NONPARAMETRIC ESTIMATION OF SURVIVOR FUNCTION IN BIVARIATE COMPETING RISK MODELS

Thesis submitted to the
Cochin University of Science and Technology
for the Award of Degree of
DOCTOR OF PHILOSOPHY
under the Faculty of Science

by

ANSA ALPHONSA ANTONY

DEPARTMENT OF STATISTICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
KOCHI-682022

SEPTEMBER 2005

CERTIFICATE

Certified that the thesis entitled 'Nonparametric Estimation of Survivor Function in Bivariate Competing Risk Models' is a bonafide record of works done by Ms. Ansa Alphonsa Antony under my guidance in the Department of Statistics, Cochin University of Science and Technology, Cochin-22, Kerala, India and that no part of it has been included anywhere previously for the award of any degree or title.

Kochi-22

1st September 2005

P.G. Sankaran

(Supervising Guide)

DECLARATION

This thesis contains no material which has been accepted for the award of any other Degree or Diploma in any University and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due references are made in the text of the thesis.

Kochi-22

1st September 2005

Ansa Alphonsa Antony

Acknowledgements

I am indebted to my guide Dr P.G. Sankaran, Senior Lecturer, Department of Statistics, Cochin University of Science and Technology for the enthusiastic guidance, encouragement and help he has extended to me during the period of my research.

The support extended to me by Dr K.R. Muraleedharan Nair, Professor and Head and Dr V.K. Ramachandran Nair, Professor and Former Head, Department of Statistics, Cochin University of Science and Technology are genuinely gratified.

I remember with profound gratitude the encouragement given to me by the faculty of the Department of Statistics, Cochin University of Science and Technology both as a M. Sc student and as a research scholar

My heartfelt thanks to Prof. J.F Lawless and Prof. Bovas Abraham, University of Waterloo, Canada who helped me with their suggestions in the initial stages of my research work.

I am grateful to Dr. T.M. Jacob, Nirmala College, Muvattupuzha and Mr K.S.D Namboothiripad, Beeta Computers, Thripunithura for the guidance and help in doing the computational work.

I remember with deep gratefulness all my former teachers.

I am thankful to the non-teaching staff of the Department of Statistics, Cochin University of Science and Technology for their aid and assistance.

The camaraderie with my colleagues and friends was refreshing and heartening. I owe to each and every one of them for their help.

The support, encouragement and blessing of my family members are deeply acknowledged.

I thank Council of Scientific and Industrial Research (CSIR), India for the financial support extended by them.

Above all, I bow before the grace of the Almighty.

Ansa Alphonsa Antony

CONTENTS

Chapter 1
PRELIMINARIES

1.1 Introduction

The data measuring the time to some event is referred to as lifetime, survival time or failure time data. The branch of Statistics that deals with modelling and analysis of lifetime data is called survival analysis. Survival analysis has become an important topic to statisticians and practitioners in areas such as biomedical, engineering and social sciences. Applications of lifetime distribution methodology range from investigations of the durability of manufactured items to studies of human diseases and their treatment. Some methods of dealing with lifetime data are quite old, but starting about 1970 the field expanded rapidly with respect to methodology, theory and fields of application.

The definition of lifetime includes a time scale and time origin, as well as a specification of the event that determines the lifetime. In some situations the events are actual deaths of individuals and lifetime is the life length measured from some particular starting point. However, in certain occasions, the lifetime denotes the event of interest, which may not be the actual death. The following examples illustrate the various types of lifetime data that arise in practical situations.

***Example* 1**: In the investigation of carcinogenic substances, laboratory animals are subjected to doses of the substance and then observed to see if they develop tumours. The variable of interest is the time to appearance of a tumour, measured from when the dose is administered.

***Example* 2**: In medical studies dealing with potentially fatal diseases, the lifetime of patients is measured from the date of diagnosis or some other starting point. In such cases, we compare treatments for a disease in terms of the lifetime distributions for patients receiving different treatments.

***Example* 3:** Manufactured items with mechanical or electronic components are often subjected to life tests in order to obtain information on durability. This involves putting items in operation, often in a laboratory setting and observing them until they fail. It is common here to refer to the lifetimes as 'failure times, since when an item ceases operating satisfactorily it is said to have 'failed.'

1.2 Basic Concepts in Univariate Set up

Let T be a nonnegative random variable representing the lifetime of individuals in some population having absolute continuous distribution function $F(t)$ with respect to a Lebesgue measure.

The survivor function of T is given by

$$S(t) = P(T > t) = 1 - F(t). \tag{1.1}$$

$S(t)$ measures the probability of an individual surviving to time t.

In some contexts involving study on lifetime of systems or manufactured items, $S(t)$ is referred to as the reliability function. $S(t)$ is a non-increasing continuous function with $S(0) = 1$ and $S(\infty) = \lim_{t\to\infty} S(t) = 0$. Occasionally, in reliability analysis, we may wish to allow $S(\infty)>0$ to consider settings where some individuals never fail.

An important basic concept associated with life distributions is the hazard function $h(t)$, defined as

$$h(t) = \lim_{\Delta t\to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t} \tag{1.2}$$

The hazard function specifies the instantaneous rate of death or failure at time t, given that the individual survives up to t. Thus $h(t)\Delta t$ is the approximate probability of death of an individual in $[t, t+\Delta t)$, given that the individual has survived up to t. The hazard function is sometimes referred to as hazard rate and

force of mortality. When the probability density function of T, $f(t)$, exists, we can write (1.2) as

$$h(t)=\frac{f(t)}{S(t)}. \tag{1.3}$$

From (1.1) and (1.3), we have

$$h(t)=-\frac{d}{dt}\log S(t) \tag{1.4}$$

which provides

$$S(t) = \exp\left(-\int_0^t h(x)dx\right). \tag{1.5}$$

From (1.5) it follows that $h(t)$ determines the distribution uniquely.

Since $f(t)=-\frac{dS(t)}{dt}$, we get,

$$f(t)=h(t)\exp\left(-\int_0^t h(x)dx\right). \tag{1.6}$$

The functions $f(t)$, $F(t)$, $S(t)$ and $h(t)$ give mathematically equivalent specifications of the distribution of T

It is also useful to define the cumulative hazard function $H(t)$ as

$$H(t)=\int_0^t h(x)dx \tag{1.7}$$

Then $H(t)$ is related to the survivor function by

$$S(t) = \exp\{-H(t)\} \tag{1.8}$$

In survival studies, many subjects fail to continue to be in the study till the event of interest occurs. This leads to incomplete data due to censored observations. The analysis of lifetime data under censoring is a major issue in survival studies.

1.3 Censoring

Censoring is inevitable in survival and reliability studies because the experimenter is unable to obtain complete information on lifetime of individuals. For

example, patients in a clinical trial may withdraw from the study, or the study may have to be terminated at a pre-fixed time point. In industrial experiments, units may break accidentally. In many situations, the removal of units prior to failure is pre-planned in order to provide savings in terms of time and cost associated with testing.

Censoring arises in lifetime data in a variety of ways. Termination of follow-up before an individual fails causes their lifetime to be right censored. In some settings it may be possible only to determine whether an individual is unfailed or failed at a succession of time points $a_1 < a_2 < \ldots < a_n$. In this case the lifetime is known only to lie in some interval $\left[a_{j-1}, a_j\right)$ and the phenomenon is known as interval censoring. Left censoring occurs if the individual is observed to fail prior to some time a_j, but the lifetime is otherwise unknown. In this case, we know only that the lifetime T belongs to the interval $\left[0, a_j\right)$ whereas for right censoring, we know only that T belongs to the interval $\left(a_j, \infty\right)$.

In both engineering and medical applications, right censoring is the most common form of censoring with lifetime data. Right censoring arises in certain situations because some individuals are still surviving at the time that the study is terminated. In some instances, individuals may move away from the study area for reasons unconnected with lifetime, so that the contact is lost. In some other instances, individuals may be withdrawn from the study because of a worsening or improving prognosis. Type I censoring, Type 2 censoring, progressive Type 2 censoring and independent random censoring are different forms of right censoring.

1.3.1 Type 1 Censoring

Type I censoring occurs when each individual has a fixed potential censoring time $Z_i > 0$ such that T_i is observed if $T_i \leq Z_i$, otherwise we only know that $T_i > Z_i$. It often arises when a study is conducted over a specified time period.

1.3.2 Type 2 Censoring

The term Type 2 censoring refers to the situation where n individuals start on study at the same time and the study terminates once r lifetimes have been observed. Here only the r smallest lifetime $t_{(1)} \leq t_{(2)} \leq \ldots \leq t_{(r)}$ in a random sample of n are observed where r is a specified integer between 1 and n.

1.3.3 Progressive Type 2 Censoring

It is a generalization of Type 2 censoring. In this case, the first r_1 failures in a life test of n items are observed and then n_1 of the remaining $n-r_1$ unfailed items are removed from the experiment, leaving $n- r_1 - n_1$ still present. When further r_2 items have failed, n_2 of the still unfailed items are removed and so on. The experiment terminates after some prearranged series of repetitions of this procedure.

Type 1 and Type 2 censorings are more prevalent in the reliability studies of engineering system. They are built into the design of the experiment to reduce the time taken for completing the study. In survival studies regarding biomedical subjects, censoring is more a part of the experimental situation than a matter of deliberate design. Such undesigned censoring occurs when some information about individual lifetime is available, but not exact lifetime. For an example of such undesigned censoring, consider the study on leukaemia patients who are followed from the start of the remission until they go out of remission. If for a given patient the study ends while the patient is still in remission, then the patient's lifetime is considered as censored. For this person, it is only known that the lifetime is not less than the period for which the person was observed. It is called the right random censoring which is the most frequent type of random censoring.

