Extending the Quandt-Ramsey Modeling to Survival Analysis

A Dissertation Presented

by

Paichuan Chen

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

August 2010

Stony Brook University

The Graduate School

Paichuan Chen

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Stephen J. Finch - Dissertation Adviser
Professor, Department of Applied Mathematics and Statistics

Nancy R. Mendell - Chairperson of Defense
Professor, Department of Applied Mathematics and Statistics

Wei Zhu
Professor, Department of Applied Mathematics and Statistics

Derek Gordon
Associate Research Professor, Department of Genetics, Rutgers University

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of Graduate School

Abstract of the Dissertation

Extending the Quandt-Ramsey Modeling to Survival Analysis

by

Paichuan Chen

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2010

The mixture of two regression regimes has been extensively studied in economics. A switching regression is often used to model a system that changes depending on some variables. The test of a mixture of regimes in hazard modeling would be seen to have fundamental importance in biostatistical research but has not been studied. A two-regime parametric mixture is proposed to model the effect of a single covariate on the event time. Typically, the Cox proportional hazards model is applied to estimate a single regime survival regression function. The mixture of two regimes model contains five parameters to be estimated; namely, two parameters to describe each regime, and one to describe the mixing proportion. A software program developed for this research finds the maximum likelihood estimates of the parameters and the likelihood ratio test of the null hypothesis of a single regime against the alternative of a mixture of two regimes. A simulation study finds an approximation to the null distribution of the test and its approximate power.

Table of Contents

Lists of Figures

List of Tables

Acknowledgements

My academic journey has finally come to the end. This has been a profound learning experience and I would like to express my sincere gratitude to all my professors, friends and family. Without their support, advice and help, my academic achievement would not have been attainable.

First of all, I deeply appreciate the help and care given to me from my academic adviser, Dr. Stephen J. Finch. After only two months under his supervision, I was able to present my research project in the Annual Baxter Memorial Seminar of the Department of Applied Mathematics and Statistics. His patience and understanding supported my transformation from a novice not cognizant of the concept or complexities in “Survival analysis” to a graduate now proficient in this subject.

Secondly, I am grateful to Dr. Nancy R. Mendell, Dr. Wei Zhu and Dr. Derek Gordon. Their expert academic mentorship provided invaluable guidance necessary to complete my dissertation.

Thirdly, I am thankful to my friends and student-colleagues in the Department of Applied Mathematics and Statistics, their enthusiasm and encouragement, fostered growth in my academic and social life.

At last, I am indebted to my family for their encouragement, financial support and love. Unfortunately, my mother could not be here at my graduation, but I am proud to say to her up in heaven, “Mom, I did it”.

Chapter 1 Introduction

The mixture of two regression regimes has been extensively studied in economics. The problem was first introduced by Quandt [1] as the switching regression (or switching regimes) problem. A switching regression is often used to model a system that changes depending on some variables. Quandt and Ramsey [2] considered the problem of estimating mixtures of normal distributions.

$$y \sim N(\mu_1, \sigma_1^2) \quad \text{with probability} \quad \pi$$

$$y \sim N(\mu_2, \sigma_2^2) \quad \text{with probability} \quad 1-\pi$$

The problem was to estimate the five parameters ($\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \pi$) from a sample on *y*, and to put into a regression setting by allowing the means μ_1 and μ_2 to be linear functions of explanatory variables i.e. $\mu_1 = \beta_1 x + \varepsilon_1$ and $\mu_2 = \beta_2 x + \varepsilon_2$. This problem is referred to as a "switching regressions" problem.

Survival models are used to analyze time to event data in biostatistics. Mixture models are used increasingly in these analyses. Yamaguchi [3] considered an accelerated failure-time regression model with an additional regression model for long term survivors (LTS) patients. Let *T* be the random variable representing time to event.

Let $f(y)$ be the conditional pdf of *Y*, given that the subject was not a LTS, and let $g(y)$ be the unconditional pdf of *Y*. When the LTS fraction was *p*, then

$$\begin{aligned} g(y) &= (1-p)f(y) \qquad \text{when} \quad y < \infty \\ g(y) &= p \qquad\qquad\qquad\qquad\; y = \infty \end{aligned}$$

Then the survival function corresponding to $g(y)$, $S_g(y)$, can be expressed using $S_f(y)$, the survival function corresponding to $f(y)$ as

$$S_g(y) = (1-p)S_f(y) + p$$

An important task in survival analysis is to investigate how differences in the survival distribution between two treatment groups depend on covariates. Greenhouse [4] discussed an application of the long term survivors (LTS) model to the analysis of clinical trails data. He introduced covariates into the LTS model by allowing functions of parameters p and θ to depend on the covariates of interest. Specifically, he used a linear logistic model for p, the cured fraction, and a log-linear model for θ. His survival function was:

$$P(T > t \mid p, \theta) = S(t \mid p, \theta) = p + (1-p)S_0(t \mid \theta),$$

where p was the cured fraction (that is, those surviving at infinity), and $S_0(t \mid \theta)$ was the surviving distribution for the fraction of the population who were not LTS.

In general, a survival function that satisfies the proportional hazards assumption is given by $S(t \mid x) = [S_0(t)]\exp(\beta' x)$, where *t* is the survival time of an individual with covariate vector x and $S_0(t)$ is the baseline survival function (Perperoglu et al. [5]).

Halabi et al. [6] considered failure time with proportional hazards and baseline exponential survival distribution with exponential and uniform censoring distributions. They generated failure time T with survival function: $S_f(t) = \exp\{-(\lambda_f t\}e^{\beta_j}\}$, $\beta_j = \log(\frac{\lambda_j}{\lambda_1})$ with $\lambda_1 = \lambda_f$, where λ_f was the hazard in the first group, and censoring times C were generated with common survival function $S_c(t) = \exp\{-\lambda_c t\}$, λ_c was the common hazard for the censored observations. They also generated censoring times C by the uniform distribution on $(0, \theta_c)$

Hu et al. [7] considered Cox proportional hazards with covariates that were measured with error $\lambda(t \mid x) = \lambda_0(t)\exp(\beta' x)$. They generated a censoring distribution C that followed an exponential distribution with mean equal to 1.

Buzas [8] considered the model with two covariates. Failure time was related to covariates (X, Z) through the hazard function: $\lambda(t \mid X, Z) = Y(t)\lambda_0(t)\exp(\beta_x^T X + \beta_z^T Z)$, where $Y(t)$ was an indicator function with 1 when $T \geq t$, and 0 otherwise. The failure time was generated exponentially with hazard $\exp\{\beta_x x\}$. Uniformly distributed censoring times were generated such that the expected proportion of censored observations is 0.5.

Kong et al. [9] considered the basic Cox proportional hazards model: $\lambda(t; z(t)) = \lambda_0(t)\exp\{\beta' z(t)\}$. For each fixed *Z*, a failure time *Y* was generated from a proportional hazards model with $\lambda_0(t) = 1$ and a relative risk of $\exp(\beta' Z)$. Type-II censoring was designed so that all individuals after the m^{th} failure were censored. Because of the specific censoring mechanism Kong et al. chose, the baseline hazard after the last failure time cannot be estimated. Hence they chose the time points before the last failure time.

In the models described above, the authors considered two groups, LTS and non-LTS, with covariates in the survival function and for the variable indicating group membership. The model I consider contained two groups, those who had fast conditional response rate and those who had slow conditional response rate. I allowed covariates for the survival function. That is, considered a mixture setting where we assumed X_i was a vector of covariates observed with a response T_i. The goal of mixtures of regressions was to describe the conditional survival distribution.

My research problem was to develop the LRT statistics that test whether there was an indication of a mixture of mechanisms with a covariate that affects the survival time.

My dissertation contains 5 chapters. Chapter 1 contains the introduction and the statement of the research. Chapter 2 of this dissertation presents the methods that I used to find the log-likelihood functions and maximum likelihood estimators. Numerical algorithms were programmed in *R* (version 2.8.0) and Microsoft Visual C++ for Windows 2000/XP. They also can be run in the UNIX operating system. This software is available upon request from me.

Chapter 3 of this dissertation gives the simulation results for the maximum likelihood estimators of single regime model, the null distribution and transformation of the LRT, and the critical values.

Chapter 4 of this dissertation presents the simulation results for the maximum likelihood estimators of mixture of two regimes model, the alternative distribution of the LRT and the power study.

Chapter 5 of this dissertation contains the conclusions and the directions for future study.

Chapter 2 Methods

2.1 Hazard Function

Let T be an exponentially distributed random variable with conditional mean $\frac{1}{\lambda e^{\beta x}}$ and conditional survival function given by $S(t \mid x) = P(T > t \mid x)$. This was the survival function for an uncensored subject with covariate *x*. Let the hazard function of T for an individual given the covariate vector x be given by $\lambda(t \mid x) = \lambda_1(t) e^{\beta' x}$, where $\lambda_1(t)$ was the baseline hazard for an individual with *x*=0, and β was $1 \times p$ vector of regression coefficients common to all individuals. The proportional hazards assumption was often used to describe the effect of *x* on the distribution of the failure time distribution of uncensored subjects (Peng, [10]). This assumption was that the hazard function of a patient with the covariate x at time t was of the form $h(t \mid x) = h_0(t)\exp(\beta x)$. For an exponentially distributed survival time, the survival function was $S(t \mid x) = P(T > t \mid x) = \exp[-\lambda t \exp(\beta x)]$, which satisfied the proportional hazard assumption .

2.2 Definition of the Independent Censoring Variables

With censored data, the survival time t_i^*, *i*=1,….,*n*, was observed only if it did not exceed the censoring time u_i; otherwise, we observed u_i. The absence of censoring indicator c_i took the value 1 if t_i was a survival time (*i.e.* $t_i = t_i^*$) and 0 if t_i was a censoring time (*i.e.* $t_i = u_i$). That is, the observed time was defined by $t_i = \min(t_i^*, u_i)$, $1 \le i \le n$. The absence of censoring indicator $c_i = 1$ when $t_i^* \le u_i$; otherwise, $c_i = 0$. (Maller and Zhou, [11])

2.3 Single Regime Model with Covariate

Let $S(t \mid \theta, x)$ be the conditional survival function which was exponential with mean $\frac{1}{\lambda e^{x\beta}}$, where $\theta = (\lambda, \beta)$, with $0 < \lambda < \infty$ and $-\infty < \beta < \infty$, t was the time to event $(t \ge 0)$, x was the covariate affecting *t* in $S(t \mid x)$, the survival function given covariate *x*. The survival function of this model with covariate was given by $S(t \mid x) = \exp(-\lambda t e^{\beta x})$.

2.3.1 Log Likelihood Function of Single Regime Model

The likelihood function was $\prod_{i=1}^{n} f(t_i \mid x_i)^{c_i} S(t_i \mid x_i)^{1-c_i}$ (Maller and Zhou [11]), where

$$f(t \mid x) = \frac{dF(t \mid x)}{dt} = \frac{d[1 - S(t \mid x)]}{dt} = \lambda e^{x\beta} \exp(-\lambda t e^{x\beta})$$

Then, the likelihood function was

$$L_n(t_i, c_i, \beta, \lambda) = \prod_{i=1}^{n} [\lambda e^{x_i\beta} \exp(-\lambda t_i e^{x_i\beta})]^{c_i} [\exp(-\lambda t_i e^{x_i\beta})]^{(1-c_i)}$$

The log-likelihood function was

$$l_n = \log(L_n) = \sum_{i=1}^{n} \{c_i[\log(\lambda) + x_i\beta - \lambda t_i e^{x_i\beta}] + [(1-c_i)(-\lambda t_i e^{x_i\beta})]\}$$

$$= \sum_{i=1}^{n} \{c_i[\log(\lambda) + x_i\beta] + (-\lambda t_i e^{x_i\beta})\}$$

2.3.2 Maximum likelihood estimators (MLE) of Single Regime Model

Since λ was bounded, there may be boundary complications when solving for the MLE. To avoid this, λ was transformed so that the transformed value was unbounded. My transformation was $\lambda = e^{\mu}$, $-\infty < \mu < \infty$

The log-likelihood function after the transformation was

$$l_n = \sum_{i=1}^{n} \{c_i[\mu + x_i\beta - t_i e^{x_i\beta+\mu}] + [(1-c_i)(-t_i e^{x_i\beta+\mu})]\}$$

$$= \sum_{i=1}^{n} c_i\mu + \sum_{i=1}^{n} c_i x_i\beta - \sum_{i=1}^{n} t_i e^{x_i\beta+\mu}$$

The first derivative of log-likelihood function with respect to μ was

$$\frac{\partial l_n}{\partial \mu} = \sum_{i=1}^{n} c_i - \sum_{i=1}^{n} t_i e^{x_i\beta+\mu}$$

Then $e^{\hat{\mu}} = \dfrac{\sum_{i=1}^{n} c_i}{\sum_{i=1}^{n} t_i e^{x_i\beta}}$

The MLE of $\hat{\lambda} = e^{\hat{\mu}}$ in this model was $\hat{\lambda} = \frac{\sum_{i=1}^{n} c_i}{\sum_{i=1}^{n} t_i e^{x_i \beta}}$.

The first derivative of the log-likelihood function with respect to β was

$$\frac{\partial l_n}{\partial \beta} = \sum_{i=1}^{n} c_i x_i - \sum_{i=1}^{n} x_i t_i e^{x_i \beta + \mu}$$

There was no closed form solution for the MLE of β, $\hat{\beta}$. Therefore, a computational algorithm was needed. I used Nelder-Mead (NM) [12] method to find $\hat{\beta}$. The details are in the section 2.10.

2.4 Mixture of Two Regimes Model with Covariate

Let $S_i(t \mid \theta, x)$, $i = 1,2$ be two conditional survival functions which were exponentially distributed with mean $\frac{1}{\lambda_i e^{x\beta_i}}$, i=1, 2, respectively, where $\theta = (\lambda_1, \lambda_2, \beta_1, \beta_2)$, with $0 < \lambda_1 < \infty$, $0 < \lambda_2 < \infty$, $-\infty < \beta_1 < \infty$, and $-\infty < \beta_2 < \infty$, t was the time to event $(t \geq 0)$, x were the covariates affecting t in $S_i(t \mid x)$, the survival function for regime i given a covariate x. The parameter π was the mixing proportion from the first exponential distribution with conditional mean equal to $\frac{1}{\lambda_1 e^{x\beta_1}}$, The conditional survival function of this mixture model with covariates was given by $S(t \mid x) = \pi S_1(t \mid \theta, x) + (1-\pi) S_2(t \mid \theta, x)$.

2.4.1 Log Likelihood Function of Mixture of Two Regimes Model

Suppose we had data $(t_i, c_{i,} x_i)$, $i = 1,2,......, n,$ where n was the number of subjects. The likelihood function for this data was

$$L_n(t_1,...,t_n,\alpha,\beta_1,\beta_2,\lambda_1,\lambda_2) = \prod_{i=1}^{n} f(t_i \mid x_i,)^{c_i} S(t_i \mid x_i)^{1-c_i}$$, where

$$f(t \mid x) = \frac{dF(t \mid x)}{dt} = \frac{d[1 - S(t \mid x)]}{dt}$$

$$= \pi\lambda_1 e^{\beta_1 x} \exp(-\lambda_1 t e^{\beta_1 x}) + (1-\pi)\lambda_2 e^{\beta_2 x} \exp(-\lambda_2 t e^{\beta_2 x})$$

That is,

$$L_n(t_1,...,t_n,\alpha,\beta_1,\beta_2,\lambda_1,\lambda_2) = \prod_{i=1}^{n} [\pi\lambda_1 e^{\beta_1 x} \exp(-\lambda_1 t_i e^{\beta_1 x}) + (1-\pi)\lambda_2 e^{\beta_2 x} \exp(-\lambda_2 t_i e^{\beta_2 x})]^{c_i}$$

$$\times [\pi \exp(-\lambda_1 t_i e^{\beta_1 x}) + (1-\pi)\exp(-\lambda_2 t_i e^{\beta_2 x})]^{1-c_i}$$

The log-likelihood function was

$$l_n = \log(L_n) = \sum_{i=1}^{n} c_i \log\{\pi\lambda_1 e^{\beta_1 x} \exp[-\lambda_1 t_i \exp(\beta_1 x_i)] + (1-\pi)\lambda_2 e^{\beta_2 x} \exp[-\lambda_2 t_i \exp(\beta_2 x_i)]\}$$

$$+ \sum_{i=1}^{n} (1-c_i)\log\{\pi \exp[-\lambda_1 t_i \exp(\beta_1 x_i)] + (1-\pi)\exp[-\lambda_2 t_i \exp(\beta_2 x_i)]\}$$

As before, I transformed the parameters that had a restricted range to parameters that range from $-\infty$ to $+\infty$ to remove numerical problems due to restrictions in range. That is, let $\lambda_1 = e^{\mu_1}$, $\lambda_2 = e^{\mu_2}$. $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$.

$$l_n = \sum_{i=1}^{n} c_i \log[\pi e^{\mu_1} e^{\beta_1 x_i} \exp(-e^{\mu_1} t_i e^{\beta_1 x_i}) + (1-\pi) e^{\mu_2} e^{\beta_2 x_i} \exp(-e^{\mu_2} t_i e^{\beta_2 x_i})]$$

$$+ \sum_{i=1}^{n} (1-c_i)\log[\pi \exp(-e^{\mu_1} t_i e^{\beta_1 x_i}) + (1-\pi)\exp(-e^{\mu_2} t_i e^{\beta_2 x_i})]$$

This reduced to

$$l_n = \sum_{i=1}^{n} c_i \log[\pi \exp(\mu_1 + \beta_1 x_i - t_i e^{\mu_1+\beta_1 x_i}) + (1-\pi)\exp(\mu_2 + \beta_2 x_i - t_i e^{\mu_2+\beta_2 x_i})]$$

$$+ \sum_{i=1}^{n} (1-c_i) \log[\pi \exp(-t_i e^{\mu_1+\beta_1 x_i}) + (1-\pi)\exp(-t_i e^{\mu_2+\beta_2 x_i})]$$

2.4.2 Maximum Likelihood Estimators (MLE) of Mixture of Two Regimes Model

There were no closed forms for the MLEs of $\lambda_1, \lambda_2, \beta_1, \beta_2$ and π. A computation algorithm was needed, which I reported as one of my dissertation results.

2.5 Censoring Parameter Calculation

I considered an exponential censoring pattern in this simulation study (Peng et al. [13]). Let *U* be the censoring time with probability density function *g(u)*, and let *T* be the failure time with probability density function *f(t)*, where *U* and *T* were independent. Let *c*, $0 < c < \infty$, be the end point of study and *r* be the censoring rate. The censoring rate *r* was defined to be:

$$P(U < T \mid \min(T,U) < c) = 1 - r .$$

Since $U < T$, this reduced to

$$P(U < T \mid U < c) = 1 - r .$$

From the definition of conditional probability,

$$\int_0^c \int_u^\infty f(t)g(u)\,dt\,du = (1-r) \cdot P(\min(T,U) < c) .$$

This equation can then be expressed as:

$$\int_0^c \int_u^\infty f(t)g(u)dtdu = (1-r)\cdot[1-P(\min(T,U)>c)],$$

$$\int_0^c \int_u^\infty f(t)g(u)dtdu = (1-r)\cdot[1-P(T>c,U>c)] \qquad (1)$$

Equation (1) was the starting point in my calculation of censoring parameters.

I used the exponential censoring distribution with 10%, 20% and 30% censoring rates in my simulation study of the null distribution. For my simulations, I set the study length to infinity.

2.5.1 Mean of Censoring Distribution of Single Regime Model

Here the random variable *T* had the exponential distribution with mean $\frac{1}{\lambda e^{\beta x}}$. Here also the censoring time random variable *U* had an exponential distribution with mean $\frac{1}{\mu}$. Ghitany et al. [14] considered the failure time *T* distribution with covariate *X*. The pdf of the failure time *T* was

$$f(t\mid x) = \frac{1}{\lambda e^{\beta x}} e^{\frac{-t}{\lambda e^{\beta x}}},$$

and the pdf of the censoring time *Y* was

$$g(u) = \frac{1}{\mu} e^{\frac{-u}{\mu}}.$$

I conditioned on the values of the covariates and let $A(x) = \lambda e^{\beta x}$. Equation (1) for this specification (which had the expected value of the censoring distribution which was possibly a function of the covariates) was

$$\int_0^c \int_u^\infty \frac{1}{A(x)} e^{\frac{-t}{A(x)}} \cdot \frac{1}{\mu} e^{\frac{-u}{\mu}} \, dtdu = (1-r) \cdot \{1 - P(T > c, U > c)\} \tag{2}$$

The left hand side of equation (2) was

$$\int_0^c e^{\frac{-y}{A(x)}} \cdot \frac{1}{\mu} e^{\frac{-u}{\mu}} \, du = \frac{A(x)}{A(x)+\mu} \left\{ 1 - e^{-\left(\frac{1}{\mu} + \frac{1}{A(x)}\right)c} \right\}.$$

The right hand side of equation (2) was

$$(1-r) \cdot \left\{ 1 - \int_c^\infty \frac{1}{A(x)} e^{\frac{-t}{A(x)}} \, dt \int_c^\infty \frac{1}{\mu} e^{\frac{-u}{\mu}} \, du \right\} = (1-r) \cdot \left\{ 1 - e^{-\left(\frac{1}{\mu} + \frac{1}{A(x)}\right)c} \right\}.$$

Then,

$$\mu = \frac{r}{1-r} A(x) = \frac{r}{1-r} \lambda e^{\beta x} \tag{3}$$

Because of the covariate x in (3), the mean of censoring distribution would change as x changed, which was not a realistic model.

Sy et al. [15] considered a long term survivors model: $S(t) = (1-p) + pS(t \mid C = 1)$, where C was the indicator variable that was 1 if the individual experienced the event and 0 otherwise. Failure time data were generated from a logistic-exponential mixture model, where $p(z) = 1/[1 + \exp\{-(b_0 + b_1 z)\}]$, $S(t \mid C = 1; z) = \exp(-\lambda(z)t)$ and $\lambda(z) = \exp(\beta_0 + \beta z)$. Censoring times U were generated from an exponential distribution with censoring rate λ_c either 0.1 or 0.4, representing mild or heavy censoring, respectively.

