LIKELIHOOD RATIO TEST FOR THE PRESENCE OF CURED INDIVIDUALS: A SIMULATION STUDY

YI LIANG

Your file Votre référence
ISBN: 0-612-93039-4
Our file Notre référence
ISBN: 0-612-93039-4

LIKELIHOOD RATIO TEST FOR THE PRESENCE OF CURED INDIVIDUALS: A SIMULATION STUDY

by

Yi Liang

A practicum report submitted to the School of Graduate Studies in partial fulfillment of the requirement for the Degree of Master of Applied Statistics

Department of Mathematics and Statistics
Memorial University of Newfoundland

May, 2002

St. John's, Newfoundland

Abstract

Likelihood ratio test plays an important role in testing for the presence of cured individuals in clinical studies. The asymptotic null distribution of the test in three commonly used mixture models is a 50-50 mixture distribution of a chi-squared distribution and the probability mass at zero under some mild conditions. In this practicum, we first study the power of the likelihood ratio test under those models via a simulation study. We find that the test is powerful for moderate large sample size. For small sample size, the result varies case by case, and the censoring distribution affects the power substantially. The Extended Generalized Gamma (EGG) mixture model is also considered, for it is more flexible than the three models mentioned above. It may be used when the data in study cannot be fit by those models. The bootstrap approach is then introduced to investigate the null distribution of the likelihood ratio test. Finally, we employ the bootstrap approach to determine the presence of cured individuals in two sets of real-life data.

Acknowledgments

I wish to thank my supervisor, Dr. Yingwei Peng, for leading me to the field of the practicum, and for his helpful comments, discussions and suggestions throughout the preparation of this practicum.

I also wish to thank my co-supervisor, Dr. Chu-In Charles Lee, for his concern, encouragement and support. He read the practicum carefully and gave many helpful comments.

I acknowledge the School of Graduate Studies and the Department of Mathematics and Statistics for the financial support in the form of Graduate Fellowship and Teaching Assistantship.

Finally, I wish to thank my wife, Chen, and my parents for their encouragement, support and love.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Cure rate estimation

In some clinical studies, the response variable of interest is the occurrence of an event, such as a disease, recurrence of a disease, or death. If we examine the response observations, usually there are a proportion of patients who do not relapse, or die, due to termination of study, or loss of follow-up. Those incomplete observations are recorded as censored ones. Among them, there might exist patients who respond favorably to the treatment in use. They are eventually free of any signs or symptoms of disease and may be considered cured. The possible observations of the patients are then time to relapse or death and censoring time.

Kersey *et al.* (1987) compared two types of bone marrow transplants, autologous and allogeneic, for a treatment of high-risk refractory acute lymphoblastic leukemia. An autologous bone marrow transplant is a transplant that uses the patient's own bone marrow, while an allogeneic bone marrow transplant uses bone marrow from someone else. In their study, the obser-

vations are relapse times of leukemia patients following transplant. The allogeneic group consists of observations on 46 patients, of whom 13 were censored, while autologous group includes 45 patients, of whom 9 were censored. The data are given in Appendix A. Since there are censored observations, we applied the Kaplan-Meier (K-M) method to estimate the survivor functions of the two groups. The Kaplan-Meier survival curves of two types of transplants are showed in Figure 1.1.

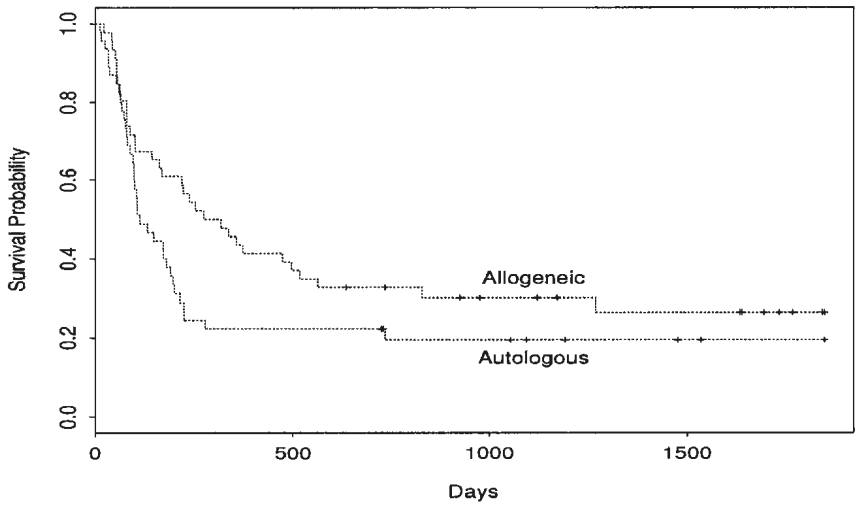

Figure 1.1: Kaplan-Meier survival curves for leukemia data

Both curves level off at a value greater than zero. This is caused by the long-term censored survival times of potentially cured patients. In evaluating the effect of the bone marrow transplant treatment, it is important to determine whether or not the cured patients exist, or the cured proportion is greater than zero. If the proportion of cured patients is substantially greater than zero, mixture models can be applied to estimate the cure rate of treatment.

1.2 Mixture models for cure rate estimation

If the cure rate is not zero, a certain fraction of the population, from which the data were collected, may never experience the particular type of failure(relapse , or death). They are cured patients, and can be regarded as from a population with degenerate distribution at infinite time. The remaining patients , who relapsed or died of the disease, come from a failure population from the cause of interest. The analysis of the data is to fit a model of the mixtures of the two distributions. Henceforth, mixture methods have been proposed for cure rate estimation. Among them, parametric mixture models have been studied and applied extensively.

A mixture model for cure rate estimation is also called a cure model. It is a model that involves the mixture of two populations, one is the uncured population G_1 and the other is the cured population G_2. For the cured population, patients never die or relapse. We can assume that the failure time for cured patients is infinity. Then G_2 follows a distribution degenerates at infinity. For parametric mixture models, the distribution of the uncured population G_1 has a known form with a few unknown parameters. The proportion of each population is unknown as well. A parametric cure model can be formulated as follows. Let t be an observation of the time of the occurrence of an event. Under the cure model, it can be viewed as arising from a superpopulation G which is a mixture of two populations G_1 and G_2 with proportions ρ_1 and ρ_2, respectively, where

$$\rho_1 + \rho_2 = 1, \qquad \rho_i \geq 0 \qquad (i = 1, 2)$$

If we denote the proportion of uncured patients by $\rho = \rho_1$, $1 - \rho = \rho_2$ is the proportion of cured patients. For all finite values of t, the density and survival functions of cured patients are, respectively, zero and one. The probability density function(p.d.f) of G can therefore be expressed in terms of the mixture form,

$$f(t;\theta,\rho) = \rho f_u(t;\theta) \tag{1.1}$$

where $f_u(t;\theta)$ is the p.d.f. of the uncured population G_1, and θ denotes the vector of all unknown parameters associated with the parametric form for the density function. The subscript u stands for the uncured population. The maximum likelihood method is commonly used to estimate the parameters ρ and θ.

Boag (1949) first employed the mixture models to estimate the cured rates of treatments for several cancers. He assumed the uncured population G_1 had a log-normal distribution, and applied the maximum likelihood method to estimate cured rates. Farewell(1982) analyzed the data from a toxicant animal experiment using a mixture model with uncured population G_1 following a Weibull distribution. Goldman (1984) conducted a survivorship analysis of clinical trials under the assumption that an unknown proportion of all the patients have constant high risk and the remaining proportion have essentially no risk. She fit (1.1) to the data with uncured population G_1 having an exponential distribution. Yamaguchi (1992), Maller and Zhou (1994), Peng *et al.* (1998) provided more examples of how the mixture models are applied in clinical trials and other fields.

1.3 Test of cured patients

Most of the published works assume that there exist cured patients in the population of interest before the mixture models were applied. If there are cured patients, the observations should come from a mixture population of uncured and cured patients. Otherwise, there are no cured patients and all observations should have been collected from a homogeneous population. Then a mixture model is not needed for modeling the observations. To investigate the presence of cured patients, the likelihood ratio test(LRT) is commonly used in the literature. The null hypothesis for testing the presence of cured patients is $H_0 : \rho = 1$, and the parameter space of cure rate ρ is $[0, 1]$. Under H_0 the cure rate is equal to 0.

Goldman (1984) first applied the likelihood ratio test for testing the presence of cured patients. She did not notice that the null hypothesis is on the boundary of the parameter space, and conducted the test under the assumption that the asymptotic null distribution of the likelihood ratio test is a χ^2 distribution with degrees of freedom equal to one.

Zhou and Maller (1992) investigated the likelihood ratio test of the presence of cured patients under the same model that Goldman considered. Their results showed that the asymptotic null distribution for this specific test is not asymptotically a χ^2 distribution. Under mild conditions, it follows a 50-50 mixture distribution of a χ^2 distribution with degrees of freedom equal to one and a degenerated distribution mass at zero.

Peng *et al.* (2001) extended the use of the likelihood ratio test to several

different types of distributions for the uncured population G_1. They examined the asymptotic null distribution of likelihood ratio test and pointed out that the distribution derived by Zhou and Maller (1992) can be used to approximate the asymptotic null distribution when G_1 has a Weibull or log-normal distribution. However, the approximation will be poor when the censoring is too light and the hazard rate is large at the upper part of the failure time distribution of G_1.

All the above research was carried out by using a specific distribution of G_1. In practice, we usually do not know the parametric form of the distribution of uncured population G_1. When we use the likelihood ratio test for actual data, we may meet the problem that the null distribution for the likelihood ratio test is unknown. Thereafter, we cannot test the presence of cured patients for the data. If we know the null distribution, such as the one proposed by Zhou and Maller(1992), another problem is the power assessment of the likelihood ratio test.

1.4 Power study of the likelihood ratio test

Power is the probability of rejecting the null hypothesis when the alternative hypothesis is true. In other words, power is the probability of correctly rejecting the null hypothesis. Given a significance level α, it is usually desirable to use a statistical test with power to be as large as possible if the sample size is controlled.

Concerning the likelihood ratio test for the presence of cured patients, there are few studies in the literature on the power of the likelihood ratio test for

the presence of cured patients. Goldman (1984) studied the power of the likelihood ratio test that she used. Because of her erroneous assumption about the asymptotic null distribution of the test statistic, her results of power are questionable. Since power is an important characteristic for evaluating the efficiency of a test, it is of statistical and practical importance to conduct a power study on the likelihood ratio test.

1.5 Organization of the practicum

This practicum is organized as follows. We first review the models and tests for testing the presence of cured patients in chapter 2. The power of a likelihood ratio test for the presence of cure patients is investigated via a simulation study in chapter 3. In chapter 4, we employ the bootstrap approach to investigate the null distribution of the LRT. Finally the results are applied to two sets of real-life data.

Chapter 2

Parametric Mixture Model and Likelihood Ratio Test

In this chapter, we introduce the commonly used parametric mixture models and discuss in details the likelihood ratio test for presence of cured individuals based on those mixture models.

We adopt the following notations throughout the whole practicum. Let T^0 be a non-negative random variable denoting the failure time of interest. The distribution function of T^0 is F^0. Let C be the censoring random variable with distribution function J. Random censorship is assumed, under which C is independent of T^0. The value observed is $T = \min(T^0, C)$. The pairs (T, D) represent the values observed, where D is a censoring indicator which equals to one if T^0 is observed and equals to zero if C is observed. Let t represent an observed value of T.

2.1 Mixture Models

The mixture model that is employed in cure rate estimation can be described as follows. Let $f(t;\theta,\rho)$ and $S(t;\theta,\rho)$ be the density and survival functions of the distribution of T. Let ρ denote the proportion of uncured patients, and $1-\rho$ be the proportion of cured patients. Then the distribution of T has a finite mixture form:

$$\begin{aligned} S(t;\theta,\rho) &= \rho S_u(t;\theta) + 1 - \rho \\ f(t;\theta,\rho) &= \rho f_u(t;\theta) \end{aligned}$$

where $f_u(t;\theta)$ and $S_u(t;\theta)$ are the density and survival function of the failure time distribution of uncured patients, θ is a collection of the unknown parameters in this distribution and the subscript u is used to represent the uncured population. The hazard function of the uncured patients is defined as

$$h_u(t) = \frac{f_u(t)}{S_u(t)}$$

The hazard function is useful, since it describes the way in which the instantaneous probability of death for a patient changes with time.