1.3.4 Independent Random Censoring

A very simple random censoring process that is often realistic is one in which each individual is assumed to have a lifetime T and a censoring time Z where T and Z are independent continuous random variables with survivor functions $S(t)$

and $G(t)$ respectively. This means that at any time t, the survival experience in the future is not statistically altered by censoring and survival experience in the past. If $G(t)$ does not depend on any of the parameters of $S(t)$, then we call it non-informative censoring process. Then, the observed variable will be (Y,δ) where $Y=\min(T,Z)$ and $\delta=1$ if $T \leq Z$ and 0 if $T>Z$ δ is called the censoring indicator. The data on n individuals consists of the pairs (Y_i,δ_i); i = 1,2...n. A comprehensive review on different types of censoring is available in Lawless (2003).

Another form of incomplete data that arises in survival or reliability studies is truncation, which arises due to the limited time span of the study or dropouts of the subjects for various reasons.

1.4 Truncation

In many life testing situations, the individuals cannot be randomly selected and followed prospectively from the time origin t = 0, but some value $u > 0$. If the selection of ith individual at time u_i requires that $T_i \geq u_i$ and the observed data for individual i consists of (u_i,t_i,δ_i) where $t_i \geq u_i$ is a lifetime or censoring time, we say that the lifetime T_i is left truncated at u_i In many occasions, at least some of the data arises chronologically before the time the individuals are selected for the study. Then the condition for being included in the data set will therefore be $T_i \leq v_i$. This is referred to as right truncation of the lifetime T_i. Truncated samples of this type arise in reliability and epidemiology (see Kalbfleisch and Lawless, 1988). For various kind of truncation, one could refer to Lawless (2003).

1.5 Estimation

One of the basic objectives in survival analysis is to estimate the survivor function $S(t)$. Two common approaches used in such contexts are parametric and non-parametric approaches. In parametric method, we assume that random variable T follows some distribution $f(t;\theta)$ where the functional form of $f(t;\theta)$ is known but the parameter θ is unknown. Continuous distributions such as exponential, Weibull,

lognormal, log logistic, Pareto and inverse Gaussian are commonly used for modelling lifetime data. For estimation of parameters, one can employ different estimation procedures such as maximum likelihood, method of moments, Bayesian techniques etc. For more details on parametric lifetime models and their estimation one may refer to Martz and Waller (1982), Sinha (1986) and Lawless (2003).

In many practical situations, the functional form of $f(t)$ is seldom known. In such situations the estimation of $f(t)$ or $S(t)$ is done using nonparametric methods. If there are no censored observations in a sample of size n, $S(t)$ can be estimated by the empirical survivor function, defined as

$$\hat{S}_{ESF}(t) = \frac{Number\ \ of\ \ observations \geq t}{n}$$

When there are censored observations, some modification is necessary. Accordingly, Kaplan and Meier (1958) defined a product-limit estimator for the survivor function $S(t)$.

1.5.1 Kaplan-Meier Estimator (Product-Limit Estimator)

Let (t_i^*, δ_i); $i = 1, 2 \ldots n$ represent a random sample of life times which may contain censored observations. Suppose that there are k $(k \leq n)$ distinct times $t_1 < t_2 < \ldots < t_k$ at which death occur and let $d_j = \sum I(t_i^* = t_j, \delta_i = 1)$, where $I(.)$ denote the usual indicator function, represent the number of deaths at t_j. Then the product limit estimator of $S(t)$ is defined as

$$\hat{S}(t) = \prod_{j:\, t_j < t} \frac{n_j - d_j}{n_j}$$

where $n_j = \sum_{i=1}^{n} I(t_i^* \geq t_j)$ is the number of individuals at risk at t_j, which is the number of individuals alive and uncensored just prior to t_j. The product-limit estimator does not change at censoring time points. The product limit estimator can

be derived as a nonparametric maximum likelihood estimator. When there are no censored observations, it reduces to the empirical survivor function.

Another approach is to develop non-parametric estimator of $S(t)$ using the estimator of cumulative hazard function. Accordingly, a nonparametric estimator of $H(t)$ was proposed by Nelson (1969) and then independently by Aalen in his doctoral thesis in 1972.

1.5.2 Nelson-Aalen Estimator

The estimator of the cumulative hazard function corresponding to (1.7) is given by the Reimann-Steiltjes integral as

$$\widehat{H}(t) = \int_0^t d\widehat{H}(u).$$

Thus the estimator of $H(t)$ is given by

$$\widehat{H}(t) = \sum_{j:t_j \le t} \frac{d_j}{n_j} \tag{1.9}$$

This is called the empirical cumulative hazard function but is more commonly known as the Nelson-Aalen estimator.

Thus, using (1.8), $S(t)$ can be estimated by

$$\tilde{S}(t) = \exp\left\{-\widehat{H}(t)\right\} \tag{1.10}$$

Both the Kaplan-Meier and Nelson-Aalen estimators possess desirable large sample properties like consistency and asymptotic normality. It is important to note that both $\widehat{H}(t)$ and $\hat{S}(t)$ are non-parametric maximum likelihood estimators. For more properties of (1.9) and (1.10), one may refer to Lawless (2003).

1.6 Competing Risk Models

In medical studies or in the analysis of industrial data, the failure of individuals or items may be attributable to more than one cause or factor. These causes (factors) in some sense compete for the failure of the experimental unit. The term competing risk refers to such situations in which a organism (or system) is exposed to two or more cause of death (or failure) but its eventual death (or failure) can be attributed to exactly one of the causes of failure and the model for lifetime in the presence of such competing risks is known as competing risk models. The competing risk models arises in public health, demography, actuarial science, industrial reliability applications and experiments in medical therapeutics. The theory of competing risks dates back to 1760 when Daniel Bernoulli studied the effect of small pox eradication on the mortality structure of the overall population. The following examples provide some situations where the competing risk data arises.

***Example* 1**: Consider the example of Hoel (1972), based on a laboratory experiment in which mice were given a dose of radiation at six weeks of age. The causes of death were recorded as Thymic Lymphoma, Reticulum Cell Sarcoma, or other. Another example of competing risk in survival analysis is from a study of breast cancer patients where the cause of death was recorded as 'cancer' or 'other' (Boag, 1949).

***Example* 2**: An example of competing risk problem in industry is the data from Hinds (1996). This data concern failure of engines fitted to heavy vehicles. Five causes of failure were identified- the cooling system, dirt contamination, mechanical failure, ignition fault and fuel fault. For each unit, the miles traveled to failure and the cause of failure are reported. There are numerous examples in industrial experiments, where items may fail due to one of several causes.

***Example* 3**: In economics, Flinn and Heckman (1983) apply a competing risks model for modelling the unemployment time, where T is the waiting time till the end of unemployment and C indexes the reason for leaving unemployment.

In the traditional analysis of these data sets, the researcher is primarily interested in the distribution of lifetimes under one specific cause of failure and all other causes are combined and treated as censored data on the basis that the causes are independent of each other. Associated with the *j*th cause of failure there is a non-negative random variable T_j which represents the observed lifetime if all causes except the *j*th were inoperative, namely the latent or conceptual lifetime (or failure time) of an organism (or a system) whose death (failure) is attributed to only the *j*th cause. If primary interest is focused on one particular cause of failure, failure from other competing risks can be viewed as a form of random right censoring. In the simultaneous presence of all *k* causes only the smallest of such non-negative random variables $T = \min(T_j)$ is in fact observable, together with the actual cause of failure. In other words, each lifetime is potentially right censored by every other lifetimes. If $F_j(t) = P(T_j \le t)$ denote the distribution function corresponding to the random variable T_j, then the survivor function of T is of the form

$$S(t) = \prod_{j=1}^{k} (1 - F_j(t)). \qquad (1.11)$$

One can also adopt a finite mixture model for the analysis of these competing risk data (see Crowder (2001)). If, for each observed failure, the cause of failure can be identified, then the data may be partitioned into separate sets for each failure mode and a lifetime distribution fitted to each mode separately. Then,

$$f(t) = p_1 f_1(t) + p_2 f_2(t) + \ldots + p_k f_k(t)$$

where $f(t)$ and $f_j(t)$ are the p.d.f's corresponding to T and T_j respectively with p_j being the probability that the cause of failure is j (j=1,2...k) For the analysis of lifetime data using finite mixture model, one could also refer to McLachlan and Peel (2000).

By assuming that the causes of failure are independent and that the life distributions belong to some known parametric family, the situation has been dealt with by Sampford (1952), David (1957), Cox (1959), Berkson and Elveback (1960),

Boardman and Kendell (1970), Herman and Patel (1971), Moeschberger and David (1971), Hoel (1972) and Moeschberger (1974). However, such parametric assumptions, especially in the context of medical studies, may be unrealistic. Even when a certain parametric form is assumed, there is no guarantee that the joint survivor function of $(T_1, T_2 \ldots T_k)$ is identifiable. Consequently, although the concept of latent lifetime variables provides a theoretical basis for discussion, other methods will be more suitable for the analysis of competing risks data.

In recent years, models have been developed to assess the lifetimes of a specific risk in the presence of other competing risk factors. The data for these competing risk models consist of the lifetime T and an indicator variable C denoting the specific cause of failure of the individual or item. The causes of failure may be assumed to be dependent or independent. In most situations, in the analysis of competing risk data, we assume that the causes of failure are independent. Even though the assumption of dependence may be more realistic, there is some concern about the identifiability of the underlying model. For analysis of lifetime data with dependent competing risks, one can refer to Aras and Deshpande (1992). See David and Moeschberger (1978) and Crowder (2001) for an exhaustive treatment of different competing risk models.