Peng et al. [13] considered a long term survivors model: $S(t; \theta, \pi) = 1 - \pi + \pi S_u(t; \theta)$. They considered three distributions as the failure time distribution of uncured patients in the mixture model: gamma distribution, Weibull distribution and log-normal distribution. The censoring distributions considered in their paper were uniform and exponential distributions. The values of parameter in the censoring distributions were determined so that the resulting censoring rates were 10%, 20% and 30% for each censoring distribution.

Following Peng et al. [10], I considered an approach that specified a single censoring distribution that had expected censoring rate close to the target value. My procedure was to use the form of equation (3) with the argument x set to a scalar multiple of the expected value of the covariate. Let the covariate value X of the failure time distribution be $U(0, 5)$. In the covariate coefficient setting $\lambda = 1$, $\beta = 1$ with exponential censoring rate at 10%, I used $1.3 \times E[x]$ as the value of the argument x in equation (3). With an appropriate multiple $k \times E[x]$, the observed censoring rates on average were approximately equal to the target censoring rates. When $\lambda = 1, \beta = 0$, the equation (3) was $\mu = \frac{r}{1-r}\lambda$. That is, the covariate x did not affect the mean of the censoring distribution. Tables 2.1 and 2.2 contain the target censoring rate, mean of censoring distribution, and consequent observed censoring rate for $\lambda = 1, \beta = 0$ and $\lambda = 1, \beta = 1$ respectively, based on 100 replications each with 500 subjects.

2.5.2 Mean of Censoring Distribution of Mixture of Two Regimes Model

The random variable T was a mixture of two exponential random variables with one mean equal to $\frac{1}{\lambda_1 e^{\beta_1 x}}$ with proportion π and the other mean equal to $\frac{1}{\lambda_2 e^{\beta_2 x}}$ with proportion $1-\pi$ so that the pdf of the failure time T given covariates x was

$$f(t \mid x) = \pi \frac{1}{\lambda_1 e^{\beta_1 x}} e^{\frac{-t}{\lambda_1 e^{\beta_1 x}}} + (1-\pi) \frac{1}{\lambda_2 e^{\beta_2 x}} e^{\frac{-t}{\lambda_2 e^{\beta_2 x}}}.$$

The censoring random variable U had the exponential distribution with mean $\frac{1}{\mu}$ and pdf

$$g(u) = \frac{1}{\mu} e^{\frac{-u}{\mu}}$$

I conditioned on the values of the covariates and let $A(x) = \lambda_1 e^{\beta_1 x}$ and $B(x) = \lambda_2 e^{\beta_2 x}$**.** Equation (1), which defined the censoring rate r, was then

$$\int_0^c \int_u^\infty \left(\pi \frac{1}{A(x)} e^{\frac{-t}{A(x)}} + (1-\pi) \frac{1}{B(x)} e^{\frac{-t}{B(x)}} \right) \cdot \frac{1}{\mu} e^{\frac{-u}{\mu}} \, dtdu = (1-r) \cdot \{1 - P(T > c, U > c)\}.$$

The left hand side was

$$\int_0^c \left(\pi e^{\frac{-u}{A(x)}} + (1-\pi) e^{\frac{-u}{B(x)}} \right) \frac{1}{\mu} e^{\frac{-u}{\mu}} \, du$$

$$= \frac{1}{\mu} \left(\pi \frac{A(x)\mu}{A(x)+\mu} + (1-\pi) \frac{B(x)\mu}{B(x)+\mu} - \pi \frac{A(x)\mu}{A(x)+\mu} e^{-\left(\frac{1}{A(x)}+\frac{1}{\mu}\right)c} - (1-\pi) \frac{B(x)\mu}{B(x)+\mu} e^{-\left(\frac{1}{B(x)}+\frac{1}{\mu}\right)c} \right)$$

$$= \pi \frac{A(x)}{A(x)+\mu} + (1-\pi) \frac{B(x)}{B(x)+\mu} - \pi \frac{A(x)}{A(x)+\mu} e^{-\left(\frac{1}{A(x)}+\frac{1}{\mu}\right)c} - (1-\pi) \frac{B(x)}{B(x)+\mu} e^{-\left(\frac{1}{B(x)}+\frac{1}{\mu}\right)c}$$

The right hand side was

$$(1-r)\cdot\left\{1-\int_c^\infty \{\pi \frac{1}{A(x)} e^{\frac{-t}{A(x)}} + (1-\pi)\frac{1}{B(x)} e^{\frac{-t}{B(x)}}\} \, dt \cdot \int_c^\infty \frac{1}{\mu} e^{\frac{-u}{\mu}} \, du\right\}$$

$$= (1-r)\cdot\left\{1-\left(\pi e^{\frac{-c}{A(x)}} + (1-\pi) e^{\frac{-c}{B(x)}}\right) e^{\frac{-c}{\mu}}\right\}$$

Then μ was a root of

$$\pi \frac{A(x)}{A(x)+\mu} + (1-\pi)\frac{B(x)}{B(x)+\mu} - \frac{A(x)}{A(x)+\mu}\pi e^{-\left(\frac{1}{A(x)}+\frac{1}{\mu}\right)c} - \frac{B(x)}{B(x)+\mu}(1-\pi) e^{-\left(\frac{1}{B(x)}+\frac{1}{\mu}\right)c}$$

$$= (1-r)\cdot\left\{1-\left(\pi e^{\frac{-c}{A(x)}} + (1-\pi) e^{\frac{-c}{B(x)}}\right) e^{\frac{-c}{\mu}}\right\}$$

Again I followed Peng et al. [11]. That is, I found a value for the mean of the censoring distributions so that the resulting censoring rate was approximately equal to the expected censoring rate. Table 2.3 contains the target censoring rate, mean of censoring distribution and observed censoring rate for $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5,$ $\pi = 0.5$ at each censoring rate. I reported the observed censoring rate as well as the expected censoring rate.

2.6 Censoring Rate Distribution

The observed fraction of censored observations *R* followed approximately normal distribution with mean close to the expected censoring rate and standard deviation 0.0186; that is, when the expected censoring rate was 10%, the observed censoring rate ranged from 0.064 to 0.134 (see Table 2.4).

2.7 Data Generation

I generated null data for six cases with $\lambda = 1,3$ and $\beta = 0,1,3$ at expected exponential censoring rates of 10%, 20% and 30%. I generated 500 replications at each setting. I considered sample sizes *n* of 500, 1000 and 2000 subjects.

The failure time t_i^* had the exponential survival distribution with mean equal to $\frac{1}{\lambda \exp(\beta x_i)}$, where the covariates x_i were from an uniform distribution *U*(0, 5), and u_i be the censoring time. The survival time $t_i = \min(t_i^*, u_i)$, $(1 \le i \le n)$. The t_i^* sample was from a single exponential distribution with λ =1, 3 and covariate coefficient $\beta = 0,1,3$ respectively.

For example, to create a sample of size 500 from an exponential distribution with $\lambda = 1$ and $\beta = 1$ and exponential censoring pattern with expected censoring rate 10%, I generated one value, t_i^*, from an exponential distribution with mean equal to $\frac{1}{e^x}$ as the failure time, where the covariate *x* was generated from an uniform distribution *U*(0,5). I then generated one value, u_i, from another independent exponential distribution with mean equal to 1.563965 (Table 2.2) as the censoring time. I then compared these two values and reported the minimum. If the value was

u_i, the observation was censored, and I set the absence of censoring indicator off, *i.e.* $c_i = 0$. I repeated this process independently 500 times.

The alternative hypothesis was that the survival time follows a mixture of two regimes model. Let t_i^* be the survival time. With probability π, I selected an observation from the first exponential distribution with mean equal to $\frac{1}{\lambda_1 \exp(\beta_1 x_i)}$, and with probability $1-\pi$, I selected an observation from the second exponential distribution with mean equal to $\frac{1}{\lambda_2 \exp(\beta_2 x_i)}$, where the covariate x was from an uniform distribution $U(0,5)$. Let u_i be the censoring time. The observed time $t_i = \min(t_i^*, u_i), \quad (1 \le i \le n)$. The sample sizes considered here were 500, 1000 and 2000 with an exponential censoring pattern at expected censoring rates of 10%, 20% and 30%. I generated alternative data for $\lambda_1 = 1, \lambda_2 = 1, \beta_2 = 0.5$ with $\beta_1 = 0.75$, 1, 1.25 and 1.5 and mixing proportion $\pi = 0.1$, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9 at each censoring rate.

For example, to create a sample of size 500 from two exponential distributions with $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5$ and $\pi = 0.5$ and exponential censoring pattern with censoring rate 10%, I generated one value for regime one from an exponential distribution with mean equal to $\frac{1}{e^{1.5x}}$, and one value for regime two from exponential distribution with mean equal to $\frac{1}{e^{0.5x}}$, where the covariate x was generated from an uniform distribution $U(0,5)$. The failure time value t_i^* was selected with probability

50% from the first regime and with probability 50% from the second regime. I then generated a value u_i from an independent exponential distribution with mean equal to 1.948705 (Table 2.3) as a censoring time. I then compared t_i^* and u_i values and selected the minimum as the reported value. If the value was u_i, the observation was censored, and I set the absence of censoring indicator off, *i.e.* $c_i = 0$. I repeated this process independently 500 times.

2.8 Random Starting Points

In order to specify the number of random starting points (RSPs), I generated 20 replications each with 500 subjects at exponential censoring rate 30%. The t_i^* sample was from a single exponential with λ=5 and covariate coefficient $\beta = 1$. The covariate sample *x* was from a uniform distribution *U*(0,5). Table 2.5 and 2.6 contain the maximum log-likelihood of single regime model and mixture of two regimes model at specified number of RSPs.

The number and the choice of RSPs were important to assure that the log-likelihood function was reasonably close to its maximum value (Caudill et al. [16]). For each set of initial starting points, we will get maximum log likelihood and maximum likelihood estimators. To avoid getting a local maximum, I compared all the values generated from each set of initial points and chose the largest of maximum log likelihood value and report the associated set of MLEs.

In Table 2.5 (null model), the difference of maximized log-likelihood function number generated from 16 RSPs and generated from 25 RSPs was less than $1e-7$. Hence, I chose 16 RSPs as the number of the RSPs for the null model. To determine the greatest value of the log-likelihood function in the null model, 4 random starting values for each λ and β were generated, as well. Let λ_i and β_j, (i, j =1,..,4), be generated from an uniform random variable U(0,1). Then combination of the 4λ values and 4β values generated 16 sets of starting values for (λ, β). In Table 2.6 (alternative model), the difference of maximized log likelihood function number generated from 48 RSPs and generated from 243 RSPs was less than $1e-7$. Hence, I chose 48 RSPs as the number of RSPs for the alternative model. To determine the greatest value of the log-likelihood function in the alternative model, 2 random starting values for each $\lambda_1, \lambda_2, \beta_1, \beta_2$ and 3 random starting values π were generated. Each of $\lambda_{1j}, \lambda_{2j}, \beta_{1l}, \beta_{2q}, \pi_w$(i, j, l, q=1,2, w=1,2,3) were generated from an uniform random variables U(0,1). The combination of the $2\lambda_1, 2\lambda_2, 2\beta_1, 2\beta_2$ and 3π values generated 48 sets of starting values for $(\lambda_1, \lambda_2, \beta_1, \beta_2, \pi)$.

2.9 The Likelihood Ratio Test (LRT)

The null hypothesis was a single regime model, in which the survival time followed an exponential distribution with mean equal to $\frac{1}{\lambda \exp(\beta x)}$. The alternative hypothesis was that the observed survival data was a mixture of two exponential regimes. One regime occurred with probability π and had mean equal to $\frac{1}{\lambda_1 \exp(\beta_1 x)}$. The other occurred with probability $1-\pi$ and had mean equal to $\frac{1}{\lambda_2 \exp(\beta_2 x)}$. The LRT statistic was equal to $-2(\log L_{H0} - \log L_{H1})$. where $\log L_{H1}$ was the log-likelihood function maximized under the alternative hypothesis and $\log L_{H0}$ was the log-likelihood function maximized under the null hypothesis. The MLEs for the LRT were calculated by using The Nelder-Mead (NM) algorithm with 16 random starting values for the null model and 48 random starting values for the alternative model.

2.10 Nelder-Mead (NM) algorithm

The Nelder Mead (NM) algorithm [10] is used to minimize a function of *n* variables. It evaluates the function at the vertices of a $(n+1)$ simplex and then iteratively uses reflection, contraction and expansion of the simplex as better points are found. A vertex is replaced by points with a better value of the function until the minimal function value is obtained. The NM algorithm uses only function values and is robust but relatively slow. It works reasonably well for non-differentiable functions. *[R version 2.8.0]*.

2.11 Software programs

I wrote programs in *R* and Microsoft Visual C++ that calculated the MLE and log-likelihood for both models. The default method was an implementation of that of Nelder and Mead (1965). I also used the NM algorithm as given in GNU Scientific Library (GSL) in Microsoft Visual C++. I set the convergence rate to be $1e-5$ and the maximum number of iteration to 1000.

The *R* program codes of calculating the mean of censoring distribution for single regime and mixture of two regimes are provided in the appendix. The Microsoft Visual C++ program code of the simulation study is provided in the appendix as well.

Chapter 3 Simulation Results for Single Regime Model

3.1 Simulation Results of Maximum Likelihood Estimators of Null Model

To check my simulation procedure, I examined the MLEs of λ and β for the single regime model. Table 3.1 presents summary statistics for the MLEs with exponential censoring. As expected, for each sample size and expected censoring rate, the MLE of λ was close to 3, the parameter used to generate the data. The MLE of β was also close to 1, the parameter used to generate the data. The mean MLEs for other settings were reported in Tables 3.2, 3.3 and 3.4.

3.2 Null Distribution Results

The null hypothesis was that the survival time followed a single exponential regime. The simulation results for the LRT were calculated by using NM algorithm with 16 $(4\lambda \times 4\beta)$ random starting values (see section 2.8) for 500 replications at each setting used (sample size of 500, 1000 and 2000, exponential censoring rate 10%, 20% and 30%, and six parameter settings).

Tables 3.2 (for expected censoring rate 10%), 3.3 (for expected censoring rate 20%) and 3.4 (for expected censoring rate 30%) contain the mean of the covariate *x*, the mean and standard deviation of the survival time *t*, the average observed censoring rate, the MLE of λ, the MLE of β, the mean of LRT, and standard deviation of LRT as well as selected percentiles of the distribution of the LRT statistic at each parameter setting and censoring rate. Finally it contains the fraction of LRT values less than 0.001. For each setting and parameter, the mean MLE was close to the parameter setting.

3.3 Modeling Null Distribution of LRT

A linear regression was run to determine which, if any, settings affected the null distribution of the LRT. The dependent variable was the mean of the LRT statistic for each sample size and each expected censoring rate (6 observations for each of nine settings of sample size and expected censoring rate). For expected censoring rate 10%, 20% and 30%, the factors λ and β were not significant for any sample size and censoring rate (data not shown).

Figure 3.1 is the scatter plot for the six $(\lambda = 1,3, \beta = 0.1,3)$ 95^{th} percentiles at sample size 2000 and expected censoring rate 10%. The values of the 95^{th} percentile seemed to lie on a horizontal plane. That is, the settings of λ and β apparently had minimal effect. Table 3.5 contains the regression results for the 95^{th} percentile at sample size 2000 and expected censoring rate 10%. The parameters λ and β were not significant factors in the regression model with p values 0.41 and 0.55 respectively. Similar results held for the other settings (data not shown). Consequently, for each sample size and expected censoring rate, I averaged the 95th and the 99th percentiles of the null distribution of the LRT for the six λ, β settings. The values are reported in Table 3.6.

The study was a 3^2 factorial experiment with *n*, the sample size and *cr*, the expected censoring rate as factors. One pair of dependent variables was the mean of the observed 95th percentiles and 99th percentiles of the null distribution of the LRT, reported in Table 3.6. Tables 3.7 and 3.8 contain the linear regression results for the mean 95th and 99th percentiles respectively. The mean percentiles were insensitive to sample size, censoring rate and their interaction.

3.4 Fraction of zero LRT

Self and Liang [17] found the asymptotic null distribution of the LRT for the mixtures. They showed that for some distributions, there was a non-zero probability of a LRT value exactly equal to 0 (*i.e.* X_0^2). Consequently, I modeled the pdf

$f(t; n\ cr, \beta)$ of the null distribution of the LRT as

$$f(t|\ n, cr, \beta) = \tau(\theta(n,cr,\beta)) + [1-\tau(\theta(n,cr,\beta))] \times g(t \mid n,cr,\beta),$$

where $\tau(\theta(n,cr,\beta))$ was the fraction of zero LRT value and $g(t \mid n,cr,\beta)$ was the

PDF of the non-zero values of the LRT. I used the fraction of LRT<0.001, called LRT_Z, (reported in the rightmost column of Tables 3.2, 3.3 and 3.4) as an estimate of $\tau(\theta(n, cr, \beta))$. That is, this was the dependent variable in a regression analysis. The independent variables were

1. n = Sample size (500, 1000 and 2000).
2. λ = Parameter λ (1 and 3)
3. β = Parameter β (0, 1 and 3)
4. cr = Expected censoring rate (10%, 20% and 30%).

I also included all two factor interactions of these variables. Table 3.9 contains the regression results. The interactions of sample size with parameter β and the interaction of parameter β with expected censoring rate were significant with *p* values 0.036 and 0.004 respectively. Other interactions were not significant with *p* values ranging from 0.15 to 0.953. None of main effects were significant with *p* values ranging from 0.31 to 0.888. Based on the hierarchical principle (Wu and Hamada [18]), I added sample size, parameter β and censoring rate to the significant interactions for my final model. The fitted model was:

$$\text{Fraction of } LRT_Z = 0.28 - 0.000006 \times n + 0.003 \times \beta + 0.122 \times cr$$
$$+ 0.00001 \times n \cdot \beta - 0.0001 \times n \cdot cr - 0.16 \times \beta \cdot cr$$

with R^2 equal to 0.432.

Figure 3.2 is the graph for fraction of LRT_Z at each sample size. From the figure, we can see the fraction of LRT_Z generally decreases as the censoring rate increases. This is consistent with Peng et al. [13] who reported that the null distribution is dependent on censoring rate.

3.5 Transformation of LRT

One observes from Tables 3.2, 3.3 and 3.4, that the *log(standard deviation of LRT)* was associated with *log(mean of LRT)*, I next calculated the mean and standard deviation of LRT_{NZ}. The slope of *log(standard deviation of* LRT_{NZ} *)* vs. *log (mean of* LRT_{NZ} *)* was 0.633(see Table 3.10). The 95% confidence interval for slope is 0.588 to 0.678. Tukey [19] suggested using the transformation $LRT^{1-0.633} = LRT^{0.367}$ and $LRT^{0.333}$, since the cube root was consistent with the confidence interval of the slope.

The linear regression results of *log(SD* $LRT^{0.367}$ *)* vs. *log(mean* $LRT^{0.367}$ *)* and *log(SD* $\sqrt[3]{LRT}$ *)* vs. *log(mean* $\sqrt[3]{LRT}$ *)* are shown in Tables 3.11 and 3.12 respectively. The *t* values were 1.49 and 0.55 for $LRT^{0.367}$ and $\sqrt[3]{LRT}$ respectively, showing that each transformation removed the association between standard deviation and mean. I chose $\sqrt[3]{LRT}$ to analyze the null distribution as it had smaller absolute *t* value and was a “simple” value.

3.6 Distribution of $\sqrt[3]{LRT}$

I examined the histogram of the $\sqrt[3]{LRT_{NZ}}$ for each setting and found the distribution to be approximately normal. Figure 3.3 is the normal Q-Q plot for $\lambda = 1, \beta = 0, cr = 10\%, n = 2000$. The points approximately lay on the line $y = x$, indicated that the distribution was similar to normal distribution. Similar results held for the other settings (data not shown). I approximated the PDF of $\sqrt[3]{LRT}$ as a mixture of zero values with probability $\tau(\theta(n, cr, \beta))$ and a normal distribution with mean $\mu(\theta(n, cr, \beta))$ and variance σ^2. The PDF of $\sqrt[3]{LRT}$ was $f_T(t) = \tau(\theta(n, cr, \beta)) + [1 - \tau(\theta(n, cr, \beta))] \times \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{[t - \mu(\theta(n, cr, \beta))]^2}{2\sigma^2}\}$, where $\theta(n, cr, \beta)$ was a function of sample size, censoring rate and covariate β. The variance was set to the pooled variance estimate of the $\sqrt[3]{LRT_{NZ}}$, which was 0.498^2.

Table 3.13 contains the linear regression results of mean $\sqrt[3]{LRT_{NZ}}$. Only the parameter β was significant with *p* value 0.031. The fitted model was:

$$\hat{\mu}(n, cr, \beta) = 1.07 - 0.028 \times \beta$$

with R^2 equal to 0.398.

Tables 3.14 (for expected censoring rate 10%), 3.15 (for expected censoring rate 20%) and 3.16 (for expected censoring rate 30%) present the summary statistics from the simulation for $\sqrt[3]{LRT}$ and the fitted values of τ, the fraction of zero values and μ, mean of $\sqrt[3]{LRT_{NZ}}$.

3.7 Critical Values for $\sqrt[3]{LRT}$

I averaged the 95th percentile of $\sqrt[3]{LRT}$ at each sample size and censoring rate as the critical values at rejection rate 0.05 and averaged the 99th percentile of $\sqrt[3]{LRT}$ as the critical values at rejection rate 0.01. These values were reported in Table 3.17. For each α and sample size *n*, there were three percentile values for expected censoring rate 10%, 20% and 30% respectively. I interpolated using these three values. For example, with $\alpha = 0.05$ and $n = 1000$, when the observed censoring rate < 10%, I used 1.727 as the critical value. For observed censoring rate >30%, I used 1.817 as the critical value. For intermediate censoring rates, I used the critical value based on linear interpolation.

For a fixed sample size, the 95th and 99th percentiles of $\sqrt[3]{LRT}$ were relatively insensitive to expected censoring rate. For example, with sample size 2000, the 99th percentiles were 2.122, 2.104, and 2.117 respectively. I used the average of these three values, 2.114 as the critical value for this sample size. Table 3.18 contains the critical values I used in my power study.

