable under study and an indicator of whether this time is an actual observed time or a censored time for each of the *n* subjects, as discussed in chapter 2.

Throughout this chapter, it is assumed that the potential censoring is unrelated to the potential event time (independent and identically distributed).

3.1.1 Maximum likelihood estimation

In statistics, maximum likelihood estimation (MLE) is a method of estimating the parameters of a statistical model given observations, by finding the parameter values that maximize the likelihood function. When performing maximum likelihood analysis on data with censored individuals, the likelihood function needs to be expanded to take into account the censored items. If the lifetime T_i is actually observed, the *i-th* subject contributes $f(t_i)$ to the likelihood, whereas if T_i is right censored, the *i-th* subject contributes $p(T > t_i) = S(t_i)$ to the likelihood.

Thus, the observed data likelihood function for a sample of size *n* is given by

$$L\,(\boldsymbol{\theta};\, \mathbf{t},\, \boldsymbol{\delta}) = \prod_{i=1}^{n} f(\boldsymbol{\theta}; t_i)^{\delta_i} S(\boldsymbol{\theta}; t_i)^{1-\delta_i} = \prod_{i:\,\delta_i=1} f(\boldsymbol{\theta}; t_i) \times \prod_{i:\,\delta_i=0} S(\boldsymbol{\theta}; t_i).$$

In practice, it is often more convenient when working with the natural logarithm of the likelihood function, called the log-likelihood function. The corresponding log-likelihood function, without the constant term, is given by

$l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) = \sum_{i:\,\delta i=1} \log[f(\boldsymbol{\theta}; t_i)] + \sum_{i:\,\delta i=0} \log[S(\boldsymbol{\theta}; t_i)].$

The method of maximum likelihood estimation find a value $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ that maximizes the log-likelihood function, i. e.,

$\widehat{\boldsymbol{\theta}} = arg_{\theta}^{max}\{l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta})\}.$

To introduce the covariate effect in our model, we link the covariate *X* to the parameter λ using a log-linear link function, i. e., $\lambda = e^{\beta_0+\beta_1 X}$. Thus, our parameter vector $\boldsymbol{\theta}$ is now defined as $\boldsymbol{\theta} = (q, \sigma, \beta_0, \beta_1)'$. Next, simplified explicit expressions for the observed data log-likelihood function are presented for the generalized gamma distribution and its special cases.

1. Weibull distribution:

$$l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{X}) = (\tfrac{1}{\sigma}) \sum_{i:\,\delta_i=1}[(\beta_0 + \beta_1 X_i) + \sum_{i:\,\delta_i=1} \log(t_i)]$$
$$- \sum_{i:\,\delta_i=1}[\exp(\beta_0 + \beta_1 X_i)\, t_i]^{1/\sigma} - \sum_{i:\,\delta_i=1} \log(\sigma t_i)$$
$$- \sum_{i:\,\delta_i=0}[\exp(\beta_0 + \beta_1 X_i)\, t_i]^{1/\sigma}.$$

2. Lognormal distribution:

$$l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{X}) = (-1) \sum_{i:\,\delta_i=1} \log(\sqrt{2\pi}\,\sigma t_i)$$
$$- \sum_{i:\,\delta_i=1} \{[\beta_0 + \beta_1 X_i + \log(t_i)]^2/(2\,\sigma^2)\}$$
$$+ \sum_{i:\,\delta_i=0} \log\{1 - \phi[\tfrac{\beta_0+\beta_1 X_i+\log(t_i)}{\sigma}]\}.$$

3. Gamma distribution:

$$l(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{X}) = \sum_{i:\,\delta_i=1} \tfrac{[\beta_0+\beta_1 X_i-\log(\sigma^2)]}{\sigma^2} - \sum_{i:\,\delta_i=1} \left[\Gamma(\tfrac{1}{\sigma^2})\right]$$
$$- \sum_{i:\,\delta_i=1} \left[\exp(\beta_0 + \beta_1 X_i)\tfrac{t_i}{\sigma^2}\right] + \sum_{i:\,\delta_i=1} \left[\left(\tfrac{1}{\sigma^2} - 1\right)\log(t_i)\right]$$

$$+\sum_{i:\,\delta_i=0} log\left[\Gamma\left(\frac{1}{\sigma^2}\right), \exp(\beta_0+\beta_1 X_i)\frac{t_i}{\sigma^2}\right] - \sum_{i:\,\delta_i=0} log\left[\Gamma\left(\frac{1}{\sigma^2}\right)\right].$$

4. Generalized gamma distribution:

$$l(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{\delta},\boldsymbol{X}) = (-1)\sum_{i:\,\delta_i=1}\left[\log\left(\frac{q}{\sigma}\right) - 2\left(q^{-2\log(q)}\right)\right] - \sum_{i:\,\delta_i=1}\log[\Gamma(q^{-2})]$$

$$+\left(\frac{1}{q\sigma}\right)\sum_{i:\,\delta_i=1}[\log(t_i) + \beta_0+\beta_1 X_i]$$

$$-(q)^{-2}\sum_{i:\,\delta_i=1}(e^{\beta_0+\beta_1 X_i}t_i)^{\frac{q}{\sigma}} - \sum_{i:\,\delta_i=1}\log(t_i)$$

$$+\sum_{i:\,\delta_i=0}\log\{\Gamma[q^{-2},\ (e^{\beta_0+\beta_1 X_i}t_i)^{\frac{q}{\sigma}}/q^2]\} - \sum_{i:\,\delta_i=0}\log[\Gamma(q^{-2})].$$

We used the "maxLik" function available in R software to carry out the maximum likelihood estimation of the parameter $\boldsymbol{\theta}$.

3.2 Simulation study: model fitting

We carried out an extensive Monte Carlo simulation study to evaluate the performance of the proposed estimation method. In this simulation study, we considered different sample sizes and censoring proportions, so that we can observe the behavior of the model under varying sample sizes and censoring rates. The total sample size was then divided into two groups with one group representing the treatment and the other group representing the control.

Two sample sizes were selected: n = 200 and n = 300 to investigate the performance of the model under small and large sample sizes. Two sets of censoring proportions were considered for the treatment (p_T) and control (p_C) groups: (p_T = 0.3, p_C = 0.2) and (p_T = 0.5, p_C = 0.4) to study the accuracy of the model under low and high censoring rates. Note that if subjects belong to the treatment group, the covariate X takes the value 1 and as such we have $\lambda_T = e^{\beta_0+\beta_1}$. Similarly, for subjects belonging to the control group, $\lambda_C = e^{\beta_0}$. Thus, on selecting values for λ_T and λ_C, the regression parameters β_0 and

β_1 can be easily determined. We chose the true values of λ_T and λ_C as 1.5 and 2, which give us the true value of $\beta_0 = \log(\lambda_c) = 0.405$ and the true value of $\beta_1 = \log(\lambda_T) - \log(\lambda_C) = 0.287$. We also chose the true value of σ as 0.5. In case of the generalized gamma distribution, we chose the true value of q as 1.2. Tables 3-1 to 3-4 present the model fitting results for the Weibull, lognormal, gamma and generalized gamma distributions, respectively.

To examine the performance of the estimates, we calculated the empirical bias and root mean square error (RMSE) of the estimates. We also calculated the coverage probabilities of the confidence interval based on the asymptotic normality of the estimates for different nominal confidence levels. All the simulations were performed using the R statistical software and the results are based on 500 Monte Carlo runs.

It is clear that the estimates coverage to the true parameter values quite accurately for all the four lifetime distributions.