There are other approaches for the analysis of competing risk data. For an approach to competing risks theory based on the Markov models and counting processes, one could refer to Aalen (1976), Fleming and Harrington (1991), and Anderson et. al (1993).

Two frameworks are often used to deal with standard competing risk settings in which a lifetime variable $T>0$ and a cause of failure $C \in \{1, 2 \ldots k\}$ can be observed for an individual:

(*i*) Cause-specific hazard $(\lambda_j(t))$ formulations, where

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P\{T < t + \Delta t, C = j \mid T \geq t\}}{\Delta t}; \qquad j = 1, 2 \ldots k \tag{1.12}$$

and

(*ii*) Cause-specific sub-distribution function $\left(F_j(t)\right)$ formulations. where

$$F_j(t) = P(T \le t, C = j); \qquad j = 1,2\ldots k. \tag{1.13}$$

1.6.1 The Cause-Specific Hazard Function

We suppose that an individual is subject to k causes of death and that for each individual we observe the time to death and the cause of death C. This process is described in terms of the cause-specific hazard functions

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t, C = j \mid T \ge t)}{\Delta t}.$$

The function $\lambda_j(t)$ was termed 'decremental forces' by the English actuary Makeham (1874) and 'cause-specific hazard function' by Prentice et.al. (1978). The function $\lambda_j(t)$ is identical to the 'force of transition' function in Aalen's Markov formulation. It is the 'forces of mortality' that an actuary would estimate from a multiple decrement table. In words, $\lambda_j(.)$ is the instantaneous rate for failure of type j at time t given the individual has survived time t and in the presence of all other failure types.

We assume that the k failure types are mutually exclusive and exhaustive so that an individual can have at most one realized lifetime. Assuming the existence of the quantities $\lambda_j(.)$, the overall failure rate or hazard function $h(t)$ is given by

$$h(t) = \sum_{j=1}^{k} \lambda_j(t). \tag{1.14}$$

Prentice et.al. (1978) emphasize that only probabilities expressible as functions of $\{\lambda_j(.)\}$ may be estimated from the observable data (T, C).

Equations (1.5) and (1.14) show that the survivor function $S(t)$ can be written in terms of cause specific hazard functions.

Let $S_j(t)$ denote cause-specific survivor function. Then,

$$S_j(t) = P(T \ge t, C = j).$$

Substituting (1.14) in (1.5), we get

$$S(t) = \exp\left(-\int_0^t \sum_{j=1}^{k} \lambda_j(x)dx \right)$$

$$= \prod_{j=1}^{k} S_j(t) \qquad (1.15)$$

1.6.2 Cause-Specific Sub-distribution Function (Cumulative Incidence Function)

The distribution function of the observable random pair (T,C) is specified by the cause-specific sub-distribution function

$$F_j(t) = P(T \le t, C = j); \qquad t > 0,\ j = 1,2\ldots k$$

Each F_j is a sub-distribution function in the sense that $F_j(+\infty) \le 1$ The importance of cause-specific sub-distribution functions is well recognized in demography, epidemiology and survival analysis. Frequently, they are the primary efficacy measures. Comparison of cause-specific sub-distribution function for different types of failure is useful in selecting appropriate treatment for a patient (Gray, 1988). The cause-specific sub-distribution functions are estimable from the data on (T,C) without making any untestable assumptions and avoid the identifiability problem inherent in competing risks (Prentice et. al., 1978). Cause-specific hazard functions are, of course, estimable from the (T,C) data but, as pointed out by Pepe and Mori (1993), they do not directly indicate the magnitude of the proportion of patients suffering each of the cause-specific endpoints. We refer to Lin (1997), Cheng et. al. (1998), Gooley et. al. (1999), Cronin and Feuer (2000) and Farley et. al. (2001) for examples involving use of cause-specific sub-distribution functions in survival analysis.

Either set of functions fully specifies the joint distribution of T and C, but they lead to different types of regression models when covariates are present. Hougaard (2000), Crowder (2001), Kalbfleisch and Prentice (2002) and Lawless (2003) provide reviews of this area.

The problem of identifiability in modelling the competing risk data in terms of the latent lifetimes is well known. Tsiatis (1975) showed that given any joint survivor function with arbitrary dependence between component variates, there exists a different joint survivor function in which the variates are independent and which reproduces the sub- distribution function $F_j(t)$. Thus, one cannot know, from observations on (T,C) alone, which of the two models is correct since they will both fit the data equally well. The problem of identifiability does not arise if the modelling of competing risk data is done in terms of the sub-distribution functions of (T,C) or cause-specific hazard rates or related quantities.

Censoring can occur both in engineering life testing and in medical follow up studies under competing risk set up as well. Many researchers have studied the problem of nonparametric estimation of survivor function of competing risk models over the past few decades. Although existing methods can adequately accommodate the presence of censored life times, the issues of multiple types of failure are not so easily handled. The latent lifetime approach or finite mixture models can also be used to deal with competing causes of failure under censoring. In the competing risks framework the identifiability problems arose because we could only observe the random vector (T,C), the occurrence of jth failure type effectively censoring the remaining latent failure times due to other causes. To overcome the identifiability problem, we suppose that each study subject has an underlying lifetime T that may be subject to censoring and along with the lifetime or censoring time T, the cause of failure (death) of each experimental unit is also observed. Then n study subjects give rise to data $(t_i,C_i,\delta_i=1)$ or $(t_i,\delta_i=0)$; $i=1,2...n$, where t_i is the observed failure time, δ_i is the censoring indicator and C_i is the cause of failure for the ith individual. In such situations the nonparametric estimation technique of Kaplan and Meier (1958) for survivor function is readily generalized to include competing risks. But it is more common to use the Nelson-Aalen estimator of the cumulative hazard function (see Lawless, 2003). Let $\delta_{ij}= I(C_i=j)$ and n_i denote the number of individuals alive and uncensored just prior to time t_i; $i = 1,2...n$,

$j = 1,2...k$. Then the estimator of the cumulative hazard function corresponding to the cause of failure j can be obtained as

$$\hat{H}_j(t) = \sum_{t_i \le t} \frac{\delta_{ij}}{n_i}; \qquad j = 1,2...k,$$

which gives the estimator of the survivor function using (1.10) and (1.14).

1.7 Some Specific Situations in Competing Risk Set up

In this section, we discuss certain phenomena like masking, missing censoring time and random left truncation that are common to lifetime data in competing risk set up.

1.7.1 Masking

Consider a computing module consisting of k components (chips) mounted on a ceramic substrate. A failure in any component causes the module to fail. If a module fails, failure analysis procedures restrict the cause of module failure to some subset of the components. If this subset consists of more than one component, it is called a masked group. Such masking can occur as a consequence of the lack of proper diagnostic equipment and cost and time constraints. The destructive nature of certain component failures makes exact diagnosis difficult resulting in masked causes of failure. In clinical trials and epidemiological studies it is not uncommon to have missing information on cause of death. Lapidus et.al. (1994) in a study of motorcycle fatalities found that 40% of death certificates are missing information. Anderson and Ryan (1998) discuss a study on colon cancer in which the cause of death is masked for 25 % of the deaths. This type of problem can also occur in animal bioanalysis (see Kodell and Chen, 1987). Examples of masked data in reliability and biomedical contexts can be found in Dinse (1986), Reiser et.al. (1995) and Flehinger et.al. (1996).

The estimation of survivor function under masking was first considered by Dinse (1982) and subsequently by Miyakawa (1984), Racine-Poon and Hoel (1984), Lo (1991), Mukherjee and Wang (1993) and Goetghebuer and Ryan (1990,1995). Ideally, in the absence of any such missingness of failure cause (cause of failure is

exactly known), we have the classical competing risk problem with observation on possibly censored lifetime T and the failure type C which would be exactly one of, say k, possible types on each individual.

1.7.2 Missing Censoring Time

There are several situations in the analysis of lifetime data in which censoring times for unfailed units are missing. For example, suppose that T is the lifetime for product in a population of manufactured units. In some situations, T is measured in calendar time and for many types of products, the producers do not know the exact date of sale for most units. Therefore the censoring time for most unfailed items is unknown. For such data, Suzuki (1985), Kalbfleisch and Lawless (1988) and Hu and Lawless (1996) estimated the distribution function of the lifetime using supplementary follow up samples of unfailed units. But there are situations in which the censoring time distribution is either known or can be estimated. Non-parametric method can be employed to estimate the distribution function in such situations (see Hu et al, 1998).

1.7.3 Random Left Truncation

Random truncation arises in lifetime data due to the limited time span of the study or dropouts of the subjects for various reasons. Random truncation models are conveniently used to model several aspects of AIDS data, such as the incubation time which is defined as the time from infection to the onset of the disease or the time from the onset of AIDS until death, or on insurance applications, the reporting lags which is the time between an accident happens and it is reported to the insurance company. In random left truncation, one observes the independent and identically distributed replicates $(T \geq L_i)$; $i = 1,2...n$ of (T,L) where T is the lifetime variable and L is another random variable called the truncating variable which is independent of T. One of the earliest examples of the random left truncation model was given by Lynden-Bell (1971) with an application in astronomy, where T refers to the brightness of celestial objects and it is only partially observable due to a preventing variable L. In survival analysis, we usually come across data subject to random left truncation along with right censoring. One may refer to Andersen et.el. (1993) for

examples of random left truncated data with right censoring in the context of survival analysis.

These features create complexity in the study of lifetime data and much of the recent development of the subject has been devoted to develop new techniques for the analysis of such data.