Chapter 4 Distribution of LRT under the Alternative

The alternative hypothesis was that the survival time follows a mixture of two exponential regimes. The LRT was calculated by using the Nelder-Mead algorithm with 48 ($2\lambda_1, 2\lambda_2, 2\beta_1, 2\beta_2, 3\pi$) random starting points (see section 2.8). I considered mixtures of two regimes with the regimes having equal λ values but different β values. I set $\lambda_1 = \lambda_2 = 1$, $\beta_2 = 0.5$ and generated 500 replications for $\beta_1 = 0.75$ and $\beta_1 = 1.5$, and 100 replications for $\beta_1 = 1$ and $\beta_1 = 1.25$ (sample size of 500, 1000 and 2000, exponential censoring rate 10%, 20% and 30%, and 9 mixing proportions, π=0.1~0.9).

4.1 Simulation Results of Maximum Likelihood Estimators of Alternative Model

To check the simulation, I examined the mean MLEs of $\lambda_1, \lambda_2, \beta_1, \beta_2$ and π in the mixture of two regimes model. Table 4.1 presents the estimated MLEs for exponential censoring when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.6$, based on 500 replications. I chose this setting because its power was near 1. As expected, for each sample size and expected censoring rate, the MLE of $\lambda_1, \lambda_2, \beta_1, \beta_2$ and π were close to 1, 1, 1.5, 0.5 and 0.6 respectively. We expected that the standard deviation of the MLEs would increase as the censoring rate increased, and that the standard deviation of the MLEs would decrease as sample size increased. Table 4.2 presents the minimum and the maximum of the MLEs at sample size 1000. At the expected censoring rate 20%, there were about 3% to 6% of the 500 $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\pi}$ values outside the range of minimum and maximum MLEs for censoring rates 10% and 30% (data not shown). These outliers caused the standard deviation of the MLEs at sample size 1000 to depart from the expected pattern.

Table 4.3 contains the means and standard deviations of survival time and the mean of the covariate x of the first and second regimes when $\lambda_1 = 1, \lambda_2 = 1, \beta_2 = 0.5, \pi = 0.6$ with $\beta_1 = 0.75$, 1, 1.25 and 1.5 at censoring rate 0%. I generated a data of sample size 1000 for each case to document each of the two regimes. When $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.6$, there were 595 subjects in the first regime ($\lambda_1 = 1, \beta_1 = 0.75$). The average and standard deviation of the survival time were 0.258 and 0.432 respectively. The average of the covariate x was 2.591. There were 405 subjects in the second regime ($\lambda_2 = 1, \beta_2 = 0.5$). The average and standard deviation of the survival time were 0.368 and 0.523 respectively. The average of the covariate x was 2.527. The difference between means of survival time was 0.11, which is about 23% of the average standard deviation. The differences between means of survival time were about 28%, 46% and 55% of the average standard deviation for $\beta_1 = 1$, $\beta_1 = 1.25$ and $\beta_1 = 1.5$ respectively.

4.2 Alternative Distribution of LRT

This study can be seen as a $3^2 \times 9^2 \times 4$ factorial experiment. The variables were:

1. n = Sample size (500, 1000 and 2000).
2. π = Mixing proportion to first regime (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90%)
3. π^2 = Mixing proportion square
4. d = Distance between β_1 and β_2 (0.25, 0.5, 0.75 and 1)
5. cr = Expected censoring rate (10%, 20% and 30%).

I also included six two factor interactions; namely, $(n \cdot \pi, n \cdot d, n \cdot cr, \pi \cdot d, \pi \cdot cr, d \cdot cr)$. I reported the distribution of $\sqrt[3]{LRT}$ in Tables 4.4 and 4.5. The standard deviations ranged from 0.244 to 0.723. These values were relatively close to 0.498, the average standard deviation of the null simulations. This suggested that the variance stabilizing property held for the alternative.

The table also reports the 1st percentile. None of these values were equal to zero. That is, the fraction of zero values observed under the alternatives was negligible. The dependent variable of this study was the mean of the LRT under the alternative. Table 4.6 contains the regression results. The main effects of sample size, mixing proportion square and the distance between β_1 and β_2 were significant with p values <0.000.

The main effect of censoring rate was significant as well with *p* value 0.014. The interaction of sample size and mixing proportion, the interaction of sample size and the distance between β_1 and β_2, and the interaction of mixing proportion with the distance between β_1 and β_2 were significant with *p* values <0.000. The interaction of mixing proportion with censoring rate was significant as well with *p* value 0.02. The main effect of the mixing proportion was not significant (*p* value=0.155). Based on the hierarchical principle, I added the mixing proportion to the final model. The fitted model was:

$$LRT = 423.31 - 0.518 \times n - 265.445 \times \pi - 805.282 \times \pi^2 - 951.905 \times d + 225.67 \times cr$$
$$+ 0.538 \times n \cdot \pi + 0.815 \times n \cdot d + 1956.614 \times \pi \cdot d - 899.382 \times \pi \cdot cr$$

with R^2 equal to 0.893. Consequently, one could model the power directly for censoring rates less than 30%.

4.3 Power Study

Table 4.4 also contains the report of the simulated power of $\sqrt[3]{LRT}$ when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ based on 500 replications, which was a setting in which the two regimes means were close. When the expected censoring rate was 10%, a sample size of 2000 was needed to have power 99% at level $\alpha = 0.01$ with mixing proportion 0.5, 0.6 or 0.7. For expected censoring rate was 20%, a sample size of

2000 was needed to have power 99% at level $\alpha = 0.01$ with mixing proportion 0.5, 0.6 or 0.7.When the expected censoring rate was 30%, the power was 95% or more at level 0.01 with mixing proportion 0.5, 0.6 or 0.7.

Table 4.5 is the report of simulation results of $\sqrt[3]{LRT}$ when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$ based on 100 replications. These two regimes were more separated than the first pair of regimes. When the expected censoring rate was 10%, a sample size of 500 had power 99% at level $\alpha = 0.01$ with mixing proportion 0.3 or greater. A sample size of 1000 was needed to have power 99% with mixing proportion 0.2. When the expected censoring rate was 20%, a sample size of 500 was needed to have power 99% at level $\alpha = 0.01$ with mixing proportion 0.4, or sample size of 1000 with mixing proportion 0.2. When the expected censoring rate was 30%, a sample size of 500 was needed to have power 99% at level $\alpha = 0.01$ with mixing proportion 0.3, or sample size of 1000 with mixing proportion 0.2.

We expected that the power would decrease as the censoring rate increased and that the power would increase as the difference between regimes increased. From Tables 4.4 and 4.5, the power was greater when the distance between the two regimes was larger. For settings $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.25, \beta_2 = 0.5$ and $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5$, most of power values were equal to 1 and the results of regression analysis were of minimal value (data not shown).

4.4 Logit(power) Linear Regression Results

For the settings $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ and $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$, the dependent variable was the *logit(power)*. There were five independent variables.

6. n = Sample size (500, 1000 and 2000).
7. π = Mixing proportion to first regime (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90%)
8. π^2 = Mixing proportion square
9. d = Distance between β_1 and β_2 (0.25 and 0.5)
10. cr = Expected censoring rate (10%, 20% and 30%).

I also included six two factor interactions; namely, $(n \cdot \pi, n \cdot d, n \cdot cr, \pi \cdot d, \pi \cdot cr, d \cdot cr)$. Because the power curve was a concave function to the mixing proportion, I included π^2 as an independent variable as well.

Table 4.6 is the linear regression results with two way interactions. The main effects of sample size, mixing proportion, mixing proportion square and the distance between β_1 and β_2 were significant with p values <0.000. The interaction of sample size with distance between β_1 and β_2 was significant with p value <0.000.

The fitted model was:

$$logit(power) = -15.069 + 0.003 \times n + 25.665 \times \pi - 21.196 \times \pi^2 + 27.693 \times d$$
$$- 0.005 \times n \cdot d$$

with R^2 equal to 0.914.

Figures 4.1, 4.2 and 4.3 are the graphs of power curves at $\alpha = 0.01$ with respect to censoring rate at each sample sizes. From the figures, we can see the power decreased as censoring rate increased in most of mixing proportion, as expected.

Figures 4.4 4.5 and 4.6 are the graphs of power curves with respect to mixing proportion at $\alpha = 0.01$. The maximum power occurred near 60-40 mixture or 70-30 mixture. The power increased as mixing proportion increased to 60-40 or 70-30 mixture and then the power decreased afterward. From the *logit(power)* fitted model, the maximum power appeared to occur at $\hat{\pi}_{Max} = \dfrac{25.665}{2 \times 21.196} = 0.606$.

To examine why the power was not symmetric with respect to mixture proportion, I generated two data of sample size 500 with same $\lambda_1, \lambda_2, \beta_1, \beta_2$ with π and $1 - \pi$. Figures 4.7 and 4.8 are the scatter plots of first and second regime data for $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.6$ with overall expected censoring rate 10% for a sample of 500 observations. The averages of time were 0.2481 and 0.5043 for first and second regime respectively. The difference in average times was 0.5043-0.2481=0.2562. The power for this parameter setting was 0.53. Figures 4.9

and 4.10 are the scatter plots of first and second regime data for $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.4$ with overall expected censoring rate 10% for a sample of 500 observations. The averages of time were 0.3481 and 0.3176 for first and second regime respectively. The difference in average times was 0.3481-0.3176=0.0305. The power for this parameter setting was 0.35. That is, when $\pi = 0.6$, the average of the difference between the two regimes was greater than the distance when $\pi = 0.4$. Therefore, the power was greater when $\pi = 0.6$. That is, the models were not symmetric in $\pi = 0.5$.

Chapter 5 Discussion and Conclusion

The software program that I developed found the maximum likelihood estimates of the parameters and the likelihood ratio test of the null hypothesis of a single regime against the alternative of a mixture of two regimes. The properties of the LRT for mixture of two regimes were determined by a simulation study.

The single regime model was distributed as an exponential function with conditional mean $\frac{1}{\lambda e^{\beta x}}$. The mean of null LRT was insensitive to the parameters λ and β at sample size 500, 1000 and 2000, and censoring rate 10%, 20% and 30%. The simulation results showed the null distribution of LRT was approximated by $\tau(\theta(n,cr,\beta))+[1-\tau(\theta(n,cr,\beta))]\times$ *g(t| n cr,* β), where $\tau(\theta(n,cr,\beta))$ was the fraction of zero LRT values (LRT_Z) and *g(t| n cr,* β) was the PDF of non-zero LRT values (LRT_{NZ}). The fraction of LRT_Z values was positively associated with the censoring rate and negatively associated with the sample size. Because of the *log(standard deviation of LRT)* was associated with *log(mean of LRT),* I studied the $\sqrt[3]{LRT}$ transformation. The mean of the non-zero $\sqrt[3]{LRT}$ values ($\sqrt[3]{LRT_{NZ}}$) was associated with the parameter β, the coefficient of the covariate affecting the survival time. Then, the pdf of $\sqrt[3]{LRT}$ was approximated by a mixture of LRT_{NZ}

and a normal distribution with mean of $\sqrt[3]{LRT_{NZ}}$ and variance 0.498^2. The null distribution of the $\sqrt[3]{LRT}$, was dependent on the sample size and censoring rate and parameter β.

The alternative model was a mixture of two regimes with the mixing proportion π from the first regime with the conditional mean equal to $\frac{1}{\lambda_1 e^{x\beta_1}}$ and 1-π from the second regime with conditional mean equal to $\frac{1}{\lambda_2 e^{x\beta_2}}$. The mean of alternative LRT was sensitive to sample size, censoring rate, mixing proportion, the distance between β_1 and β_2, and the censoring rate. When the distance between β_1 and β_2 was 0.25, a sample size of 2000 was needed to have power 99% at level $\alpha = 0.01$ for the expected censoring rates 10% and 20%, and to have power 95% for the expected censoring rate was 30% with mixing proportion 0.5, 0.6 or 0.7. When the distance between β_1 and β_2 increased to 0.5, a sample size of 500 was needed to have power 99% at level $\alpha = 0.01$ with mixing proportion greater than 0.3. When the distance between β_1 and β_2 increased to 0.75 or greater, the powers were near 1 in almost all cases (3 censoring rates, 9 mixing proportions and 3 sample sizes).

The standard deviations of $\sqrt[3]{LRT}$ under the alternative ranged from 0.24 to 0.72. These values were relatively close to 0.5, the average standard deviation of the null simulations. This suggested that the variance stabilizing property held for the alternative. The power was relatively insensitive to the censoring rate. The power increased as the sample size increased and the distance between two regimes increased. From the power curves (figures 4.1, 4.2 and 4.3) and the *logit(power)* fitted regression model, the maximum power occurred for an approximate 60-40 mixture.

An extension of this dissertation would be to consider the uniform censoring pattern with censoring rates 10%, 20% and 30%, and compare the results with the exponential censoring pattern. Additionally, we may introduce other covariates that affect the group membership. That is, we might consider the mixing proportion $\pi(z) = \frac{e^{\alpha z}}{1+e^{\alpha z}}$, $-\infty < \alpha < \infty$, where z is the covariate that affects the group membership. Finally, we can extend the mixture mechanism to a mixture of three or more regimes with covariates, and finite study length.

Figure 3.1 Scatter plot of the 95th percentile of the LRT for n=2000 at expected censoring rate 10%

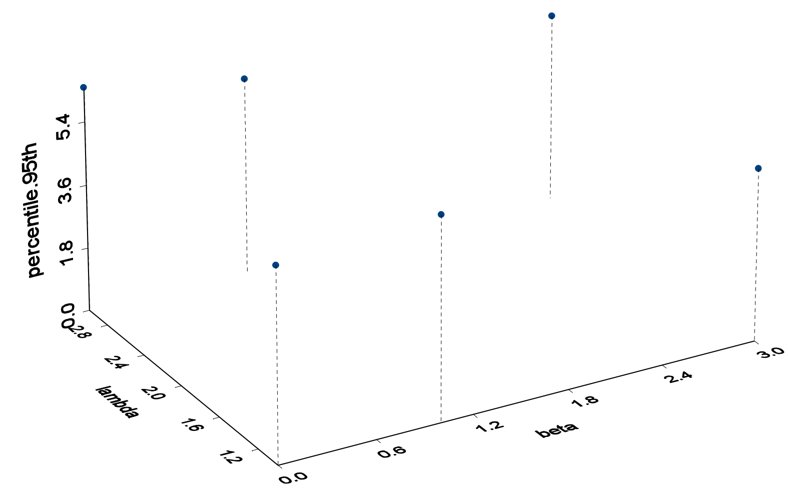


Figure 3.2 Fraction of LRT_Z

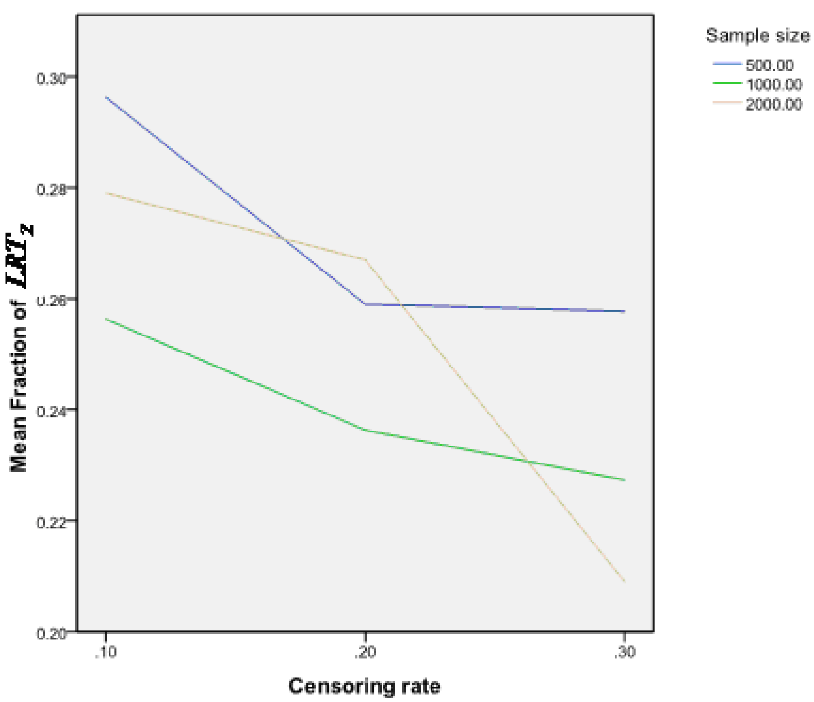
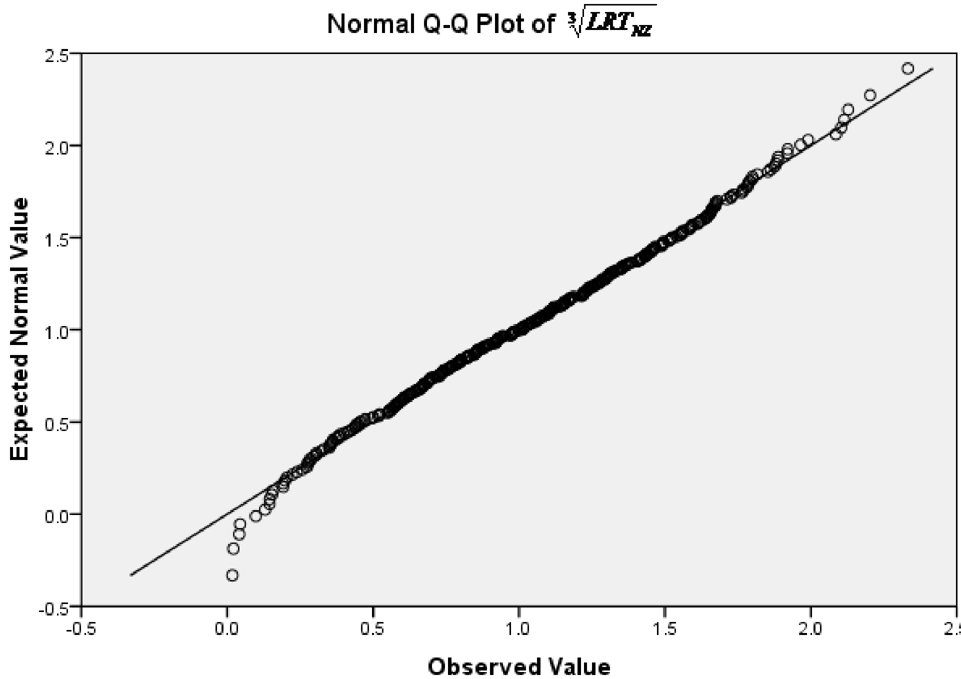


Figure 3.3 Normal Q-Q plot of $\sqrt[3]{LRT_{NZ}}$ when $\lambda = 1, \beta = 0, cr = 10\%, n = 2000$



Note: 500 observations

Figure 4.1, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to censoring rate, sample size 500 at $\alpha = 0.01$

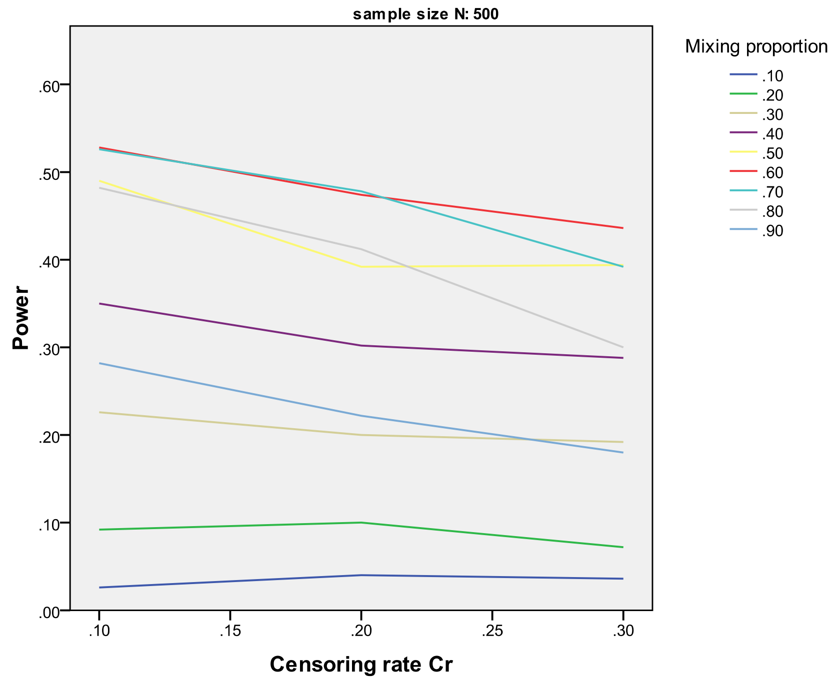


Figure 4.2, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to censoring rate, sample size 1000 at $\alpha = 0.01$

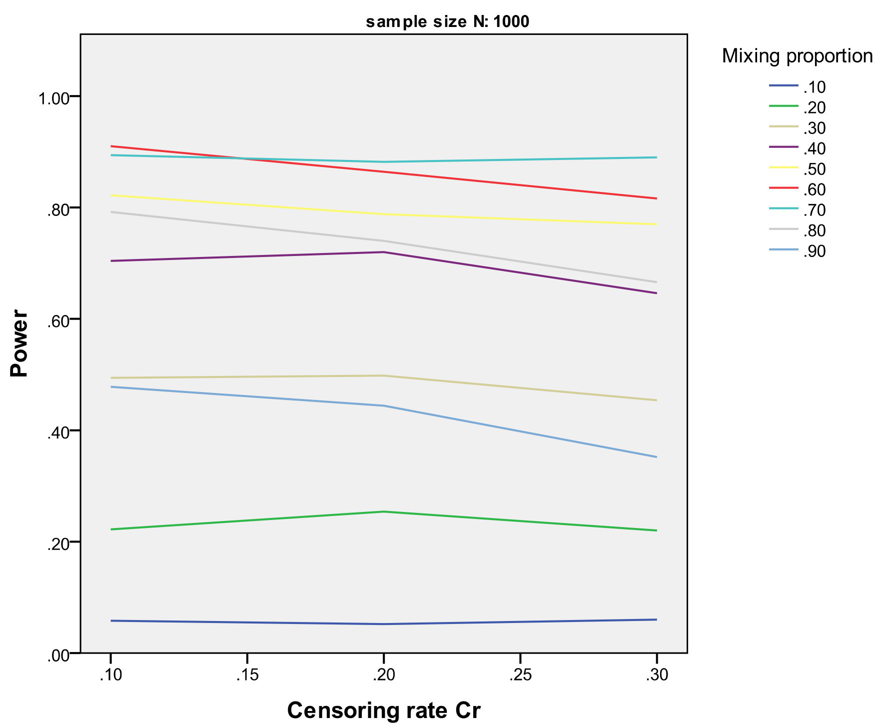


Figure 4.3, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to censoring rate, sample size 2000 at $\alpha = 0.01$