2.2 Some Parametric Mixture Models

Several distributions have been considered in the mixture models for cure rate estimation. They include the exponential, Weibull, log-normal, gamma, or extended generalized gamma (EGG) distributions.

Exponential Mixture Model

In this model, the uncured population G_1 has an exponential distribution. The exponential distribution has a constant hazard function

$$h_u(t) = \lambda, \lambda > 0$$

The survivor and density functions for the exponential mixture model are, respectively,

$$\begin{aligned} S(t) &= \rho e^{-\lambda t} + 1 - \rho \\ f(t) &= \rho \lambda e^{-\lambda t} \end{aligned}$$

Weibull Mixture Model

In this model, the uncured population G_1 has a Weibull distribution. The Weibull distribution, an important generalization of the exponential distribution, has the hazard function

$$h_u(t) = \lambda p (\lambda t)^{p-1}$$

with $\lambda, p > 0$. This hazard is monotone decreasing for $p < 1$, increasing for $p > 1$, and reduces to the exponential hazard if $p = 1$. The survivor and density functions for the Weibull mixture model are, respectively,

$$S(t) = \rho \exp[-(\lambda t)^p] + 1 - \rho$$

$$f(t) = \rho \lambda p (\lambda t)^{p-1} \exp[-(\lambda t)^p]$$

Log-normal Mixture Model

In this model, the uncured population G_1 has a log-normal distribution. The hazard function of log-normal $h_u(t)$, which has the value 0 at $t = 0$, increases

to a maximum and then decreases, approaching zero as t becomes large. The survivor and density functions for log-normal mixture model are, respectively,

$$S(t) = \rho(1 - \Phi(p \log \lambda t)) + 1 - \rho$$

$$f(t) = \rho(2\pi)^{-1/2} p t^{-1} \exp\left[\frac{-(p \log \lambda t)^2}{2}\right]$$

where Φ is the incomplete normal integral

$$\Phi(w) = \int_{-\infty}^{w} \phi(u)\, du$$

Gamma Mixture Model

In this model, the uncured population G_1 has a gamma distribution. As the other generalization of the exponential distribution, the gamma distribution has the density function

$$f_u(t) = \frac{\lambda(\lambda t)^{q-1} \exp(-\lambda t)}{\Gamma(q)}$$

where $q, \lambda > 0$. The survivor and hazard functions of the gamma distribution involve the incomplete gamma integral

$$I_q(s) = \frac{\int_0^s x^{q-1} e^{-x} dx}{\Gamma(q)}$$

and are, respectively,

$$S_u(t) = 1 - I_q(\lambda t)$$

$$h_u(t) = \frac{\lambda(\lambda t)^{q-1} \exp(-\lambda t)\Gamma(q)^{-1}}{1 - I_q(\lambda t)}$$

The hazard function is monotone increasing from 0 when $q > 1$, monotone decreasing from ∞ when $q < 1$, and either case approaches λ as t becomes

large. When $q = 1$, it is exponential distribution. The survivor and density functions for the gamma mixture model are, respectively,

$$S(t) = \rho(1 - I_q(\lambda t)) + 1 - \rho$$

$$f(t) = \frac{\rho\lambda(\lambda t)^{q-1}\exp(-\lambda t)}{\Gamma(q)}$$

Extended Generalized Gamma(EGG) Mixture Model

In this model, the uncured population G_1 follows an Extended Generalized Gamma(EGG) distribution. The survivor and density functions of the EGG distribution are

$$S_u(t;q,\lambda,p) = \begin{cases} \frac{|q|}{\Gamma((q^{-2})}(q^{-2})^{q^{-2}}\int_0^{\lambda t} px^{p/q-1}\exp(-q^{-2}x^{pq})\,dx & \text{when } q \neq 0 \\ 1 - \int_0^{p\log\lambda t}\frac{1}{\sqrt{2\pi}}\exp(\frac{-x^2}{2})\,dx & \text{when } q = 0 \end{cases}$$

$$f_u(t;q,\lambda,p) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})}(q^{-2})^{q^{-2}}pt^{-1}\exp[q^{-2}(qp\log\lambda t - e^{qp\log\lambda t})] & \text{when } q \neq 0 \\ (2\pi)^{-1/2}pt^{-1}\exp(-(p\log\lambda t)^2/2) & \text{when } q = 0 \end{cases}$$

The survivor and density functions for the EGG mixture model can be constructed accordingly as follows.

$$\begin{aligned} S(t) &= \rho S_u(t;q,\lambda,p) + 1 - \rho \\ f(t) &= \rho f_u(t;q,\lambda,p) \end{aligned}$$

The EGG mixture model is flexible and includes as special cases all of the mixture models mentioned in the preceding sections. The exponential ($p = q = 1$), Weibull ($q = 1$) and log-normal ($q = 0$) mixture models are all evident. In addition, the gamma mixture model appears as $p = 1$. This can be obtained by employing the transformation: $\log\lambda X = q(\log\lambda T + \log q^{-2})$ and replacing the parameter q^{-2} by $k = q^{-2}$. The relationship among those mixture models is shown in Figure 2.1.

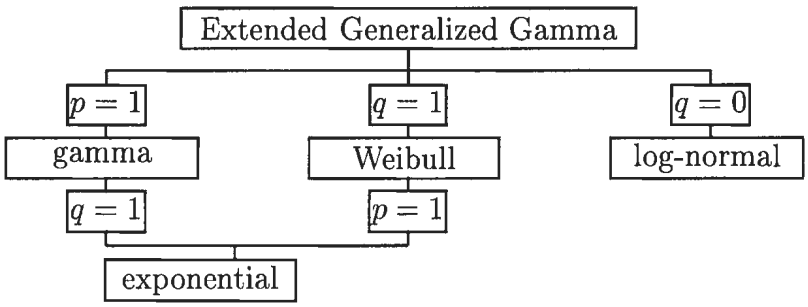

Figure 2.1: Relationship among the mixture models

Actual survival data can come from any positive value distributions, and the failure time distribution of uncured patients may not be one of the four distributions: exponential, Weibull, log-normal and gamma. Peng *et al.* (1998) suggested that a generalized F distribution can be applied to the mixture model for cure rate estimation. The generalized F distribution is flexible and contains almost all the known and commonly used positive distributions as special cases. But there are computational difficulties in carrying out the generalized F mixture model. We consider the EGG instead of the generalized F, for it involves fewer parameters and is computationally easier.

2.3 Likelihood Ratio Test and Its Null Distribution

When we fit the data with parametric mixture models discussed in the previous sections, we need to know whether there are cured individuals or not. Therefore, we need to test the presence of cured individuals. The hypothesis

of no cured individuals is $H_0 : \rho = 1$ in the mixture models, and the parameter space of cure rate ρ is $[0, 1]$.

The likelihood ratio test (LRT) can be used to test this hypothesis. The likelihood function $L(t; \theta, \rho)$ is as follows.

$$L(t; \theta, \rho) = \prod_{i=1}^{n} [f(t_i; \theta, \rho)]^{d_i} [S(t_i; \theta, \rho)]^{1-d_i}$$

where $t = (t_1, \cdots, t_n)$. Here t_is are the values of the observations, and n is the number of the observations. d_i is the censoring indicator. It is 0 if t_i is censored. Otherwise it is 1.

The likelihood ratio test statistic is given by

$$d_n = -2[l_0(\theta) - l_1(\theta, \rho)] \tag{2.1}$$

where $l_1(\theta, \rho)$ is the natural logarithm of the likelihood function maximized under the hypothesis of a mixture model and $l_0(\theta)$ is the natural logarithm of the likelihood function maximized under the hypothesis of no cured individuals.

The forms of the likelihood ratio test statistic for those mixture models, which are discussed above, can be presented as follows.

For exponential mixture model,

$$d_n = 2 \max_{\lambda, \rho} (\sum_{i=1}^{n} (d_i(\log \rho + \log \lambda - \lambda t_i) + (1 - d_i) \log(\rho e^{-\lambda t_i} + 1 - \rho)))$$

$$-2(\sum_{i=1}^{n} d_i)(\log(\sum_{i=1}^{n} d_i) - \log(\sum_{i=1}^{n} t_i) - 1)$$

For Weibull mixture model,

$$d_n = 2 \max_{\lambda, p, \rho} (\sum_{i=1}^{n} (d_i(\log \rho + \log \lambda + \log p + (p - 1) \log \lambda t_i - (\lambda t_i)^p)$$

$$+(1-d_i)\log(\rho\exp(-(\lambda t_i)^p)+1-\rho)))$$

$$-2\max_{\lambda,p}(\sum_{i=1}^{n}(d_i(\log\lambda+\log p+(p-1)\log\lambda t_i)-(\lambda t_i)^p))$$

For log-normal mixture model,

$$d_n=2\max_{\lambda,p,\rho}(\sum_{i=1}^{n}(d_i(\log\rho-\frac{1}{2}\log 2\pi+\log p-\log t_i-\frac{(p\log\lambda t_i)^2}{2})$$

$$+(1-d_i)\log(\rho(1-\Phi(p\log\lambda t_i))+1-\rho)))$$

$$-2\max_{\lambda,p}(\sum_{i=1}^{n}(d_i(-\frac{1}{2}\log 2\pi+\log p-\log t_i-\frac{(p\log\lambda t_i)^2}{2})+(1-d_i)\log(1-\Phi(p\log\lambda t_i))))$$

For gamma mixture model,

$$d_n=2\max_{\lambda,q,\rho}(\sum_{i=1}^{n}(d_i(\log\rho+\log\lambda+(q-1)\log\lambda t_i-\lambda t_i-\log(\Gamma(q)))$$

$$+(1-d_i)\log(\rho(1-I_q(\lambda t_i))+1-\rho)))$$

$$-2\max_{\lambda,q}(\sum_{i=1}^{n}(d_i(\log\lambda+(q-1)\log\lambda t_i-\lambda t_i-\log(\Gamma(q)))+(1-d_i)\log(1-I_q(\lambda t_i))))$$

For extended generalized gamma(EGG) mixture model, if $q=0$, d_n is the same as the one in the log-normal mixture model; if $q\neq 0$, then

$$d_n=\max_{\lambda,p,q,\rho}(\sum_{i=1}^{n}(d_i(\log\rho+\log(|q|)-\log(\Gamma(q^{-2}))+q^{-2}\log(q^{-2})$$

$$+\log p-\log t_i+q^{-2}(qp\log\lambda t_i-e^{qp\log\lambda t_i}))$$

$$+(1-d_i)\log(\rho(\frac{|q|}{\Gamma((q^{-2})}(q^{-2})^{q^{-2}}\int_0^{\lambda t_i}px^{p/q-1}\exp(-q^{-2}x^{pq})\,dx)+1-\rho)))$$

$$-2\max_{\lambda,p,q}(\sum_{i=1}^{n}(d_i(\log(|q|)-\log(\Gamma(q^{-2}))+q^{-2}\log(q^{-2})+\log p-\log t_i+q^{-2}(qp\log\lambda t_i-e^{qp\log\lambda t_i}))$$

$$+(1-d_i)(\log(|q|)-\log(\Gamma(q^{-2}))+q^{-2}\log(q^{-2})+\log(\int_0^{\lambda t_i}px^{p/q-1}\exp(-q^{-2}x^{pq})\,dx))))$$

Here d_n does not follow a χ^2 distribution under H_0. This is due to the non-standard situation that the value of the parameter ρ in H_0 is on a boundary of

$[0, 1]$, the parameter space of ρ. Consequently, the asymptotic null distribution of d_n needs to be worked out.

Goldman (1984) first used the likelihood ratio test to determine the presence of cured patients using an exponential mixture model. She assumed that, under H_0, d_n has an asymptotic χ^2 distribution. Maller and Zhou (1995) proved that, under mild conditions, the asymptotic null distribution of d_n in the exponential mixture model follows a mixture distribution. Their results showed that if the moment generating function of C, given by $E\{e^{(\theta+\xi)C}\}$, is finite for some $\xi > 0$, the asymptotic null distribution of the LRT statistic d_n will follow a 50-50 mixture of a χ^2 distribution and probability mass at zero

$$P(d_n < t) \to \frac{1}{2} + \frac{1}{2}P(\chi_1^2 < t) \tag{2.2}$$

Vu *et al.* (1996) pointed out that this holds for the gamma mixture model under mild conditions. Peng *et al.* (2001) studied the properties of the distribution of d_n with the failure time distribution of uncured patients as the log-normal or the Weibull distribution via a simulation study. Their results showed that under some certain conditions the empirical null distribution of the likelihood ratio test for the Weibull or log-normal mixture model agrees with that for the gamma mixture model. But caution is needed in this case to determine the presence of cured individuals. There is no reported study on the null distribution of LRT under the EGG mixture model. The power assessment of the likelihood ratio test for each case is also unknown.