Table 3-1 Estimates, bias, RMSE, standard error (s. e.) and coverage probability (c. p.) for Weibull lifetime

n (p_T, p_C)	Parameter	Estimate	Bias	RMSE	s. e.	c. p.	
						90%	95%
200 (0.3, 0.2)	β_0	0.4076	0.0022	0.0538	0.0562	0.9020	0.9530
	β_1	0.2883	0.0006	0.0708	0.0815	0.9450	0.9730
	σ	0.4954	-0.0046	0.0141	0.0308	0.8900	0.9390
300 (0.3, 0.2)	β_0	0.4077	0.0022	0.0412	0.0460	0.8950	0.9470
	β_1	0.2881	0.0004	0.0573	0.0667	0.9450	0.9770
	σ	0.4975	-0.0025	0.0200	0.0252	0.8900	0.9430
200 (0.5, 0.4)	β_0	0.4107	0.0052	0.0332	0.0702	0.9010	0.9430
	β_1	0.2849	-0.0028	0.0592	0.1061	0.9530	0.9720
	σ	0.4948	-0.0052	0.0224	0.0380	0.8800	0.9290
300 (0.5, 0.4)	β_0	0.4071	0.0016	0.0469	0.0573	0.8920	0.9380
	β_1	0.2891	0.0014	0.0687	0.0863	0.9390	0.9710
	σ	0.4943	-0.0057	0.0241	0.0309	0.8890	0.9340

Table 3-2 Estimates, bias, RMSE, standard error (s. e.) and coverage probability (c. p.) for lognormal lifetime

n (p_T, p_C)	Parameter	Estimate	Bias	RMSE	s. e.	c. p.	
						90%	95%
200 (0.3, 0.2)	β_0	0.4046	-0.0009	0.0480	0.0535	0.8970	0.9560
	β_1	0.2889	0.0012	0.0469	0.0773	0.9520	0.9750
	σ	0.4956	-0.0044	0.0224	0.0283	0.8820	0.9290
300 (0.3, 0.2)	β_0	0.4049	-0.0006	0.0100	0.0437	0.8990	0.9420
	β_1	0.2863	-0.0014	0.0245	0.0632	0.9550	0.9770
	σ	0.4969	-0.0031	0.0283	0.0232	0.8880	0.9350
200 (0.5, 0.4)	β_0	0.4054	-0.0008	0.0374	0.0636	0.9070	0.9580
	β_1	0.2905	0.0028	0.0539	0.0921	0.9610	0.9800
	σ	0.4957	-0.0043	0.0316	0.0359	0.8950	0.9420
300 (0.5, 0.4)	β_0	0.4034	-0.0021	0.0265	0.0520	0.8890	0.9470
	β_1	0.2869	-0.0008	0.0480	0.0754	0.9450	0.9710
	σ	0.4973	-0.0027	0.0100	0.0295	0.8910	0.9520

Table 3-3 Estimates, bias, RMSE, standard error (s. e.) and coverage probability (c. p.) for gamma lifetime

n (p_T, p_C)	Parameter	Estimate	Bias	RMSE	s. e.	c. p.	
						90%	95%
200 (0.3, 0.2)	β_0	0.4098	0.0043	0.0300	0.0562	0.9200	0.9560
	β_1	0.2865	-0.0012	0.0810	0.0815	0.9500	0.9720
	σ	0.4949	0.0050	0.0255	0.0308	0.8720	0.9360
300 (0.3, 0.2)	β_0	0.4051	-0.0004	0.0436	0.0460	0.8840	0.9420
	β_1	0.2899	0.0022	0.0550	0.0667	0.9520	0.9740
	σ	0.4977	-0.0023	0.0261	0.0252	0.8780	0.9220
200 (0.5, 0.4)	β_0	0.4079	0.0024	0.0524	0.0702	0.9000	0.9520
	β_1	0.2851	-0.0025	0.0848	0.1061	0.9480	0.9700
	σ	0.4963	-0.0036	0.0200	0.0380	0.8860	0.9360
300 (0.5, 0.4)	β_0	0.4068	0.0013	0.0373	0.0573	0.9080	0.9500
	β_1	0.2853	-0.0023	0.0485	0.0863	0.9620	0.9780
	σ	0.4980	-0.0019	0.0245	0.0309	0.8920	0.9540

Table 3-4 Estimates, bias, RMSE, standard error (s. e.) and coverage probability (c. p.) for generalized gamma lifetime

n (p_T, p_C)	Parameter	Estimate	Bias	RMSE	s. e.	c. p.	
						90%	95%
200 (0.3, 0.2)	β_0	0.3957	-0.0098	0.0868	0.0835	0.8980	0.9360
	β_1	0.2893	0.0016	0.0818	0.0820	0.9520	0.9760
	σ	0.4862	-0.0138	0.0531	0.0478	0.8780	0.9300
	q	1.2424	0.0424	0.2986	0.2584	0.8960	0.9540
300 (0.3, 0.2)	β_0	0.4027	-0.0028	0.0650	0.0677	0.9040	0.9540
	β_1	0.2909	0.0033	0.0724	0.0673	0.9400	0.9620
	σ	0.4919	-0.0081	0.0398	0.0384	0.8740	0.9420
	q	1.2298	0.0298	0.2011	0.1987	0.9040	0.9660
200 (0.5, 0.4)	β_0	0.3997	-0.0057	0.0983	0.0985	0.9160	0.9720
	β_1	0.2866	-0.0011	0.1048	0.0974	0.9220	0.9640
	σ	0.4808	-0.0191	0.0673	0.0638	0.9020	0.9480
	q	1.2765	0.0765	0.3668	0.3311	0.9220	0.9740
300 (0.5, 0.4)	β_0	0.4009	-0.0045	0.0843	0.0793	0.8760	0.9360
	β_1	0.2925	0.0048	0.0845	0.0801	0.9160	0.9520
	σ	0.4902	-0.0098	0.0534	0.0509	0.8900	0.9500
	q	1.2428	0.0428	0.2788	0.2505	0.8960	0.9560

Note that the bias and standard errors decrease with an increase in sample size. Furthermore, the coverage probabilities are close to the nominal levels. Thus, we can conclude that the estimation method performs very well.

Chapter 4 Model discrimination within the generalized gamma family

Model selection/discrimination plays an important role in selecting a particular model within a family of models or a set of candidate models that provide the best fit to a given data. We have seen that the generalized gamma family is quite flexible and includes several commonly used distributions as special cases, which enables us to select a simple distribution within this family that provide an adequate fit to the survival data. This selection can be carried out using a likelihood ratio test or information-based criteria, namely the AIC and BIC.

4.1 Likelihood ratio test

In this method, for a given distribution of the lifetime, we investigate the performance of the likelihood ratio test in testing the null hypothesis that the lifetime distribution can be described by one of the Weibull (H_0: $q = 1$), lognormal (H_0: $q = 0$), and gamma (H_0: $q = \sigma$) distributions *versus* an alternative hypothesis that the lifetime distribution can be described by a member of generalized gamma family other than the one described in the null hypothesis. The likelihood ratio test statistic is defined as:

$\Lambda = -2\left(\hat{l}_0 - \hat{l}\right)$, where $\hat{l}_0$ and $\hat{l}$ denote the maximized log-likelihood values under H_0 and ($H_0 \cup H_1$), respectively.

Note that under the standard likelihood theory, the asymptotic distribution of Λ under H_0 is a central chi-square distribution with one degree of freedom. However, if the parameter q lie on the boundary of the parameter space, the asymptotic null distribution of Λ is a 50-50 mixture of a point mass at zero and a chi-square distribution with one degree of freedom (mixture chi-square distribution). In case of testing for H_0: $q = 0$ (lognormal), we

note that the parameter (q) lie on the boundary of the parameter space ($q \geq 0$) and the asymptotic null distribution of Λ in this case is given by

$$P[\Lambda \leq \eta] = \frac{1}{2} + \frac{1}{2} P[\chi_1{}^2 \leq \eta],$$

where $\chi_1{}^2$ denotes a chi-square random variable with one degree of freedom. [16, 17]

For each computer generated dataset, we calculated the likelihood ratio test statistic of the fitted Weibull, lognormal and gamma models. Based on 1000 datasets for each simulation and nominal significance level of 5%, we computed the observed significance levels and power values of the likelihood ratio test, which were determined by the rejection rates of the null hypothesis. The results and discussion are presented in subsection 4.3.