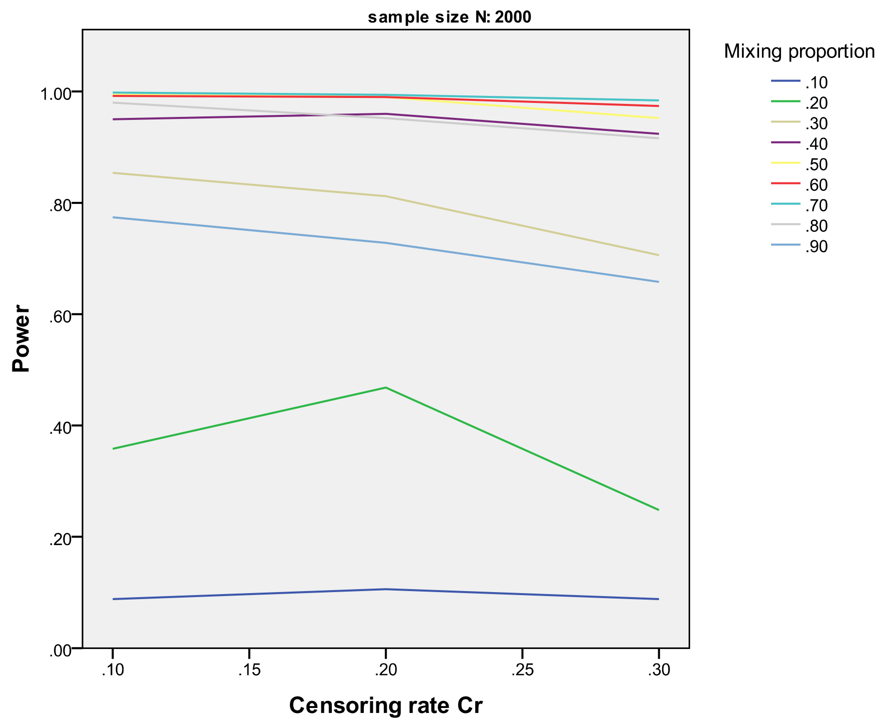


Figure 4.4, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to mixing proportion, sample size 500 at $\alpha = 0.01$

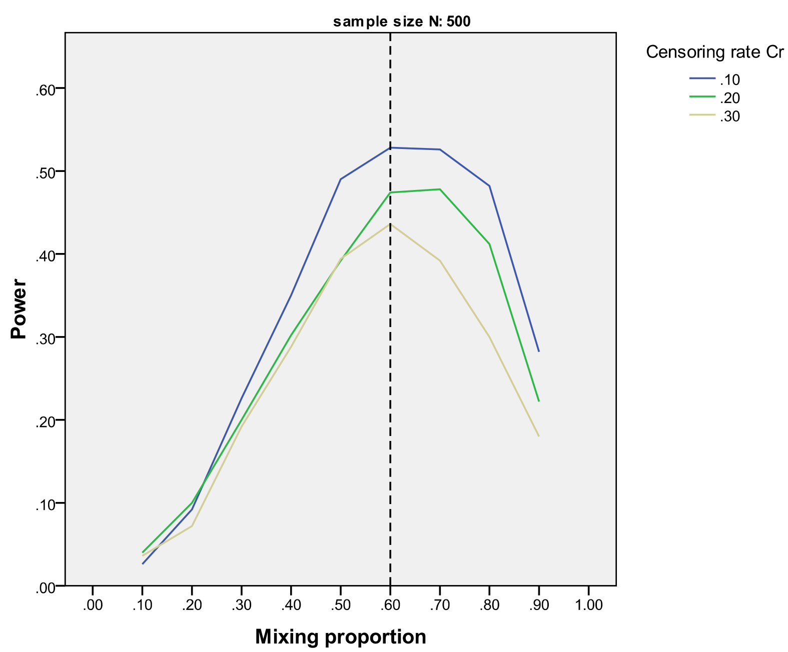


Figure 4.5, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to mixing proportion, sample size 1000 at $\alpha = 0.01$

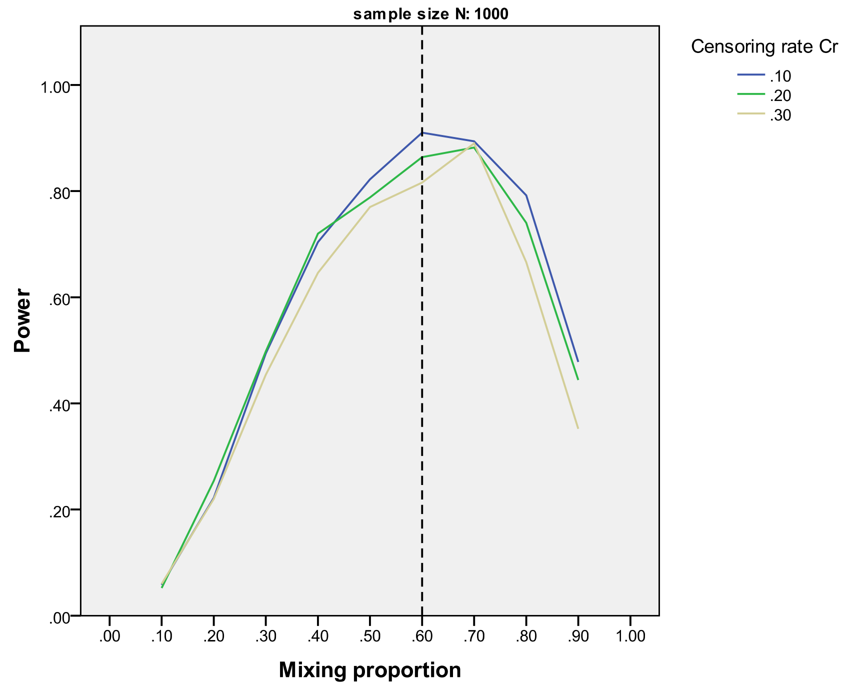


Figure 4.6, Power curves when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ with respect to mixing proportion, sample size 2000 at $\alpha = 0.01$

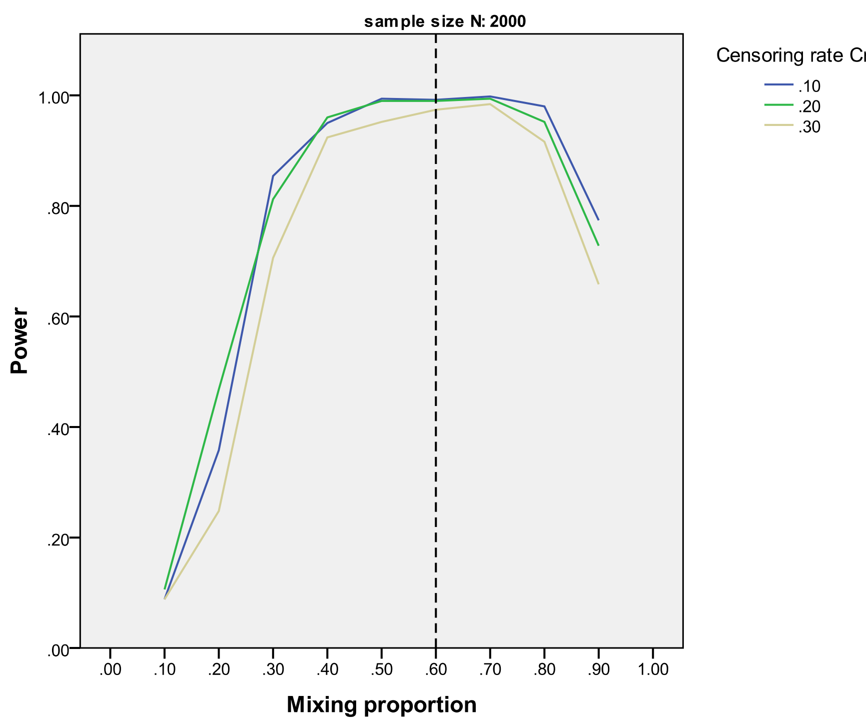


Figure 4.7 Scatter plot for first regime data when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.6$ at 10% censoring rate

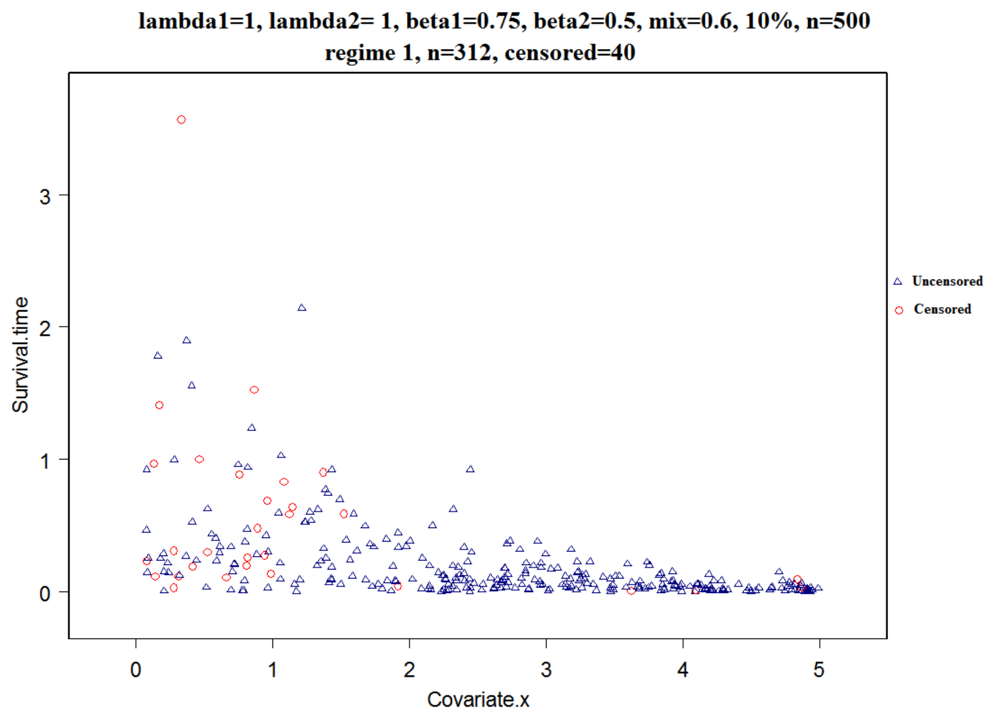


Figure 4.8 Scatter plot for second regime data when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.6$ at 10% censoring rate

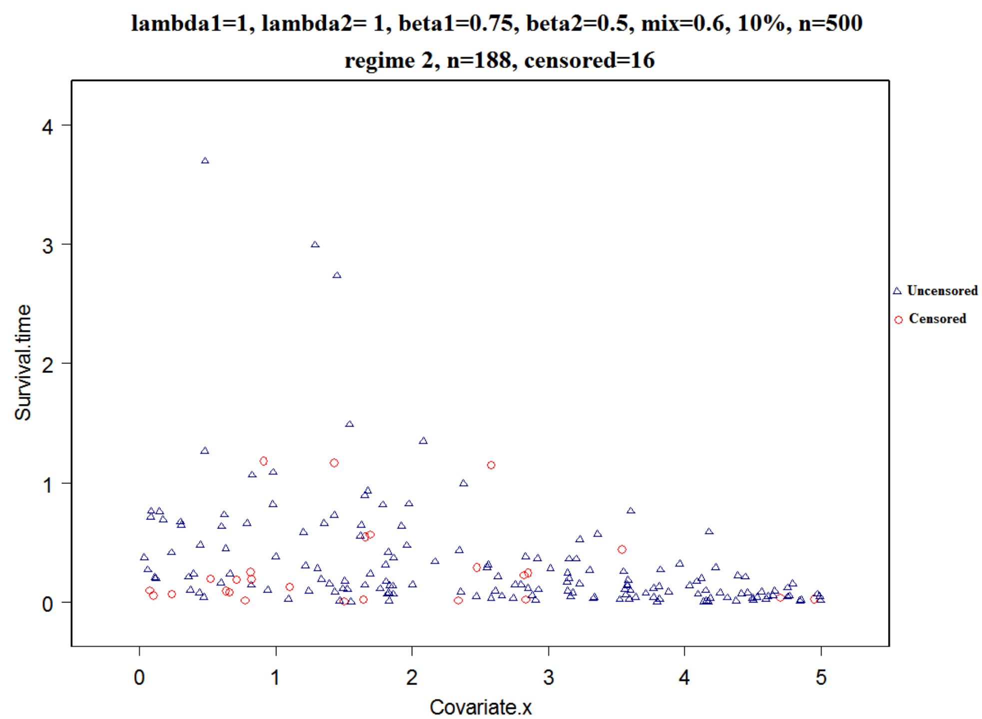


Figure 4.9 Scatter plot for first regime data when
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.4$ at 10% censoring rate

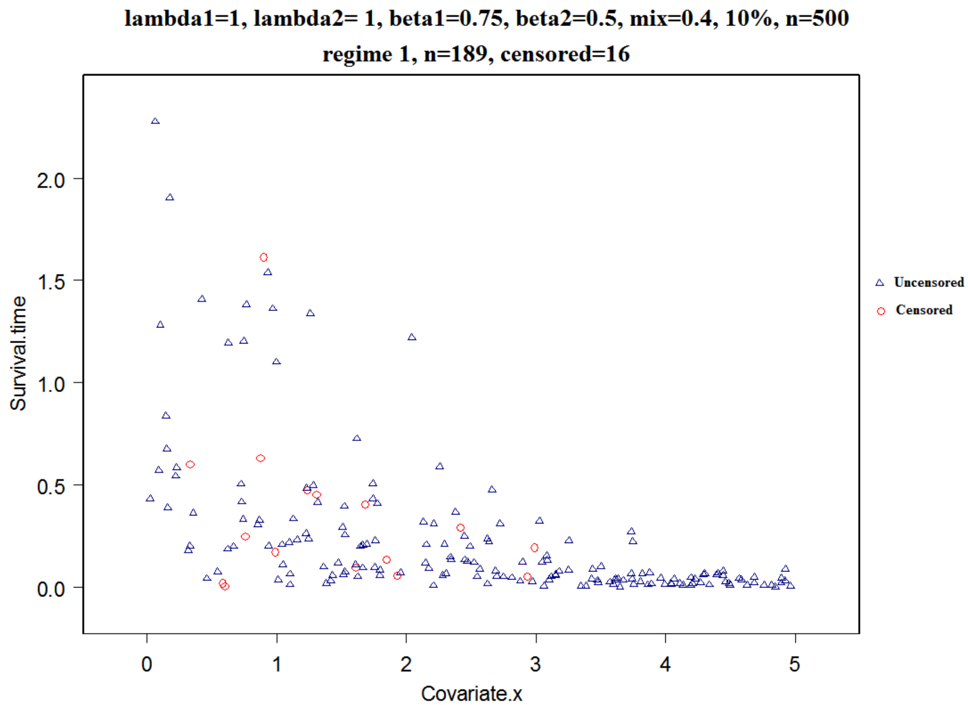


Figure 4.10 Scatter plot for Second regime data when
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.4$ at 10% censoring rate

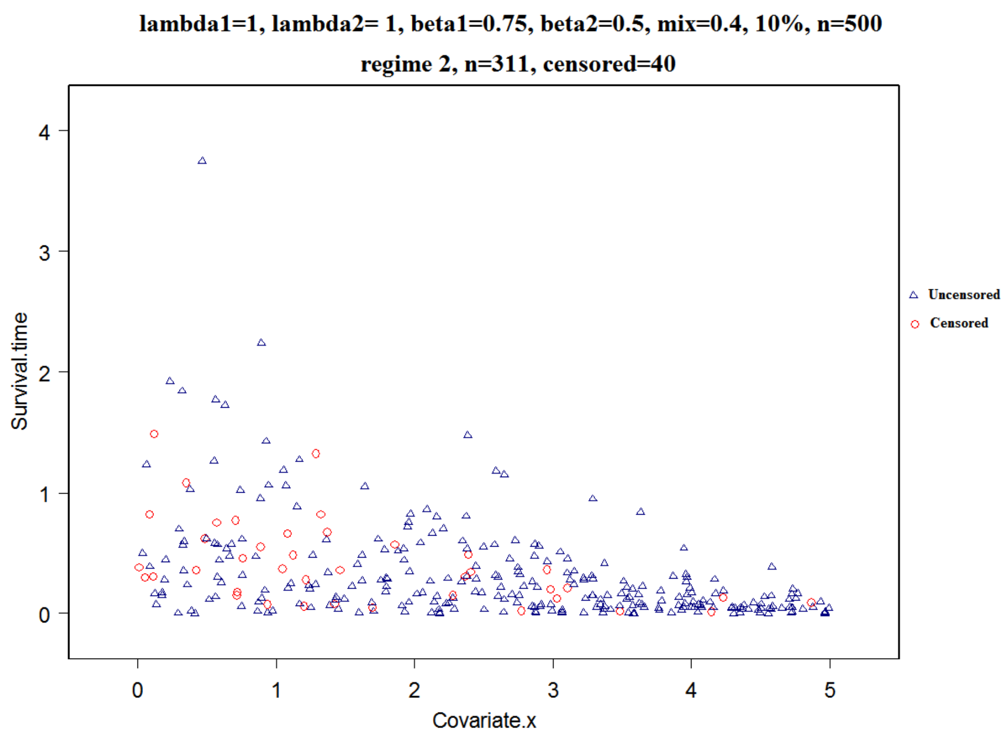


Table 2.1 Means of exponential censoring distribution of single regime model when $\lambda = 1, \beta = 0$

Target censoring rate	Mean of censoring distribution	Average observed censoring rate	95% confidence interval	
			Lower	Upper
10%	9.0000	10.13%*	9.87%	10.39%
20%	4.0000	19.89%	19.47%	20.31%
30%	2.3333	30.21%	29.83%	30.58%
40%	1.5000	40.03%	39.64%	40.41%
50%	1.0000	50.43%	49.99%	50.87%
60%	0.6667	59.88%	59.48%	60.29%
70%	0.4286	70.04%	69.62%	70.46%
80%	0.2500	79.93%	79.55%	80.31%
90%	0.1111	90.10%	89.83%	90.37%

Base on 100 replications, 500 subjects

Note*: 10.13% is the average of 100 censoring rates that ranging from 6.4% to 13.4%.