1.8 Multivariate Set up

Rarely are medical investigators interested in a single outcome. A glance through medical literature reveals that it is the rule rather than the exception to have multiple response variables. Consequently, there has been an explosion in the statistical methodology to handle so-called multivariate lifetime data. Multivariate lifetime data arise when each study subject may experience several events or when there exists some grouping of subjects, which induces dependence among lifetimes of the same group. The sequence of tumour recurrence, the occurrence of blindness in the left and right eyes and the onset of a genetic disease among family members are some examples of such situations in biomedical research. Another example of this situation is provided by a genetic study examining the age at death of parent and children. Wei, et.al. (1989) considered estimation of lifetime distribution of tumour recurrence among patients with bladder cancer. Later, Ichida et.al. (1993) and Klein and Moeschberger (1997) considered estimation of the joint distribution of time to wound excision and time to wound infection in a population of burn victims.

In many scientific investigations, each study subject can potentially experience more than one event. Medical examples of such serial events include the recurrence of a given illness, such as episodes and the progression of disease through successive stages such as HIV infection $\rightarrow$AIDS $\rightarrow$ death. In most cases, we are interested in the duration between two successive states or events, called gap times, which are measured from the same time origin. Lin et.al. (1999) provide simple nonparametric approach for estimating the joint and marginal distributions of the gap times. For a survey on nonparametric estimation of survivor function of gap times on censored observations, one can refer to Anderson et.al. (1993). A semi-parametric

study unifying the recurrent event data with competing risks is given in DeMasi (2000), Bandeen-Roche and Liang (2002) and Kalbfleisch and Prentice (2002).

Most of the early works in competing risk literature assumed independence among causes of failure. However, one often has dependent causes of failure in many physical situations where the lifetime of an individual failing from one cause may be correlated with the lifetime of the same individual failing from a different cause. Thus, in order to allow for such dependencies, the joint distribution of the life times associated with an individual must be multivariate in nature. Towards this end, a general framework based on absolutely continuous multivariate distributions useful in lifetime studies has been provided by Moeschberger and David (1971).

It is natural to seek a nonparametric estimator of the multivariate survivor function for the analysis of lifetime data since parametric assumptions for lifetime data, especially in medical studies, is not realistic. Similar to the role played by the Kaplan-Meier estimator for univariate lifetime data, such an estimator could form the basis for the display of lifetime data for comparison among samples. Unfortunately, the multivariate survivor function estimation problem is yet to be completely solved. There are methods for estimation of multivariate survivor functions under the assumption of independent censoring. The nonparametric maximum likelihood and self-consistency principle do not lead to a consistent estimator of the survivor function for continuous right censored multivariate lifetime data (see Efron (1967) and Turnbull (1976)). There are many possible strongly consistent nonparametric estimators for the multivariate survivor function, but an estimator that is computationally convenient with attractive moderate and large sample efficiency properties have yet to be developed. Different methods of estimation of multivariate survivor function available in literature are discussed in Section 1.9.2.

For simplicity, we confine our discussion to the analysis of bivariate data, but they can directly be extended to the multivariate set up.

1.9 Bivariate Set up

In many practical situations one may have paired lifetime data. For example, times to death or times to initial contraction of disease may be of interest for littermate pairs of rats or for twin studies in humans. The time to deterioration level or the time to reaction of a treatment may be of interest in pairs of lungs, kidneys, eyes or ears of humans. In reliability, one may be interested in the distribution of the life lengths of a particular pair of components in a system. Mathematically, let $T = (T_1, T_2)$ be a non-negative random vector admitting an absolute continuous distribution function $F(t_1, t_2)$ with respect to a Lebesgue measure.

Then the survivor function of T, denoted by $S(t_1, t_2)$, is given by

$$S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$$

which is related to $F(t_1, t_2)$ as $S(t_1, t_2) = 1 - F(t_1, \infty) - F(\infty, t_2) + F(t_1, t_2)$ If the density function of T, $f(t_1, t_2)$, exists, we have

$$f(t_1, t_2) = \frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2}.$$

1.9.1 Bivariate Hazard Function

In the univariate case, it is well known that the hazard function determines the survivor function uniquely. In the bivariate set up however, one can define the hazard function in more than one way. The first definition of bivariate hazard function was given by Basu (1971).

Basu (1971) defined the bivariate hazard function as a scalar quantity given by

$$r(t_1, t_2) = \frac{f(t_1, t_2)}{S(t_1, t_2)}$$

It is important to note that $r(t_1, t_2)$, in general, does not determine the bivariate distribution uniquely.

A second approach to defining bivariate hazard function is provided by Johnson and Kotz (1975). Johnson and Kotz defined bivariate hazard function as a vector given by

$$h(t_1,t_2) = \left(h_1(t_1,t_2), h_2(t_1,t_2)\right) \tag{1.16}$$

where $h_i(t_1,t_2) = -\dfrac{\partial \log S(t_1,t_2)}{\partial t_i}, \qquad i = 1,2.$

$h_1(t_1,t_2)$ is the instantaneous rate of a failure of T_1 at time t_1 given that T_1 was alive at the time $T_1 = t_1$ and that T_2 survived beyond time $T_2 = t_2$. The meaning of $h_2(t_1,t_2)$ is similar.

When the components $h_i(t_1,t_2)$ exist and are continuous in an open set containing $R_2^+ = \{(t_1,t_2) \mid t_i > 0, i = 1,2\}$ by choosing a path orthogonal to the axis connecting $(0,0)$ and (t_1,t_2) in R_2^+, we have the representation from Galambos and Kotz (1978) as an extension of the one-dimensional relationship (1.5). Accordingly, $S(t_1,t_2)$ can be determined from (1.16) as

$$S(t_1,t_2) = \exp\left\{-\int_0^{t_1} h_1(u,0)du - \int_0^{t_2} h_2(t_1,u)du\right\} \tag{1.17}$$

or alternatively

$$S(t_1,t_2) = \exp\left\{-\int_0^{t_1} h_1(u,t_2)du - \int_0^{t_2} h_2(u,0)du\right\} \tag{1.18}$$

Thus, the vector $h(t_1,t_2)$ uniquely determines the distribution of T through (1.17) and (1.18).

Dabrowska (1988) provides a representation of bivariate survivor function in terms of cumulative hazard function which is a vector of three components that correspond to double and single failures. The cumulative hazard vector is defined as

$$\Lambda(t_1,t_2) = \left(\Lambda_{10}(t_1,t_2), \Lambda_{01}(t_1,t_2), \Lambda_{11}(t_1,t_2)\right)$$

where $\Lambda_{10}(dt_1,t_2)=\frac{-S(dt_1,t_2)}{S(t_1^-,t_2)}$, $\Lambda_{01}(t_1,dt_2)=\frac{-S(t_1,dt_2)}{S(t_1,t_2^-)}$, $\Lambda_{11}(dt_1,dt_2)=\frac{S(dt_1,dt_2)}{S(t_1^-,t_2^-)}$

and $\Lambda_{10}(0,t_2)=\Lambda_{01}(t_1,0)=\Lambda_{11}(0,0)=0$.

The bivariate survivor function is uniquely represented using $\Lambda(t_1,t_2)$ as

$$S(t_1,t_2) = \prod_{u\le t_1}(1-\Lambda_{10}(du,0))\prod_{v\le t_2}(1-\Lambda_{01}(0,dv))\prod_{\substack{u\le t_1\\ v\le t_2}}(1-L(du,dv))$$

where $L(du,dv) = \frac{\Lambda_{10}(du,v^-)\Lambda_{01}(u^-,dv)-\Lambda_{11}(du,dv)}{(1-\Lambda_{10}(du,v^-))(1-\Lambda_{01}(u^-,dv))}$.

Dabrowska(1988) also provided an extension of the above representation to the censored set up.

Cox (1972), Marshall (1975), Shanbag and Kotz (1987) and Basu and Sun(1997) have also discussed different versions of hazard function in the bivariate (multivariate) set up.

1.9.2 Estimation of Bivariate Survivor Function

As mentioned earlier, in survival analysis, estimation of the survivor function is one of the main problems of interest. Parametric estimation procedures such as maximum likelihood and Bayesian technique can be applied for the estimation of $S(t_1,t_2)$ by assuming that $T = (T_1,T_2)$ follows some bivariate distribution with unknown parameters. However, nonparametric methods for the estimation of $S(t_1,t_2)$ have become very popular within survival analysis for several reasons. One reason is that lifetime data often have some features that are not easily explained by parametric models. For example, human lifetime data show a decreasing hazard in the first five years of life, then it has a constant hazard and finally it has an increasing hazard. This cannot be fully explained by parametric models. Accordingly, nonparametric approach is common in survival analysis. If there are no censored observations, we can estimate the bivariate survivor function

by the empirical survivor function using the data (T_{1i}, T_{2i}); i =1,2...n, which is given by

$$S^{ESF}(t_1, t_2) = \sum_{i=1}^{n} \frac{I(T_{1i} > t_1, T_{2i} > t_2)}{n}.$$

But in both reliability studies and medical follow-up studies, the phenomenon of censoring is very common. Censoring occurs when the experimental unit is removed from the study before both components have been observed to fail. The censoring may arise for a number of reasons. The items may be withdrawn due to a change in health status or contamination. They may be censored by death from a cause unrelated to the study. The censoring may also be due to sequential entry into an experiment subsequently ended at a fixed time. The problem of estimation in such situations has received considerable attention in statistical literature. Suppose that the pair of lifetime variables $T = (T_1, T_2)$ is subject to random right censoring. Let $Z = (Z_1, Z_2)$ be the censoring vector associated with $T = (T_1, T_2)$. Under the bivariate right censoring, we observe $Y = (Y_1, Y_2)$ where $Y_1 = \min(T_1, Z_1)$ and $Y_2 = \min(T_2, Z_2)$. The censoring indicator is $\delta = (\delta_1, \delta_2)$ where $\delta_1 = I(T_1 = Y_1)$ and $\delta_2 = I(T_2 = Y_2)$. This censoring is often encountered in several situations such as twin studies, studies on diseases of the right and left eyes and studies where two recurrence times of a certain disease are recorded.