4.2 Information-based criteria

In choosing a criterion for model selection, we should accept the fact that models only approximate the reality. Given a set of data, the objective is to determine which of the candidate models best approximates the data, which involves trying to minimize the loss of information. Because the field of information theory is used to quantify or measure the expected value of information, the information-theoretic approach is used to derive the two most commonly used criteria in model selection: the AIC and the BIC. [18, 19]

The AIC is defined as

$$AIC = -2l(\widehat{\boldsymbol{\theta}}) + 2k,$$

where $l(\widehat{\boldsymbol{\theta}})$ is the value of maximized log-likelihood function, and k is the number of estimated parameters in the candidate model. The quantity $2k$ represents the penalty term and a model with minimum value of AIC is considered to be the best among a set of candidate models.

The BIC, proposed by Schwarz, [20] is another model selection criterion based on information theory but defined within a Bayesian context. The BIC is defined as

$$BIC = -2l(\widehat{\boldsymbol{\theta}}) + klog(n),$$

where *n* is the sample size. Compared to AIC, the BIC has a greater penalty for adding parameters to the model. As in the case of AIC, a lower value of BIC implies better fit.

It is to be noted that the AIC and BIC do not involve a formal test of hypothesis, as in the case of likelihood ratio test, so it can tell nothing about the quality of the model in an absolute sense.

For each computer generated dataset, we calculated the AIC and BIC values of the fitted models within the generalized gamma family for a given true model. Based on 1000 datasets in each situation, we calculated the observed selection rates for each of the two criteria used. The results and discussions are presented in the following subsection.

4.3 Simulation results and discussions

The results for the likelihood ratio test and information-based criteria are shown in Tables 4-1 and 4-2, respectively.

From Table 4-1, it can be seen that the chi-square and mixture chi-square distributions provide good approximations to the null distribution of the likelihood ratio test statistic as the observed levels are very close to the nominal level. The observed powers, may vary in different situations. When Weibull distribution is the true lifetime distribution, the test has remarkably high power to reject the lognormal distribution compared to gamma distribution, for any simulation setting. When lognormal distribution is the true lifetime distribution, the test has almost 100% power to reject the Weibull distribution and a moderate power to reject the gamma. When the true lifetime distribution is gamma, the

Table 4-1 Observed levels (in bold) and powers of likelihood ratio test

Fitted model	True generalized gamma model		
	Weibull	Lognormal	Gamma
n = 200, (p_T, p_C) = (0.3, 0.2)			
Weibull	**0.047**	0.999	0.672
Lognormal	0.999	**0.051**	0.756
Gamma	0.724	0.762	**0.056**
n = 200, (p_T, p_C) = (0.5, 0.4)			
Weibull	**0.056**	0.984	0.514
Lognormal	0.987	**0.050**	0.666
Gamma	0.575	0.624	**0.052**
n = 300, (p_T, p_C) = (0.3, 0.2)			
Weibull	**0.048**	1.000	0.846
Lognormal	1.000	**0.050**	0.896
Gamma	0.847	0.911	**0.064**
n = 300, (p_T, p_C) = (0.5, 0.4)			
Weibull	**0.052**	0.999	0.723
Lognormal	0.997	**0.044**	0.800
Gamma	0.747	0.814	**0.053**

power to reject the Weibull and lognormal distributions are moderate. As expected, the power to reject the wrong model increases with an increasing in sample size and decrease in censoring proportion. Based on these observations, we can conclude that the likelihood ratio test can distinctly discriminate between the lognormal and Weibull models with very high power.

Table 4-2 Observed selection rates based on AIC*

Fitted model	True generalized gamma model		
	Weibull	Lognormal	Gamma
	n = 200, (p_T, p_C) = (0.3, 0.2)		
Weibull	**0.909**	0.000	0.111
Lognormal	0.000	**0.877**	0.105
Gamma	0.091	0.123	**0.784**
	n = 200, (p_T, p_C) = (0.5, 0.4)		
Weibull	**0.849**	0.001	0.165
Lognormal	0.003	**0.845**	0.145
Gamma	0.148	0.154	**0.690**
	n = 300, (p_T, p_C) = (0.3, 0.2)		
Weibull	**0.933**	0.000	0.080
Lognormal	0.000	**0.912**	0.070
Gamma	0.067	0.088	**0.850**
	n = 300, (p_T, p_C) = (0.5, 0.4)		
Weibull	**0.908**	0.001	0.122
Lognormal	0.000	**0.871**	0.109
Gamma	0.092	0.128	**0.769**

**The observed selection rates based on BIC turn out to be exactly the same as AIC*

From Table 4-2**,** we first note that when the true lifetime distributions are either Weibull or lognormal, the selection criteria performs well in selecting the correct model. In both these cases, although there is a small chance of selecting gamma, the rates of Weibull selecting lognormal and lognormal selecting Weibull are nearly zero. In addition, the correct selection rates increase with an increase in sample size and with a decrease in censoring

proportion. When the true lifetime distribution is gamma, the selection criteria performs moderately in selecting the correct model.

Chapter 5 Illustrative example

In this chapter, we demonstrate an application of proposed methodology to a data on lung cancer. The data is taken from Loprinzi, C. L., et al. [21] and represents survival in patients with advanced lung cancer from the North Central Cancer Treatment Group. Table 5-1 lists the data variables and the corresponding descriptions. The total number of subjects in this data is 228, and the percentage of censored observations is 28%. In addition, the observed time has mean = 305.2 days and standard deviation= 210.6. The data can be easily accessed in the package "survival" of R statistical software.

Table 5-1 Variables and descriptions for lung cancer data

Variable	Description
time	survival time in days
status	censoring status: 1=censored, 2=observed
sex	male=1, female=2

Using sex as the only covariate, we fit different lifetime distributions to the data and the maximum likelihood estimate of the model parameters are presented in Table 5-2.

Table 5-2 Estimates of model parameters for different lifetime distributions

Model	Parameter estimates				Maximized likelihood function
	β_0	β_1	σ	q	
Weibull ($q = 1$)	-6.2797	0.3956	0.7551	-	-1148.652
Lognormal ($q = 0$)	-6.018	0.5714	1.0711	-	-1162.616
Gamma ($q = \sigma$)	-6.2332	0.4279	0.8054	-	-1149.202
Generalized gamma	-6.2834	0.3931	0.7513	1.0156	**-1148.649**

Based on the results of Table 5-2, it is clear that the Weibull distribution provides a similar fit to that of the generalized gamma distribution. Next, in Table 5-3, we present the AIC and BIC values for different lifetime distributions.

Table 5-3 AIC and BIC values for different lifetime distributions

Model	AIC	BIC
Weibull	**2303.304**	**2313.592**
Lognormal	2331.232	2341.52
Gamma	2304.404	2314.692
Generalized gamma	2305.298	2319.015

It is clear from Table 5-3 that both the AIC and BIC values for the Weibull distribution are the minimum. Thus, based on the information criteria, Weibull should be considered as our working model.

We also carried out the likelihood ratio test to test for the suitability of Weibull, lognormal and gamma models and the results are presented in Table 5-4.