Table 2.2 Means of exponential censoring distribution of single regime model when $\lambda = 1, \beta = 1$

Target censoring rate	x value	Mean of censoring distribution	Average observed censoring rate	95% confidence interval	
				Lower	Upper
10%	$0.7 \times E[x]$	1.563965	9.76%	9.49%	10.02%
20%	$0.77 \times E[x]$	0.583503	19.82%	19.49%	20.15%
30%	$0.85 \times E[x]$	0.278677	30.00%	29.65%	30.35%
40%	$0.92 \times E[x]$	0.150388	39.61%	39.09%	40.12%
50%	$1.00 \times E[x]$	0.082085	50.04%	49.67%	50.41%
60%	$1.07 \times E[x]$	0.045938	59.65%	59.22%	60.07%
70%	$1.14 \times E[x]$	0.02479	69.49%	69.05%	69.92%
80%	$1.22 \times E[x]$	0.01184	79.73%	79.36%	80.09%
90%	$1.28 \times E[x]$	0.004529	89.68%	89.38%	89.97%

Base on 100 replications, 500 subjects

x~U(0,5)

Table 2.3 Means of exponential censoring distribution of mixture of two regimes model when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.5$

Target censoring rate	Mean of censoring distribution	Average observed censoring rate	95% confidence interval	
			Lower	Lower
10%	1.948705	10.11%	9.83%	10.38%
20%	0.7858606	20.18%	19.86%	20.49%
30%	0.3890864	30.07%	29.69%	30.44%
40%	0.2118107	39.99%	39.56%	40.41%
50%	0.1153447	49.45%	49.00%	49.89%
60%	0.05648397	60.36%	59.94%	60.77%
70%	0.02473104	69.83%	69.42%	70.23%
80%	0.00759614	80.50%	80.17%	80.82%
90%	0.00189027	89.78%	89.51%	90.04%

Base on 100 replications, 500 subjects

Table 2.4 Range of observed censoring rate

Excepted censoring rate	Range of observed censoring rate
10%	6.4%-13.4%
20%	16.6~24.8%
30%	25%~34.2%
40%	34.6%~45%
50%	45.2~55.4%
60%	55.2%~65.4%
70%	65.6%~75.8%
80%	75.2%~86%
90%	86.6%~92.4%

Base on 500 replications, 500 subjects

Table 2.5 Maximum sum of log-likelihood of single regime model for selected numbers of random starting points (RSPs)

replication	maxsum_log H_0		difference between	maxsum_log H_0	difference between	maxsum_log H_0	difference between
	4 RSPs	9 RSPs	9 RSPs and 4 RSPs	16 RSPs	16 RSPs and 9 RSPs	25 RSPs	25 RSPs and 16 RSPs
1	598.4370921	598.4370921	0.0000000	598.4370926	0.0000005	598.4370926	0.0000000
2	575.0743335	575.0743339	0.0000004	575.0743339	0.0000000	575.0743339	0.0000000
3	613.3633173	613.3633177	0.0000004	613.3633180	0.0000002	613.3633180	0.0000000
4	668.0320371	668.0320371	0.0000000	668.0320372	0.0000002	668.0320373	0.0000000
5	665.8972986	665.8972994	0.0000008	665.8972998	0.0000004	665.8972998	0.0000000
6	756.8317544	756.8317553	0.0000008	756.8317556	0.0000003	756.8317556	0.0000000
7	638.8000922	638.8000922	0.0000000	638.8000923	0.0000001	638.8000923	0.0000000
8	628.0533392	628.0533394	0.0000001	628.0533395	0.0000001	628.0533395	0.0000000
9	658.2649248	658.2649252	0.0000003	658.2649252	0.0000000	658.2649252	0.0000000
10	543.0309485	543.0309485	0.0000000	543.0309487	0.0000002	543.0309487	0.0000000
11	651.9480187	651.9480187	0.0000000	651.9480187	0.0000000	651.9480187	0.0000000
12	637.9529252	637.9529252	0.0000000	637.9529252	0.0000000	637.9529252	0.0000000
13	684.4505419	684.4505426	0.0000007	684.4505426	0.0000000	684.4505426	0.0000000
14	529.4179365	529.4179368	0.0000003	529.4179370	0.0000003	529.4179370	0.0000000
15	607.6943776	607.6943781	0.0000004	607.6943783	0.0000002	607.6943783	0.0000000
16	696.5356968	696.5356968	0.0000000	696.5356968	0.0000000	696.5356968	0.0000000
17	659.7506218	659.7506222	0.0000005	659.7506226	0.0000004	659.7506227	0.0000000
18	688.3480868	688.3480875	0.0000007	688.3480875	0.0000000	688.3480875	0.0000000

Table 2.5 Maximum sum of log-likelihood of single regime model for selected numbers of random starting points (RSPs) (continued)

replication	maxsum_log H_0		difference between	maxsum_log H_0	difference between	maxsum_log H_0	difference between
	4 RSPs	9 RSPs	9 RSPs and 4 RSPs	16 RSPs	16 RSPs and 9 RSPs	25 RSPs	25 RSPs and 16 RSPs
19	604.6785784	604.6785784	0.0000000	604.6785784	0.0000000	604.6785784	0.0000000
20	614.5405300	614.5405300	0.0000000	614.5405300	0.0000000	614.5405300	0.0000000
Ave.			0.0000003		0.0000001		0.0000000

Note: Based on 500 subjects in each replication

Table 2.6 Maximum sum of log-likelihood of mixture of two regimes model for selected numbers of random starting points (RSPs)

replication	maxsum_log H_1		difference between	maxsum_log H_1	difference between	maxsum_log H_1	difference between
	1 RSP	32 RSPs	32 RSPs and 1 RSP	48 RSPs	48 RSPs and 32 RSPs	243 RSPs	243 RSPs and 32 RSPs
1	598.4370921	598.4370921	0.0000000	598.4370926	0.0000005	598.4370926	0.0000000
2	575.0743335	575.0743339	0.0000004	575.0743339	0.0000000	575.0743339	0.0000000
3	613.3633173	613.3633177	0.0000004	613.3633180	0.0000002	613.3633180	0.0000000
4	668.0320371	668.0320371	0.0000000	668.0320372	0.0000002	668.0320373	0.0000000
5	665.8972986	665.8972994	0.0000008	665.8972998	0.0000004	665.8972998	0.0000000
6	756.8317544	756.8317553	0.0000008	756.8317556	0.0000003	756.8317556	0.0000000
7	638.8000922	638.8000922	0.0000000	638.8000923	0.0000001	638.8000923	0.0000000
8	628.0533392	628.0533394	0.0000001	628.0533395	0.0000001	628.0533395	0.0000000
9	658.2649248	658.2649252	0.0000003	658.2649252	0.0000000	658.2649252	0.0000000
10	543.0309485	543.0309485	0.0000000	543.0309487	0.0000002	543.0309487	0.0000000
11	651.9480187	651.9480187	0.0000000	651.9480187	0.0000000	651.9480187	0.0000000
12	637.9529252	637.9529252	0.0000000	637.9529252	0.0000000	637.9529252	0.0000000
13	684.4505419	684.4505426	0.0000007	684.4505426	0.0000000	684.4505426	0.0000000
14	529.4179365	529.4179368	0.0000003	529.4179370	0.0000003	529.4179370	0.0000000
15	607.6943776	607.6943781	0.0000004	607.6943783	0.0000002	607.6943783	0.0000000
16	696.5356968	696.5356968	0.0000000	696.5356968	0.0000000	696.5356968	0.0000000
17	659.7506218	659.7506222	0.0000005	659.7506226	0.0000004	659.7506227	0.0000000
18	688.3480868	688.3480875	0.0000007	688.3480875	0.0000000	688.3480875	0.0000000

Table 2.6 Maximum sum of log-likelihood of mixture of two regimes model for selected numbers of random starting points (RSPs) (continued)

replication	maxsum_log H_1		difference between	maxsum_log H_1	difference between	maxsum_log H_1	difference between
	1 RSP	32 RSPs	32 RSPs and 1 RSP	48 RSPs	48 RSPs and 32 RSPs	243 RSPs	243 RSPs and 32 RSPs
19	606.7065635	606.7729594	0.0663958	606.7730034	0.0000441	606.7730034	0.0000000
20	615.3679828	615.3863405	0.0183577	615.3871822	0.0008417	615.3871822	0.0000000
Ave.			0.0042379		0.0000444		0.0000000

Note: Based on 500 subjects in each replication

Table 3.1 Summary statistics for simulated MLE when $\lambda = 3, \beta = 1$ in single regime model

n	Average observed censoring rate	Parameters	Average MLE	SD	Percentile of MLE		
					25%	50%	75%
500	10.18%	λ	2.995	0.303	2.758	2.984	3.185
		β	1.001	0.033	0.980	1.001	1.024
	20.00%	λ	3.033	0.340	2.819	3.009	3.230
		β	1.000	0.037	0.978	0.999	1.026
	29.88%	λ	3.004	0.430	2.724	3.006	3.266
		β	1.003	0.044	0.972	1.003	1.033
1000	10.20%	λ	3.008	0.213	2.862	3.003	3.143
		β	1.000	0.023	0.984	1.000	1.015
	20.09%	λ	3.004	0.250	2.843	2.988	3.161
		β	1.002	0.026	0.984	1.001	1.020
	30.11%	λ	3.001	0.275	2.809	2.989	3.180
		β	1.001	0.029	0.981	0.999	1.018
2000	10.19%	λ	3.012	0.156	2.906	3.009	3.110
		β	1.000	0.017	0.989	1.000	1.011
	20.07%	λ	3.004	0.173	2.888	3.001	3.115
		β	1.000	0.018	0.988	1.001	1.013
	29.96%	λ	3.010	0.199	2.876	3.002	3.123
		β	1.000	0.020	0.987	1.000	1.013

Based on 500 replications in each setting.

Table 3.2 Summary statistics of simulation results of LRT when sampling from single regime at exponential censoring rate 10% (Null distribution)

n	λ	β	Mean x	Mean t	SD of t	Observed censoring rate	$\hat{\lambda}$	$\hat{\beta}$	Mean LRT	SD of LRT	Percentiles of LRT 50	75	95	99	Fraction of LRT< 0.001
500	1	0	2.50	0.90	0.90	10.04%	1.00	0.00	1.54	2.18	0.62	2.44	5.67	10.80	24.0%
		1	2.50	0.15	0.27	10.03%	1.01	1.00	1.44	2.36	0.42	1.89	6.32	10.38	32.8%
		3	2.50	0.03	0.08	10.07%	0.98	3.00	1.44	2.00	0.59	2.03	5.76	8.85	28.4%
	3	0	2.50	0.30	0.30	9.97%	3.03	0.00	1.25	1.84	0.43	1.87	4.95	7.94	30.0%
		1	2.50	0.05	0.09	10.18%	3.00	1.00	1.44	2.05	0.52	2.21	5.70	9.58	30.4%
		3	2.50	0.01	0.03	10.12%	3.04	3.00	1.52	2.35	0.37	2.33	6.58	10.89	32.2%
1000	1	0	2.50	0.90	0.90	9.95%	1.00	0.00	1.51	2.15	0.61	2.38	5.52	9.89	30.6%
		1	2.50	0.15	0.28	10.12%	1.01	1.00	1.59	2.32	0.60	2.35	5.87	9.59	26.4%
		3	2.50	0.03	0.08	10.01%	0.99	3.00	1.14	1.74	0.34	1.59	4.96	7.59	28.8%
	3	0	2.50	0.30	0.30	9.93%	3.01	0.00	1.58	2.30	0.65	2.29	6.38	11.59	25.0%
		1	2.50	0.05	0.09	10.21%	3.01	1.00	1.23	1.88	0.42	1.79	4.67	8.67	27.8%
		3	2.50	0.01	0.03	10.07%	3.00	3.00	1.11	1.73	0.47	1.47	3.78	8.50	15.2%
2000	1	0	2.50	0.90	0.90	10.01%	1.00	0.00	1.39	1.92	0.57	2.06	5.50	9.32	24.2%
		1	2.50	0.15	0.28	10.08%	1.00	1.00	1.45	2.08	0.57	2.03	5.78	9.75	25.2%
		3	2.50	0.03	0.08	10.04%	1.00	3.00	1.21	1.87	0.38	1.73	4.91	8.32	25.2%
	3	0	2.50	0.30	0.30	10.00%	3.00	0.00	1.51	2.22	0.55	2.18	6.36	9.94	27.2%
		1	2.50	0.05	0.09	10.19%	3.01	1.00	1.43	2.20	0.44	2.08	5.58	10.47	23.0%
		3	2.50	0.01	0.03	10.00%	3.01	3.00	1.09	1.96	0.05	1.25	5.37	9.60	42.6%

Based on 500 replications in each setting.

Table 3.3 Summary statistics of simulation results of LRT when sampling from single regime at exponential censoring rate 20% (Null distribution)

n	λ	β	Mean x	Mean t	SD of t	Observed censoring rate	$\hat{\lambda}$	$\hat{\beta}$	Mean LRT	SD of LRT	Percentiles of LRT 50	75	95	99	Fraction of LRT< 0.001
500	1	0	2.50	0.80	0.80	19.96%	1.01	0.00	1.60	2.33	0.44	2.29	6.57	9.26	28.0%
		1	2.50	0.11	0.19	20.00%	1.01	1.00	1.49	2.15	0.55	2.14	5.83	9.76	26.4%
		3	2.50	0.01	0.03	20.23%	1.01	3.00	1.40	2.17	0.46	1.83	6.02	9.38	21.6%
	3	0	2.49	0.27	0.27	20.11%	3.02	0.00	1.47	2.12	0.54	2.14	6.11	9.39	27.6%
		1	2.50	0.04	0.06	20.00%	3.03	1.00	1.60	2.28	0.57	2.33	6.71	9.96	30.6%
		3	2.50	0.00	0.01	20.08%	3.02	3.00	1.77	2.52	0.70	2.63	6.71	10.91	21.2%
1000	1	0	2.50	0.80	0.80	19.97%	1.00	0.00	1.61	2.28	0.66	2.37	5.98	10.73	24.6%
		1	2.50	0.11	0.19	20.12%	1.01	1.00	1.57	2.21	0.64	2.23	6.40	9.50	25.2%
		3	2.50	0.01	0.03	20.08%	1.00	3.00	1.18	1.74	0.43	1.65	4.90	7.49	18.8%
	3	0	2.50	0.27	0.27	20.02%	3.00	0.00	1.21	1.78	0.40	1.87	4.44	8.34	29.0%
		1	2.50	0.04	0.06	20.09%	3.00	1.00	1.49	2.06	0.53	2.21	5.75	9.40	25.2%
		3	2.50	0.00	0.01	20.08%	3.02	3.00	1.51	1.97	0.66	2.35	5.91	8.73	19.0%
2000	1	0	2.50	0.80	0.80	19.93%	1.00	0.00	1.65	2.20	0.74	2.55	6.37	9.39	23.0%
		1	2.50	0.11	0.19	20.09%	1.00	1.00	1.31	1.90	0.50	1.83	5.53	9.34	26.0%
		3	2.50	0.01	0.03	20.09%	1.00	3.00	1.19	1.87	0.37	1.70	5.29	7.83	23.8%
	3	0	2.50	0.27	0.27	19.95%	3.02	-0.01	1.43	2.17	0.41	2.12	5.69	10.11	29.2%
		1	2.50	0.04	0.06	20.07%	3.00	1.00	1.13	1.75	0.33	1.68	4.62	7.86	28.2%
		3	2.50	0.00	0.01	20.06%	3.01	3.00	1.65	2.54	0.53	2.41	7.45	11.71	30.0%

Based on 500 replications in each setting.

Table 3.4 Summary statistics of simulation results of LRT when sampling from single regime at exponential censoring rate 30% (Null distribution)

n	λ	β	Mean x	Mean t	SD of t	Observed censoring rate	$\hat{\lambda}$	$\hat{\beta}$	Mean LRT	SD of LRT	Percentiles of LRT 50	75	95	99	Fraction of LRT< 0.001
500	1	0	2.50	0.70	0.70	30.03%	1.00	0.00	1.38	2.07	0.39	1.98	5.28	10.25	29.2%
		1	2.49	0.08	0.13	30.15%	1.00	1.00	1.38	2.36	0.41	1.99	5.30	11.22	31.6%
		3	2.50	0.00	0.01	29.94%	1.03	3.00	1.56	2.08	0.75	2.28	5.98	8.96	15.6%
	3	0	2.50	0.23	0.23	30.06%	3.02	0.00	1.50	2.09	0.59	2.09	5.84	9.16	26.4%
		1	2.51	0.03	0.04	29.88%	3.00	1.00	1.34	1.96	0.49	1.79	5.69	8.71	26.6%
		3	2.50	0.00	0.00	29.80%	3.03	3.00	1.16	1.95	0.33	1.27	5.37	10.08	25.2%
1000	1	0	2.50	0.70	0.70	30.01%	1.00	0.00	1.74	2.30	0.72	2.77	6.75	10.33	25.8%
		1	2.50	0.08	0.13	30.03%	1.01	1.00	1.41	2.08	0.55	1.95	6.06	8.83	25.2%
		3	2.50	0.00	0.01	29.99%	1.00	3.00	1.34	2.08	0.51	1.72	6.02	9.92	16.8%
	3	0	2.50	0.23	0.23	29.80%	3.03	0.00	1.66	2.24	0.62	2.76	6.32	9.47	27.0%
		1	2.50	0.03	0.04	30.11%	3.00	1.00	1.42	2.03	0.62	1.98	5.65	8.43	23.6%
		3	2.41	0.00	0.00	30.04%	2.98	3.00	1.37	1.97	0.56	1.97	5.24	8.85	18.0%
2000	1	0	2.50	0.70	0.70	29.98%	1.00	0.00	1.48	2.16	0.65	2.02	6.23	10.20	21.0%
		1	2.50	0.08	0.13	29.94%	1.01	1.00	1.70	2.25	0.86	2.55	6.54	11.68	21.2%
		3	2.50	0.00	0.01	29.92%	1.00	3.00	1.46	2.00	0.60	2.12	5.54	8.31	17.2%
	3	0	2.50	0.23	0.23	30.03%	3.00	0.00	1.24	1.81	0.46	1.98	4.98	8.21	27.0%
		1	2.50	0.03	0.04	29.97%	3.01	1.00	1.54	2.03	0.75	2.41	5.57	8.37	20.2%
		3	2.50	0.00	0.00	29.94%	3.00	3.00	1.45	2.11	0.66	2.06	5.46	10.52	18.8%

Based on 500 replications in each setting.

Table 3.5 Linear regression analysis results for the 95th percentile at n=2000 and expected censoring rate 10%

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.471	.501		10.915	.008
	lambda	.233	.224	.529	1.037	.408
	beta	-.199	.275	-.563	-.723	.545
	interaction	-.033	.123	-.235	-.272	.811

a. Dependent Variable: 95th percentile

b. R square: 0.758

Table 3.6 Mean values of selected percentiles of null distribution of LRT averaged over 18 (λ, β) settings

Expected censoring rate	Percentile	n	Mean	SD	Low 95% CI	Up 95% CI
0.1	95th	500	5.829	0.571	5.230	6.428
		1000	5.196	0.926	4.225	6.167
		2000	5.583	0.482	5.077	6.089
	99th	500	9.740	1.176	8.506	10.975
		1000	9.304	1.387	7.848	10.760
		2000	9.565	0.721	8.808	10.321
0.2	95th	500	6.324	0.385	5.920	6.727
		1000	5.563	0.741	4.784	6.341
		2000	5.826	0.977	4.800	6.852
	99th	500	9.777	0.616	9.131	10.423
		1000	9.031	1.115	7.860	10.201
		2000	9.372	1.463	7.838	10.908
0.3	95th	500	5.577	0.302	5.260	5.893
		1000	6.008	0.525	5.457	6.558
		2000	5.718	0.568	5.122	6.314
	99th	500	9.730	0.955	8.727	10.732
		1000	9.304	0.729	8.540	10.068
		2000	9.548	1.458	8.017	11.078

Based on 500 replications for each setting.

Table 3.7 Linear regression analysis results for nine means of the 95th percentile of null distribution of LRT

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	5.764	.692		8.329	.000
	expected censoring rate CR	.433	3.203	.118	.135	.898
	sample size N	.000	.001	-.463	-.425	.689
	N*CR	.001	.002	.343	.257	.807

a. Dependent Variable: 95th percentile

b. R square: 0.152

Table 3.8 Linear regression analysis results for nine means of the 99th percentile of null distribution of LRT

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	9.610	.571		16.817	.000
	expected censoring rate CR	-.008	2.645	-.003	-.003	.998
	sample size N	-9.843E-5	.000	-.259	-.228	.829
	N*CR	-3.214E-5	.002	-.022	-.016	.988

a. Dependent Variable: 99th percentile

b. R square: 0.076

Table 3.9 Linear regression analysis results with two way interaction for fraction of LRT_Z

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.303	.051		5.939	.000
	Sample Size N	-3.082E-5	.000	-.392	-1.027	.310
	Lambda L	-.011	.018	-.234	-.626	.535
	Beta B	-.002	.016	-.059	-.142	.888
	Censoring rate CR	.114	.204	.190	.560	.578
	N*L	1.262E-5	.000	.468	1.465	.150
	N*B	1.497E-5	.000	.564	2.168	.036
	N*CR	-.0001	.000	-.400	-1.126	.266
	L*B	.002	.004	.156	.575	.568
	L*CR	.004	.066	.021	.059	.953
	B*CR	-.160	.053	-.946	-3.029	.004

a. Dependent Variable: Fraction of LRT_Z

a. R square: 0.485

Table 3.10 Linear regression analysis results log(SD LRT_{NZ}) vs. log(mean LRT_{NZ})

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1 (Constant)	.165	.016		10.286	.000
log(mean LRT_{NZ})	.633	.056	.845	11.401	.000

a. Dependent Variable: log(SD LRT_{NZ})

Table 3.11 Linear regression analysis results log(SD $LRT^{0.367}$) vs. log(mean $LRT^{0.367}$)

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1 (Constant)	.557	.065		8.516	.000
log(mean $LRT^{0.367}$	.122	.082	.202	1.489	.142

a. Dependent Variable: log(SD $LRT^{0.367}$)

Table 3.12 Linear regression analysis results log(SD $\sqrt[3]{LRT}$) vs. log(mean $\sqrt[3]{LRT}$)

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1 (Constant)	.583	.064		9.087	.000
log(mean $\sqrt[3]{LRT}$)	.045	.082	.076	.550	.585

a. Dependent Variable: log(SD $\sqrt[3]{LRT}$)

Table 3.13 Linear regression analysis results of mean $\sqrt[3]{LRT_{NZ}}$ with two way interactions

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.158	.058		20.050	.000
	Sample Size N	-4.495E-5	.000	-.510	-1.322	.193
	Lambda L	-.020	.021	-.358	-.948	.348
	Beta B	-.041	.019	-.939	-2.226	.031
	Censoring rate CR	-.261	.231	-.388	-1.127	.266
	N*L	-4.714E-7	.000	-.016	-.048	.962
	N*B	1.393E-6	.000	.047	.178	.860
	N*CR	.000	.000	.479	1.333	.189
	L*B	.007	.005	.404	1.471	.148
	L*CR	.049	.075	.240	.652	.518
	B*CR	-.012	.060	-.064	-.202	.841

b. Dependent Variable: mean $\sqrt[3]{LRT_{NZ}}$

c. R square: 0.472

Table 3.14 Summary statistics of $\sqrt[3]{LRT}$ and the fitted values of τ and μ at exponential censoring rate 10% (500 replications)

n	λ	β	Observed censoring fraction	Observed fraction LRT_Z	Fitted fraction LRT_Z (τ)	Observed mean $\sqrt[3]{LRT_{NZ}}$	Fitted mean $\sqrt[3]{LRT_{NZ}}$ (μ)	Observed SD $\sqrt[3]{LRT_{NZ}}$	Percentiles of $\sqrt[3]{LRT}$ 90	95	99
500	1	0	0.100	0.240	0.273	1.054	1.070	0.504	1.637	1.784	2.210
		1	0.100	0.328	0.276	1.080	1.042	0.506	1.634	1.849	2.182
		3	0.101	0.284	0.295	1.079	0.986	0.471	1.593	1.792	2.069
	3	0	0.100	0.300	0.269	1.021	1.070	0.476	1.558	1.704	1.995
		1	0.102	0.304	0.277	1.081	1.042	0.490	1.609	1.786	2.124
		3	0.101	0.322	0.292	1.049	0.986	0.562	1.629	1.874	2.216
1000	1	0	0.100	0.306	0.251	1.107	1.070	0.489	1.611	1.767	2.146
		1	0.101	0.264	0.269	1.081	1.042	0.510	1.665	1.804	2.124
		3	0.100	0.288	0.292	0.938	0.986	0.503	1.555	1.706	1.965
	3	0	0.099	0.250	0.257	1.086	1.070	0.498	1.644	1.855	2.263
		1	0.102	0.278	0.267	0.998	1.042	0.474	1.528	1.671	2.055
		3	0.101	0.152	0.308	0.889	0.986	0.451	1.418	1.558	2.041
2000	1	0	0.100	0.242	0.228	1.042	1.070	0.467	1.597	1.765	2.104
		1	0.101	0.252	0.249	1.050	1.042	0.488	1.601	1.795	2.136
		3	0.100	0.252	0.295	0.958	0.986	0.490	1.522	1.699	2.026
	3	0	0.100	0.272	0.221	1.075	1.070	0.499	1.662	1.853	2.150
		1	0.102	0.230	0.254	1.010	1.042	0.510	1.618	1.773	2.187
		3	0.100	0.426	0.257	0.965	0.986	0.554	1.540	1.751	2.125