Chapter 3

The Power of Likelihood Ratio Test

In this chapter, we will conduct a simulation study to investigate the power of the likelihood ratio test. Simulated data are generated from a known population that includes cured individuals. We will apply the likelihood ratio test for testing the presence of cured individuals. The distribution in (2.2) is used as the asymptotic null distribution to specify the rejection region, and henceforth the power can be computed.

3.1 Simulation Studies

In our study we do not restrict the distribution of failure times of uncured individuals to an exponential or gamma distribution, since the log-normal and Weibull distributions are common and useful in survival analysis. We consider three distributions as the failure time distribution of uncured individuals in the mixture models: (1) gamma distribution with the density function $f_u(t;\theta) = t^{\theta-1}e^{-t}/\Gamma(\theta)$; (2) Weibull distribution with the density function

$f_u(t;\theta) = \theta t^{\theta-1}\exp(-t^\theta)$; (3) log-normal distribution with the density function $f_u(t;\theta) = \theta(\sqrt{2\pi}t)^{-1}\exp(-0.5\theta^2\log^2 t)$. The value of θ in the gamma distribution is fixed at 0.5, 1.0, 1.5, 2.5, in which $\theta = 1$ corresponds to the exponential distribution. For Weibull and log-normal distributions, θ is fixed at 0.5, 1.5, 2.5.

We restrict the censoring distributions to uniform and exponential distributions because these two distributions are usually considered in the literature. Here we assume that we know the distribution of the failure time variable T^1 of the uncured individuals and the type of the distribution of censoring variable C^1 of the uncured individuals. By definition, the censoring rate of the uncured individuals is $P(T^1 > C^1)$. The values of parameters in the censoring distributions can be determined so that the resulting censoring rate for uncured individuals equals to 30 per cent. For example, if T^1 follows a gamma distribution with parameter $\theta = 1.5$, and C^1 follows an exponential distribution with parameter λ, to achieve the 30 per cent censoring for uncured individuals, $P(T^1 > C^1)$ will equal to 0.30. That is,

$$P(T^1 > C^1) = \int\int_{x_1>x_2} \frac{x_1^{1.5-1}}{\Gamma(1.5)} e^{-x_1} dx_1 \lambda e^{-\lambda x_2} dx_2 = 0.30$$

or

$$1 - (\lambda + 1)^{1.5} = 0.30$$

From the equation, $\lambda \approx 0.2684$. This means, we need to choose λ, the parameter of the censoring distribution, to be 0.2684 if we want to have 30 per cent censoring in the uncured individuals.

Three sample sizes are considered: 50, 100, and 200. The power of testing

the presence of cured individuals is usually important when the actual cure rate ranges between 5% and 10%. The larger the actual cure rate, the easier it will be to detect the presence of cured individuals. Henceforth, we fix the cure rate at 3%, 5%, 7%, 10%, 15%. Given one of the combinations of three failure time distributions of uncured individuals, five cure rates and two censoring distributions, a sample can be generated with a given sample size. Different sample sizes are used so as to find the effect of sample size on the power, and different cure rates are set to examine the effect of cure rate on the likelihood ratio test.

For the sample from one of the three failure time distributions of uncured individuals, we use the corresponding mixture model and the EGG mixture model to fit. That is, if a sample is generated from the gamma distribution, we will fit the sample with the gamma mixture model or the EGG mixture model. The reason why the EGG mixture model is used to fit the data is that we want to compare the effect of power under different models. The result may be useful for model selection. For each model, the test statistic d_n can be computed by using (2.1)

Let the significant level be equal to 0.05. From the 50-50 mixture distribution in (2.2), we can obtain a critical value under this level. The power of the likelihood ratio test is estimated as a proportion of d_n greater than the critical value based on 500 simulation samples. The results for different simulated data are listed below.

3.2 Simulation with uncured individuals following gamma distribution

Table 3.1 presents the results when the censoring time follows an exponential distribution and the simulated data are fit by the gamma mixture model and the EGG mixture model in turn. The graphical description of power for this case is in Figure 3.1.

We can see that the powers are large for $\theta = 1.5, 2.5$ by fitting with the gamma mixture model. The power is at least 0.76 when the sample size is 200 and the true cure rate of the superpopulation ranges from 5% to 10%. But for $\theta = 0.5$, the power is quite small for all the discussed situations. All the estimated values are no more than 0.50, regardless of the sample size in use. A heuristic explanation is that, when θ is larger than 1.0, the hazard function increases, which means that uncured patients are likely to relapse or die early. Therefore, long-term survivors may be regarded as cured individuals. When θ is less than 1.0, the hazard function decreases. Long-term survivors are more likely to be censored uncured individuals.

As the sample size decreases from 200 to 50, the power decreases dramatically for large θ. However, for $\theta = 0.5$, the change is not substantial. When the sample size is 50, the values of power for large θ range from 0.37 to 0.71 with cure rate varying from 5% to 10%. If the actual distribution of failure times of uncured individuals is a gamma distribution with θ larger than 1.0, the likelihood ratio test will be effective for detecting the presence of cured individuals under the condition that the sample size is not quite small.

Table 3.1: Estimated power of the LRT based on 500 samples under exponential censoring distributions with 30% censoring in uncured individuals

	sample		cure rate				
Model	size	θ	3%	5%	7%	10%	15%
gamma	50	0.5	0.13	0.20	0.20	0.31	0.31
		1.0	0.21	0.28	0.33	0.46	0.53
		1.5	0.24	0.37	0.46	0.56	0.61
		2.5	0.33	0.48	0.56	0.71	0.84
	100	0.5	0.16	0.25	0.35	0.44	0.45
		1.0	0.29	0.43	0.56	0.65	0.76
		1.5	0.36	0.56	0.64	0.76	0.87
		2.5	0.51	0.70	0.80	0.88	0.96
	200	0.5	0.17	0.29	0.37	0.45	0.50
		1.0	0.45	0.62	0.78	0.88	0.92
		1.5	0.57	0.76	0.89	0.94	0.96
		2.5	0.73	0.90	0.95	0.99	1.00
EGG	50	0.5	0.13	0.14	0.14	0.19	0.19
		1.0	0.15	0.20	0.24	0.29	0.32
		1.5	0.17	0.25	0.29	0.33	0.37
		2.5	0.26	0.33	0.37	0.48	0.51
	100	0.5	0.16	0.16	0.17	0.25	0.27
		1.0	0.30	0.49	0.57	0.68	0.81
		1.5	0.36	0.56	0.64	0.76	0.87
		2.5	0.40	0.61	0.67	0.81	0.92
	200	0.5	0.17	0.21	0.22	0.30	0.30
		1.0	0.47	0.61	0.65	0.78	0.83
		1.5	0.50	0.67	0.72	0.85	0.92
		2.5	0.68	0.83	0.89	0.93	0.97

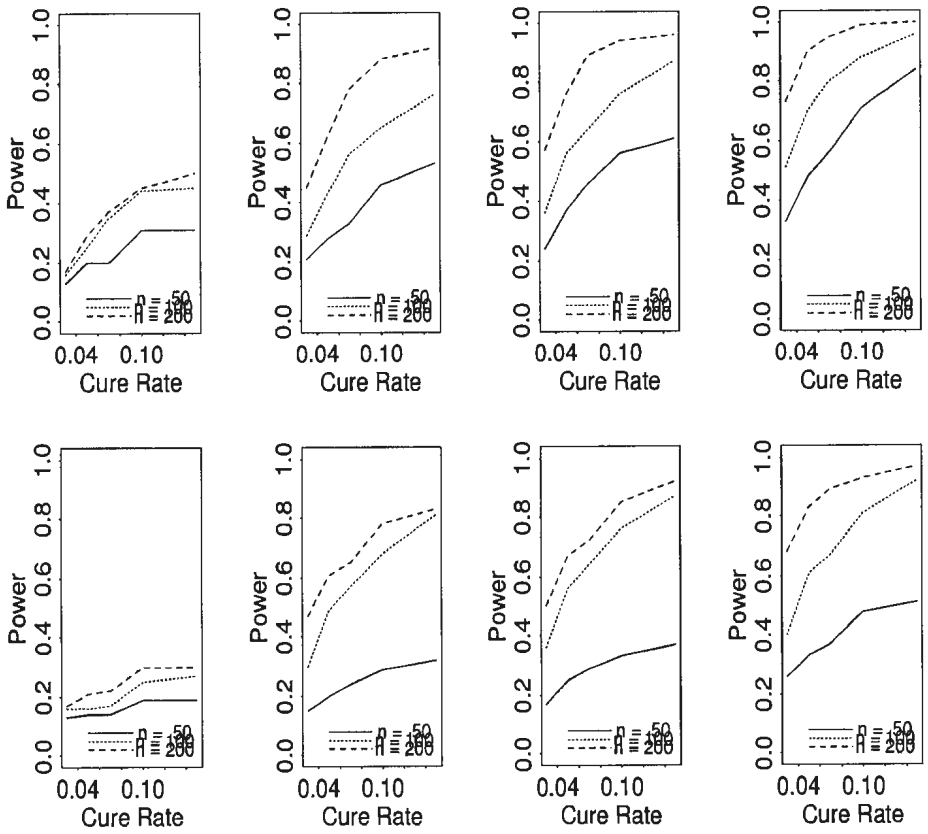

Figure 3.1: Power plot with uncured individuals following gamma distribution and exponential censoring (the first row fit by the gamma mixture, and the second row fit by the EGG mixture; from the left to the right: $\theta = 0.5$, 1.0, 1.5, 2.5, respectively)

If we fit the EGG mixture model to the data, all the values of power are slightly less than those fit by the gamma mixture model. We may consider the EGG mixture model as an instead of the gamma mixture model when the gamma assumption is validated adequately.

Table 3.2 presents the results when the censoring time follows a uniform distribution and the simulated data are fit by the gamma mixture model and the EGG mixture model in turn. The graphical description of power for this

case is in Figure 3.2.

If the data are fit by the gamma mixture model, all the estimated values of power are quite small except for the case that $\theta = 2.5$ and sample size equals to 200. When θ is greater than 1.0, the power increases rapidly as the value of θ increases. For θ equal to 0.5, the power is extremely low. All values are no more than 0.19.

For sample size less than 50, the values are no more than 0.39. As the sample size increases, the power increases significantly. But sample size only has a strong effect on the power if θ is greater than 1.0. If the actual distribution of failure times of uncured individuals is a gamma distribution with θ larger than 1.0, a relative large sample size will be needed to employ the test.

The values calculated by fitting the data with the EGG mixture model are much less than those fitting with the gamma mixture model. Those values are even less than 0.50. Therefore, the result with the replacement of EGG mixture model for gamma mixture model is poor.

3.3 Simulation with uncured individuals following Weibull distribution

Table 3.3 presents the results when the censoring time follows an exponential distribution and the simulated data are fit by the Weibull mixture model and the EGG mixture model in turn. The graphical description of power for this case is in Figure 3.3.