Table 5-4 The p-value of likelihood-ratio test for different models

Fit model	Λ	*p*-value
Weibull	0.006	0.94
Lognormal	27.934	< 0.001
Gamma	1.106	0.29

It is clear that the assumption of lognormal model is rejected by the likelihood ratio test and at 5% level of significance, both the Weibull and gamma models turn out to provide adequate fit to the data. Note that the AIC and BIC values for the Weibull and gamma models are quite close to each other.

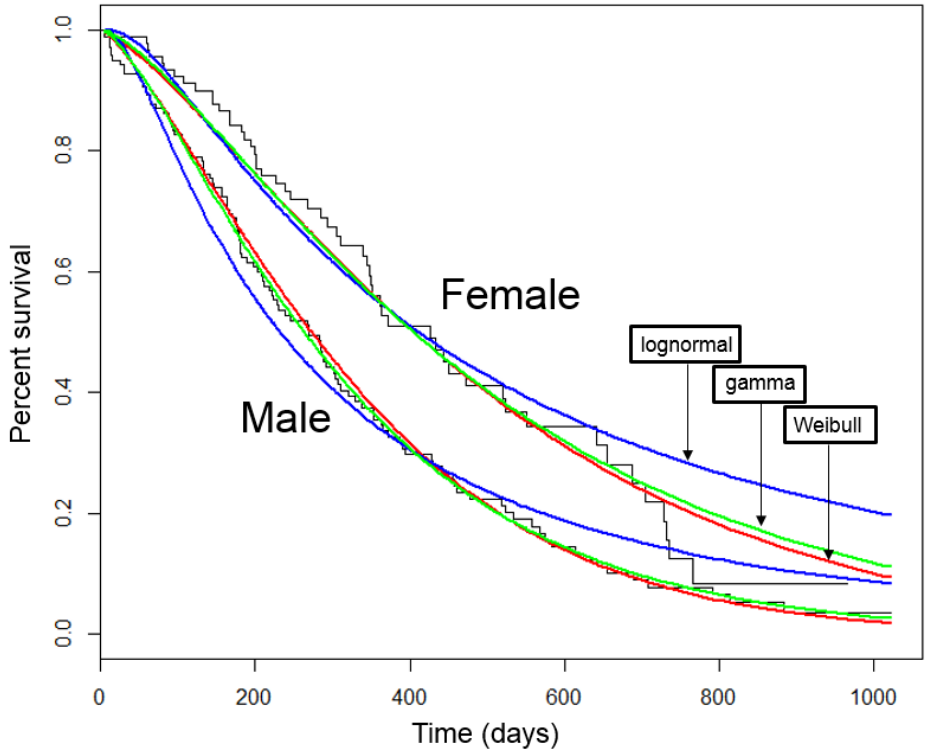

Figure 5-1 Kaplan-Meier curves stratified by sex: male and female, respectively. And the fitted curves of survival function for Weibull, lognormal, and gamma distributions.

In Figure 5-1, we present the Kaplan-Meier estimates of the survival function and superimpose the estimated survival functions from different lifetime distributions. The plots are stratified by patient's gender. It is clear that the Weibull distribution provides a closer concordance to the Kaplan-Meier curves which illustrates that the Weibull provides the best fit.

Chapter 6 Conclusion

In this thesis, we have considered analysis of right censored survival data assuming a flexible distribution to model the lifetime. The flexibility of the generalized gamma distribution enable us to carry out a formal test of hypothesis to check for the suitability of a particular distribution within its family. We also studied the effect of covariates on the lifetime. Our simulation study results show that the maximum likelihood estimation procedure works very well. Our study also show that the likelihood ratio test can distinctly discriminate between Weibull and lognormal lifetimes, whereas the information-based criteria can select the correct model with high accuracy.

Appendix A

R codes developed in this study

Appendix A I. R codes for data generation

1. For Weibull distribution

```
cen_rate=function(alpha,lambda,sigma,prop){
t=rexp(1000, rate=1)
(sum(exp(-((lambda*t/alpha)^(1/sigma))))/length(t))-prop
}# generalize the rate for exponial data distribution
data_Wei=function(n, lam_c, lam_t, sig, pt, pc){
x=rbinom(n, size=1, prob=0.5)
Y=rep(NA,n)
T=rep(NA,n)
d=rep(NA,n)
C=rep(NA,n)
for(k in 1:n){
if(x[k]==1){
Y[k]=rweibull(1,shape=1/sig, scale=1/lam_t)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_t, sigma=sig, prop=pt)$root)
T[k]=min(Y[k],C[k])
}else{
Y[k]=rweibull(1,shape=1/sig,scale=1/lam_c)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_c, sigma=sig, prop=pc)$root)
T[k]=min(Y[k],C[k]) } }#end of for
for(j in 1:n){
if(T[j]==Y[j]){
d[j]=1
}else{
d[j]=0 } }
return(data.frame(T,x,d)) }#generalize the data for Weibull distribution
```

2. For lognormal distribution

```
 cen_rate=function(alpha,lambda,sigma,prop){
t<-rexp(1000, rate=1)
(sum(1-pnorm((1/sigma)*log(lambda*t/alpha), mean=0, sd=1, lower.tail=TRUE,log.p=FALSE))/length(t))-prop}
data_Lognormal=function(n, lam_c, lam_t, sig, pt, pc){
x=rbinom(n,size=1,prob=0.5)
Y=rep(NA,n)
T=rep(NA,n)
d=rep(NA,n)
C=rep(NA,n)
for(k in 1:n){
if(x[k]==1){
Y[k]=rlnorm(1,meanlog=log(1/lam_t),sdlog=sig)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_t,sigma=sig,prop=pt)$root)
T[k]=min(Y[k],C[k])
}else{
Y[k]= rlnorm(1,meanlog=log(1/lam_c),sdlog=sig)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_c,sigma=sig,prop=pc)$root)
T[k]=min(Y[k],C[k]) } }#end of for
for(j in 1:n){
if(T[j]==Y[j]){
d[j]=1
}else{
d[j]=0 } }
return(data.frame(T,x,d)) }
```

3. For gamma distribution

```
cen_rate=function(alpha,lambda,sigma,prop){
t=rexp(1000,rate=1)
((sum(Rgamma(sigma^(-2), (lambda*t*sigma^(-2))/alpha, lower=FALSE, log=FALSE)))/length(t))-prop}
data_Gamma=function(n, lam_c, lam_t, sig, pt, pc){
x=rbinom(n,size=1,prob=0.5)
Y=rep(NA,n)
T=rep(NA,n)
d=rep(NA,n)
C=rep(NA,n)
for(k in 1:n){
if(x[k]==1){
Y[k]=rgamma(1,shape=sig^(-2),scale=(sig^(2)/lam_t))
C[k]=rexp(1,rate=uniroot(cen_rate,c(0, 2),
lambda=lam_t,sigma=sig,prop=pt)$root)
T[k]=min(Y[k],C[k])} else{
Y[k]=rgamma(1,shape=sig^(-2),scale=(sig^(2)/lam_c))
C[k]=rexp(1,rate=uniroot(cen_rate,c(0, 2),
lambda=lam_c,sigma=sig,prop=pc)$root)
T[k]=min(Y[k],C[k]) } }#end of for
for(j in 1:n){
if(T[j]==Y[j]){
d[j]=1} else{
d[j]=0 } }
return(data.frame(T,x,d))
}###end of data generation Gamma
```