The problem of estimating the bivariate survivor function under random right censoring is firstly addressed successfully in Campbell (1981). Tsai et.al. (1986) suggested an estimation procedure for bivariate survivor function using the estimation of conditional survivor functions. Campbell (1981) and Hanley and Parnes (1983) have studied non-parametric maximum likelihood estimation for the survivor function, but it does not have closed form expression and it is not unique. Later Dabrowska (1988) extended the Kaplan and Meier (1958) estimator to the bivariate set up using the product integral representation of survivor function. A semiparametric estimation for dependent multivariate lifetime data when the marginal distributions of the failure times follow proportional hazards model is done

by Cai and Prentice (1995). Non-parametric estimators of survivor function have also been proposed by Campbell and Foldes (1982), Burke (1988), Pruitt (1991), Prentice and Cai (1992), van der Laan (1996) and Wang and Wells (1997) among others, whereas for example, Oakes (1989) and Wang and Wells (1999) proposed semi parametric estimators. Dabrowska (1988) and Prentice and Cai (1992) have developed computationally convenient estimators with good moderate sample performance, but these estimators are in general not nonparametrically efficient and, in particular, since they use Kaplan-Meier marginals, they do not address the auxiliary data problem. Unlike the univariate Kaplan-Meier (1958) estimator, which has the usual optimal properties such as consistency and asymptotic normality, estimators of the bivariate survivor function proposed in literature have some unsatisfactory features and are in general quite complex (see Gill, 1992). Many of the estimators are not proper bivariate distributions, have non-explicit formulae, do not behave well in practice or depend heavily on the choice of smoothing parameters. van der Laan et.al. (2002) proposed a locally efficient estimator for multivariate survivor function when all the component lifetimes are censored by a common variable independent of the lifetimes. Akritas and van Keilegom (2003) obtained path-independent bivariate survivor function through the estimation of marginal and conditional distributions. For different estimation procedures for bivariate survivor function under censoring, one may refer to van der Laan (1997), Oakes (2001), Kalbfleisch and Prentice (2002) and Lawless (2003).

Thus there has been much research on analyzing bivariate lifetimes, but very little has accommodated failures that occur in the presence of competing failure process. But, in several studies, the situation where each component of the bivariate lifetime vector $T = (T_1, T_2)$ has more than one cause of failure is common. Further, the lifetime vector T may be subject to random censoring. The problem of nonparametric estimation of cause-specific sub-distribution function and survivor function in bivariate set up in the presence of more than one cause of failure is not yet addressed. Apart from censoring, there are other features such as masking, missing censoring time and random left truncation that can occur in competing risk set up, which makes the estimation of survivor function and cause-specific

distribution function problematic. The analysis of bivariate (multivariate) lifetime data in such situations is not addressed so far. Motivated by this, in the present study, we undertake the problem of nonparametric estimation of bivariate survivor function and cause-specific sub-distribution functions in competing risk set up.

1.10 Present Study

So far, in the bivariate set up, the analysis of lifetime (failure time) data with multiple causes of failure is done by treating each cause of failure separately, with failures from other causes considered as independent censoring. This approach is unrealistic in many situations. For example, in the analysis of mortality data on married couples one would be interested to compare the hazards for the same cause of death as well as to check whether death due to one cause is more important for the partners' risk of death from other causes. In reliability analysis, one often has systems with more than one component and many systems, subsystems and components have more than one cause of failure. Design of high-reliability systems generally requires that the individual system components have extremely high reliability even after long periods of time. Knowledge of the failure behaviour of a component can lead to savings in its cost of production and maintenance and, in some cases, to the preservation of human life. For the purpose of improving reliability, it is necessary to identify the cause of failure down to the component level. By treating each cause of failure separately with failures from other causes considered as independent censoring, the analysis of lifetime data would be incomplete. Motivated by this, we introduce a new approach for the analysis of bivariate competing risk data using the bivariate vector hazard rate of Johnson and Kotz (1975).

The thesis is organized into seven chapters. After this introductory chapter, in Chapter 2, we consider a system having two study objects with more than one cause of failure for each study object. Based on the vector hazard rate of Johnson and Kotz (1975), we develop a nonparametric estimator for survivor function under independent censoring. We also propose an estimator for the cause-specific sub-distribution functions. Asymptotic properties of the estimators are discussed. Finally,

we illustrate the procedure with a real data and the performance of the estimators is studied using a simulated data.

The problem of missing failure type may arise with bivariate lifetime data as well. Consider a situation in which a system consists of k components and each component is subject to more than one cause of failure. Due to inadequacy in the diagnostic mechanism or reluctance to report any specific cause of failure (disease), the exact cause of failure cannot be identified easily. In such situations where the cause of failure is masked, when failure of a component occurs, test procedures restrict the cause to a set of possible types containing the true type. In Chapter 3, we develop a nonparametric estimator for the bivariate survivor function of competing risk models under masked causes of failure based on the vector hazard rate of Johnson and Kotz (1975). Asymptotic properties of the estimator are established. A simulation study is carried out to assess the performance of the estimator. We also illustrate the method with a real data set. The proposed estimator is an extension of the estimator for the survivor function of masked data in the univariate set up, given by Dewanji and Sengupta (2003).

There are situations in the analysis of lifetime or failure time data under the competing risk set up where the censoring times of unfailed units are missing. Chapter 4 deals with the problem of missing censoring times in both univariate and bivariate competing risk set up. The maximum likelihood and simple moment estimators of cause-specific density function and distribution function in both univariate and bivariate competing risk set up are discussed. A simulation study is also conducted to observe the asymptotic behaviour of the estimators.

Left truncation in survival analysis means that an individual is included for study only if its lifetime is larger than some value. In practice, we come across lifetime data with left truncation and right censoring. The estimation of the bivariate survivor function and cause-specific sub-distribution function for such data under competing risk set up is not yet considered. In Chapter 5, we consider the bivariate truncation model where both components of the lifetime vector are subject to random left truncation and right censoring and each component is exposed to more than one cause of failure. For the bivariate random left truncated and right censored competing

risk lifetime data, we develop a nonparametric estimator for the bivariate survivor function based on the vector hazard rate of Johnson and Kotz (1975). We also develop nonparametric estimators for the cause-specific sub-distribution function. The asymptotic properties of the estimators are discussed. A simulation study, assessing the empirical behaviour of the estimators is also presented.

Dependence relations between random variables are one of the most widely studied subjects in Probability and Statistics. There are several global summary measures such as Karl Pearson's coefficient of correlation, Kendall's τ and Spearman's rank correlation coefficient, that are commonly used to study the dependence among random variables. Although it is customary to compute a correlation coefficient, the dependence between a pair of continuous random variables is often more complex than a single scalar dependence measure can reflect. Therefore, a global summary statistic such as the correlation coefficient will not convey the dependence structure. Accordingly, various local dependence measures are developed in literature. Clayton (1978) introduced a measure of local dependence based on the hazard rate for continuous distributions. Later, Oakes (1989) defined a measure of dependence, which is the conditional version of Kendall's concordance measure. Fan et.al. (1998) proposed local dependence measure, which is the weighted average of the above two concordance measures. Later, Fan et.al. (2000) proposed a class of local dependence measures to study the association between the variables. In Chapter 6, we present a simple local dependence measure for bivariate lifetime data, based on the covariance function of residual lifetime variables. It is shown that zero correlation between residual lifetime variables implies independence among the variables. We propose a nonparametric estimator for the local dependence measure and study its asymptotic properties. Further, a test for independence between the variables is developed. The method is illustrated with a real data set.

Finally, Chapter 7 summarizes major conclusions of the present study. A brief discussion of future works that can be carried out in bivariate (multivariate) competing risk set up is also presented.

Chapter 2

BIVARIATE COMPETING RISK MODELS

2.1 Introduction

We consider the situation in which each experimental unit has two component lifetimes (T_1, T_2) and each of the pair (T_1, T_2) is subject to multiple causes of failure. We denote the associate causes of failure by (C_1, C_2) where C_j is the cause of failure for T_j, j = 1,2. For example, times (T_1, T_2) could represent the ages at death of twins (Hougaard et.al., 1992) or lifetimes of components of a two-component system. With humans the primary cause of death may be classified as cancer, heart disease or other causes. With mechanical parts, failures may be classified according to their root cause. As mentioned in Chapter 1, there is substantial literature on each of competing risks and bivariate (multivariate) lifetime distributions, but there seems to be little discussion of the bivariate (multivariate) competing risk setting. For various estimation procedures of bivariate survivor function under censoring in non-competing risk set up, one could refer to van der Laan (1997), Kalbfleisch and Prentice (2002) and Lawless (2003). However, the estimation of cause-specific sub-distribution in bivariate (multivariate) competing risk set up is not considered so far in literature. Motivated by this, in Section 2.2, we consider lifetime involving pairs of study individuals with more than one possible cause of failure for each individual. Under the assumption that lifetime and censoring time are independent, nonparametric estimation of survivor function using the vector hazard function of Johnson and Kotz (1975) is carried out in Section 2.3. Further, we develop nonparametric estimators for cause-specific sub-distribution functions. In Section 2.4, we discuss various asymptotic properties of the estimators. An illustration of the procedure is presented in Section 2.5. To study the empirical behaviour of the estimators, a simulation study is carried out in Section 2.6. Finally, Section 2.7 gives a conclusion for the chapter.