The power of the test is large when θ is larger than 1.0. The values of power

Table 3.2: Estimated power of the LRT based on 500 samples under uniform censoring distributions with 30% censoring in uncured individuals

Model	sample size	θ	cure rate 3%	5%	7%	10%	15%
gamma	50	0.5	0.07	0.07	0.11	0.12	0.11
		1.0	0.07	0.11	0.14	0.17	0.20
		1.5	0.07	0.13	0.18	0.19	0.24
		2.5	0.16	0.24	0.26	0.32	0.39
	100	0.5	0.07	0.10	0.11	0.12	0.16
		1.0	0.10	0.16	0.17	0.23	0.34
		1.5	0.15	0.17	0.20	0.32	0.35
		2.5	0.23	0.33	0.43	0.49	0.57
	200	0.5	0.09	0.10	0.12	0.18	0.19
		1.0	0.12	0.16	0.24	0.28	0.40
		1.5	0.17	0.29	0.37	0.46	0.58
		2.5	0.32	0.51	0.62	0.71	0.82
EGG	50	0.5	0.13	0.14	0.14	0.14	0.13
		1.0	0.09	0.12	0.12	0.14	0.14
		1.5	0.09	0.12	0.13	0.15	0.16
		2.5	0.09	0.12	0.16	0.16	0.18
	100	0.5	0.13	0.11	0.10	0.14	0.15
		1.0	0.12	0.13	0.13	0.14	0.15
		1.5	0.12	0.13	0.13	0.18	0.15
		2.5	0.13	0.21	0.22	0.25	0.28
	200	0.5	0.11	0.13	0.11	0.12	0.13
		1.0	0.12	0.14	0.15	0.17	0.21
		1.5	0.12	0.16	0.20	0.23	0.30
		2.5	0.20	0.30	0.35	0.39	0.45

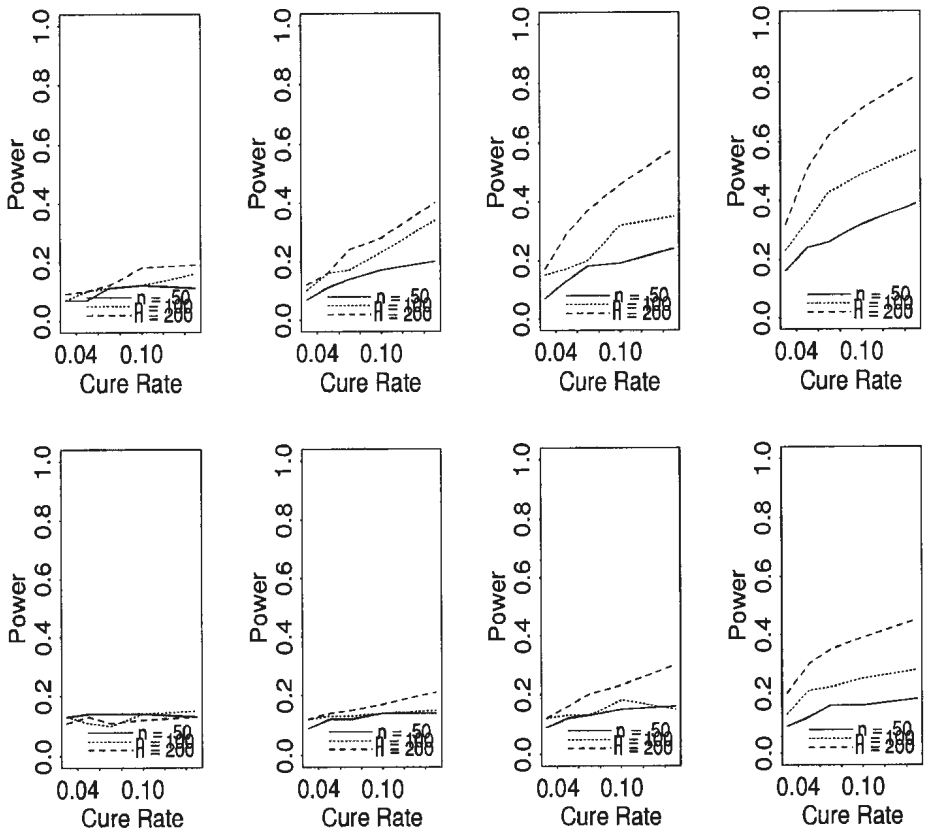

Figure 3.2: Power plot with uncured individuals following gamma distribution and uniform censoring (the first row fit by the gamma mixture, and the second row fit by the EGG mixture; form the left to the right, $\theta = 0.5$, 1.0, 1.5, 2.5, respectively)

for cure rate no less than 5% are at least 0.63 when the sample size is greater than or equal to 100. But for θ equals to 0.5, the power is extremely low. All values are at most 0.29 even for sample size as large as 200. Furthermore, the power will be small if θ is less than 1.0. This is due to the characteristic of Weibull distribution hazard function. That is, the hazard function is monotone increasing if $\theta > 1$, decreasing if $\theta < 1$, and constant for $\theta = 1$. Long-term survivors are more likely identified as cured individuals for $\theta > 1$ than for

$\theta < 1$. The power should be large for $\theta > 1$.

The power is small for $\theta = 0.5$, regardless of the sample size. If $theta > 1$, the power increases rapidly as the sample size increases from 50 to 200. For this case, the sample size does not substantially affect the power for small θ. But the effect of the sample size is apparent for relatively large θ.

The power computed by fitting the data with the EGG mixture model are quite similar to that fitting with the Weibull mixture model. The EGG mixture model can be considered to fit the data so as to test the presence of cured individuals when the Weibull assumption can be verified.

Table 3.4 presents the results when the censoring time follows a uniform distribution and the simulated data are fit by the Weibull mixture model and the EGG mixture model in turn. The graphical description of power for this case is in Figure 3.4.

When the data are fit by the Weibull mixture model, the power is large only for moderate large sample size($n > 100$) and large θ($\theta > 2.5$). For suggested cure rate ranging from 5% to 10%, the power is at least 0.92. If θ is small, the power is small as well. The power increases rapidly as the value of θ increases.

The sample size only affects the power substantially for large value of θ. If θ is at least larger than 2.5, even a small sample size will result in a large power. For $\theta = 0.5$, the sample size almost has no effect on the power.

The power under the EGG mixture fitting is relatively poor compared to that under the Weibull mixture fitting. Flexible model can provide a better fitting to the data, but it is less effective for testing the cured individuals in this case.

Table 3.3: Estimated power of the LRT based on 500 samples under exponential censoring with 30% censoring in uncured individuals

	sample		cure rate				
Model	size	θ	3%	5%	7%	10%	15%
Weibull	50	0.5	0.10	0.11	0.13	0.13	0.16
		1.5	0.31	0.46	0.52	0.65	0.74
		2.5	0.48	0.64	0.75	0.84	0.92
	100	0.5	0.07	0.11	0.13	0.16	0.22
		1.5	0.47	0.63	0.75	0.86	0.93
		2.5	0.70	0.86	0.92	0.99	1.00
	200	0.5	0.11	0.15	0.20	0.25	0.29
		1.5	0.72	0.89	0.93	0.98	1.00
		2.5	0.92	0.98	1.00	1.00	1.00
EGG	50	0.5	0.09	0.12	0.15	0.19	0.19
		1.5	0.27	0.35	0.40	0.61	0.73
		2.5	0.43	0.59	0.72	0.78	0.87
	100	0.5	0.09	0.12	0.15	0.16	0.16
		1.5	0.44	0.60	0.66	0.78	0.80
		2.5	0.67	0.82	0.92	0.98	1.00
	200	0.5	0.10	0.14	0.11	0.12	0.30
		1.5	0.67	0.81	0.92	0.97	1.00
		2.5	0.90	0.98	0.99	1.00	1.00

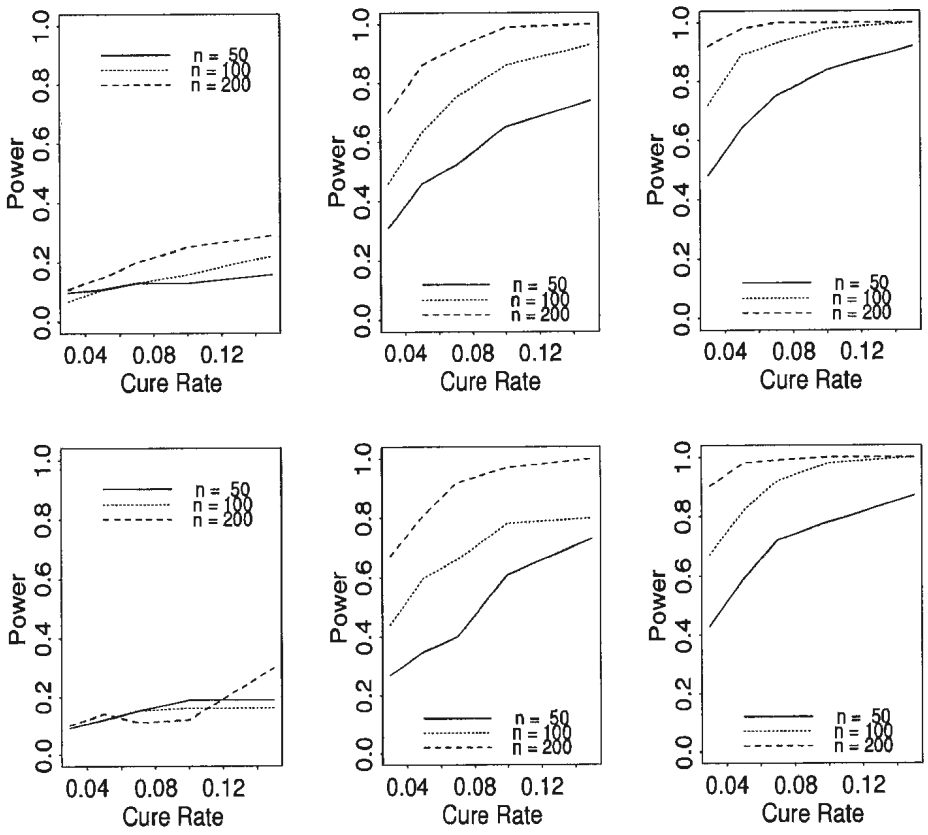

Figure 3.3: Power plot with uncured individuals following Weibull distribution and exponential censoring (the first row fit by the Weibull mixture, and the second row fit by the EGG mixture; from the left to the right: $\theta = 0.5$, 1.5, 2.5, respectively)

3.4 Simulation with uncured individuals following log-normal distribution

Table 3.5 presents the results of the power when the censoring time follows exponential distribution and the simulated data are fit by the log-normal mixture model and the EGG mixture model in turn. The graphical description of

Table 3.4: Estimated power of the LRT based on 500 samples under uniform censoring distributions with 30% censoring in uncured individuals

	sample		cure rate				
Model	size	θ	3%	5%	7%	10%	15%
Weibull	50	0.5	0.08	0.09	0.08	0.10	0.11
		1.5	0.11	0.18	0.20	0.23	0.30
		2.5	0.35	0.49	0.60	0.66	0.76
	100	0.5	0.07	0.09	0.10	0.11	0.13
		1.5	0.17	0.23	0.33	0.35	0.43
		2.5	0.55	0.75	0.82	0.88	0.92
	200	0.5	0.10	0.09	0.10	0.14	0.17
		1.5	0.21	0.38	0.48	0.59	0.68
		2.5	0.78	0.92	0.98	0.99	1.00
EGG	50	0.5	0.11	0.13	0.11	0.13	0.15
		1.5	0.09	0.13	0.15	0.17	0.20
		2.5	0.23	0.31	0.42	0.44	0.47
	100	0.5	0.12	0.15	0.11	0.14	0.14
		1.5	0.12	0.20	0.20	0.21	0.25
		2.5	0.46	0.60	0.69	0.74	0.74
	200	0.5	0.11	0.12	0.11	0.13	0.13
		1.5	0.18	0.24	0.32	0.34	0.34
		2.5	0.74	0.86	0.92	0.96	0.96

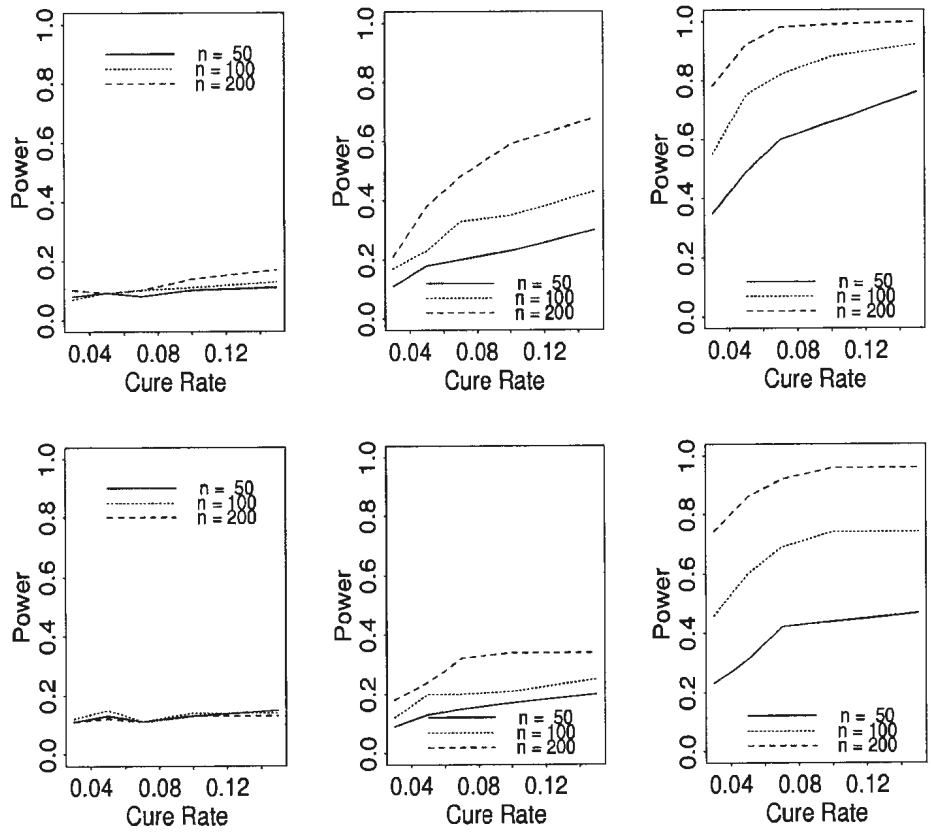

Figure 3.4: Power plot with uncured individuals following Weibull distribution and uniform censoring (the first row fit by the Weibull mixture, and the second row fit by the EGG mixture; from the left to the right: $\theta = 0.5$, 1.5, 2.5, respectively)

power for this case is in Figure 3.5.