Note: $\tau = 0.28 - 0.000006 \times n + 0.003 \times \hat{\beta} + 0.122 \times cr + 0.00001 \times n \cdot \hat{\beta} - 0.0001 \times n \cdot cr - 0.16 \times \hat{\beta} \cdot cr$

$\mu = 1.07 - 0.028 \times \hat{\beta}$

Table 3.15 Summary statistics of $\sqrt[3]{LRT}$ and the fitted values of τ and μ at exponential censoring rate 20% (500 replications)

n	λ	β	Observed censoring fraction	Observed fraction LRT_Z	Fitted fraction LRT_Z (τ)	Observed mean $\sqrt[3]{LRT_{NZ}}$	Fitted mean $\sqrt[3]{LRT_{NZ}}$ (μ)	Observed SD $\sqrt[3]{LRT_{NZ}}$	Percentiles of $\sqrt[3]{LRT}$		
									90	95	99
500	1	0	0.1996	0.28	0.278	1.069	1.070	0.544	1.737	1.873	2.100
		1	0.2	0.264	0.288	1.042	1.042	0.522	1.622	1.800	2.137
		3	0.2023	0.216	0.308	0.984	0.986	0.511	1.618	1.819	2.109
	3	0	0.2011	0.276	0.279	1.059	1.070	0.505	1.653	1.828	2.110
		1	0.2	0.306	0.284	1.124	1.042	0.505	1.699	1.886	2.152
		3	0.2008	0.212	0.308	1.050	0.986	0.556	1.722	1.886	2.218
1000	1	0	0.1997	0.246	0.266	1.071	1.070	0.519	1.669	1.815	2.206
		1	0.2012	0.252	0.278	1.069	1.042	0.509	1.661	1.857	2.118
		3	0.2008	0.188	0.313	0.917	0.986	0.481	1.505	1.698	1.956
	3	0	0.2002	0.29	0.260	1.017	1.070	0.454	1.539	1.643	2.028
		1	0.2009	0.252	0.278	1.066	1.042	0.490	1.641	1.791	2.110
		3	0.2008	0.19	0.312	0.999	0.986	0.524	1.643	1.808	2.059
2000	1	0	0.1993	0.23	0.239	1.088	1.070	0.505	1.696	1.854	2.110
		1	0.2009	0.26	0.255	1.023	1.042	0.467	1.563	1.769	2.106
		3	0.2009	0.238	0.306	0.936	0.986	0.492	1.517	1.743	1.985
	3	0	0.1995	0.292	0.224	1.053	1.070	0.513	1.653	1.786	2.162
		1	0.2007	0.282	0.250	0.961	1.042	0.478	1.499	1.665	1.988
		3	0.2006	0.3	0.292	1.078	0.986	0.559	1.675	1.953	2.271

Note: $\tau = 0.28 - 0.000006 \times n + 0.003 \times \hat{\beta} + 0.122 \times cr + 0.00001 \times n \cdot \hat{\beta} - 0.0001 \times n \cdot cr - 0.16 \times \hat{\beta} \cdot cr$

$\mu = 1.07 - 0.028 \times \hat{\beta}$

Table 3.16Summary statistics of $\sqrt[3]{LRT}$ and the fitted values of τ and μ at exponential censoring rate 30% (500 replications)

n	λ	β	Observed censoring fraction	Observed fraction LRT_Z	Fitted fraction LRT_Z (τ)	Observed mean $\sqrt[3]{LRT_{NZ}}$	Fitted mean $\sqrt[3]{LRT_{NZ}}$ (μ)	Observed SD $\sqrt[3]{LRT_{NZ}}$	Percentiles of $\sqrt[3]{LRT}$ 90	95	99
500	1	0	0.3003	0.292	0.285	1.043	1.070	0.505	1.602	1.741	2.172
		1	0.3015	0.316	0.291	1.064	1.042	0.494	1.566	1.743	2.239
		3	0.2994	0.156	0.322	0.986	0.986	0.519	1.634	1.815	2.077
	3	0	0.3006	0.264	0.288	1.074	1.070	0.492	1.668	1.801	2.092
		1	0.2988	0.266	0.295	1.033	1.042	0.473	1.588	1.785	2.058
		3	0.298	0.252	0.313	0.943	0.986	0.481	1.521	1.752	2.160
1000	1	0	0.3001	0.258	0.272	1.117	1.070	0.522	1.732	1.890	2.178
		1	0.3003	0.252	0.286	1.034	1.042	0.486	1.632	1.823	2.067
		3	0.2999	0.168	0.325	0.939	0.986	0.503	1.578	1.819	2.148
	3	0	0.298	0.27	0.270	1.121	1.070	0.503	1.699	1.849	2.115
		1	0.3011	0.236	0.289	1.043	1.042	0.471	1.595	1.781	2.035
		3	0.3004	0.18	0.323	0.974	0.986	0.492	1.582	1.737	2.069
2000	1	0	0.2998	0.21	0.253	1.033	1.070	0.491	1.578	1.840	2.169
		1	0.2994	0.212	0.275	1.107	1.042	0.487	1.647	1.870	2.269
		3	0.2992	0.172	0.331	0.991	0.986	0.501	1.616	1.769	2.025
	3	0	0.3003	0.27	0.238	1.006	1.070	0.465	1.531	1.708	2.018
		1	0.2997	0.202	0.277	1.060	1.042	0.479	1.607	1.772	2.030
		3	0.2994	0.188	0.327	1.019	0.986	0.473	1.569	1.760	2.191

Note: $\tau = 0.28 - 0.000006 \times n + 0.003 \times \hat{\beta} + 0.122 \times cr + 0.00001 \times n \cdot \hat{\beta} - 0.0001 \times n \cdot cr - 0.16 \times \hat{\beta} \cdot cr$

$\mu = 1.07 - 0.028 \times \hat{\beta}$

Table 3.17 Mean values of selected percentiles of $\sqrt[3]{LRT}$ averaged over (λ, β) settings

Expected censoring rate	Percentile	n	Mean	SD	Low 95% CI	Up 95% CI
10%	95th	500	1.798	0.024	1.736	1.860
		1000	1.727	0.043	1.616	1.838
		2000	1.773	0.021	1.719	1.826
	99th	500	2.133	0.036	2.040	2.225
		1000	2.099	0.042	1.991	2.207
		2000	2.122	0.022	2.065	2.179
20%	95th	500	1.849	0.015	1.809	1.888
		1000	1.769	0.033	1.684	1.854
		2000	1.795	0.040	1.691	1.898
	99th	500	2.138	0.018	2.091	2.184
		1000	2.079	0.035	1.990	2.169
		2000	2.104	0.044	1.990	2.218
30%	95th	500	1.773	0.013	1.739	1.806
		1000	1.817	0.022	1.761	1.872
		2000	1.787	0.024	1.726	1.849
	99th	500	2.133	0.028	2.060	2.205
		1000	2.102	0.022	2.045	2.159
		2000	2.117	0.044	2.005	2.230

Based on 500 replications for each setting.

Table 3.18 Critical values of $\sqrt[3]{LRT}$

	Sample size		
	500	1000	2000
$\alpha = 0.05$	1.807	1.771	1.785
$\alpha = 0.01$	2.135	2.093	2.114

Table 4.1 Summary statistics for simulated MLE when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.6$ (mixture of two regimes model)

n	Observed censoring rate	Parameters	Mean MLE	SD MLE	Percentile of MLE		
					25	50	75
500	10.02%	λ_1	1.002	0.205	0.851	0.976	1.116
		λ_2	1.068	0.272	0.878	1.032	1.183
		β_1	1.505	0.076	1.463	1.508	1.55
		β_2	0.495	0.086	0.446	0.498	0.54
		π	0.6	0.033	0.578	0.601	0.623
	19.99%	λ_1	1.006	0.223	0.834	0.992	1.138
		λ_2	1.054	0.303	0.831	0.996	1.231
		β_1	1.505	0.064	1.46	1.502	1.55
		β_2	0.498	0.082	0.443	0.499	0.555
		π	0.599	0.032	0.576	0.601	0.622
	30.34%	λ_1	1.007	0.248	0.837	0.983	1.155
		λ_2	1.055	0.445	0.792	0.991	1.243
		β_1	1.505	0.073	1.457	1.507	1.548
		β_2	0.502	0.1	0.436	0.499	0.566
		π	0.603	0.035	0.578	0.604	0.628
1000	10.14%	λ_1	1.018	0.135	0.919	1.001	1.112
		λ_2	1.02	0.167	0.907	1	1.118
		β_1	1.498	0.04	1.47	1.499	1.527
		β_2	0.496	0.048	0.466	0.5	0.526
		π	0.599	0.023	0.585	0.6	0.615
	19.88%	λ_1	1.066	0.377	0.915	1.014	1.118
		λ_2	1.129	0.558	0.893	1.018	1.195
		β_1	1.456	0.206	1.463	1.496	1.529
		β_2	0.522	0.144	0.456	0.497	0.547
		π	0.595	0.032	0.579	0.597	0.615
	30.26%	λ_1	1.015	0.167	0.898	1.007	1.123
		λ_2	1.035	0.261	0.85	1.012	1.187
		β_1	1.5	0.048	1.467	1.497	1.534
		β_2	0.499	0.069	0.451	0.501	0.541
		π	0.6	0.023	0.585	0.601	0.615
2000	10.16%	λ_1	1.003	0.098	0.934	0.999	1.073
		λ_2	1.006	0.109	0.925	1	1.075
		β_1	1.501	0.03	1.48	1.5	1.521
		β_2	0.5	0.033	0.477	0.502	0.523
		π	0.6	0.017	0.588	0.601	0.612

Based on 500 replications for each setting.

Table 4.1 Summary statistics for simulated MLE when
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.6$ (mixture of two regimes model) (continued)

n	Observed censoring rate	Parameters	Mean MLE	SD	MLE Percentile of MLE 25	50	75
2000	19.93%	λ_1	0.994	0.098	0.93	0.994	1.061
		λ_2	1.006	0.143	0.904	0.996	1.091
		β_1	1.503	0.03	1.481	1.501	1.523
		β_2	0.502	0.041	0.475	0.5	0.531
		π	0.6	0.016	0.589	0.6	0.611
	30.25%	λ_1	0.995	0.133	0.904	0.994	1.068
		λ_2	1.037	0.2	0.897	1.019	1.166
		β_1	1.504	0.038	1.48	1.502	1.529
		β_2	0.495	0.053	0.457	0.496	0.528
		π	0.6	0.017	0.588	0.6	0.611

Based on 500 replications for each setting.

Table 4.2 The minimum and maximum of MLEs when
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.6$, sample size 1000

		$\hat{\lambda}_1$		$\hat{\lambda}_2$		$\hat{\beta}_1$		$\hat{\beta}_2$		$\hat{\pi}$	
		Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Expected censoring rate	10%	0.707	1.428	0.657	1.852	1.403	1.602	0.321	0.658	0.521	0.663
	20%	0.591	4.499	0.291	4.658	0.328	1.662	0.152	1.431	0.500	0.710
	30%	0.602	1.664	0.448	2.094	1.359	1.656	0.292	0.724	0.531	0.666

Based on 500 replications for each setting.

Table 4.3 Mean and standard deviation of survival time and covariate x of first and second regimes at 0% censoring rate

		First regime		Second regime	
		mean	SD	mean	SD
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5, \pi = 0.6$ n= 595 vs. 405	t	0.258	0.432	0.368	0.523
	x	2.591		2.527	
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5, \pi = 0.6$ n= 606 vs. 394	t	0.222	0.431	0.346	0.459
	x	2.508		2.586	
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.25, \beta_2 = 0.5, \pi = 0.6$ n= 611 vs. 389	t	0.159	0.346	0.346	0.470
	x	2.510		2.540	
$\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1.5, \beta_2 = 0.5, \pi = 0.6$ n= 612 vs. 388	t	0.129	0.290	0.358	0.540
	x	2.460		2.432	

Based on sample size: 1000

Table 4.4 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
0.1	500	2.512	0.288	0.397	10.5%	2.278	1.029	0.584	0.512	0.215	1.267	0.481	0.125	0.435	0.604	1.285	0.026	0.114
		2.497	0.245	0.319	20.0%	1.970	0.973	0.521	0.530	0.176	1.232	0.516	0.109	0.305	0.489	1.279	0.040	0.230
		2.509	0.200	0.250	29.9%	1.706	1.060	0.702	0.507	0.231	1.251	0.536	0.126	0.315	0.524	1.291	0.036	0.156
	1000	2.501	0.290	0.402	10.3%	1.630	1.006	0.545	0.516	0.242	1.349	0.518	0.071	0.338	0.633	1.400	0.058	0.216
		2.506	0.240	0.321	20.3%	2.029	1.018	0.781	0.511	0.165	1.325	0.499	0.160	0.438	0.660	1.335	0.052	0.200
		2.505	0.199	0.251	29.9%	2.395	1.026	0.547	0.513	0.239	1.312	0.521	0.191	0.389	0.579	1.346	0.060	0.192
	2000	2.496	0.289	0.400	10.3%	1.270	1.009	0.708	0.514	0.131	1.449	0.513	0.131	0.533	0.730	1.523	0.088	0.278
		2.496	0.244	0.321	20.2%	0.947	1.060	0.773	0.493	0.191	1.560	0.483	0.222	0.635	0.884	1.621	0.106	0.358
		2.494	0.199	0.252	29.9%	1.114	1.055	0.700	0.497	0.220	1.420	0.530	0.205	0.481	0.699	1.469	0.088	0.262
0.2	500	2.492	0.265	0.385	10.3%	1.588	1.007	0.633	0.528	0.232	1.509	0.506	0.257	0.640	0.794	1.543	0.092	0.300
		2.499	0.220	0.308	19.9%	1.661	1.059	0.664	0.525	0.224	1.518	0.522	0.103	0.528	0.799	1.568	0.100	0.302
		2.510	0.172	0.234	29.9%	2.001	1.046	0.659	0.530	0.241	1.486	0.511	0.181	0.574	0.786	1.538	0.072	0.274
	1000	2.496	0.269	0.391	10.0%	2.023	1.021	0.649	0.519	0.233	1.739	0.526	0.340	0.798	1.043	1.791	0.222	0.514
		2.503	0.220	0.307	19.8%	1.676	1.033	0.686	0.517	0.233	1.751	0.484	0.403	0.909	1.174	1.760	0.254	0.496
		2.502	0.173	0.234	29.8%	2.254	1.038	0.704	0.520	0.211	1.699	0.519	0.249	0.766	0.994	1.737	0.220	0.470
	2000	2.501	0.266	0.391	10.1%	1.443	1.023	0.543	0.535	0.259	1.907	0.503	0.419	1.028	1.268	1.954	0.358	0.630
		2.500	0.219	0.306	20.0%	1.266	1.028	0.759	0.508	0.217	2.050	0.454	0.746	1.213	1.487	2.080	0.468	0.762
		2.499	0.173	0.234	29.9%	1.820	1.005	0.254	0.550	0.168	1.698	0.543	0.202	0.667	0.989	1.738	0.248	0.466
0.3	500	2.497	0.236	0.369	10.5%	1.728	1.042	0.678	0.546	0.247	1.714	0.533	0.223	0.748	1.004	1.760	0.226	0.462
		2.508	0.193	0.290	20.2%	1.428	1.059	0.729	0.541	0.249	1.719	0.497	0.472	0.757	1.032	1.759	0.200	0.470
		2.489	0.158	0.220	30.2%	1.672	1.052	0.695	0.551	0.272	1.717	0.498	0.522	0.829	1.018	1.744	0.192	0.450

Table 4.4 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$ 1	5	10	50	power α =001	power $\alpha = 0.05$
0.3	1000	2.498	0.243	0.379	10.5%	1.248	1.041	0.705	0.536	0.285	2.053	0.465	0.735	1.241	1.447	2.082	0.494	0.736
		2.497	0.198	0.291	20.4%	1.236	1.032	0.730	0.534	0.267	2.048	0.515	0.762	1.037	1.273	2.089	0.498	0.742
		2.501	0.157	0.222	30.1%	1.277	1.040	0.705	0.540	0.279	1.978	0.508	0.502	1.005	1.277	2.043	0.454	0.708
	2000	2.503	0.244	0.375	10.4%	1.002	1.047	0.734	0.515	0.302	2.582	0.442	1.460	1.793	2.003	2.602	0.854	0.952
		2.499	0.196	0.292	20.3%	1.078	1.018	0.736	0.524	0.290	2.475	0.432	1.435	1.756	1.904	2.483	0.812	0.942
		2.500	0.156	0.221	30.2%	1.175	1.009	0.661	0.547	0.309	2.330	0.487	0.947	1.447	1.702	2.378	0.706	0.870
0.4	500	2.495	0.211	0.351	10.2%	1.358	1.037	0.662	0.575	0.300	1.931	0.519	0.582	1.046	1.245	1.965	0.350	0.622
		2.510	0.169	0.272	20.0%	1.339	1.031	0.672	0.573	0.285	1.887	0.478	0.525	0.994	1.284	1.922	0.302	0.624
		2.492	0.132	0.203	30.0%	1.593	1.089	0.749	0.570	0.271	1.838	0.511	0.281	0.960	1.181	1.888	0.288	0.560
	1000	2.499	0.217	0.359	9.9%	1.196	1.032	0.701	0.560	0.322	2.330	0.469	1.168	1.498	1.742	2.361	0.704	0.888
		2.493	0.172	0.273	20.1%	1.154	1.034	0.702	0.570	0.319	2.332	0.439	1.130	1.611	1.773	2.348	0.720	0.902
		2.500	0.133	0.203	30.1%	1.170	1.061	0.676	0.570	0.302	2.221	0.482	0.878	1.346	1.575	2.258	0.646	0.830
	2000	2.502	0.215	0.358	10.0%	1.010	1.042	0.665	0.559	0.362	2.890	0.444	1.824	2.111	2.328	2.889	0.950	0.992
		2.499	0.173	0.274	20.0%	0.998	1.040	0.708	0.552	0.350	2.850	0.401	1.759	2.188	2.358	2.866	0.960	0.990
		2.499	0.132	0.203	30.0%	1.173	1.044	0.707	0.558	0.342	2.747	0.424	1.581	2.005	2.192	2.769	0.924	0.984
0.5	500	2.501	0.192	0.348	9.8%	0.996	1.354	0.668	0.589	0.586	2.092	0.480	0.686	1.234	1.457	2.117	0.490	0.760
		2.499	0.150	0.254	19.9%	1.068	1.295	0.623	0.661	0.648	1.945	0.546	0.341	0.904	1.217	2.011	0.392	0.642
		2.498	0.109	0.182	30.1%	0.871	1.839	0.698	0.486	0.688	1.949	0.536	0.496	0.949	1.259	1.991	0.394	0.660
	1000	2.506	0.188	0.339	10.0%	1.001	1.242	0.603	0.644	0.585	2.535	0.463	1.487	1.803	1.935	2.542	0.822	0.956
		2.498	0.146	0.252	20.0%	1.008	1.256	0.586	0.673	0.615	2.455	0.454	1.467	1.688	1.842	2.475	0.788	0.932
		2.496	0.109	0.184	30.0%	0.913	1.414	0.751	0.495	0.580	2.401	0.476	1.134	1.523	1.730	2.421	0.770	0.884

Table 4.4 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
0.5	2000	2.499	0.191	0.344	9.9%	0.994	1.059	0.730	0.521	0.508	3.212	0.395	2.144	2.540	2.710	3.230	0.994	0.998
		2.497	0.145	0.251	19.9%	0.978	1.128	0.632	0.617	0.556	3.109	0.412	2.085	2.443	2.580	3.096	0.990	1.000
		2.498	0.111	0.185	30.0%	0.899	1.375	0.744	0.470	0.584	2.939	0.446	1.642	2.115	2.337	2.978	0.952	0.982
0.6	500	2.509	0.163	0.318	10.1%	1.013	1.369	0.675	0.559	0.688	2.168	0.496	1.034	1.314	1.520	2.169	0.528	0.772
		2.500	0.123	0.232	19.6%	1.029	1.565	0.658	0.617	0.703	2.089	0.496	0.739	1.196	1.422	2.084	0.474	0.734
		2.504	0.090	0.162	29.8%	0.995	1.398	0.672	0.563	0.714	1.995	0.533	0.525	0.991	1.326	2.045	0.436	0.662
	1000	2.500	0.170	0.326	10.3%	1.009	1.144	0.685	0.567	0.656	2.711	0.448	1.579	1.951	2.158	2.726	0.910	0.980
		2.491	0.124	0.235	19.6%	1.025	1.132	0.677	0.587	0.666	2.586	0.441	1.560	1.860	2.020	2.613	0.864	0.956
		2.494	0.090	0.164	30.2%	1.025	1.164	0.673	0.591	0.669	2.477	0.478	1.149	1.580	1.880	2.475	0.816	0.924
	2000	2.503	0.167	0.322	10.2%	1.001	1.065	0.694	0.556	0.637	3.366	0.460	2.149	2.575	2.731	3.382	0.992	1.000
		2.498	0.122	0.233	19.7%	0.996	1.150	0.699	0.549	0.645	3.245	0.429	2.060	2.511	2.681	3.267	0.990	1.000
		2.497	0.090	0.164	30.0%	0.998	1.086	0.692	0.571	0.639	3.070	0.440	1.900	2.349	2.533	3.081	0.974	0.996
0.7	500	2.500	0.148	0.305	10.1%	1.030	1.414	0.695	0.556	0.721	2.170	0.534	0.739	1.256	1.489	2.161	0.526	0.764
		2.507	0.100	0.212	20.0%	0.994	1.475	0.698	0.582	0.733	2.070	0.509	0.815	1.174	1.397	2.102	0.478	0.716
		2.509	0.074	0.148	30.6%	0.998	1.358	0.706	0.602	0.725	2.004	0.526	0.615	1.118	1.315	2.019	0.392	0.644
	1000	2.497	0.149	0.307	10.2%	0.992	1.220	0.722	0.519	0.701	2.656	0.482	1.458	1.771	2.039	2.662	0.894	0.950
		2.496	0.104	0.212	20.0%	0.994	1.209	0.725	0.515	0.727	2.619	0.477	1.323	1.848	2.039	2.610	0.882	0.964
		2.508	0.074	0.148	30.2%	0.999	1.214	0.717	0.574	0.718	2.474	0.495	1.291	1.597	1.794	2.494	0.890	0.912
	2000	2.505	0.146	0.307	10.0%	0.991	1.101	0.741	0.508	0.686	3.384	0.402	2.487	2.774	2.872	3.371	0.998	1.000
		2.498	0.105	0.215	20.0%	0.984	1.119	0.741	0.512	0.691	3.252	0.425	2.178	2.524	2.675	3.247	0.994	1.000
		2.495	0.075	0.149	30.1%	0.994	1.166	0.726	0.533	0.705	3.036	0.451	1.835	2.323	2.465	3.039	0.984	0.994