2.2 Survivor Function and Cause-Specific Sub-distribution Function

Let $T = (T_1, T_2)$ be a pair of non negative random variables defined on a probability space $(\Omega, \mathbb{F}, P)$. Assume that T has absolutely continuous distribution function $F(t_1, t_2)$. The variables T_1 and T_2 are thought of as lifetimes of married couples, failure times of a two-component system etc. Let $Z = (Z_1, Z_2)$ be a pair of random censoring times. Under the bivariate right censoring, the observable variables are given by $Y = (Y_1, Y_2)$ and $\delta = (\delta_1, \delta_2)$ where $Y_i = \min(T_i, Z_i)$ and $\delta_i = I(T_i = Y_i)$; $i = 1,2$.

Let $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$, $G(t_1, t_2) = P(Z_1 > t_1, Z_2 > t_2)$ and $H(t_1, t_2) = P(Y_1 > t_1, Y_2 > t_2)$ be the survivor functions of T, Z and Y respectively. We assume that the failure mechanism and censoring mechanism are independent. Then T and Z are independent. Thus we obtain

$$H(t_1, t_2) = G(t_1, t_2) S(t_1, t_2). \tag{2.1}$$

Let $C = (C_1, C_2)$ be a set of causes corresponding to $T = (T_1, T_2)$ Suppose that there are γ_1 causes of failure for T_1 and γ_2 causes of failure for T_2

Through out this thesis, we use the notations given in Dabrowska (1988). For example, we mean $F(dt) = dF(t)$.

2.2.1 Survivor Function and Hazard Function

Now we consider bivariate cumulative hazard function

$$\Lambda(t_1, t_2) = (\Lambda_1(t_1, t_2), \Lambda_2(t_1, t_2))$$

where

$$\Lambda_1(dt_1, t_2) = \frac{P(T_1 \in dt_1, T_2 > t_2)}{P(T_1 \geq t_1, T_2 > t_2)} = \frac{-S(dt_1, t_2)}{S(t_1^-, t_2)}$$

and

$$\Lambda_2(t_1, dt_2) = \frac{P(T_1 > t_1, T_2 \in dt_2)}{P(T_1 > t_1, T_2 \geq t_2)} = \frac{-S(t_1, dt_2)}{S(t_1, t_2)}$$

with $\Lambda_1(0,t_2) = \Lambda_2(t_1,0)=0$.

When $T = (T_1,T_2)$ has a joint density function $f(t_1,t_2)$, we have $\Lambda_1(dt_1,t_2)$ $=h_1(t_1,t_2)dt_1$ and $\Lambda_2(t_1,dt_2) = h_2(t_1,t_2)dt_2$ with

$$h_1(t_1,t_2) = \lim_{\Delta t_1 \to 0} \frac{1}{\Delta t_1} P(T_1 \le t_1 + \Delta t_1 \mid T_1 \ge t_1, T_2 > t_2) \tag{2.2}$$

and

$$h_2(t_1,t_2) = \lim_{\Delta t_2 \to 0} \frac{1}{\Delta t_2} P(T_2 \le t_2 + \Delta t_2 \mid T_1 > t_1, T_2 \ge t_2). \tag{2.3}$$

Thus $h_1(t_1,t_2)$ is the instantaneous rate of a failure of T_1 at time t_1 given that T_1 was alive at the time $T_1=t_1^-$ and that T_2 survived beyond time $T_2=t_2$. The meaning of $h_2(t_1,t_2)$ is similar.

From (1.17) and (1.18),

$$S(t_1,t_2) = \exp\{-\Lambda_1(t_1,0) - \Lambda_2(t_1,t_2)\} \tag{2.4}$$

and

$$S(t_1,t_2) = \exp\{-\Lambda_1(t_1,t_2) - \Lambda_2(0,t_2)\} \tag{2.5}$$

Thus (2.4) and (2.5) provides a representation of the bivariate survivor function $S(t_1,t_2)$ in terms of $\Lambda_1(t_1,t_2)$ and $\Lambda_2(t_1,t_2)$.

The cause-specific hazard functions corresponding to (2.2) and (2.3) are given by

$$h_{1i}(t_1,t_2) = \lim_{\Delta t_1 \to 0} \frac{1}{\Delta t_1} P(T_1 \le t_1 + \Delta t_1, C_1 = i \mid T_1 \ge t_1, T_2 > t_2);\ i = 1,2\ldots\ \gamma_1 \tag{2.6}$$

and

$$h_{2j}(t_1,t_2) = \lim_{\Delta t_2 \to 0} \frac{1}{\Delta t_2} P(T_2 \le t_2 + \Delta t_2, C_2 = j \mid T_1 > t_1, T_2 \ge t_2);\ j = 1,2\ldots\ \gamma_2. \tag{2.7}$$

The cause-specific cumulative hazard functions for T_1 and T_2 are respectively given by

$$\Lambda_i^{(1)}(t_1,t_2) = \int_0^{t_1} h_{1i}(u,t_2)\,du; \qquad i = 1,2\ldots\ \gamma_1 \tag{2.8}$$

and

$$\Lambda_j^{(2)}(t_1,t_2) = \int_0^{t_2} h_{2j}(t_1,u)\,du; \qquad j = 1,2\ldots\ \gamma_2. \tag{2.9}$$

Assuming that failure type C_k must be a unique element of $\{1,2...\gamma_k\}$, using (2.8) and (2.9), the cumulative hazard function $\Lambda_k(t_1,t_2)$, k =1,2 is given by

$$\Lambda_1(t_1,t_2) = \sum_{i=1}^{\gamma_1} \Lambda_i^{(1)}(t_1,t_2) = \sum_{i=1}^{\gamma_1} \int_0^{t_1} h_{1i}(u,t_2)\,du \tag{2.10}$$

and

$$\Lambda_2(t_1,t_2) = \sum_{j=1}^{\gamma_2} \Lambda_j^{(2)}(t_1,t_2) = \sum_{j=1}^{\gamma_2} \int_0^{t_2} h_{2j}(t_1,u)\,du. \tag{2.11}$$

Thus, by (2.10) and (2.11), (2.4) and (2.5) can be written as

$$S(t_1,t_2) = \exp\left\{-\sum_{i=1}^{\gamma_1} \int_0^{t_1} h_{1i}(u,0)\,du - \sum_{j=1}^{\gamma_2} \int_0^{t_2} h_{2j}(t_1,u)\,du\right\} \tag{2.12}$$

and

$$S(t_1,t_2) = \exp\left\{-\sum_{i=1}^{\gamma_1} \int_0^{t_1} h_{1i}(u,t_2)\,du - \sum_{j=1}^{\gamma_2} \int_0^{t_2} h_{2j}(0,u)\,du\right\}, \tag{2.13}$$

which provides a representation of $S(t_1,t_2)$ in terms of cause-specific hazard functions.

We denote

$$F_{1i}^{(1)}(t_1,t_2) = P(T_1 \geq t_1, T_2 > t_2, C_1 = i), \tag{2.14}$$

$$F_{2j}^{(2)}(t_1,t_2) = P(T_1 > t_1, T_2 \geq t_2, C_2 = j), \tag{2.15}$$

$$p_{1i}(t_1,t_2) = P(Y_1 \geq t_1, Y_2 > t_2, \delta_1 = 1, C_1 = i) \tag{2.16}$$

and

$$p_{2j}(t_1,t_2) = P(Y_1 > t_1, Y_2 \geq t_2, \delta_2 = 1, C_2 = j); \quad i = 1,2\ldots\ \gamma_1 \quad j = 1,2\ldots\ \gamma_2. \tag{2.17}$$

From (2.1), (2.16) and (2.17), we obtain

$$p_{1i}(dt_1,t_2) = G(t_1^-,t_2)\, F_{1i}^{(1)}(dt_1,t_2); \qquad i = 1, 2\ldots\ \gamma_1 \tag{2.18}$$

and

$$p_{2j}(t_1,dt_2) = G(t_1^-,t_2)\, F_{2j}^{(2)}(t_1,dt_2); \qquad j = 1,2\ldots\ \gamma_2. \tag{2.19}$$

Thus, from (2.6) and (2.14), we get

$$\begin{aligned} h_{1i}(t_1,t_2) &= \frac{P(T_1 \in dt_1, T_2 > t_2, C_1 = i)}{P(T_1 \geq t_1, T_2 > t_2)} \\ &= \frac{F_{1i}^{(1)}(dt_1,t_2)}{S(t_1^-,t_2)} \end{aligned} \tag{2.20}$$

Using (2.1) and (2.18), we can write (2.20) as

$$\begin{aligned} h_{1i}(t_1,t_2) &= \frac{F_{1i}^{(1)}(dt_1,t_2)G(t_1^-,t_2)}{H(t_1^-,t_2)} \\ &= \frac{p_{1i}(dt_1,t_2)}{H(t_1^-,t_2)}; \qquad i = 1, 2\ldots\ \gamma_1. \end{aligned} \tag{2.21}$$

Similarly, using (2.1),(2.7),(2.15) and (2.19)

$$h_{2j}(t_1,t_2) = \frac{p_{2j}(t_1,dt_2)}{H(t_1,t_2)}; \qquad j = 1,2\ldots\ \gamma_2. \tag{2.22}$$