If the data are fit by the log-normal mixture model, the power is large for $\theta > 1$ with the cure rate varying from 5% to 10%. When $\theta = 0.5$, the power is small. All estimated values are no more than 0.37.

The sample size affects greatly the power when the $\theta > 1$. The power increases quickly as the sample size increases. For small θ, such as $\theta < 1$, the sample size does not affect the power substantially.

Table 3.5: Estimated power of the LRT based on 500 samples under exponential censoring distributions with 30% censoring in uncured individuals

	sample		cure rate				
Model	size	θ	3%	5%	7%	10%	15%
log-normal	50	0.5	0.08	0.10	0.11	0.12	0.17
		1.5	0.24	0.35	0.44	0.57	0.71
		2.5	0.42	0.56	0.70	0.81	0.87
	100	0.5	0.10	0.14	0.16	0.19	0.25
		1.5	0.42	0.55	0.67	0.81	0.89
		2.5	0.63	0.80	0.90	0.96	0.99
	200	0.5	0.13	0.17	0.24	0.28	0.37
		1.5	0.63	0.81	0.88	0.97	0.99
		2.5	0.85	0.96	0.99	1.00	1.00
EGG	50	0.5	0.06	0.06	0.07	0.08	0.09
		1.5	0.16	0.23	0.27	0.30	0.38
		2.5	0.36	0.50	0.60	0.68	0.80
	100	0.5	0.06	0.07	0.08	0.08	0.10
		1.5	0.25	0.37	0.45	0.55	0.63
		2.5	0.55	0.76	0.87	0.89	0.97
	200	0.5	0.08	0.10	0.10	0.11	0.18
		1.5	0.44	0.62	0.74	0.80	0.89
		2.5	0.81	0.92	0.97	0.99	1.00

The power under the EGG mixture model is similar to that under the log-normal mixture model. Both models can be selected to fit the data in this case.

Table 3.6 presents the results when the censoring time follows a uniform distribution and the simulated data are fit by the log-normal mixture model and the EGG mixture model in turn. The graphical description of power for this case is in Figure 3.6.

The power is large only for $\theta = 2.5$ and sample size equal to 200 while cure rate ranging from 5% to 10%. When $\theta = 0.5$, the power is small. All estimated

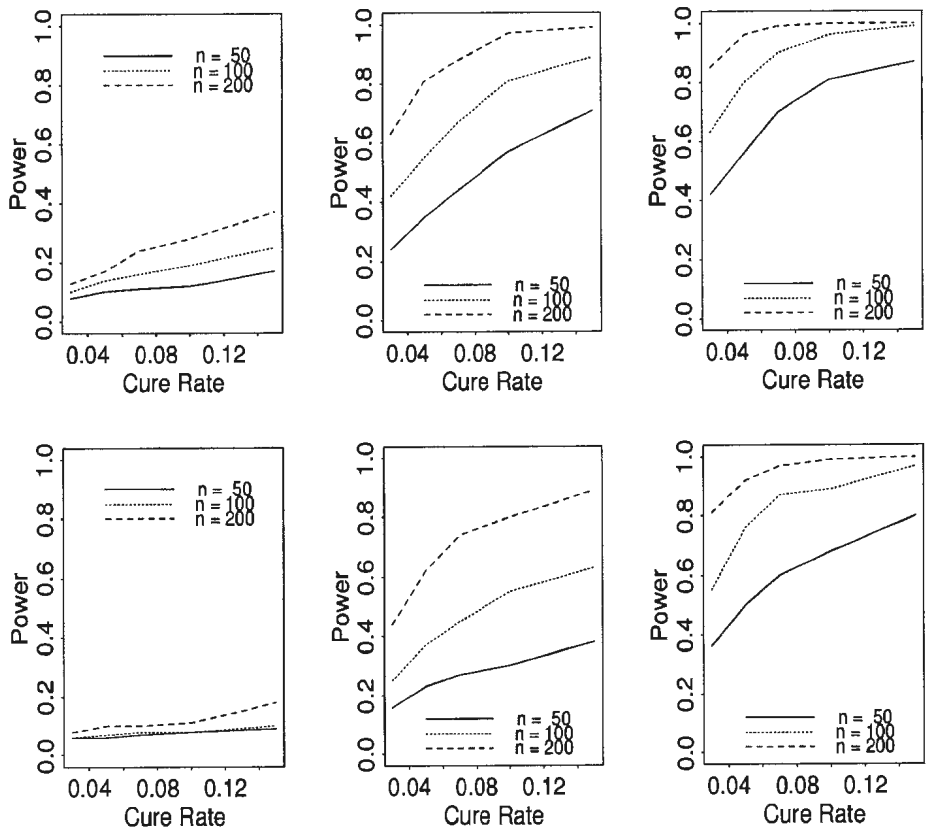

Figure 3.5: Power plot with uncured individuals following log-normal distribution and exponential censoring (the first row fit by the log-normal mixture, and the second row fit by the EGG mixture; from the left to the right: $\theta =$ 0.5, 1.5, 2.5, respectively)

values are less than 0.37.

The power increases quickly as the sample size increases when the $\theta > 1$. For small θ, such as $\theta < 1$, the sample size does not affect the power substantially.

The power under the EGG mixture model is better than that under the log-normal mixture model. We may consider a more flexible model in this case.

Table 3.6: Estimated power of the LRT under uniform censoring with 30% censored in uncured patients

Model	sample size	θ	cure rate 3%	5%	7%	10%	15%
log-normal	50	0.5	0.06	0.06	0.07	0.08	0.09
		1.5	0.08	0.09	0.11	0.12	0.17
		2.5	0.18	0.27	0.28	0.35	0.40
	100	0.5	0.06	0.07	0.08	0.08	0.10
		1.5	0.09	0.12	0.16	0.16	0.20
		2.5	0.24	0.39	0.46	0.54	0.66
	200	0.5	0.08	0.08	0.09	0.11	0.11
		1.5	0.14	0.17	0.22	0.31	0.31
		2.5	0.44	0.64	0.73	0.83	0.90
EGG	50	0.5	0.09	0.12	0.12	0.12	0.16
		1.5	0.07	0.11	0.12	0.18	0.26
		2.5	0.23	0.37	0.42	0.51	0.63
	100	0.5	0.10	0.13	0.19	0.21	0.26
		1.5	0.08	0.17	0.23	0.31	0.44
		2.5	0.40	0.56	0.70	0.83	0.90
	200	0.5	0.11	0.18	0.24	.029	0.37
		1.5	0.16	0.26	.040	0.61	0.74
		2.5	0.61	0.81	0.92	0.97	0.99

3.5 Simulation with nonparametric model

In this section, we introduce a nonparametric statistical test for testing the presence of cured individuals. The nonparametric test allows us to test for cured individuals without taking a parametric assumption. We compare it to the likelihood ratio test discussed in the previous sections.

The nonparametric test is developed on the estimation of the cure rate proposed by Maller and Zhou (1992). They proved that, under some mild

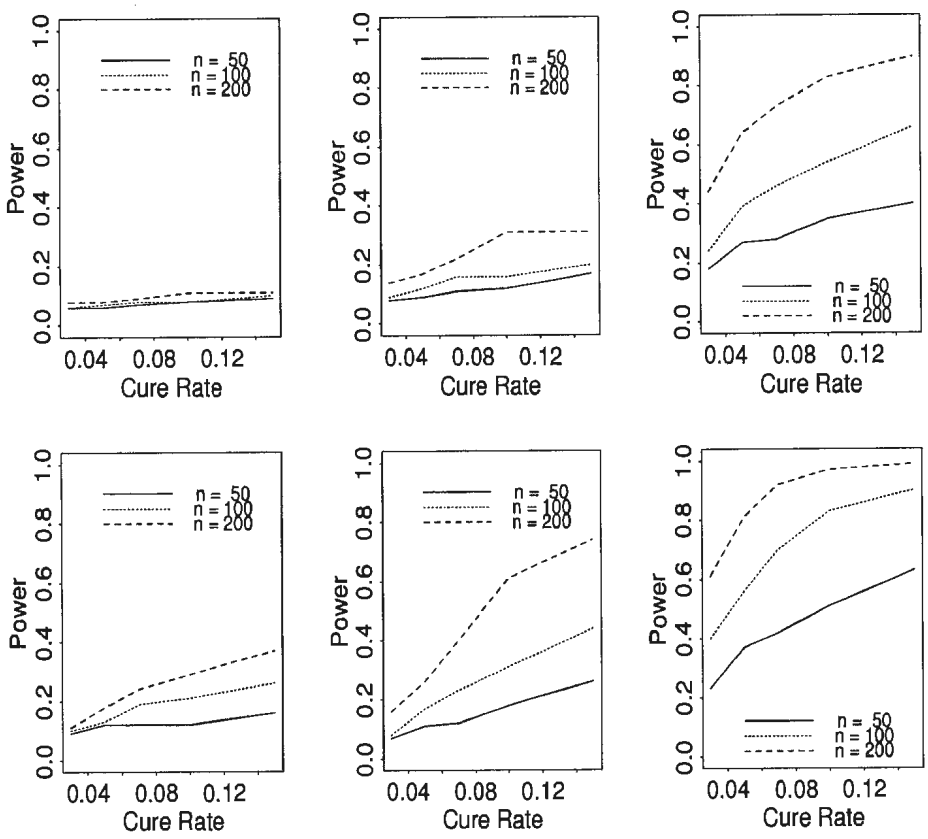

Figure 3.6: Power plot with uncured individuals following log-normal distribution and uniform censoring (the first row fit by the log-normal mixture, and the second row fit by the EGG mixture; from the left to the right: $\theta = 0.5$, 1.5, 2.5, respectively)

conditions,

$$\hat{F}(t_n) \to \rho$$

Here $\hat{F}(t)$ is the K-M cumulative distribution function estimator from a sample of size n, and t_n is the maximum observed failure or censored time. The largest observation is censored, which suggests a level-off of the K-M curve, if and only if $\hat{F}(t_n) < 1$, or $\hat{S}(t_n) = 1 - \hat{F}(t_n) > 0$.

If $0 < \rho < 1$, the estimator of ρ is consistent and asymptotically normal under certain conditions on the censoring mechanism. By using the asymptotic normality property of $\hat{S}(t_n)$, we can construct a 95% confidence interval for it. That is, $\hat{S}(t_n) \pm 1.96\hat{\sigma}_n$, where $\hat{\sigma}_n$, the standard error of $\hat{S}(t_n)$, is given by Greenwood's formula. If the lower bound of the 95% confidence interval of $\hat{S}(t_n)$ is greater than 0, we may consider that cured individuals exist.

We only consider gamma distribution as the failure time distribution of uncured individuals in the mixture models. The density distribution of the gamma distribution is: $f_u(t;\theta) = t^{\theta-1}e^{-t}/\Gamma(\theta)$. The parameter θ is fixed at 0.5, 1.0, 1.5 and 2.5. The censoring distribution may be the uniform or exponential distribution. The values of parameters in the censoring distribution is determined so that the resulting censoring rate for uncured individuals equals to 30 per cent. The cure rate takes a value from $3\%, 5\%, 7\%, 10\%$ and 15%. And the sample size is fixed to 200. Given one of the combinations of the failure time distribution of uncured individuals, five cure rates and two censoring distributions, a sample can be generated with fixed sample size 200.