Table 4.4 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 0.75, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
0.8	500	2.508	0.122	0.283	10.1%	0.993	1.544	0.739	0.486	0.773	2.053	0.578	0.380	0.870	1.296	2.112	0.482	0.722
		2.494	0.087	0.195	20.4%	0.986	1.573	0.738	0.491	0.761	1.971	0.557	0.598	0.953	1.224	2.005	0.412	0.622
		2.494	0.056	0.128	30.1%	1.014	1.613	0.728	0.505	0.769	1.846	0.568	0.257	0.825	1.062	1.874	0.300	0.560
	1000	2.499	0.121	0.281	10.2%	0.992	1.286	0.741	0.517	0.760	2.530	0.568	0.861	1.582	1.816	2.561	0.792	0.918
		2.490	0.088	0.194	20.5%	0.977	1.436	0.743	0.494	0.775	2.442	0.541	0.921	1.506	1.712	2.507	0.740	0.888
		2.506	0.056	0.127	30.3%	0.969	1.476	0.744	0.483	0.776	2.285	0.530	0.896	1.365	1.591	2.285	0.666	0.838
	2000	2.500	0.121	0.281	10.2%	0.987	1.184	0.752	0.490	0.769	3.187	0.467	1.834	2.365	2.578	3.197	0.980	0.996
		2.498	0.088	0.197	20.3%	0.982	1.196	0.749	0.508	0.769	2.970	0.500	1.845	2.141	2.328	2.988	0.952	0.992
		2.499	0.055	0.127	30.4%	0.997	1.177	0.748	0.515	0.758	2.814	0.489	1.598	1.903	2.166	2.843	0.916	0.982
0.9	500	2.490	0.102	0.262	10.1%	1.000	1.547	0.742	0.526	0.834	1.762	0.597	0.174	0.698	0.993	1.766	0.282	0.472
		2.501	0.064	0.166	19.8%	0.982	1.546	0.750	0.540	0.814	1.646	0.637	0.104	0.559	0.810	1.647	0.222	0.406
		2.495	0.040	0.106	30.3%	1.010	1.278	0.739	0.691	0.810	1.571	0.596	0.177	0.610	0.809	1.599	0.180	0.360
	1000	2.502	0.099	0.259	10.2%	0.981	1.340	0.746	0.569	0.810	1.999	0.649	0.215	0.762	1.125	2.060	0.478	0.654
		2.502	0.065	0.169	19.9%	0.993	1.284	0.748	0.531	0.826	2.026	0.578	0.379	1.072	1.338	2.015	0.444	0.674
		2.499	0.039	0.109	30.3%	0.978	1.264	0.754	0.557	0.824	1.830	0.630	0.267	0.672	1.017	1.854	0.352	0.564
	2000	2.500	0.099	0.260	10.1%	0.992	1.251	0.750	0.528	0.815	2.564	0.613	1.015	1.491	1.772	2.621	0.774	0.898
		2.505	0.066	0.171	20.2%	0.980	1.350	0.755	0.487	0.842	2.457	0.606	0.867	1.384	1.670	2.481	0.728	0.866
		2.503	0.040	0.109	29.8%	0.986	1.248	0.755	0.521	0.840	2.324	0.588	0.782	1.238	1.557	2.362	0.658	0.832

Note: Based on 500 replications for each setting.

The level of significance 0.01, the critical values were used 2.135 for n=500, 2.093 for n=1000 and 2.114 for n=2000 respectively.
The level of significance 0.05, the critical values were used 1.807 for n=500, 1.771 for n=1000 and 1.785 for n=2000 respectively.

Table 4.5 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
0.1	500	2.503	0.297	0.402	10.4%	7.871	1.035	0.907	0.496	0.170	1.896	0.529	0.010	0.994	1.346	1.928	0.350	0.600
		2.501	0.249	0.322	20.0%	3.006	1.036	0.871	0.503	0.181	1.711	0.525	0.010	0.662	1.022	1.790	0.210	0.470
		2.506	0.204	0.253	29.9%	2.399	1.058	0.894	0.494	0.1177	1.828	0.647	0.009	0.653	0.995	1.894	0.270	0.580
	1000	2.505	0.296	0.402	10.3%	1.308	1.044	1.051	0.494	0.152	2.084	0.571	0.014	1.053	1.358	2.134	0.560	0.770
		2.494	0.251	0.325	20.3%	1.567	1.051	0.923	0.494	0.152	2.136	0.537	0.005	1.171	1.456	2.212	0.610	0.810
		2.500	0.206	0.254	29.9%	5.023	1.042	1.121	0.492	0.143	2.207	0.486	0.430	1.306	1.617	2.269	0.650	0.860
	2000	2.502	0.297	0.403	10.3%	1.109	1.024	1.013	0.492	0.112	2.655	0.484	0.013	1.939	2.157	2.720	0.920	0.960
		2.505	0.248	0.321	20.2%	1.114	1.024	0.975	0.496	0.131	2.650	0.536	0.011	1.594	2.081	2.688	0.880	0.950
		2.497	0.207	0.256	29.9%	1.372	1.033	1.045	0.490	0.119	2.650	0.454	0.018	2.126	2.169	2.650	0.960	0.990
0.2	500	2.498	0.281	0.391	10.3%	1.129	1.075	1.017	0.483	0.225	2.604	0.441	1.462	1.713	1.857	2.635	0.860	0.920
		2.502	0.235	0.310	19.9%	1.737	1.047	0.995	0.501	0.214	2.522	0.462	1.294	1.678	1.939	2.522	0.800	0.940
		2.499	0.194	0.244	29.9%	1.327	1.045	1.004	0.500	0.229	2.510	0.461	0.984	1.658	1.949	2.499	0.800	0.950
	1000	2.501	0.281	0.396	10.0%	1.275	1.033	0.992	0.489	0.212	3.268	0.378	2.308	2.601	2.831	3.276	1.000	1.000
		2.509	0.233	0.310	19.8%	1.237	1.028	1.032	0.498	0.194	3.159	0.348	2.305	2.634	2.726	3.158	1.000	1.000
		2.511	0.191	0.241	29.8%	1.087	1.048	1.042	0.490	0.214	3.117	0.394	2.105	2.561	2.645	3.077	1.000	1.000
	2000	2.502	0.281	0.396	10.1%	1.116	1.010	0.990	0.497	0.202	3.987	0.379	3.015	3.356	3.457	4.006	1.000	1.000
		2.500	0.233	0.311	20.0%	0.947	1.051	1.036	0.487	0.211	4.024	0.319	2.909	3.525	3.635	4.022	1.000	1.000
		2.498	0.192	0.244	29.9%	1.034	1.036	1.022	0.492	0.205	3.918	0.296	3.154	3.351	3.522	3.912	1.000	1.000
0.3	500	2.499	0.265	0.382	10.5%	1.191	1.040	0.991	0.506	0.299	3.216	0.404	2.315	2.469	2.571	3.270	1.000	1.000
		2.493	0.219	0.297	20.2%	1.028	1.096	1.028	0.489	0.302	3.132	0.412	2.079	2.531	2.613	3.109	0.980	1.000
		2.507	0.174	0.227	30.2%	1.642	1.091	1.034	0.496	0.311	3.101	0.381	1.752	2.516	2.632	3.085	0.990	0.990

Table 4.5 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
0.3	1000	2.490	0.266	0.386	10.5%	1.050	1.023	0.997	0.497	0.302	3.941	0.336	3.160	3.415	3.516	3.926	1.000	1.000
		2.506	0.218	0.297	20.4%	1.020	1.036	1.028	0.494	0.295	3.931	0.346	3.103	3.369	3.475	3.961	1.000	1.000
		2.511	0.175	0.226	30.1%	1.091	1.020	1.004	0.500	0.306	3.817	0.349	3.000	3.279	3.391	3.773	1.000	1.000
	2000	2.500	0.263	0.379	10.4%	0.981	1.032	1.014	0.490	0.310	5.088	0.297	4.300	4.456	4.654	5.106	1.000	1.000
		2.499	0.218	0.299	20.3%	1.028	1.022	1.004	0.498	0.298	4.868	0.279	4.230	4.442	4.520	4.862	1.000	1.000
		2.501	0.175	0.226	30.2%	1.082	1.011	0.988	0.499	0.300	4.786	0.331	3.928	4.258	4.315	4.796	1.000	1.000
0.4	500	2.500	0.249	0.372	10.2%	1.178	1.040	0.937	0.548	0.396	3.674	0.374	2.441	2.955	3.168	3.641	1.000	1.000
		2.500	0.205	0.285	20.0%	0.958	1.101	0.983	0.539	0.384	3.600	0.376	2.547	2.941	3.080	3.635	1.000	1.000
		2.496	0.162	0.214	30.0%	0.985	1.156	0.991	0.523	0.393	3.460	0.355	2.420	2.716	2.942	3.535	1.000	1.000
	1000	2.500	0.249	0.373	9.9%	1.050	1.010	0.980	0.521	0.396	4.603	0.327	3.380	4.042	4.208	4.613	1.000	1.000
		2.502	0.203	0.289	20.1%	1.025	1.027	0.987	0.520	0.400	4.528	0.292	3.804	4.070	4.100	4.515	1.000	1.000
		2.499	0.162	0.216	30.1%	0.941	1.086	0.999	0.515	0.401	4.451	0.340	3.446	3.918	4.069	4.401	1.000	1.000
	2000	2.495	0.248	0.370	10.0%	1.037	1.017	0.994	0.497	0.401	5.785	0.305	5.011	5.198	5.355	5.816	1.000	1.000
		2.498	0.204	0.287	20.0%	1.013	1.018	0.999	0.505	0.396	5.652	0.304	4.938	5.171	5.234	5.644	1.000	1.000
		2.499	0.163	0.219	30.0%	1.049	0.989	0.988	0.509	0.403	5.550	0.316	4.761	4.974	5.175	5.553	1.000	1.000
0.5	500	2.495	0.235	0.358	9.8%	1.060	1.063	0.885	0.599	0.517	4.032	0.383	2.872	3.420	3.519	4.033	1.000	1.000
		2.510	0.187	0.270	19.9%	1.111	0.995	0.561	0.954	0.502	3.923	0.347	3.108	3.240	3.455	3.922	1.000	1.000
		2.502	0.147	0.200	30.1%	1.030	1.133	0.987	0.516	0.511	3.872	0.320	3.221	3.406	3.453	3.860	1.000	1.000
	1000	2.493	0.235	0.361	10.0%	1.044	1.001	0.938	0.568	0.493	5.068	0.327	4.418	4.534	4.614	5.097	1.000	1.000
		2.507	0.187	0.273	20.0%	1.044	1.020	0.537	0.969	0.503	5.047	0.331	4.154	4.424	4.597	5.054	1.000	1.000
		2.502	0.148	0.203	30.0%	1.011	1.036	0.863	0.646	0.514	4.825	0.290	4.114	4.277	4.416	4.834	1.000	1.000

Table 4.5 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$				power α =001	power $\alpha = 0.05$
													1	5	10	50		
		2.496	0.233	0.360	9.9%	1.017	1.004	0.974	0.530	0.501	6.434	0.278	5.653	5.875	6.085	6.429	1.000	1.000
0.5	2000	2.501	0.188	0.272	19.9%	1.003	1.016	1.005	0.499	0.496	6.269	0.244	5.604	5.861	5.948	6.268	1.000	1.000
		2.500	0.148	0.202	30.0%	1.007	1.023	1.005	0.496	0.503	6.150	0.291	5.285	5.701	5.781	6.171	1.000	1.000
		2.501	0.217	0.344	10.1%	0.944	1.153	0.975	0.511	0.620	4.340	0.334	3.629	3.737	3.899	4.303	1.000	1.000
	500	2.503	0.172	0.259	19.6%	1.034	1.133	0.968	0.532	0.609	4.171	0.339	3.340	3.573	3.805	4.146	1.000	1.000
		2.507	0.133	0.186	29.8%	1.033	1.158	0.962	0.537	0.607	4.128	0.326	3.247	3.560	3.688	4.141	1.000	1.000
		2.499	0.216	0.344	10.3%	0.997	1.038	0.994	0.507	0.603	5.434	0.360	4.555	4.821	4.920	5.420	1.000	1.000
0.6	1000	2.492	0.175	0.261	19.6%	0.971	1.087	0.989	0.514	0.594	5.269	0.351	4.346	4.609	4.789	5.291	1.000	1.000
		2.499	0.134	0.187	30.2%	0.985	1.040	0.995	0.509	0.604	5.109	0.344	4.377	4.511	4.647	5.126	1.000	1.000
		2.501	0.216	0.343	10.2%	0.998	1.026	0.997	0.497	0.603	6.920	0.288	6.269	6.458	6.523	6.963	1.000	1.000
	2000	2.502	0.175	0.262	19.7%	0.999	0.999	1.000	0.510	0.600	6.732	0.292	5.963	6.217	6.335	6.750	1.000	1.000
		2.500	0.133	0.186	30.0%	1.004	1.039	0.995	0.502	0.604	6.449	0.323	5.737	5.858	6.010	6.460	1.000	1.000
		2.503	0.200	0.330	10.1%	0.968	1.283	1.023	0.479	0.683	4.459	0.357	3.577	3.919	3.999	4.473	1.000	1.000
	500	2.500	0.158	0.241	20.0%	0.999	1.167	1.004	0.498	0.692	4.376	0.385	2.935	3.772	3.929	4.396	1.000	1.000
		2.496	0.120	0.169	30.6%	0.961	1.289	1.023	0.488	0.680	4.107	0.410	3.182	3.377	3.516	4.100	1.000	1.000
		2.506	0.200	0.329	10.2%	1.015	1.008	1.002	0.511	0.690	5.596	0.350	4.849	5.055	5.170	5.543	1.000	1.000
0.7	1000	2.500	0.157	0.240	20.0%	1.006	1.085	1.004	0.495	0.698	5.461	0.375	4.133	4.852	5.057	5.430	1.000	1.000
		2.502	0.120	0.171	30.2%	0.981	1.146	1.009	0.479	0.698	5.297	0.380	4.509	4.651	4.807	5.310	1.000	1.000
		2.497	0.201	0.330	10.0%	0.999	1.040	1.003	0.491	0.701	7.167	0.324	6.354	6.654	6.743	7.147	1.000	1.000
	2000	2.505	0.157	0.240	20.0%	1.020	1.003	0.999	0.508	0.699	6.927	0.319	6.149	6.397	6.523	6.907	1.000	1.000
		2.504	0.121	0.172	30.1%	0.989	1.047	1.007	0.502	0.697	6.640	0.311	5.813	5.987	6.131	6.703	1.000	1.000

Table 4.5 Simulation results of $\sqrt[3]{LRT}$ with summary statistics when $\lambda_1 = 1, \lambda_2 = 1, \beta_1 = 1, \beta_2 = 0.5$ and $\pi = 0.1 \sim 0.9$ (continued)

π	n	mean x	mean t	SD t	observed censoring rate	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\pi}$	Mean $\sqrt[3]{LRT}$	SD $\sqrt[3]{LRT}$	Percentile of $\sqrt[3]{LRT}$ 1	5	10	50	power α =001	power α =0.05
0.8	500	2.499	0.184	0.309	10.1%	1.003	1.293	1.007	0.465	0.787	4.481	0.413	3.506	3.740	3.800	4.511	1.000	1.000
		2.493	0.142	0.224	20.4%	0.984	1.473	1.012	0.462	0.788	4.212	0.452	3.119	3.427	3.609	4.220	1.000	1.000
		2.499	0.107	0.154	30.1%	0.984	1.648	1.014	0.523	0.795	3.931	0.501	2.408	3.051	3.347	3.891	1.000	1.000
	1000	2.499	0.185	0.315	10.2%	0.996	1.107	1.002	0.489	0.800	5.595	0.399	4.368	4.942	5.128	5.596	1.000	1.000
		2.500	0.142	0.223	20.5%	0.980	1.252	1.010	0.477	0.794	5.373	0.446	4.195	4.412	4.772	5.406	1.000	1.000
		2.502	0.107	0.155	30.3%	0.976	1.826	1.010	0.456	0.800	5.166	0.450	4.210	4.358	4.596	5.151	1.000	1.000
	2000	2.499	0.184	0.311	10.2%	0.983	1.062	1.008	0.494	0.797	7.057	0.359	6.206	6.412	6.555	7.078	1.000	1.000
		2.499	0.143	0.224	20.3%	1.004	1.011	0.999	0.515	0.797	6.730	0.399	5.871	6.029	6.253	6.690	1.000	1.000
		2.500	0.107	0.156	30.4%	0.986	1.111	1.006	0.499	0.799	6.427	0.363	5.534	5.798	5.960	6.404	1.000	1.000
0.9	500	2.491	0.167	0.291	10.1%	1.007	2.549	1.006	0.459	0.874	3.869	0.670	2.015	2.632	2.941	3.871	0.990	1.000
		2.506	0.128	0.206	19.8%	0.972	4.036	1.021	0.443	0.875	3.729	0.723	2.203	2.483	2.685	3.827	1.000	1.000
		2.501	0.094	0.142	30.3%	0.968	12.640	1.013	0.467	0.880	3.449	0.641	1.701	2.317	2.651	3.452	0.980	0.990
	1000	2.498	0.169	0.297	10.2%	0.985	1.604	1.008	0.492	0.895	4.989	0.541	3.340	3.997	4.244	5.010	1.000	1.000
		2.503	0.128	0.206	19.9%	1.000	1.420	1.004	0.480	0.890	4.684	0.555	3.497	3.854	3.969	4.676	1.000	1.000
		2.490	0.096	0.142	30.3%	0.973	1.674	1.013	0.517	0.890	4.352	0.575	2.810	3.392	3.506	4.417	1.000	1.000
	2000	2.495	0.169	0.298	10.1%	0.996	1.228	1.003	0.486	0.897	6.303	0.479	4.268	5.568	5.772	6.278	1.000	1.000
		2.500	0.128	0.206	20.2%	1.004	1.188	0.999	0.500	0.897	5.972	0.487	4.548	5.097	5.299	6.052	1.000	1.000
		2.506	0.095	0.142	29.8%	1.004	1.268	1.000	0.495	0.899	5.657	0.543	4.294	4.760	5.018	5.624	1.000	1.000

Note: Based on 500 replications for each setting.

The level of significance 0.01, the critical values were used 2.135 for n=500, 2.093 for n=1000 and 2.114 for n=2000 respectively.
The level of significance 0.05, the critical values were used 1.807 for n=500, 1.771 for n=1000 and 1.785 for n=2000 respectively.