4. For generalized gamma distribution

```
cen_rate=function(alpha, lambda, sigma, Q, prop){
t=rexp(1000, rate=1)
((sum(Rgamma(Q^(-2), (Q^(-2))*((lambda*(t/alpha))^(Q/sigma)), lower=FALSE)))/length(t))-prop}
rGen=function(n, lambda, sigma, Q)  {
G=rgamma(n,shape=Q^(-2), scale=(Q^(2))/(lambda^(Q/sigma)))
V=rep(NA, n)
for (i in 1:n) {
V[i]=G[i]^(sigma/Q) }
return(V) }
data_GenGamma=function(n, lam_c, lam_t, sig, q, pt, pc){
x=rbinom(n,size=1,prob=0.5)
Y=rep(NA,n)
T=rep(NA,n)
d=rep(NA,n)
C=rep(NA,n)
for(k in 1:n){
if(x[k]==1){
Y[k]=rGen(1, lambda=lam_t, sigma=sig, Q=q)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_t,sigma=sig,Q=q, prop=pt)$root)
T[k]=min(Y[k],C[k])} else{
Y[k]=rGen(1, lambda=lam_c, sigma=sig, Q=q)
C[k]=rexp(1,rate=uniroot(cen_rate,c(0,10),
lambda=lam_c,sigma=sig, Q=q, prop=pc)$root)
T[k]=min(Y[k],C[k]) } }#end of for
for(j in 1:n){
if(T[j]==Y[j]){
d[j]=1} else{
d[j]=0 } }
return(data.frame(T,x,d)) }
```

Appendix A II. R codes for parameter estimation

1. For Weibull distribution

```
Complex_Wei=function (n, lam_c, lam_t, sig, pt, pc, m){
b0_est=rep(NA, m)
b1_est=rep(NA, m)
sig_est=rep(NA, m)
sd_b0=rep(NA, m)
sd_b1=rep(NA, m)
sd_sig=rep(NA, m)
temp_b0=rep(NA, m)
temp_b1=rep(NA, m)
temp_sig=rep(NA, m)
L_b0_95=rep(NA, m)
U_b0_95=rep(NA, m)
L_b1_95=rep(NA, m)
U_b1_95=rep(NA, m)
L_sig_95=rep(NA, m)
U_sig_95=rep(NA, m)
L_b0_90=rep(NA, m)
U_b0_90=rep(NA, m)
L_b1_90=rep(NA, m)
U_b1_90=rep(NA, m)
L_sig_90=rep(NA, m)
U_sig_90=rep(NA, m)
count_b0_95=0
count_b1_95=0
count_sig_95=0
count_b0_90=0
count_b1_90=0
count_sig_90=0
for (r in 1:m) {
data=data_Wei(n, lam_c, lam_t, sig, pt, pc)
data_obs=data[data$d==1,]
data_cens=data[data$d==0,]
obs_t=data_obs$T
obs_x=data_obs$x
cens_t=data_cens$T
cens_x=data_cens$x
NR_Wei=function(param=c(beta0,beta1,sigma)){
((1/param[3])*((sum(param[1]+(param[2]*obs_x)))+(sum(log(obs_t)))))-
sum((obs_t*exp(param[1]+(param[2]*obs_x)))^(1/param[3]))-sum(log(obs_t*param[3]))-
sum((cens_t*exp(param[1]+(param[2]*cens_x)))^(1/param[3]))
}#estimation function for test
est=maxLik(NR_Wei,start=c(0.35,0.25,0.4),method="NR")
b0_est[r]=coef(est)[1]
b1_est[r]=coef(est)[2]
sig_est[r]=coef(est)[3]
sd_b0[r]=stdEr(est)[1]
sd_b1[r]=stdEr(est)[2]
sd_sig[r]=stdEr(est)[3]
temp_b0[r]=b0_est[r]-log(lam_c)
temp_b1[r]=b1_est[r]-(log(lam_t)-log(lam_c))
temp_sig[r]=sig_est[r]-sig
L_b0_95[r]=b0_est[r]-1.96*sd_b0[r]
U_b0_95[r]=b0_est[r]+1.96*sd_b0[r]
L_b1_95[r]=b1_est[r]-1.96*sd_b1[r]
U_b1_95[r]=b1_est[r]+1.96*sd_b1[r]
L_sig_95[r]=sig_est[r]-1.96*sd_sig[r]
U_sig_95[r]=sig_est[r]+1.96*sd_sig[r]
```

```
L_b0_90[r]=b0_est[r]-1.645*sd_b0[r]
U_b0_90[r]=b0_est[r]+1.645*sd_b0[r]
L_b1_90[r]=b1_est[r]-1.645*sd_b1[r]
U_b1_90[r]=b1_est[r]+1.645*sd_b1[r]
L_sig_90[r]=sig_est[r]-1.645*sd_sig[r]
U_sig_90[r]=sig_est[r]+1.645*sd_sig[r]
if ( (L_b0_95[r]<log(lam_c)) & (log(lam_c)<U_b0_95[r]) ){
count_b0_95=count_b0_95+1 }
if ( (L_b1_95[r]<(log(lam_t)-log(lam_c))) & ((log(lam_t)-log(lam_c))<U_b0_95[r]) ){
count_b1_95=count_b1_95+1 }
if ( (L_sig_95[r]<sig) & (sig<U_sig_95[r]) ){
count_sig_95=count_sig_95+1 }
if ( (L_b0_90[r]<log(lam_c)) & (log(lam_c)<U_b0_90[r]) ){
count_b0_90=count_b0_90+1 }
if ( (L_b1_90[r]<(log(lam_t)-log(lam_c))) & ((log(lam_t)-log(lam_c))<U_b0_90[r]) ){
count_b1_90=count_b1_90+1 }
if ( (L_sig_90[r]<sig) & (sig<U_sig_90[r]) ){
count_sig_90=count_sig_90+1 }}####end of for
ave_b0=sum(b0_est)/m
ave_b1=sum(b1_est)/m
ave_sig=sum(sig_est)/m
Rmse_b0=sqrt((1/(m-1))*sum((temp_b0[r])^2))
Rmse_b1=sqrt((1/(m-1))*sum((temp_b1[r])^2))
Rmse_sig=sqrt((1/(m-1))*sum((temp_sig[r])^2))
bias_b0=sum(temp_b0)/m
bias_b1=sum(temp_b1)/m
bias_sig=sum(temp_sig)/m
cp_b0_95=count_b0_95/m
cp_b1_95=count_b1_95/m
cp_sig_95=count_sig_95/m
cp_b0_90=count_b0_90/m
cp_b1_90=count_b1_90/m
cp_sig_90=count_sig_90/m
ave_sd_b0=(1/m)*(sum(sd_b0))
ave_sd_b1=(1/m)*(sum(sd_b1))
ave_sd_sig=(1/m)*(sum(sd_sig))
return(c(ave_b0, ave_b1, ave_sig, bias_b0, bias_b1, bias_sig, Rmse_b0, Rmse_b1, Rmse_sig, cp_b0_95,
cp_b1_95, cp_sig_95, cp_b0_90, cp_b1_90, cp_sig_90, ave_sd_b0, ave_sd_b1, ave_sd_sig)) }# end of
function complex_Wei
```

The codes for lognormal and gamma distributions are the same as Weibull distribution,

except the log-likelihood function.