For each sample, it may be believed to contain cured individuals if the largest observation of the sample is censored and the lower bound of the 95% confidence interval of the K-M estimation of the survival function at the largest observation is greater than 0. The power is estimated by the proportion of this kind of samples based on 500 simulation samples. The estimated nonparametric results are shown in Table 3.7. From the table we can see that no matter how large the cure rate is, the power changes slightly within 0.29 and 0.50. The censoring distribution seems to have no effect on the calculations. The

Table 3.7: Estimated power of the LRT based on 500 samples under nonparametric model with 30% censoring in uncured individuals

		cure rate				
censoring	θ	3%	5%	7%	10%	15%
exponential	0.5	0.36	0.37	0.38	0.38	0.46
	1.0	0.33	0.35	0.37	0.38	0.47
	1.5	0.35	0.33	0.36	0.44	0.44
	2.5	0.33	0.36	0.37	0.41	0.44
uniform	0.5	0.32	0.38	0.37	0.37	0.50
	1.0	0.33	0.34	0.37	0.38	0.48
	1.5	0.29	0.36	0.36	0.35	0.45
	2.5	0.29	0.38	0.42	0.49	0.47

nonparametric method seems not effective for testing the presence of cured individuals.

3.6 Summary

From the above discussion, we can see that the likelihood ratio test is usually not powerful for small sample sizes such as 50 or less. It is often powerful for large sample sizes such as 200 or more. The EGG mixture model performs well for testing the presence of cured individuals in three types of data. If the failure time distribution of uncured individuals is unknown, the EGG mixture model may be used to fit the data of interest.

Censoring distribution strongly affects the power of the test. The power is usually lower for uniform censoring than for exponential censoring. We should examine the censoring distribution carefully before we conduct the test. Also the censoring rate of the data of interest cannot be too light(10% or below).

We find that the test is powerful even if the hazard rate at the upper tail

of the uncured failure time distribution is large. But from Peng *et at.* (2001), the approximation to the null distribution in this case is not acceptable. The critical value is smaller than the one from the suggested distribution (2.2). Therefore, the LRT is powerful regardless of magnitude of the hazard rate at the upper tail of the uncured failure time distribution.

The parametric mixture model will be more efficient than the nonparametric mixture model, since we assume that the failure time distribution and censoring distribution have the parametric forms. If we have little knowledge about the the failure time and censoring distributions, a more flexible model, such as EGG mixture model, may be employed to conduct the test. The nonparametric approach does not seem to be a good choice for testing the presence of cured individuals.

In this practicum, sample size determination has not been considered. We know that the power increases as the cure rate, or the parameter involved, or the sample size increases. With the modern computer facilities, we can find the sample size even though we do not have an explicit form of the power function .

Chapter 4

Bootstrapping on the Likelihood Ratio Test for the presence of cured individuals

Maller and Zhou (1995) proved that, under mild conditions and the exponential assumption, the asymptotic null distribution of the likelihood ratio test statistic d_n follows a 50-50 mixture distribution which is showed in (2.2). Vu *et al.* (1996) extended the result to the gamma mixture model. Peng *el al.* (2001) pointed out that the 50-50 mixture distribution can also be used to approximate the asymptotic null distribution of the likelihood ratio test in the Weibull and log-normal mixture models when the censoring rate is not too light. However, there are two problems here. One is, as showed by Peng *et al.* (2001), that the null distribution of the likelihood ratio test deviates significantly from the 50-50 mixture distribution under moderate sample sizes when the censoring rate is small or the hazard rate is large. Another problem is that there are no existing results about the asymptotic null distribution of LRT if the failure time distribution is other than gamma, log-normal or Weibull.

In this chapter, we investigate the use of bootstrap method to estimate the null distribution of the likelihood ratio test.

4.1 Bootstrapping the Likelihood Ratio Test

Assessing the null distribution of the likelihood ratio test when H_0 is in the boundary was first introduced by McLachlan (1987) for the problem of assessing the number of components in a normal mixture. He highlighted the role of the bootstrap for the assessment of the null distribution for the likelihood ratio test of a single normal density versus a mixture of two normal densities in the univariate case.

Let $t_1, \ldots, t_n$ denote a random sample of size n. We regard this sample as the original sample. It is believed that the sample comes from a population with a finite mixture distribution of g components. The likelihood ratio test is applied to test the number of components g in the situation of $g = g_1$ under H_0 versus $g = g_2$ under H_1. The likelihood ratio test statistic d_n can be bootstrapped as follows. A bootstrap sample is generated from the mixture density $f(t, \hat{\theta})$ of components g_2, where $\hat{\theta}$ is taken to be the maximum likelihood estimate of θ based on the original sample. The value of d_n is computed for the bootstrap sample after fitting mixture models to it for $g = g_1$ and $g = g_2$ in turn. This process is repeated independently K times, and the replicated values of d_n evaluated from the successive bootstrap samples can be used to assess the bootstrap, and hence the true, null distribution of d_n. The original and subsequent bootstrap values of d_n can be treated as the realizations of a random sample of size $K + 1$, and the probability that a specified member

is greater than j of the others is $1 - j/(K+1)$. The jth order statistic of the K replications can then be taken as an estimate of the quantile of order $j/(K+1)$, and the achieved level of significance can be assessed by reference with respect to the ordered bootstrap replications of d_n.

The likelihood ratio test which rejects null hypothesis will approximately have size

$$\alpha = 1 - j/(K+1) \tag{4.1}$$

if d_n for the original data is greater than the jth smallest of its K bootstrap replications.

This result can be applied to the likelihood ratio test for testing the presence of cured individuals. Under null hypothesis, the number of mixture is 1, and the number of mixture under alternative is 2. Following McLachlan's approach, a number of replications of d_n can be generated and then be used as the approximation of null distribution. To determine the size of the test, we can adjust the number of replications according to the formula in (4.1).

For the cure rate estimation, it is common that there are censored observations in the original data. However, McLachlan's approach does not address censoring in data. To apply his approach, we need to find a method of resampling from censored data.

4.2 Bootstrapping for censored data

In this section, we discuss how to obtain bootstrap samples from censored data. We employ the method proposed by Hinkley *et al.* (1997) and Efron *et*

al. (1993). The idea of their method is given as follows.

Suppose that the data available are a random sample $(t_1, d_1), \ldots, (t_n, d_m)$, and that censoring occurs at random. Since the censoring variable C which has a distribution function J, is independent of T^0, the knowledge of quantities $c_1, \ldots, c_m$ alone will tell us nothing about F^0. They would in effect be ancillary statistics. This suggests that simulations should be conditional on the pattern of censorship, so far as practicable. To allow for the censoring pattern, we argue that the only values of c_i known exactly are those t_i with $d_i = 0$. The observed values of the remaining observations are lower bounds for the censoring variables, and they can be treated as censored observations of the censoring variable. Since d_i is an indicator of failure time and therefore $1 - d_i$ is an indicator of censoring, the K-M estimate of the censoring survival function $1 - J$ may be written as

$$1 - \hat{J}(t) = \prod_{i:t_i \leq t} \left(\frac{m - i}{m + 1 - i} \right)^{1 - d_i}$$

J can then be estimated. For the failure time distribution F^0, Hinkley *et al.* also suggested that the K-M estimate can be applied to approximate the survival function $1 - F^0(t)$ when there is little information about the distribution of failure time. We use parametric model to fit the data, and F^0 should then be estimated in a parametric form. For example, we use the EGG mixture model to fit the data. F^0 can be estimated by $H(\hat{\theta})$, where H is a distribution function of the EGG family and $\hat{\theta}$ is the maximum likelihood estimate of the parameters when the original sample is fit by the EGG mixture model.

The algorithm for sampling scheme is as follows.

1. generate $t_1^{0*}, \ldots, t_m^{0*}$ independently from $H(\hat{\theta})$;

2. generate $c_1^{0*}, \ldots, c_m^{0*}$ from $\hat{J}$;

3. set $t_j^* = \min(t_j^{0*}, c_j^*)$, for $j = 1, \ldots, m$.

Each time we carry out the scheme, we can obtain a resample from the original sample. If we repeat the process K times, we will generate K resamples from the original sample.

With the generated resamples, we can estimate the null distribution of the likelihood ratio test and then conduct the test for presence of cured individuals in the data.

4.3 Examples

In this section, we will apply the method discussed in the previous sections to the data either from clinical trials or from other fields.

4.3.1 Leukemia Data

The leukemia data have been introduced in Chapter 1. Mallor and Zhou (1994) fit the exponential mixture models to these two groups, allogeneic and autologous, separately. They concluded that there were cured patients in the allogeneic group, but were not sure about the presence of cured patients in the autologous group since the exponential fit to this group is poor. They pointed out that the Weibull mixture model also provided poor fit to the data of autologous group, with its almost uniform initial failure rate. Peng *et al.* (2001) fit the autologous group using the log-normal mixture model. Their

result showed that the log-normal mixture model provided a good fit to the failure times in this group. Based on the log-normal mixture model, the test statistic d_n is 30.169. This value is greater than 2.31 and 2.71, which are, respectively, the simulated 95th percentile from the empirical null distribution of 1000 samples and the 95th percentile from 50-50 mixture distribution in (2.2). They concluded that the cure rate in this group was nonzero, that is, there were cured patients.

To use the bootstrap method, the exponential mixture model is used to

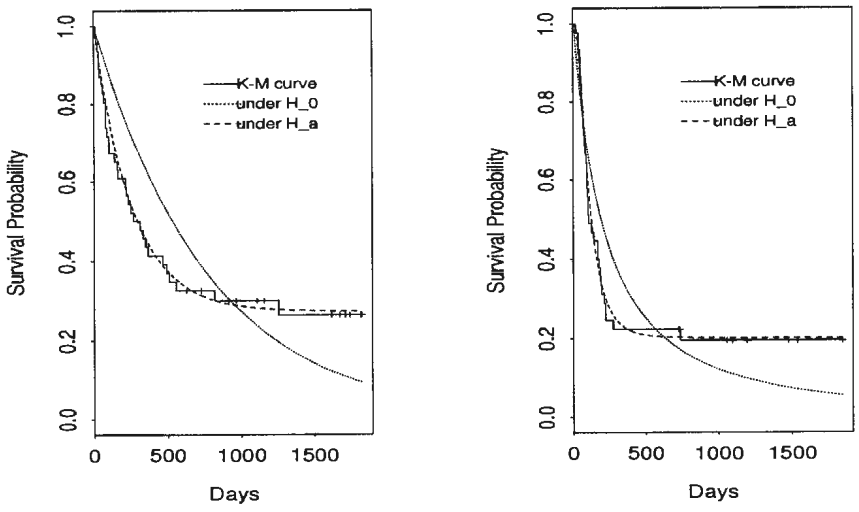

Figure 4.1: Leukemia data: the left, allogeneic group fit by exponential mixture model; the right, autologous group fit by log-normal mixture model

fit the allogeneic group and the log-normal mixture model to the autologous group. The bootstrap critical value is 2.55 for the allogeneic group and 0.51 for the autologous group. Comparing those values to the values of d_n computed by Maller and Zhou (1995) and Peng *et al.*(2001), we consider that there is

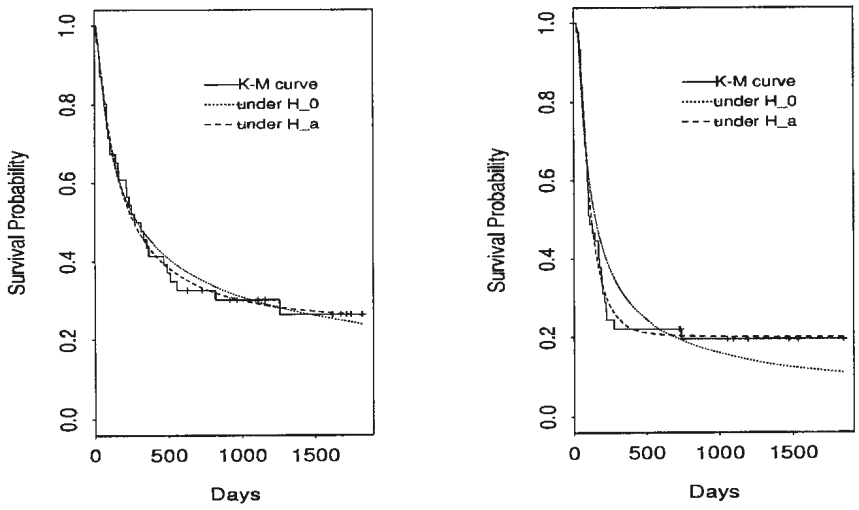

Figure 4.2: Leukemia data: the left, allogeneic group; the right, autologous group. Both are fit by EGG mixture model

strong evidence for the presence of cured patients in each group.