Table 4.6 Linear regression analysis results for mean of alternative LRT with two way interactions

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	288.681	100.136		2.883	.004
	sample size N	-.467	.052	-.657	-9.032	.000
	mixing proportion π	-265.445	186.081	-.155	-1.427	.155
	mixing proportion square π^2	-805.282	138.864	-.481	-5.799	.000
	distance between betas D	-831.853	109.712	-.525	-7.582	.000
	censoring rate CR	898.817	362.157	.166	2.482	.014
	N*π	.538	.050	.561	10.654	.000
	N*D	.815	.047	.988	17.489	.000
	N*CR	-.255	.160	-.095	-1.601	.110
	π*D	1956.614	112.563	.995	17.382	.000
	π*CR	-899.382	385.332	-.140	-2.334	.020
	D*CR	-600.261	355.954	-.107	-1.686	.093

a. Dependent Variable: mean LRT

b. R square 0.895

Table 4.7 *Logit(power)* linear regression analysis results of power at $\alpha = 0.01$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-14.113	1.182		-11.937	.000
	sample size N	.004	.001	.672	6.485	.000
	mixing proportion π	25.560	1.978	1.881	12.924	.000
	mixing proportion square π^2	-21.196	1.398	-1.599	-15.161	.000
	censoring rate CR	-3.044	4.160	-.071	-.732	.465
	distance between betas D	23.987	2.470	.854	9.711	.000
	N* π	.000	.001	-.071	-1.055	.293
	N*CR	.000	.002	-.029	-.388	.699
	N*D	-.005	.001	-.416	-4.861	.000
	π *CR	-2.951	3.880	-.058	-.761	.448
	π *D	3.521	2.534	.120	1.389	.167
	CR*D	9.730	8.014	.113	1.214	.227

a. Dependent Variable: logit(power)

b. R square: 0.919

Bibliography

[1] Richard E. Quandt, New Approach to Estimating Switching Regressions. Journal of American Statistical Association, Vol. 67 No. 338 (Jun., 1972) pp.306-310

[2] Richard E. Quandt, James B. Ramsey, Estimating Mixture of Normal Distributions and Switching Regressions. Journal of the American Statistical Association, Vol. 73, No. 364, (Dec., 1978), pp. 730-738

[3] Kazuo Yamaguchi, Accelerated Failure-Time Regression Models with a Regression Model of Surviving Fraction: An Application to the Analysis of "Permanent Employment" in Japan, Journal of the American Statistical Association, Vol. 87, No. 418. (Jun., 1992), pp. 284-292

[4] Joel B. Greenhouse, Applications of a Mixture Survival Model with Covariates to the Analysis of a Depression Prevention Trail. Statistics in Medicine, Vol. 15, 2077-2094 (1996)

[5] Aris Perperoglou, Antonis Keramopoullos, Hans C. van Houwelingen, Approaches in Modeling Long-term Survival: An Application to Breast Cancer. Statistics in Medicine, (2007), 26:2666-2685

[6] Susan Halabi, Bahadur Singh, Sample Size Determination for Comparing Several Survival Curves with Unequal Allocations. Statistics in Medicine. Statist. Med. 2004; 23:1793–1815

[7] Ping Hu, Anastasios A. Tsiatis, Marie Davidian, Estimating the Parameters in the Cox Model When Covariate Variables are Measured with Error. Biometrics, Vol. 54, No. 4, (Dec., 1998)

[8] Jeffery S. Buzas, Unbiased Scores in Proportional Hazards Regression with Covariate Measurement Error. Journal of Statistical Planning and Inference, 67 (1998) 247-257

[9] Fan Hui Kong, Weihua Huang, Xiaoming Li, Estimating Survival Curves Under Proportional Hazards Model with Covariate Measurement Errors. Journal of Statistics. Vol. 25: 573-587 (1998)

[10] Yingwei Peng, Fitting Semiparametric Cure Models. Computational Statistics and Data Analysis 41 (2003) 481-490

[11] Ross Maller, Xian Zhou, Survival Analysis with Long-Term Survivors. John Wiley & Sons (1996)

[12] Nelder, J.A., Mead, R., A Simplex Method for Function Minimization. The Computer Journal, 7 (4), pp.308-313 (1965)

[13] Yingwei Peng, Keith B.G. Dear, K.C. Carriere, Testing for Presence of Cured Patients: a Simulation Study. Statistics in Medicine 20 (2001) 1783-1796

[14] M. E. Ghitany, R. A. Maller, S. Zhou, Exponential Mixture Models with Long-Term Survivors and Covariates. Journal of Multivariate Analysis, 49, 218-214 (1994)

[15] Judy P. Sy, Jeremy M. G. Taylor, Estimation in a Cox Proportional Hazards Cure Model. Biometrics, Vol. 56, No. 1 (Mar., 2000), pp. 227-236

[16] Steven B. Caudill, Ram N. Acharya, Maximum Likelihood Estimation of a Mixture of Normal Regressions: Starting Values and Singularities. Communications in Statistics – Simulation and Computation, 27(3), 667-674 (1998)

[17] Steven G. Self, Kung-Yee Liang, Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. Journal of American Statistical Association, Vol. 82, No. 398 (Jun., 1987), pp. 605-610

[18] C. F. Jeff Wu, Michael Hamada, Experiments, Planning, Analysis and Parameter Design Optimization. Wiley Interscience (2000)

[19] John W. Tukey, Exploratory Data Analysis, Addison-Wesley (1997)

Appendix

Calculating the mean of censoring distribution for single regime model in *R*

```
n <- 500; # number of subjects
rep <-100; # number of replications
x<-matrix(0,rep,n); # covariate x
t<-matrix(0,rep,n); # survival time
c<-matrix(0,rep,n); # absence of censoring indicator c=1 uncensored, c=0 censored.
u<-matrix(0,rep,n); # censoring time
tt<-matrix(0,rep,n); # failure time

lambda0 <- 3; # initial parameter setting
beta0<-1; # initial parameter setting
cr<-0.20; # expected censoring rate
coe<- 0.78; # adjusted coefficient of E[x]

for ( k    in 1: rep)
    {
    x[k,]<- runif(n, min=0, max=5); # x~Uniform(0,5)
    u[k,]<-rexp( n,rate=lambda0*exp(coe*2.5*beta0)*(cr/(1-cr))); #censoring distn.
    for (i in 1:n)
        {
        tt[k,i]<-rexp( 1,rate=lambda0*exp(x[k,i]*beta0)); #failure time distn.
        t[k,i]<-min(tt[k,i],u[k,i]);
        if (tt[k,i] <= u[k,i]) c[k,i] <- 1 else c[k,i] <- 0;
        }
    }

1-sum(c)/(n*rep); # average of observed censoring rate
(1/(lambda0*exp(coe*2.5*beta0)*(cr/(1-cr)))); # mean of censoring distribution
```

Calculating the mean of censoring distribution for mixture of two regimes model in *R*

```
n <- 500; # number of subjects
rep <-100; # number of replicates
x<-matrix(0,rep,n); # covariate x
select<-matrix(0,rep,n); # criterion of mixing proportion
t<-matrix(0,rep,n); # survival time
c<-matrix(0,rep,n); # absence of censoring indicator c=1 uncensored, c=0 otherwise.
u<-matrix(0,rep,n); # censoring time
left<-matrix(0,rep,n); # first regime
right<-matrix(0,rep,n); # second regime
u<-matrix(0,rep,n); # censoring time
tt<-matrix(0,rep,n); # failure time

lam1 <- 1; # initial parameter setting
lam2 <- 1; # initial parameter setting
be1 <- 1.5; # initial parameter setting
be2 <- 0.5; # initial parameter setting
m <- 0.5; # initial parameter setting, mixing proportion
cr<- 0.3; # expected censoring rate
coe<- 0.85; # adjusted coefficient of E[x]

for ( k   in 1: rep)
    {
    select[k,] <- runif (n, 0, 1);
    x[k,]<-runif(n, min=0, max=5);
    u[k,]<-rexp(n, rate= (m*(lam1*exp(be1*coe*2.5))+
            (1-m)*(lam2*exp(be2*coe*2.5)))*cr/(1-cr)); #censoring distn.

    for (i in 1:n)
        {
        left[k,i] <- rexp(1, rate=lam1*exp(x[k,i]*be1)); #failure time of first regime
        right[k,i] <- rexp(1, rate=lam2*exp(x[k,i]*be2)); #failure time of second
                  regime
        if (select[k,i] <= m) tt[k,i] <- left[k,i] else tt[k,i] <- right[k,i];
        }
    for (i in 1:n)
        {
```

```
        t[k,i]<-min(tt[k,i],u[k,i]);
        if (tt[k,i] <= u[k,i]) c[k,i] <- 1 else c[k,i] <- 0;
        }
    }

1-(sum(c)/(n*rep));    # average of observed censoring rate
1/((m*(lam1*exp(be1*coe*2.5))+(1-m)*(lam2*exp(be2*coe*2.5)))*cr/(1-cr)); #
mean of censoring distribution
```

Microsoft Visual C++ code for the simulation study

```
const int num_sub = 500;         // number of subjects
const int num_rep = 500;       // number of replicates

// initial parameters values
// null parameters setting
const double lambda_h0 = 3;
const double beta_h0 = 1;
const double mean_censor_dist_h0 = 0.1896988;

/*
// alternative parameters setting
const double lambda1_h1 = 1;
const double lambda2_h1 = 1;
const double beta1_h1 = 0.75;
const double beta2_h1 = 0.5;
const double mix_h1 = 0.6;
const double mean_censor_dist_h1 = 0.5580538;
*/

// number of random starting points
const int num_of_init_mu=4;
const int num_of_init_be=4;

const int num_of_init_mu1=2;
const int num_of_init_mu2=2;
const int num_of_init_be1=2;
const int num_of_init_be2=2;
const int num_of_init_alp=3;

/* range of covariate x */
const double uniform_min = 0;
const double uniform_max = 5;

// single regime model function
double my_f0(const gsl_vector *v, void *params)
{
     power_data *my_pwr_data;
```

```
	my_pwr_data = (power_data*)params;
	double sumlog0=0;

	gsl_vector* covar_t = my_pwr_data->covar_t_st;
	gsl_vector* covar_c = my_pwr_data->covar_c_st;
	gsl_vector* covar_x = my_pwr_data->covar_x_st;
	int n = my_pwr_data->size;

	double mu = gsl_vector_get(v, 0);
	double beta = gsl_vector_get(v, 1);
	double lambda, logitem0;
	lambda=exp(mu);

for (int i = 0; i < n; i++)
	{
		logitem0 = gsl_vector_get(covar_c,i)*(log(lambda)+beta*
gsl_vector_get(covar_x,i))-lambda* gsl_vector_get(covar_t,i)* exp(beta*
gsl_vector_get(covar_x,i));
		sumlog0 = sumlog0 + logitem0;
	}

	return -sumlog0;
}

//mixture of two regimes model function
double my_f1(const gsl_vector *v, void *params)
{
	power_data *my_pwr_data;
	my_pwr_data = (power_data*)params;
	double sumlog1=0;

	gsl_vector* covar_t = my_pwr_data->covar_t_st;
	gsl_vector* covar_c = my_pwr_data->covar_c_st;
	gsl_vector* covar_x = my_pwr_data->covar_x_st;

	int n = my_pwr_data->size;

	double mu1 = gsl_vector_get(v, 0);
```

```
	double mu2 = gsl_vector_get(v, 1);
	double beta1 = gsl_vector_get(v, 2);
	double beta2 = gsl_vector_get(v, 3);
	double alpha = gsl_vector_get(v, 4);
	double lambda1, lambda2, mix, x1, x2, logiteml;
	lambda1=exp(mu1);
	lambda2=exp(mu2);
	mix= exp(alpha)/(1+exp(alpha));
	for (int i = 0; i < n; i++)
	{
	x1= mix*exp(-lambda1*gsl_vector_get(covar_t,i)*exp(beta1*gsl_vector_get(covar_x,i)));
	x2=
(1-mix)*exp(-lambda2*gsl_vector_get(covar_t,i)*exp(beta2*gsl_vector_get(covar_x,i)));

	logiteml=gsl_vector_get(covar_c,i)*log(x1*lambda1*exp(beta1*gsl_vector_get(covar_x,i))
+x2*lambda2*exp(beta2*gsl_vector_get(covar_x,i)))+(1-gsl_vector_get(covar_c,i))*log(x1+x2) ;
		sumlogl = sumlogl+logiteml;
	}

	return -sumlogl;
}

	//random number generator
	const gsl_rng_type * T;
	gsl_rng * r;
	gsl_rng_env_setup();
	T = gsl_rng_default;
	r = gsl_rng_alloc (T);
	gsl_rng_set (r, (unsigned)time(0));

	//data generation
	//null data generation
	for (int row = 0; row < num_rep; row++)
	{
		for (int col = 0; col < num_sub; col++)
		{
			gsl_matrix_set(covar_x,row,col,gsl_ran_flat (r, uniform_min, uniform_max));
			gsl_matrix_set(failure_dist,row,col, gsl_ran_exponential(r,
```

```
1/(lambda_h0*exp(gsl_matrix_get(covar_x,row,col)*beta_h0))));
            }
        }

        for (int row = 0; row < num_rep; row++)
        {
            for (int col = 0; col < num_sub; col++)
            {
            gsl_matrix_set (censor_dist,row,col, gsl_ran_exponential(r, mean_censor_dist_h0));
            gsl_matrix_set(covar_t,row,col, min( gsl_matrix_get(failure_dist, row, col),
gsl_matrix_get(censor_dist, row, col)));

                if (gsl_matrix_get(failure_dist,row,col) <
gsl_matrix_get(censor_dist,row,col))
                {
                    gsl_matrix_set(covar_c,row,col,1);
                }
                else
                {
                    gsl_matrix_set(covar_c,row,col,0);
                }
            }
        }

/*
        //alternative data generation
        gsl_matrix* select = gsl_matrix_calloc(num_rep,num_sub);

        for (int row = 0; row < num_rep; row++)
        {
            for (int col = 0; col < num_sub; col++)
            {
                gsl_matrix_set(select,row,col,gsl_ran_flat (r, 0, 1));
                gsl_matrix_set(covar_x,row,col,gsl_ran_flat (r, uniform_min, uniform_max));
                gsl_matrix_set(first,row,col,gsl_ran_exponential(r,
1/(lambda1_h1*exp(gsl_matrix_get(covar_x,row,col)*beta1_h1))));
                gsl_matrix_set(second,row,col,gsl_ran_exponential(r,
1/(lambda2_h1*exp(gsl_matrix_get(covar_x,row,col)*beta2_h1))));
```

```
			if (gsl_matrix_get(select,row,col) < mix_h1)
				{
				gsl_matrix_set(failure_dist,row,col, gsl_matrix_get(first, row, col));
				}
			else
				{
				gsl_matrix_set(failure_dist,row,col, gsl_matrix_get(second, row, col));
				}
		}
	}

	for (int row = 0; row < num_rep; row++)
	{
		for (int col = 0; col < num_sub; col++)
		{
		gsl_matrix_set (censor_dist,row,col, gsl_ran_exponential(r, mean_censor_dist_h1));

		gsl_matrix_set(covar_t,row,col, min( gsl_matrix_get(failure_dist, row, col),
gsl_matrix_get(censor_dist, row, col)));

			if (gsl_matrix_get(failure_dist,row,col) <
gsl_matrix_get(censor_dist,row,col))
			{
				gsl_matrix_set(covar_c,row,col,1);
			}
			else
			{
				gsl_matrix_set(covar_c,row,col,0);
			}
		}
	}
*/
	// RSPs initial values ~Uni(0,1)
		for (int p = 0; p < num_of_init_mu; p++)
		{
			gsl_vector_set(mus,p,gsl_ran_flat (r, 0, 1));
		}
```

```
            for (int p = 0; p < num_of_init_be; p++)
            {
                 gsl_vector_set(betas,p,gsl_ran_flat (r, 0, 1));
            }

            for (int p = 0; p < num_of_init_mu1; p++)
            {
                 gsl_vector_set(mu1s,p,gsl_ran_flat (r, 0, 1));
            }

            for (int p = 0; p < num_of_init_mu2; p++)
            {
                 gsl_vector_set(mu2s,p,gsl_ran_flat (r, 0, 1));
            }

            for (int p = 0; p < num_of_init_be1; p++)
            {
                 gsl_vector_set(beta1s,p,gsl_ran_flat (r, 0, 1));
            }

            for (int p = 0; p < num_of_init_be2; p++)
            {
                 gsl_vector_set(beta2s,p,gsl_ran_flat (r, 0, 1));
            }

            for (int p = 0; p < num_of_init_alp; p++)
            {
                 gsl_vector_set(alphas,p,gsl_ran_flat (r, 0, 1));
            }

// simulations Nelder-Mead algorithm

       for(int w=0 ; w < num_rep; w++)
       {
            power_data my_pwr_data;

            gsl_matrix_get_row(covar_t_st,covar_t,w);
```

```
gsl_matrix_get_row(covar_c_st,covar_c,w);
gsl_matrix_get_row(covar_x_st,covar_x,w);

my_pwr_data.covar_t_st = covar_t_st;
my_pwr_data.covar_c_st = covar_c_st;
my_pwr_data.covar_x_st = covar_x_st;
my_pwr_data.size = num_sub;

for (int ll = 0; ll < num_of_init_mu; ll++)
{
     for (int mm = 0; mm < num_of_init_be; mm++)
     {
                    gsl_vector_set(xx, 0, gsl_vector_get(mus,ll));
                    gsl_vector_set(xx, 1, gsl_vector_get(betas,mm));

                    minex_func.f = &my_f0;
                    minex_func.n=np;
                    minex_func.params = (void *)&my_pwr_data;
                    gsl_multimin_fminimizer_set(s, &minex_func, xx, ss);

                    iter = 0;
                    int status;
                    double size;

                    do
                    {
                         iter++;
                         status = gsl_multimin_fminimizer_iterate(s);

                         if(status)
                         break;

                         size = gsl_multimin_fminimizer_size (s);
                         status = gsl_multimin_test_size (size, 1e-5);

                    }
                    while (status == GSL_CONTINUE && iter < 1000);
```

```
                                gsl_vector_set(L0_result,w,-s->fval);

gsl_vector_set(lambda_result,w,exp(gsl_vector_get(s->x,0)));
                                gsl_vector_set(beta_result,w,gsl_vector_get(s->x,1));

                            }
                        }

    for (int ll = 0; ll < num_of_init_mu1; ll++)
    {
        for (int mm = 0; mm < num_of_init_mu2; mm++)
        {
            for (int nn = 0; nn < num_of_init_be1; nn++)
            {
                for (int pp = 0; pp < num_of_init_be2; pp++)
                {
                    for (int qq = 0; qq < num_of_init_alp; qq++)
                    {

                        gsl_vector_set(xx2, 0, gsl_vector_get(mu1s,ll));
                        gsl_vector_set(xx2, 1, gsl_vector_get(mu2s,mm));
                        gsl_vector_set(xx2, 2, gsl_vector_get(beta1s,nn));
                        gsl_vector_set(xx2, 3, gsl_vector_get(beta2s,pp));
                        gsl_vector_set(xx2, 4, gsl_vector_get(alphas,qq));

                        minex_func.f = &my_f1;
                        minex_func.n=np2;
                        minex_func.params = (void *)&my_pwr_data;
                        gsl_multimin_fminimizer_set(s2, &minex_func, xx2, ss2);

                        iter2 = 0;
                        int status;
                        double size2;

                        do
                        {
                            iter2++;
                            status = gsl_multimin_fminimizer_iterate(s2);
```

```
                        if(status)
                        break;

                        size2 = gsl_multimin_fminimizer_size (s2);
                        status = gsl_multimin_test_size (size2, 1e-5);

                    }
                    while (status == GSL_CONTINUE && iter2 < 1000);

                    gsl_vector_set(L1_result,w,-s2->fval);
            gsl_vector_set(lambda1_result,w,exp(gsl_vector_get(s2->x,0)));
            gsl_vector_set(lambda2_result,w,exp(gsl_vector_get(s2->x,1)));
            gsl_vector_set(beta1_result,w,gsl_vector_get(s2->x,2));
            gsl_vector_set(beta2_result,w,gsl_vector_get(s2->x,3));
gsl_vector_set(mix_result,w,exp(gsl_vector_get(s2->x,4))/(1+exp(gsl_vector_get(s2->x,4))));

                }
            }
        }
    }
}

for(int i=0;i<num_rep;i++)
{
    gsl_vector_set(maxf0,i,gsl_vector_get(L0_result,i));
    gsl_vector_set(max1,i,gsl_vector_get(lambda_result,i));
    gsl_vector_set(max2,i,gsl_vector_get(beta_result,i));

    gsl_vector_set(maxf1,i,gsl_vector_get(L1_result,i));
    gsl_vector_set(max3,i,gsl_vector_get(lambda1_result,i));
    gsl_vector_set(max4,i,gsl_vector_get(lambda2_result,i));
    gsl_vector_set(max5,i,gsl_vector_get(beta1_result,i));
    gsl_vector_set(max6,i,gsl_vector_get(beta2_result,i));
    gsl_vector_set(max7,i,gsl_vector_get(mix_result,i));

/* output format
if (gsl_vector_get(max7,i) < 0.5)
```

```
                {

                        gsl_vector_set(real_max3,i, gsl_vector_get(max3, i));
                        gsl_vector_set(real_max4,i, gsl_vector_get(max4, i));
                        gsl_vector_set(real_max5,i, gsl_vector_get(max5, i));
                        gsl_vector_set(real_max6,i, gsl_vector_get(max6, i));
                        gsl_vector_set(real_max7,i, gsl_vector_get(max7, i));

                }
                else
                {
                        gsl_vector_set(real_max3,i, gsl_vector_get(max4, i));
                        gsl_vector_set(real_max4,i, gsl_vector_get(max3, i));
                        gsl_vector_set(real_max5,i, gsl_vector_get(max6, i));
                        gsl_vector_set(real_max6,i, gsl_vector_get(max5, i));
                        gsl_vector_set(real_max7,i, 1-gsl_vector_get(max7, i));
                }
*/
}

double covar_t_tmp[num_sub];
double covar_c_tmp[num_sub];
double covar_x_tmp[num_sub];

for(int i=0;i<num_sub ; i++)
{
        covar_t_tmp[i]=gsl_matrix_get(covar_t,w,i);
        covar_c_tmp[i]=gsl_matrix_get(covar_c,w,i);
        covar_x_tmp[i]=gsl_matrix_get(covar_x,w,i);
}

gsl_vector_set(mean_covar_t,w, gsl_stats_mean(covar_t_tmp,1,num_sub));
gsl_vector_set(mean_covar_x,w, gsl_stats_mean(covar_x_tmp,1,num_sub));
gsl_vector_set(mean_covar_c,w, gsl_stats_mean(covar_c_tmp,1,num_sub));
gsl_vector_set(sd_covar_t,w, gsl_stats_sd(covar_t_tmp,1,num_sub));


}//w
```

```
FILE *ofp;

ofp=fopen("result", "w");

fprintf(ofp, " mean_x\t mean_t\t sd_t\t observed_censoring_farction\t sumlog0\t lambdahat\t
betahat\t  sumlog1\t lambda1hat\t lambda2hat\t beta1hat\t beta2hat\t mixhat\t LRT\n");

for (int f = 0; f < num_rep; f++)
{
        //fprintf(ofp, "%f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\n",
gsl_vector_get(mean_covar_x,f), gsl_vector_get(mean_covar_t,f), gsl_vector_get(sd_covar_t,f),
1-gsl_vector_get(mean_covar_c,f), gsl_vector_get(maxf0,f), gsl_vector_get(max1,f),
gsl_vector_get(max2,f), gsl_vector_get(maxf1,f), gsl_vector_get(real_max3,f),
gsl_vector_get(real_max4,f), gsl_vector_get(real_max5,f), gsl_vector_get(real_max6,f),
gsl_vector_get(real_max7,f), ((-2) * gsl_vector_get(maxf0,f) - (-2) *
gsl_vector_get(maxf1,f)));

        fprintf(ofp, "%f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\t %f\n",
gsl_vector_get(mean_covar_x,f), gsl_vector_get(mean_covar_t,f), gsl_vector_get(sd_covar_t,f),
1-gsl_vector_get(mean_covar_c,f), gsl_vector_get(maxf0,f), gsl_vector_get(max1,f),
gsl_vector_get(max2,f), gsl_vector_get(maxf1,f), gsl_vector_get(max3,f),
gsl_vector_get(max4,f), gsl_vector_get(max5,f), gsl_vector_get(max6,f),
gsl_vector_get(max7,f), ((-2) * gsl_vector_get(maxf0,f) - (-2) * gsl_vector_get(maxf1,f)));

    }

    fclose(ofp);
        return 1;
}
```