Therefore, from (2.10), (2.11), (2.21) and (2.22), we obtain

$$\Lambda_1(t_1,t_2) = \sum_{i=1}^{\gamma_1}\int_0^{t_1} h_{1i}(u,t_2)\,du = -\sum_{i=1}^{\gamma_1}\int_0^{t_1}\frac{p_{1i}(du,t_2)}{H(u^-,t_2)}\,du \tag{2.23}$$

and

$$\Lambda_2(t_1,t_2) = \sum_{j=1}^{\gamma_2}\int_0^{t_2} h_{2j}(t_1,u)\,du = -\sum_{j=1}^{\gamma_2}\int_0^{t_2}\frac{p_{2j}(t_1,du)}{H(t_1,u^-)} \tag{2.24}$$

2.2.2 Cause-Specific Sub-distribution Function

As mentioned in Section 1.7.2, in many practical situations, the cause-specific sub-distribution function is preferred over cause-specific hazard function. In the bivariate competing risk set up, we define the cause specific hazard function as

$$\Lambda_{ij}(dt_1,dt_2)=\frac{P(T_1\in dt_1,T_2\in dt_2,C_1=i,C_2=j)}{P(T_1\geq t_1,T_2\geq t_2)};\ i=1,2\ldots\gamma_1;j=1,2\ldots\ \gamma_2. \qquad (2.25)$$

We can write (2.25) as

$$\Lambda_{ij}(dt_1,dt_2)=\frac{F_{ij}(dt_1,dt_2)}{S(t_1^-,t_2)} \qquad (2.26)$$

where $F_{ij}(t_1,t_2)$ denote the cause-specific sub-distribution function given by

$$F_{ij}(t_1,t_2)=P(T_1\leq t_1,T_2\leq t_2,C_1=i,C_2=j);\ i=1,2\ldots\gamma_1;j=1,2\ldots\ \gamma_2. \qquad (2.27)$$

This quantity measures the probability for the failure of both the study subjects (T_1,T_2) prior to (t_1,t_2) due to the causes (i,j). In mortality studies this is helpful to compare whether death of one cause is important for the partners risk of death of other causes.

From (2.26) and (2.27), it follows that

$$F_{ij}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}S(u^-,v^-)\Lambda_{ij}(du,dv);\qquad i=1,2\ldots\gamma_1\ ;j=1,2\ldots\ \gamma_2 \qquad (2.28)$$

Let $D=\{(t_1,t_2):H(t_1,t_2)>0\}$

Define

$$F_{ij}^{*}(t_1,t_2)=P(T_1\leq t_1,T_2\leq t_2,\delta_1=1,\delta_2=1,C_1=i,C_2=j). \qquad (2.29)$$

Then, from the independent censoring assumption, we get

$$F_{ij}(t_1,t_2)=\frac{F_{ij}^{*}(t_1,t_2)}{G(t_1^-,t_2)}=\frac{S(t_1,t_2^-)F_{ij}^{*}(t_1,t_2)}{H(t_1,t_2)} \qquad (2.30)$$

From (2.26) we have

$$\Lambda_{ij}(dt_1,dt_2)=\frac{F_{ij}^{*}(dt_1,dt_2)}{G(t_1,t_2)S(t_1^{-},t_2^{-})}=\frac{F_{ij}^{*}(dt_1,dt_2)}{H(t_1^{-},t_2)} \text{ for all } (t_1,t_2)\in D \qquad (2.31)$$

Hence, from (2.28) and (2.31), for all $(t_1,t_2)\in D$, we obtain

$$F_{ij}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{F_{ij}^{*}(du,dv)}{G(u^{-},v^{-})} \qquad (2.32)$$

and

$$F_{ij}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{S(u^{-},v^{-})F_{ij}^{*}(du,dv)}{H(u^{-},v^{-})}; \qquad i=1,2\ldots\gamma_1\ ; j=1,2\ldots\ \gamma_2. \qquad (2.33)$$

2.3 Nonparametric Estimation

In this section we give non-parametric estimators of $S(t_1,t_2)$ and $F_{ij}(t_1,t_2)$ Now suppose that $Y_u=(Y_{1u},Y_{2u}),\delta_u=(\delta_{1u},\delta_{2u})$ $u=1,2\ldots n$ be an independent and identically distributed (i.i.d) sample, each (Y_u,δ_u) having the same distribution as (Y,δ) and the corresponding failure cause pair is $C_u=(C_{1u},C_{2u})$.

2.3.1 Estimation of Bivariate Survivor Function

To estimate the bivariate survivor function $S(t_1,t_2)$, define

$$\widehat{H}(t_1,t_2)=\frac{1}{n}\sum_{u=1}^{n}I(Y_{1u}>t_1,Y_{2u}>t_2),$$

$$\hat{p}_{1i}(t_1,t_2)=\frac{1}{n}\sum_{u=1}^{n}I(Y_{1u}\geq t_1,Y_{2u}>t_2,\delta_{1u}=1,C_1=i)$$

and

$$\hat{p}_{2j}(t_1,t_2)=\frac{1}{n}\sum_{u=1}^{n}I(Y_{1u}>t_1,Y_{2u}\geq t_2,\delta_{2u}=1,C_2=j) \qquad i=1,2\ldots\gamma_1\ ; j=1,2\ldots\ \gamma_2,$$

with $I(.)$ as the usual indicator function.

Then from (2.23) and (2.24), the estimators of $\Lambda_1(t_1,t_2)$ and $\Lambda_2(t_1,t_2)$ are obtained as

$$\hat{\Lambda}_1(t_1,t_2) = -\sum_{i=1}^{\gamma_1}\int_0^{t_1}\frac{\hat{p}_{1i}(du,t_2)}{\hat{H}(u^-,t_2)}du \tag{2.34}$$

and

$$\hat{\Lambda}_2(t_1,t_2) = -\sum_{j=1}^{\gamma_2}\int_0^{t_2}\frac{\hat{p}_{2j}(t_1,du)}{\hat{H}(t_1,u^-)} \tag{2.35}$$

for all t_1,t_2 such that $\hat{H}(t_1^-,t_2) > 0$ and 0 otherwise.

Thus from (2.4), (2.5), (2.34) and (2.35), the estimator of $S(t_1,t_2)$ is obtained as

$$\hat{S}_1(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1(t_1,0)-\hat{\Lambda}_2(t_1,t_2)\right\} \tag{2.36}$$

and

$$\hat{S}_2(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1(t_1,t_2)-\hat{\Lambda}_2(0,t_2)\right\} \tag{2.37}$$

The estimator of $S(t_1,t_2)$ obtained by (2.36) and (2.37) may be different. To get a unique estimator, we follow the approach given in Akritas and van Keilegom (2003). Since the estimator of the bivariate distribution should not depend on which variable we consider as the conditioning variable, the proposed estimator for $S(t_1,t_2)$ is a convex combination of two expressions (2.36) and (2.37). Thus the estimator for $S(t_1,t_2)$ is given by

$$\hat{S}_a(t_1,t_2) = a(t_1,t_2)\hat{S}_1(t_1,t_2)+(1-a(t_1,t_2))\hat{S}_2(t_1,t_2). \tag{2.38}$$

Now the question is how to choose $a(t_1,t_2)$. Choose the weight $a(t_1,t_2)$ in such a way that the mean squared error (MSE) of $\hat{S}_a(t_1,t_2)$ is minimum. As given in Section 3 of Akritas and van Keilegom (2003), one can obtain $a(t_1,t_2)$ which minimized the MSE as

$$a(t_1,t_2) = \frac{\sigma_{22}-\sigma_{12}+\mu_2^2-\mu_1\mu_2}{\sigma_{11}+\sigma_{22}-2\sigma_{12}+\mu_1^2+\mu_2^2-2\mu_1\mu_2}$$

where σ_{ij} is the asymptotic covariance between S_i and S_j and μ_i is the asymptotic bias of S_i; i,j = 1,2.To ensure that $\hat{S}_a(t_1,t_2)$ belongs to the interval [0,1], we replace $a(t_1,t_2)$ by $\min\left[1,\max\left\{a(t_1,t_2),0\right\}\right]$.

Remark 2.1 When the system has only one study object, both (2.34) and (2.35) reduces to the estimator of the cumulative hazard function in the univariate case, given in Kalbfleisch and Prentice (2002, page 255).

Remark 2.2 The extension to multivariate set up is direct as the survivor function $S(t_1,t_2...t_k)$ of $(T_1,T_2...T_k)$ can be uniquely represented by

$$S(t_1,t_2...t_k)=\exp\left\{-\Lambda_1(t_1,0...0)-\Lambda_2(t_1,t_2,0...0)-...-\Lambda_k(t_1,t_2...t_k)\right\}$$

where

$$\Lambda_j(t_1,...t_j,0...0)=\int_0^{t_j}h_j(t_1,...t_{j-1},u,0...0)du$$

with

$$h_j(t_1,t_2...t_k)=-\frac{\partial\log S(t_1,...t_k)}{\partial t_j};\qquad j=1,2...k.$$

2.3.2 Estimation of Cause-Specific Sub-distribution Function

Our next objective is to estimate cause-specific sub-distribution functions (2.27) nonparametrically. This could be approached by maximum likelihood considering four different counting processes

(*i*) $I(T_1\in dt_1,T_2\in dt_2,C_1=i,C_2=j)$ and $n_{ij}(t_1,t_2)=\sum_u I(T_{1u}\in dt_1,T_{2u}\in dt_2,C_{1u}=i,C_{2u}=j)$,

(*ii*) $I(T_1\in dt_1,T_2>t_2,C_1=i)$ and $n_{i\cdot}(t_1,t_2)=\sum_u I(T_{1u}\in dt_1,T_{2u}>t_2,C_{1u}=i)$,

(*iii*) $I(T_1>t_1,T_2\in dt_2,C_2=j)$ and $n_{\cdot j}(t_1,t_2)=\sum_u I(T_{1u}>t_1,T_{2u}\in dt_2,C_{2u}=j)$,

(*iv*) $I(T_1>t_1,T_2>t_2)$ and $n_{\cdot\cdot}(t_1,t_2)=\sum_u I(T_{1u}>t_1,T_{2u}>t_2)$

However, nonparametric maximum likelihood estimation of bivariate survivor function $S(t_1,t_2)$ even in the non-competing risk case is problematic, for example the maximum likelihood estimator is not unique (see Lawless, 2003, page 500). In this section, we consider simple non-parametric estimators of the functions (2.27) using (2.32) and (2.33).