2. For lognormal distribution, the log-likelihood function is:

```
NR_Lognormal=function(param=c(beta0,beta1,sigma))   {
(-1)*(sum(log(sqrt(2*pi)*param[3]*obs_t)))-(1/(2*(param[3]^2)))*(sum((param[1]+param[2]*obs_x+log(obs_t))^2))
+(sum(log(1-pnorm((param[1]+param[2]*cens_x+log(cens_t))/param[3])))) }
est=maxLik(NR_Lognormal,start=c(0.35,0.25,0.4), method="NR")
```

3. For gamma distribution, the log-likelihood function is:

```
NR_GenGamma=function(param=c(beta0,beta1,sigma), QQ) {
lamt=exp(param[1]+(param[2]*obs_x))
lamc=exp(param[1]+(param[2]*cens_x))
l = (length(obs_t)*(log(QQ)-log(param[3])-2*(QQ^(-2))*log(QQ)))-(length(obs_t)*log(gamma(QQ^(-2))))+
((1/(QQ*param[3]))*sum(log(lamt*obs_t)))-((QQ^(-2))*sum((lamt*obs_t)^(QQ/param[3])))-sum(log(obs_t))+
sum(log(lgamma(1/(QQ^2),((lamc*cens_t)^(QQ/param[3]))/(QQ^2), lower=F)))-
(length(cens_t)*log(gamma(1/(QQ^2))))
return(l) }
est=maxLik(NR_GenGamma, start=c(0.4, 0.3, 0.6), QQ=q, method="NR")
```

4. For generalized gamma distribution

```
Complex_GenGamma=function (n, lam_c, lam_t, sig, q, pt, pc, m){
b0_est=rep(NA, m)
b1_est=rep(NA, m)
sig_est=rep(NA, m)
Q_est=rep(NA, m)
sd_b0=rep(NA, m)
sd_b1=rep(NA, m)
sd_sig=rep(NA, m)
sd_Q=rep(NA, m)
temp_b0=rep(NA, m)
temp_b1=rep(NA, m)
temp_sig=rep(NA, m)
temp_Q=rep(NA, m)
L_b0_95=rep(NA, m)
U_b0_95=rep(NA, m)
L_b1_95=rep(NA, m)
U_b1_95=rep(NA, m)
L_sig_95=rep(NA, m)
U_sig_95=rep(NA, m)
L_Q_95=rep(NA, m)
U_Q_95=rep(NA, m)
L_b0_90=rep(NA, m)
U_b0_90=rep(NA, m)
L_b1_90=rep(NA, m)
U_b1_90=rep(NA, m)
L_sig_90=rep(NA, m)
U_sig_90=rep(NA, m)
L_Q_90=rep(NA, m)
U_Q_90=rep(NA, m)
count_b0_95=0
count_b1_95=0
count_sig_95=0
count_Q_95=0
count_b0_90=0
count_b1_90=0
count_sig_90=0
count_Q_90=0
for (r in 1:m) {
data=data_GenGamma(n, lam_c, lam_t, sig, q, pt, pc)
data_obs=data[data$d==1,]
data_cens=data[data$d==0,]
obs_t=data_obs$T
obs_x=data_obs$x
cens_t=data_cens$T
cens_x=data_cens$x
NR_GenGamma=function(param=c(beta0,beta1,sigma, Q)) {
lamt=exp(param[1]+(param[2]*obs_x))
lamc=exp(param[1]+(param[2]*cens_x))
l = (length(obs_t)*(log(param[4])-log(param[3])-2*(param[4]^(-2))*log(param[4])))-
(length(obs_t)*log(gamma(param[4]^(-2))))+
((1/(param[4]*param[3]))*sum(log(lamt*obs_t)))-((param[4]^(-2))*sum((lamt*obs_t)^(param[4]/param[3])))-
sum(log(obs_t))+
sum(log(lgamma(1/(param[4]^2),((lamc*cens_t)^(param[4]/param[3]))/(param[4]^2), lower=F)))-
(length(cens_t)*log(gamma(1/(param[4]^2))))
return(l) }
est=maxLik(NR_GenGamma, start=c(0.4, 0.3, 0.6, 0.6), method="NR")
b0_est[r]=coef(est)[1]
b1_est[r]=coef(est)[2]
sig_est[r]=coef(est)[3]
Q_est[r]=coef(est)[4]
```

```
sd_b0[r]=stdEr(est)[1]
sd_b1[r]=stdEr(est)[2]
sd_sig[r]=stdEr(est)[3]
sd_Q[r]=stdEr(est)[4]
temp_b0[r]=b0_est[r]-log(lam_c)
temp_b1[r]=b1_est[r]-(log(lam_t)-log(lam_c))
temp_sig[r]=sig_est[r]-sig
temp_Q[r]=Q_est[r]-q
L_b0_95[r]=b0_est[r]-1.96*sd_b0[r]
U_b0_95[r]=b0_est[r]+1.96*sd_b0[r]
L_b1_95[r]=b1_est[r]-1.96*sd_b1[r]
U_b1_95[r]=b1_est[r]+1.96*sd_b1[r]
L_sig_95[r]=sig_est[r]-1.96*sd_sig[r]
U_sig_95[r]=sig_est[r]+1.96*sd_sig[r]
L_Q_95[r]=Q_est[r]-1.96*sd_Q[r]
U_Q_95[r]=Q_est[r]+1.96*sd_Q[r]
L_b0_90[r]=b0_est[r]-1.645*sd_b0[r]
U_b0_90[r]=b0_est[r]+1.645*sd_b0[r]
L_b1_90[r]=b1_est[r]-1.645*sd_b1[r]
U_b1_90[r]=b1_est[r]+1.645*sd_b1[r]
L_sig_90[r]=sig_est[r]-1.645*sd_sig[r]
U_sig_90[r]=sig_est[r]+1.645*sd_sig[r]
L_Q_90[r]=Q_est[r]-1.645*sd_Q[r]
U_Q_90[r]=Q_est[r]+1.645*sd_Q[r]
if ( (L_b0_95[r]<log(lam_c)) & (log(lam_c)<U_b0_95[r]) ){
count_b0_95=count_b0_95+1 }
if ( (L_b1_95[r]<(log(lam_t)-log(lam_c))) & ((log(lam_t)-log(lam_c))<U_b0_95[r]) ){
count_b1_95=count_b1_95+1 }
if ( (L_sig_95[r]<sig) & (sig<U_sig_95[r]) ){
count_sig_95=count_sig_95+1 }
if ( (L_Q_95[r]<q) & (q<U_Q_95[r]) ){
count_Q_95=count_Q_95+1 }
if ( (L_b0_90[r]<log(lam_c)) & (log(lam_c)<U_b0_90[r]) ){
count_b0_90=count_b0_90+1 }
if ( (L_b1_90[r]<(log(lam_t)-log(lam_c))) & ((log(lam_t)-log(lam_c))<U_b0_90[r]) ){
count_b1_90=count_b1_90+1 }
if ( (L_sig_90[r]<sig) & (sig<U_sig_90[r]) ){
count_sig_90=count_sig_90+1 }
if ( (L_Q_90[r]<q) & (q<U_Q_90[r]) ){
count_Q_90=count_Q_90+1 } }####end of for
ave_b0=sum(b0_est)/m
ave_b1=sum(b1_est)/m
ave_sig=sum(sig_est)/m
ave_Q=sum(Q_est)/m
Rmse_b0=sqrt((1/(m-1))*sum((temp_b0)^2))
Rmse_b1=sqrt((1/(m-1))*sum((temp_b1)^2))
Rmse_sig=sqrt((1/(m-1))*sum((temp_sig)^2))
Rmse_Q=sqrt((1/(m-1))*sum((temp_Q)^2))
bias_b0=sum(temp_b0)/m
bias_b1=sum(temp_b1)/m
bias_sig=sum(temp_sig)/m
bias_Q=sum(temp_Q)/m
cp_b0_95=count_b0_95/m
cp_b1_95=count_b1_95/m
cp_sig_95=count_sig_95/m
cp_Q_95=count_Q_95/m
cp_b0_90=count_b0_90/m
cp_b1_90=count_b1_90/m
cp_sig_90=count_sig_90/m
cp_Q_90=count_Q_90/m
ave_sd_b0=(1/m)*(sum(sd_b0))
ave_sd_b1=(1/m)*(sum(sd_b1))
```

```
ave_sd_sig=(1/m)*(sum(sd_sig))
ave_sd_Q=(1/m)*(sum(sd_Q))
return(c(ave_b0, ave_b1, ave_sig, ave_Q, bias_b0, bias_b1, bias_sig, bias_Q, Rmse_b0, Rmse_b1, Rmse_sig,
Rmse_Q, cp_b0_95, cp_b1_95, cp_sig_95, cp_Q_95, cp_b0_90, cp_b1_90, cp_sig_90, cp_Q_90, ave_sd_b0,
ave_sd_b1, ave_sd_sig, ave_sd_Q)) }# end of function complex_GenGamma
```