To consider a more flexible model, the EGG mixture model is used to model the two groups separately. Though the null distribution of d_n under the EGG assumption is unknown, we can calculate the approximate critical value for the likelihood ratio test by using bootstrap method. The null hypothesis is considered using $K = 1000$ replications of d_n to construct a test with significant level of 0.05. The critical value is then determined by using the approach in section 4.1.

For $\alpha = 0.05$, the bootstrap critical value is 2.58 for the autologous group and 2.72 for the allogeneic group. If we fit the data with the EGG mixture model, d_n is 11.93 and 0.77, respectively, for the autologous and allogeneic group. Henceforth, we may believe that there were cured patients in the au-

tologous group, but there is not enough evidence of cured patients in the allogeneic group. This is not the same as what Maller and Zhou (1995) concluded. The EGG mixture model, with two extra parameters, is more flexible than the exponential mixture model. The flexibility of the model causes the identification problem for cured individuals. The long-term survivors tend to be treated as censored uncured individuals in the flexible mixture model.

4.3.2 Prison and Arrest Data

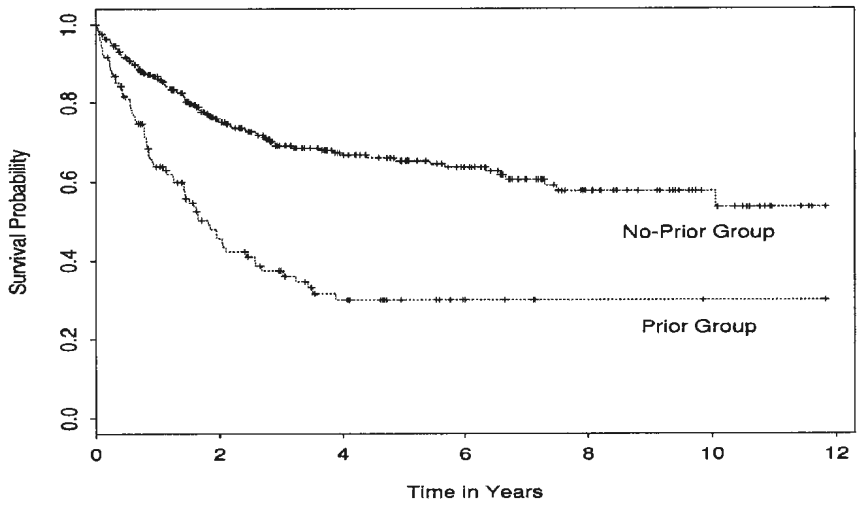

Figure 4.3: Recidivism data

The data in this example were discussed first by Broadhurst and Maller (1990). They are two groups of prisoners who either had convictions or did not have convictions prior to the serious sexual offence. The event of interest is a return to prison for any offence following a prisoner's release from prison for a serious sexual offence. The failure time is the time taken for that return to

prison, if it occurs. Releases not returning to prison within the limit of follow-up time represent censored observations. All data are graphically described in Figure 4.1.

There are obvious level-offs for both groups. There are potentially cured

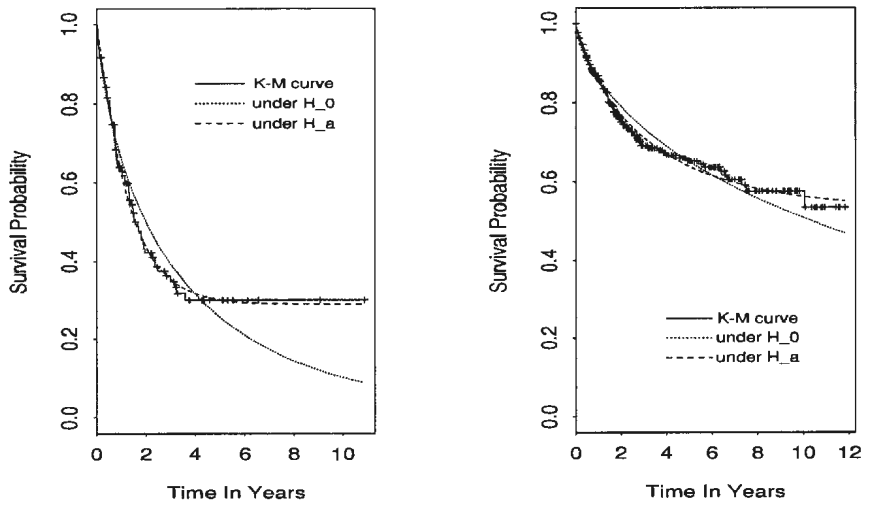

Figure 4.4: Recidivism data: the left, prior group; the right, no-prior group. Both are fit by Weibull mixture model

individuals, who will not return to prison for any offence, in each group. The likelihood ratio test is applied to test the presence of cured individuals. However, we do not know the distributions of the failure times of uncured individuals for each group. Broadhurst and Maller (1990) fit two groups with the Weibull mixture models.

If we fit two groups with the Weibull mixture models, the value of d_n is 13.63 for the prior group and 6.09 for the no-prior group. The bootstrap critical values are, respectively, 2.28 and 1.89. Therefore, there is evidence that cured

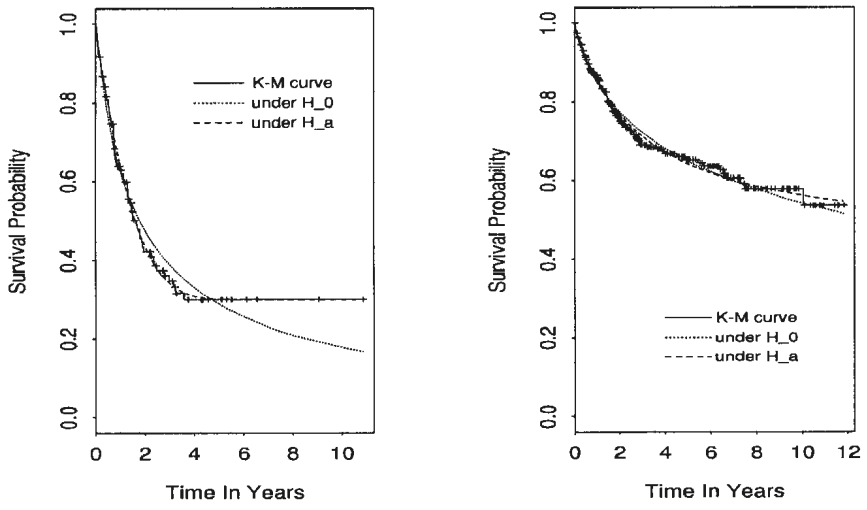

Figure 4.5: Recidivism data: the left, prior group; the right, no-prior group. Both are fit by EGG mixture model

individuals exist in each group. This confirms the result drawn by Broadhurst and Maller (1990).

If we fit the EGG mixture models to these two groups, d_n will be 6.63 for the prior group and 1.23 for the no-prior group. The bootstrap critical value is, respectively 1.59 and 2.72. There is evidence that there are cured individuals in the prior group, while there is no cured individual in the no-prior group. The explanation for the no-prior group is that we use more flexible model to fit the data. There is one more extra parameter in the EGG mixture model, which help to identify long-term survivors as censored uncured individuals.

Bibliography

1. Boag, J. W. 'Maximum likelihood estimates of the proportion of patients cured by cancer therapy', *Journal of the Royal Statistical Society*, **11**, 15-53(1949)

2. Broadhurst, R. G. and Maller, R. A. 'The recidivism of prisoners released for the first time: Reconsidering the effectiveness question', *Australian and New Zealand Journal of Criminology*, **23**, 88-103(1990)

3. Efron, B. and Tibshirani, R. J. 'An introduction to the bootstrap', Chapman and Hall, (1993)

4. Farewell, V. T. 'The use of mixture models for the analysis of survival data with long-term survivors', *Biometrics*, **38**, 1041-1046(1982)

5. Farewell, V. T. 'Mixture models in survival analysis: are they worth the risk?', *Canadian Journal of Statistics*, **14**, 257-262(1982)

6. Goldman, A.I. 'Survivorship analysis when cure is a possibility: A Monte Carlo Study', *Statistics in Medicine*, **3**, 153-163(1984)

7. Hinkley, D. V. and Davison, A. C. 'Bootstrap methods and their application', Cambridge University Press, (1997)

8. Kersey, J. H., Weisdorf, D., Nesbit, M. B., LeBien, T. W., Woods, W. G., McGlave, P. B., Kim, T., Vallera, D. A., Goldman, A. I., Bostrom, B., Hurd, D. and Ramsay, N. K. C. 'Comparison of autologous and allgeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukaemia', *The New England Journal of Medicine*, **317**, 461-467(1987)

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

10. Lawless, J. F. 'Statistical Models and Methods for Lifetime Data', Wiley, New York(1982)

11. Maller, R. A. and Zhou, S. 'Estimating the proportion of immunes in a censored sample', *Biometrika*, **79**, 731-9(1992)

12. Maller, R. A. and Zhou, S. 'Testing for Sufficient Follow-up and Outliers in Survival Data', *Journal of the American Statistical Association*, **89**, 1449-1506(1994)

13. Maller, R. A. and Zhou, S. 'Testing for the Presence of Immune or Cured Individuals in Censored Survival Data', *Biometrics*, **51**, 1197-1205(1995)

14. Maller, R. A. and Zhou, S. 'Survival Analysis with Long-Term Survivors', Wiley, Chichester, New York, London(1996)

15. McLachlan, G. J. 'On bootstrapping the likelihood ratio test statistic for

the number of components in a normal mixture', *Applied Statistics*, **36**, 318-324(1987)

16. Peng, Y., Dear, K. G. and Denham, J. W. 'A generalized F mixture model for cure rate estimation', *Statistics in Medicine*, **17**, 813-830(1998)

17. Peng, Y., Dear, K. G. and Carriere, K. C. 'Testing for the presence of cured patients: A Simulation Study', *Statistics in Medicine*, **20**, 1783-1796(2001)

18. Vu, H. T. V., Maller, R. A. and Zhou, S. 'Asymptotic properties of a class of mixture models for failure data: the interior and boundary cases', Techinical report 33, the University of Western Australia, (1996)

19. Yamaguchi, K. "Accelerated failure time regression models with a regression model of survival fraction: An application to the analysis of 'permanent employment' in Japan", *Journal of the American Statistical Association*, **87**, 284-292(1992)

20. Zhou, S. and Maller, R. A. 'The likelihood ratio test for the presence of immunes in a censored sample', *Statistics*, **27**, 181-201(1995)

Appendix A

Data Sets

Leukemia Data

Data are listed by failure time (in days) and censored status (0 = censored observation). There are 46 patients in the allogeneic group and 45 patients in the autolgous group.

Allogeneic Group

11	1	14	1	23	1	31	1	32	1	35	1	51	1	59	1
62	1	78	1	78	1	79	1	87	1	99	1	100	1	141	1
160	1	166	1	216	1	219	1	235	1	250	1	270	1	313	1
332	1	352	1	368	1	468	1	491	1	511	1	557	1	628	0
726	0	819	1	915	0	966	0	1109	0	1158	0	1256	1	1614	0
1619	0	1674	0	1712	0	1745	0	1820	0	1825	0				

Autolgous Group

21	1	40	1	42	1	50	1	53	1	54	1	56	1	61	1
64	1	67	1	73	1	76	1	79	1	81	1	88	1	95	1
98	1	98	1	99	1	104	1	105	1	106	1	112	1	131	1
147	1	171	1	172	1	179	1	189	1	195	1	199	1	213	1
223	1	224	1	277	1	724	0	729	0	734	1	1053	0	1094	0
1192	0	1475	0	1535	0	1535	0	1845	0						

Prison and Arrest Data

Data are listed by failure time (in years) and censored status (0 = censored observation). There are 296 prisoners with no prior convictions and 121 with prior convictions.