Appendix A III. R codes for model discrimination

1. Likelihood ratio test

```
Likelihood_Wei=function(n, lam_c, lam_t, sig, q, pt, pc, m) {
L_Wei=rep(NA, m)
L_LN=rep(NA, m)
L_Gamma=rep(NA, m)
L_GenGamma=rep(NA, m)
Teststat_Wei=rep(NA, m)
Teststat_LN=rep(NA, m)
Teststat_Gamma=rep(NA, m)
P_Wei=rep(NA, m)
P_LN=rep(NA, m)
P_Gamma=rep(NA, m)
count_Wei=0
count_LN=0
count_Gamma=0
for (r in 1:m) {
data=data_Wei(n, lam_c, lam_t, sig, pt, pc)
data_obs=data[data$d==1,]
data_cens=data[data$d==0,]
obs_t=data_obs$T
obs_x=data_obs$x
cens_t=data_cens$T
cens_x=data_cens$x
print(r)
NR_Wei=function(param=c(beta0,beta1,sigma)){
(-1)*(((1/param[3])*((sum(param[1]+(param[2]*obs_x)))+(sum(log(obs_t)))))-
sum((obs_t*exp(param[1]+(param[2]*obs_x)))^(1/param[3]))-sum(log(obs_t*param[3]))-
sum((cens_t*exp(param[1]+(param[2]*cens_x)))^(1/param[3]))) }#estimation function for test
est_Wei=optim(par=c(0.35, 0.25, 0.4), fn=NR_Wei, method="Nelder-Mead")
L_Wei[r]=(-1)*est_Wei$value
NR_LN=function(param=c(beta0,beta1,sigma)){
(-1)*((-1)*(sum(log(sqrt(2*pi)*param[3]*obs_t)))-
(1/(2*(param[3]^2)))*(sum((param[1]+param[2]*obs_x+log(obs_t))^2))+
(sum(log(1-pnorm((param[1]+param[2]*cens_x+log(cens_t))/param[3])))) }#estimation function for test
est_LN=optim(par=c(0.35,0.25,0.4), fn=NR_LN, method="Nelder-Mead")
L_LN[r]=(-1)*est_LN$value
NR_Gamma=function(param=c(beta0,beta1,sigma), QQ) {
lamt=exp(param[1]+(param[2]*obs_x))
lamc=exp(param[1]+(param[2]*cens_x))
l = (-1)*((length(obs_t)*(log(QQ)-log(param[3])-2*(QQ^(-2))*log(QQ)))-(length(obs_t)*log(gamma(QQ^(-2))))+
((1/(QQ*param[3]))*sum(log(lamt*obs_t)))-((QQ^(-2))*sum((lamt*obs_t)^(QQ/param[3])))-sum(log(obs_t))+
sum(log(lgamma(1/(QQ^2),((lamc*cens_t)^(QQ/param[3]))/(QQ^2), lower=F)))-
(length(cens_t)*log(gamma(1/(QQ^2)))))
return(l) }
est_Gamma=optim(par=c(0.35, 0.25, 0.4), QQ=sig, fn=NR_Gamma, method="Nelder-Mead")
L_Gamma[r]=(-1)*est_Gamma$value
NR_GenGamma=function(param=c(beta0,beta1,sigma, Q)) {
lamt=exp(param[1]+(param[2]*obs_x))
lamc=exp(param[1]+(param[2]*cens_x))
l = (-1)*((length(obs_t)*(log(param[4])-log(param[3])-2*(param[4]^(-2))*log(param[4])))-
(length(obs_t)*log(gamma(param[4]^(-2))))+
((1/(param[4]*param[3]))*sum(log(lamt*obs_t)))-((param[4]^(-2))*sum((lamt*obs_t)^(param[4]/param[3])))-
sum(log(obs_t))+
sum(log(lgamma(1/(param[4]^2),((lamc*cens_t)^(param[4]/param[3]))/(param[4]^2), lower=F)))-
(length(cens_t)*log(gamma(1/(param[4]^2)))))
return(l) }
```

```
est_GenGamma=optim(par=c(0.35, 0.25, 0.4, 0.1), fn=NR_GenGamma, method="Nelder-Mead")
L_GenGamma[r]=(-1)*est_GenGamma$value
Teststat_Wei[r]=(-2)*(L_Wei[r]-L_GenGamma[r])
Teststat_LN[r]=(-2)*(L_LN[r]-L_GenGamma[r])
Teststat_Gamma[r]=(-2)*(L_Gamma[r]-L_GenGamma[r])
P_Wei[r]=pchisq(Teststat_Wei[r], lower.tail=FALSE, df=1)
P_LN[r]=1/2-(1/2)*pchisq(Teststat_LN[r], lower.tail=TRUE, df=1)
P_Gamma[r]=pchisq(Teststat_Gamma[r], lower.tail=FALSE, df=1)
if (P_Wei[r]<0.05) {
count_Wei=count_Wei+1 }#end of if
if (P_LN[r]<0.05) {
count_LN=count_LN+1 }#end of if
if (P_Gamma[r]<0.05) {
count_Gamma=count_Gamma+1 }#end of if
prop_Wei=count_Wei/m
prop_LN=count_LN/m
prop_Gamma=count_Gamma/m }##end of for
return(c(prop_Wei, prop_LN, prop_Gamma)) }#end of the function
f=Likelihood_Wei(200, 1.5, 2, 0.5, 0.7, 0.3, 0.2, 100)
```