No-Prior Group

0.0027	0	0.0082	1	0.0192	1	0.0219	1	0.0247	1	0.0466	1	0.0767	1
0.0959	1	0.0986	0	0.1151	1	0.1260	1	0.1342	1	0.1425	1	0.1562	0
0.1589	0	0.1699	0	0.2274	1	0.2329	1	0.2438	1	0.2521	1	0.2740	1
0.2849	0	0.3151	0	0.3205	1	0.3315	1	0.3562	1	0.3644	1	0.3836	0
0.3890	1	0.3973	1	0.4082	1	0.4192	1	0.4822	0	0.4877	1	0.5041	0
0.5068	1	0.5260	1	0.5507	0	0.5562	1	0.5616	1	0.5973	1	0.6329	0
0.6356	0	0.6767	1	0.6822	1	0.6986	1	0.7068	0	0.7123	1	0.7205	0
0.7288	1	0.7315	0	0.7507	0	0.7644	1	0.7699	0	0.7863	0	0.8301	1
0.8493	0	0.8521	0	0.8740	0	0.9123	1	0.9260	0	0.9370	0	0.9562	0
0.9973	0	1.0055	1	1.0082	1	1.0110	0	1.0219	1	1.0548	0	1.0658	0
1.0822	1	1.0877	0	1.1068	1	1.1151	1	1.1233	1	1.1945	1	1.2192	1
1.2219	0	1.2247	0	1.2384	0	1.2411	0	1.2438	0	1.2521	0	1.2630	0
1.2849	1	1.3123	1	1.3753	0	1.3918	0	1.3973	1	1.4192	1	1.4274	1
1.4411	1	1.4466	1	1.4466	1	1.4606	0	1.4712	0	1.4822	0	1.4959	0
1.5068	1	1.5151	0	1.5288	0	1.5315	0	1.5534	1	1.5945	0	1.6110	0
1.6247	1	1.6329	0	1.6575	0	1.6740	1	1.6767	1	1.6959	1	1.7041	0
1.7342	0	1.7370	0	1.7425	0	1.7479	0	1.7890	1	1.8037	0	1.8164	1
1.8219	0	1.8411	1	1.8438	0	1.8685	0	1.8740	0	1.8849	0	1.8904	0
1.9206	1	1.9370	0	1.9425	0	1.9452	0	1.9507	0	1.9644	1	2.0027	1
2.0356	0	2.0877	0	2.1014	1	2.1096	0	2.1288	0	2.1726	1	2.1973	1
2.2548	0	2.3123	0	2.3397	0	2.3616	0	2.3644	1	2.4356	1	2.4685	0
2.4904	0	2.5151	0	2.5781	1	2.6247	1	2.6301	0	2.7068	0	2.7315	1
2.7370	0	2.7452	1	2.7589	0	2.7945	0	2.8082	0	2.8137	1	2.8192	0
2.8384	0	2.8630	1	2.9041	1	2.9068	0	2.9370	0	2.9671	0	3.0712	0
3.1507	0	3.1644	0	3.1808	1	3.2164	0	3.2438	0	3.3370	0	3.3452	0
3.3753	0	3.4658	0	3.4685	0	3.5918	0	3.6548	1	3.6630	0	3.7014	0
3.7233	0	3.7260	0	3.7425	0	3.7507	0	3.8192	0	3.8548	1	3.8767	0
3.9151	0	3.9452	0	4.0110	1	4.0137	0	4.0959	0	4.2082	0	4.2219	0

No-Prior Group(continued)

4.2904 0	4.3753 0	4.3836 0	4.3973 1	4.5863 0	4.7288 0	4.7781 0
4.7890 0	4.8329 0	4.8575 1	4.9397 0	4.9589 0	5.0192 0	5.0219 0
5.0384 0	5.0548 0	5.0658 0	5.0849 0	5.2137 0	5.3507 0	5.4247 1
5.5315 0	5.6137 0	5.6658 1	5.7178 0	5.8658 0	5.9397 0	5.9918 0
6.0000 0	6.0767 0	6.1123 0	6.1151 0	6.1589 0	6.2630 0	6.3260 0
6.3397 0	6.3425 1	6.4247 0	6.5397 0	6.5781 1	6.5805 0	6.5945 0
6.6137 0	6.6192 0	6.6630 1	6.7151 0	6.7178 0	6.7315 0	6.7671 0
6.8630 0	6.9781 0	7.0027 0	7.1616 0	7.1918 0	7.2301 0	7.2438 0
7.2575 0	7.2849 0	7.2959 1	7.4438 0	7.4904 1	7.5205 0	7.5726 0
7.6219 0	7.8849 0	7.9178 0	7.9562 0	7.9589 0	8.0630 0	8.0822 0
8.1918 0	8.2082 0	8.4110 0	8.4658 0	8.6329 0	8.8055 0	9.1205 0
9.1644 0	9.3178 0	9.3699 0	9.4055 0	9.4630 0	9.6274 0	9.6877 0
9.7425 0	9.8329 0	10.0575 1	10.0904 0	10.3671 0	10.4932 0	10.5616 0
10.6000 0	10.7753 0	10.8493 0	10.9315 0	10.9616 0	11.4301 0	11.5562 0
11.6137 0	11.8274 0					

Prior Group

0.0219 1	0.0329 1	0.0466 1	0.0575 1	0.0630 1	0.0822 1	0.0822 1
0.0959 1	0.0959 1	0.1205 1	0.1699 0	0.1781 1	0.2000 1	0.2027 1
0.2164 1	0.2301 1	0.2630 1	0.2849 0	0.2877 1	0.2959 1	0.3397 1
0.3753 0	0.3836 1	0.3973 1	0.4082 1	0.4192 0	0.4466 1	0.5123 1
0.5151 1	0.5260 1	0.5425 1	0.5616 1	0.5863 1	0.6055 1	0.6329 0
0.6603 0	0.6685 0	0.6904 0	0.7123 1	0.7178 1	0.7205 1	0.7260 1
0.7507 1	0.7507 1	0.7644 1	0.7753 0	0.7863 1	0.7863 1	0.8000 1
0.8411 1	0.8575 1	0.8986 0	0.9534 0	0.9562 0	0.9616 0	0.9918 0
0.9973 0	1.0082 1	1.0438 0	1.0548 1	1.1479 1	1.1507 1	1.2082 0
1.2740 0	1.3041 1	1.3096 1	1.3205 1	1.3370 1	1.3397 0	1.3918 1
1.4411 0	1.4630 1	1.4904 1	1.4959 0	1.5178 1	1.5205 1	1.5726 0
1.6685 1	1.6932 1	1.7973 1	1.7973 1	1.8767 1	1.8959 1	1.9452 1
2.2164 0	2.2438 1	2.2658 0	2.3671 1	2.3753 1	2.4493 0	2.4712 1
2.7233 0	2.7534 0	2.7918 1	2.8219 0	2.9836 1	3.1151 0	3.1562 1
3.2055 0	3.2110 0	3.2329 1	3.2630 0	3.5836 1	3.7425 0	3.7562 0
3.7808 0	4.2548 0	4.2877 0	4.3123 0	4.3452 0	4.5589 0	5.0822 0
5.1233 0	5.3014 0	5.4877 0	5.5260 0	6.1178 0	6.5479 0	6.5671 0
9.0685 0	10.8822 0					

Appendix B

S-Plus Programs

The program for computing the power of the likelihood ratio test via a simulation study:

```
# failure time distribution for uncured individuals -- Gamma
# censoring distribution -- exponential
# n1 -- the number of samples
# n2 -- sample size
# n3 -- censoring rate in a sample
# n4 -- suggested cure rate of the population
# n5 -- the parameter of the failure time distribution
powersimu <- function(n1, n2, n3, n4, n5)
{
# define some useful functions
my.gam0 <- function(theta, delta, x) {
   -sum(delta * (theta[1] * log(theta[2]) + (theta[1] - 1) * log(x)
             - theta[2] * x - log(gamma(theta[1])))
      + (1 - delta) * log(1 - pgamma(x, theta[1], theta[2])))
}
my.gam1 <- function(theta, delta, x) {
   -sum(delta * (log(theta[1]) + theta[2] * log(theta[3])
      + (theta[2] - 1) * log(x) - theta[3] * x - log(gamma(theta[2])))
      + (1 - delta) * log(theta[1] * (1 - pgamma(x, theta[2], theta[3]))
      + 1 - theta[1]))
}
f.gam.exp2 <- function(x, y, z) {
   f.gam.exp1 <- function(x1, x2) {
      fu1 <- function(x, lumda) {
         x^(y - 1) * exp(-x) * (1 - exp(-lumda * x))/gamma(y)
      }
      return(integrate(f = fu1, lower = 0, upper = Inf,
                       lumda = x)$integral)
   }
   return(f.gam.exp1(x, y) - z)
}
# simulate samples and calculate the power
n3 <- n3/(1 - n4)
n4 <- 1 - n4
a2 <- uniroot(f.gam.exp2, interval = c(0.1, 100), y = n5, z = n3)
power1 <- 0
j <- 1
for (j in 1:n1)
   {
   x <- rgamma(n2, n5)
   count001 <- rbinom(n2, 1, n4)
   y <- rexp(n2, a2$root)
   for (i in 1:n2)
```

```
      {
      if (count001[i] < 1) x[i] <- Inf
      }
   delta <- rep(0, n2)
   for (i in 1:n2)
      {
      if (x[i] < y[i])  delta[i] <- 1
      x[i] <- min(x[i], y[i])
      }
   value001 <- nlminb(start= c(0.01, 0.01), obj = my.gam0,
      lower = c(0.001, 0.001), upper = c(10.0, 10.0),
      delta = delta, x = x)
   l0 <- (-1)*value001$objective
   value002 <- nlminb(start= c(0.01, 0.01, 0.01), obj = my.gam1,
      lower = c(0, 0.001, 0.001), upper = c(1, 10.0, 10.0),
      delta = delta, x = x)
   l1 <- (-1)*value002$objective
   d.n <- (-2*(l0 - l1))
   reject.value <- qchisq(0.9, 1)
   if (d.n > reject.value) power1 <- power1 + 1
   j <- j + 1
   }
power1 <- power1/n1
power1
}
```

The program for investigating the estimate of the null distribution of the likelihood ratio test with bootstrap method via a real-life data set:

```
# the original data set -- prior (Prior Group Data)
# n1 -- the number of samples
bootvalue <- function(n1) {
   # the function to generate bootstrap samples
   prior01 <- function(n2) {
      surv.b1 <- survfit(Surv(time, 1 - cens), data = prior)
      time1 <- summary(surv.b1)$time
      value1 <- summary(surv.b1)$surv
      cumdist1 <- 1 - value1
      dat1 <- array(rep(0, n2*length(prior$time)*2), dim=c(n2,
      length(prior$time), 2))
      for (k in 1:n2) {
      # sampling form censoring distribution G(x)
         samp1.unif <- runif(length(prior$time))
         samp1.cens <- rep(0, length(prior$time))
         for (i in 1:length(prior$time)) {
            if (samp1.unif[i] < cumdist1[1]) samp1.cens[i] <- time1[1]
            for (j in 2:length(value1)) {
               if (samp1.unif[i] < cumdist1[j] && samp1.unif[i]
                     > cumdist1[j - 1]) samp1.cens[i] <- time1[j]
            }
         }
      # sampling from failure time distribution F(x)
         dist1.real <- nlminb(start = c(0.71, 1.1, 1.3), obj = my.web1,
                        lower = c(0, 0, 0), upper = c(1, Inf, Inf),
                        delta = prior$cens, x = prior$time)
         samp1.fail <- rweibull(length(prior$time), dist1.real$para[2],
                        dist1.real$para[3])
      # generate a sample
         xx1 <- rep(0, length(prior$time))
         delta1 <- rep(0, length(prior$time))
         for (i in 1:length(prior$time)) {
            xx1[i] <- min(samp1.cens[i], samp1.fail[i])
            if (samp1.fail[i] < samp1.cens[i])
              delta1[i] <- 1
         }
         dat1[k, ,1] <- xx1
         dat1[k, ,2] <- delta1
      }
   dat1
   }
```

```
# determine the size of the likelihood ratio test
   size <- 0.05
# generate n1 bootstrap samples
   p1 <- prior01(n1)
   p2 <- rep(0, n1)
# calculate the value of -2log_lumbda
   for (i in 1:n1) {
      value001 <- nlminb(start= c(0.745, 3.28), obj = my.web0,
                  lower = c(0.001, 0.001), upper = c(100.0, 100.0),
                  delta = p1[i, ,2], x = p1[i, ,1])
      l0 <- (-1)*value001$objective
      value002 <- nlminb(start= c(0.71, 1.1, 0.3), obj = my.web1,
                  lower = c(0, 0.001, 0.001), upper = c(1, 100.0, 100.0),
                  delta = p1[i, ,2], x = p1[i, ,1])
      l1 <- (-1)*value002$objective
      p2[i] <- (-2*(l0 - l1))
   }
# return a value as the (1 - size)th percentile
   p2 <-  sort(p2)
   reject <- p2[(1 - size)*(n1 + 1)]
   reject
}
```