2. Information-based criteria

```
Choose_Wei=function (n, lam_c, lam_t, sig, pt, pc, m){
L_Wei=rep(NA, m)
AIC_Wei=rep(NA, m)
BIC_Wei=rep(NA, m)
count_Wei_AIC=0
count_Wei_BIC=0
L_Lognormal=rep(NA, m)
AIC_Lognormal=rep(NA, m)
BIC_Lognormal=rep(NA, m)
count_Lognormal_AIC=0
count_Lognormal_BIC=0
L_Gamma=rep(NA, m)
AIC_Gamma=rep(NA, m)
BIC_Gamma=rep(NA, m)
count_Gamma_AIC=0
count_Gamma_BIC=0
for (r in 1:m) {
data=data_Wei(n, lam_c, lam_t, sig, pt, pc)
data_obs=data[data$d==1,]
data_cens=data[data$d==0,]
obs_t=data_obs$T
obs_x=data_obs$x
cens_t=data_cens$T
cens_x=data_cens$x
NR_Wei=function(param=c(beta0,beta1,sigma)){
((1/param[3])*((sum(param[1]+(param[2]*obs_x)))+(sum(log(obs_t)))))-
sum((obs_t*exp(param[1]+(param[2]*obs_x)))^(1/param[3]))-sum(log(obs_t*param[3]))-
sum((cens_t*exp(param[1]+(param[2]*cens_x)))^(1/param[3])) }#estimation function for test
est_Wei=maxLik(NR_Wei,start=c(0.35,0.25,0.4), method="NR")
L_Wei[r]=est_Wei$maximum
AIC_Wei[r]=(-2)*L_Wei[r]+2*3
BIC_Wei[r]=(-2)*L_Wei[r]+3*log(n)
NR_Lognormal=function(param=c(beta0,beta1,sigma))   {
(-1)*(sum(log(sqrt(2*pi)*param[3]*obs_t)))-(1/(2*(param[3]^2)))*(sum((param[1]+param[2]*obs_x+log(obs_t))^2))
+(sum(log(1-pnorm((param[1]+param[2]*cens_x+log(cens_t))/param[3])))) }
est_Lognormal=maxLik(NR_Lognormal,start=c(0.35,0.25,0.4), method="NR")
L_Lognormal[r]=est_Lognormal$maximum
AIC_Lognormal[r]=(-2)*L_Lognormal[r]+2*3
BIC_Lognormal[r]=(-2)*L_Lognormal[r]+3*log(n)
NR_GenGamma=function(param=c(beta0,beta1,sigma), QQ) {
```

```
lamt=exp(param[1]+(param[2]*obs_x))
lamc=exp(param[1]+(param[2]*cens_x))
l = (length(obs_t)*(log(QQ)-log(param[3])-2*(QQ^(-2))*log(QQ)))-(length(obs_t)*log(gamma(QQ^(-2))))+
((1/(QQ*param[3]))*sum(log(lamt*obs_t)))-((QQ^(-2))*sum((lamt*obs_t)^(QQ/param[3])))-sum(log(obs_t))+
sum(log(lgamma(1/(QQ^2),((lamc*cens_t)^(QQ/param[3]))/(QQ^2), lower=F)))-
(length(cens_t)*log(gamma(1/(QQ^2))))
return(l) }
est_Gamma=maxLik(NR_GenGamma, start=c(0.35, 0.25, 0.4), QQ=sig, method="NR")
L_Gamma[r]=est_Gamma$maximum
AIC_Gamma[r]=(-2)*L_Gamma[r]+2*3
BIC_Gamma[r]=(-2)*L_Gamma[r]+3*log(n)
if ((AIC_Wei[r]<AIC_Lognormal[r]) & (AIC_Wei[r]<AIC_Gamma[r])) {
        count_Wei_AIC=count_Wei_AIC+1 }
if ((AIC_Lognormal[r]<AIC_Wei[r]) & (AIC_Lognormal[r]<AIC_Gamma[r])) {
        count_Lognormal_AIC=count_Lognormal_AIC+1 }
if ((AIC_Gamma[r]<AIC_Wei[r]) & (AIC_Gamma[r]<AIC_Lognormal[r])) {
        count_Gamma_AIC=count_Gamma_AIC+1 }
if ((BIC_Wei[r]<BIC_Lognormal[r]) & (BIC_Wei[r]<BIC_Gamma[r])) {
        count_Wei_BIC=count_Wei_BIC+1}
if ((BIC_Lognormal[r]<BIC_Wei[r]) & (BIC_Lognormal[r]<BIC_Gamma[r])) {
        count_Lognormal_BIC=count_Lognormal_BIC+1}
if ((BIC_Gamma[r]<BIC_Wei[r]) & (BIC_Gamma[r]<BIC_Lognormal[r])) {
        count_Gamma_BIC=count_Gamma_BIC+1} }#end of for
return(c(count_Wei_AIC, count_Lognormal_AIC, count_Gamma_AIC, count_Wei_BIC, count_Lognormal_BIC,
count_Gamma_BIC)) }#end of the code
```

References

[1] John, P. K., and Melvin, L. M., Survival analysis techniques for censored and truncated data, 2nd ed. New York: Springer, 2003.

[2] Curtis, J. P., Wang, Y., Portnay, E. L., Masoudi, F. A., Havranek, E. P., and Krumholz, H. M., Aspirin, ibuprofen, and mortality after myocardial infarction: retrospective cohort study, theBMJ, 327,1322, 2003.

[3] Evans, M., Hastings, N., and Peacock, B., Statistical Distributions, 3rd ed. New York: Wiley, 2000.

[4] Amoroso, L. R., Around the income curve, Annals of Mathematica, series 2 (1), 123-159, 1925.

[5] Johnson, N.L., Kotz, S., and Balakrishnan, N., Continuous Univariate Distributions, Vol. 1, 2nd edition, Wiley and Sons, Inc., New York, 1994.

[6] Stacy, E. W., A generalization of the gamma distribution, Annals of Mathematical Statistics, 33 (3), 1187-1192, 1962.

[7] Agarwal, S. K., and Al-Saleh, J. A., Generalized gamma type distribution and its hazard rate function, Communication in Statistics-Theory and Methods, 30 (2), 309-318, 2001.

[8] Balakrishnan, N., and Peng, Y., Generalized gamma frailty model, Statistics in Medicine, 25 (16), 2797-2816, 2006.

[9] Cordeiro, G. M., Ortega, E. M. M. and Silva, G. O., The exponentiated generalized gamma distribution with application to lifetime data, Journal of Statistical Computation and Simulation, 81 (7), 827-842, 2012.

[10] Balakrishnan, N. and Pal, S., An EM algorithm for 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, 2015.

[11] Harter, H. L., Maximum-likelihood estimation of the parameters of a four parameter generalized gamma population from complete and censored samples, Technometrics, 9, 159-165, 1967.

[12] Hirose, H., Maximum likelihood parameters estimation by model augmentation with application to the extended four-parameters Generalized Gamma distribution, Department of control Engineering and Science Kyushu Institute of technology, Fukuoka, 820-850, Japan, 1999.

[13] Stacy, E. W., and Mihram, G. A., Parameter estimation for a generalized gamma distribution, Technometrics, 7, 349-358, 1965.

[14] Stacy, E. W., Quasimaximum likelihood estimators for two-parameter gamma distribution, IBM Journal Research and Development, 17, 115-124, 1973.

[15] Hwang, T., and Huang, P., On new moment estimation of parameters of the Generalized Gamma distribution using its characterization, Taiwanese journal of Mathematics, 10 (4), 1083-1093, 2006.

[16] Self, S. G., and Liang, K. Y., Asymptotic properties of maximum likelihood estimator and likelihood ratio test under nonstandard conditions, Journal of the American Statistical Association, 82, 605-610, 1987.

[17] Claesken, G., Nguti, R., and Jassen, P., One-side tests in shared frailty models, Test, 17, 69-82, 2008.

[18] Fabozzi, F. J., Focardi, S. M., Rachev, S. T. and Arshanapalli, B. G., The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications, John Wiley & Sons, 2014.

[19] Akaike, H., A New Look at the Statistical Model Identification, I.E.E.E. Transactions on Automatic Control, 19, 716-723, 1974.

[20] Schwarz, G., Estimating the Dimension of a Model, Annals of Statistics, 6, 461–464, 1978.

[21] Loprinzi, C. L., Laurie, J. A., Wieand, H. S., Krook, J. E., Novotny, P. J., Kugler, J. W., Bartel, J., Law, M., Bateman, M., and Klatt, N. E., Prospective evaluation of prognostic variables from patient-completed questionnaires, Journal of Clinical Oncology, 12 (3), 601-607, 1994.

Biographical Information

Hongbo Yu moved to Texas, USA from China in Jan, 2015. He studied mathematics in University of Texas at Arlington (UTA) and earned a Master of Science degree. Before joining in UTA, he earned a Bachelor of material science, and a Ph. D in physics from Zhejiang University, and Chinese Academy of Sciences, respectively.

He has more than 10 years working experience as Post-doc (Bilkent University, Turkey and University of Michigan, USA), research engineer (Georgia Institute of Technology, USA) and R&D director (Enraytek Optoelectronics Co., Ltd, Shanghai, China) in both academic and industrial fields. He published more than 50 scientific papers and more than 10 patents in research areas of III-nitride based semiconductor growth and device fabrication.

Hongbo is passionate about reading, running, and popular music. Currently, he resides in Keller, TX with his wife and two daughters.