An unbiased estimate of $F_{ij}^{*}(t_1,t_2)$ is given by

$$\widehat{F}_{ij}(t_1,t_2)=\frac{1}{n}\sum_{u=1}^{n} I(Y_{1u}\le t_1, Y_{2u}\le t_2, \delta_{1u}=1, \delta_{2u}=1, C_{1u}=i, C_{2u}=j)$$

Based on (2.32) and (2.33), we therefore suggest two simple estimators for $F_{ij}(t_1,t_2)$ as

$$\widehat{F}_{ij}^{(1)}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{\widehat{F}_{ij}^{*}(du_1,du_2)}{\widehat{G}(u_1,u_2)} \tag{2.39}$$

for (t_1,t_2) such that $\widehat{G}(t_1,t_2)>0$ and 0 otherwise

and

$$\widehat{F}_{ij}^{(2)}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{\widehat{S}_u(u_1^{-},u_2)\widehat{F}_{ij}^{*}(du_1,du_2)}{\widehat{H}(u_1,u_2)} \quad i=1,2\ldots\gamma_1\ ;\ j=1,2\ldots\gamma_2 \tag{2.40}$$

for (t_1,t_2) such that $\widehat{H}(t_1,t_2)>0$ and 0 otherwise.

The estimator $\widehat{G}(t_1,t_2)$ is the Burke (1988) estimator for the survivor function of the censoring variables. The Burke (1988) estimator of $\widehat{G}(t_1,t_2)$ is defined when $t_1\le \max\{T_1\}$ and $t_2\le \max\{T_2 : T_1>t_1\}$ The estimators (2.39) and (2.40) are defined for continuous time settings as well as discrete with the terms in (2.39) and (2.40) being non-zero only for observed failure times (t_1,t_2).

2.4 Properties of Estimators

In this section we prove the consistency and weak convergence of the estimators of $S(t_1,t_2)$ and $F_{ij}(t_1,t_2)$

Theorem 2.1

$\underset{D}{Sup}\left\|\widehat{H}-H\right\|\to 0$ almost surely. In other words, the estimator $H(t_1,t_2)$ is strongly consistent.

The proof follows from Lemma A.2 of Kulkarni and Rattihalli (2002,p.913)

Theorem 2.2

Suppose that T and Z are independent. Then $\underset{D}{Sup}\left\|\hat{\Lambda}_k-\Lambda_k\right\|\to 0$ almost surely; k =1,2.

Proof

For k =1, from (2.10),

$$\underset{D}{Sup}\left\|\hat{\Lambda}_1-\Lambda_1\right\| = \underset{D}{Sup}\left\|\sum_{i=1}^{\gamma_1}\widehat{\Lambda}_i^{(1)}-\sum_{i=1}^{\gamma_1}\Lambda_i^{(1)}\right\|$$

$$\leq \sum_{i=1}^{\gamma_1}\underset{D}{Sup}\left\|\widehat{\Lambda}_i^{(1)}-\Lambda_i^{(1)}\right\| \qquad (2.41)$$

For fixed i, $\Lambda_i^{(1)}$ is nothing but Λ_{10} of Dabrowska(1988). Therefore, the rest of the proof follows from the equation (2.41) and Lemma 4.1 of Dabrowska (1988). The proof for $\widehat{\Lambda}_2$ is similar.

Theorem 2.3

Under the assumptions of Theorem 2.2, $\underset{D}{Sup}\left\|\hat{S}_a-S\right\|\to 0$ almost surely.

Proof

From (2.38), we can have

$$\underset{D}{Sup}\left\|\hat{S}_a-S\right\| \leq \underset{D}{Sup}\left\|\hat{S}_1-S\right\| + \underset{D}{Sup}\left\|\hat{S}_2-S\right\|. \qquad (2.42)$$

From (2.4) and (2.36) we have

$$\underset{D}{Sup}\left\|\log \hat{S}_1(t_1,t_2)-\log S(t_1,t_2)\right\| = \underset{D}{Sup}\left\|\hat{\Lambda}_1(t_1,0)-\Lambda_1(t_1,0)+\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right\|$$

$$\leq \underset{D}{Sup}\left\|\hat{\Lambda}_1(t_1,0)-\Lambda_1(t_1,0)\right\|+\underset{D}{Sup}\left\|\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right\| \qquad (2.43)$$

From Theorem 2.2, (2.43) implies that $\underset{D}{Sup}\,\|\log \hat{S}_1-\log S\|\to 0$ almost surely which shows that $\underset{D}{Sup}\left\|\hat{S}_1-S\right\|\to 0$ almost surely. The proof for $\hat{S}_2$ is analogous. Thus (2.42) implies that $\underset{D}{Sup}\left\|\hat{S}_a-S\right\|\to 0$ almost surely.

Theorem 2.4

Under the assumptions of Theorem 4.2 $\underset{D}{Sup}\left\|\hat{F}_{ij}^{(1)}-F_{ij}\right\|\to 0$ almost surely for every $i=1,2\ldots\ \gamma_1$ and $j=1,2\ldots\ \gamma_2$.

Proof

From (2.32) and (2.39),

$$\underset{D}{Sup}\left\|\hat{F}_{ij}^{(1)}-F_{ij}\right\| = \underset{D}{Sup}\left\|\int_0^{t_1}\int_0^{t_2}\left(\hat{G}^{-1}d^2\hat{F}_{ij}^{*}-G^{-1}d^2F_{ij}^{*}\right)\right\|$$

$$\leq \int_0^{t_1}\int_0^{t_2}\underset{D}{Sup}\left\|\frac{\left(Gd^2\hat{F}_{ij}^{*}-\hat{G}d^2F_{ij}^{*}\right)}{\hat{G}G}\right\| \qquad (2.44)$$

where $d^2F_{ij}^{*}(t_1,t_2)=F_{ij}^{*}(dt_1,dt_2)$.

Since $\underset{D}{Sup}\left\|\hat{G}-G\right\|\to 0$ (see Horvath, 1983) and $0<G(t_1,t_2)<1$, (2.44) reduces to

$$\underset{D}{Sup}\left\|\hat{F}_{ij}^{(1)}-F_{ij}\right\| \leq \int_0^{t_1}\int_0^{t_2}\left[\frac{1}{G^2}\underset{D}{Sup}\left\|G\left(d^2\hat{F}_{ij}-d^2F_{ij}^{*}\right)\right\|+\frac{1}{G^2}\underset{D}{Sup}\left\|\left(G-\hat{G}\right)d^2F_{ij}^{*}\right\|\right]. \qquad (2.45)$$

Since $\underset{D}{Sup}\left\|\hat{F}_{ij}^{(1)}-F_{ij}\right\|\to 0$ and $\underset{D}{Sup}\left\|\hat{G}-G\right\|\to 0$, (2.45) becomes,

$$\underset{D}{Sup}\left\|\widehat{F}_{ij}^{(1)}-F_{ij}\right\| \le \int_0^{t_1}\int_0^{t_2}\frac{1}{G}\underset{D}{Sup}\left\|d^2\widehat{F}_{ij}^{\bullet}-d^2F_{ij}^{\bullet}\right\|+\int_0^{t_1}\int_0^{t_2}\frac{1}{G^2}\underset{D}{Sup}\left\|G-\widehat{G}\right\|d^2F^{\bullet}$$

which converges to zero with probability one. This completes the proof.

Theorem 2.5

Under the assumptions of Theorem 4.2 $\underset{D}{Sup}\left\|\widehat{F}_{ij}^{(2)}-F_{ij}\right\| \to 0$ almost surely for every i =1,2... γ_1 and j =1,2... γ_2.

Proof

The proof follows as in Theorem 2.4 since $H(t_1,t_2)=G(t_1,t_2)S(t_1,t_2)$, $\underset{D}{Sup}\left\|\hat{S}_a-S\right\|\to 0$ and $\underset{D}{Sup}\left\|\widehat{H}-H\right\|\to 0$.

Theorem 2.6

Under the assumptions of Theorem 2.2, for all $(t_1,t_2)\in D$, $\sqrt{n}\left(\hat{S}_a(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance given by (2.47).

Proof

Let $S_i(t_i\mid t_j)$ be the conditional survivor function of T_i given $T_j>t_j$, i, j=1,2, $i\neq j$ Let $S_i(t_i)$ be the marginal survivor function of T_i i,=1,2. Since $S(t_1,t_2)=S_i(t_i\mid t_j)\,S_i(t_i)$; i, j=1,2, $i\neq j$, $S_1(t_1)=\exp\{-\Lambda_1(t_1,0)\}$ and $S_2(t_2)=\exp\{-\Lambda_2(0,t_2)\}$, we have

$\hat{S}_1(t_1,t_2)=\hat{S}_2(t_2\mid t_1)\hat{S}_1(t_1)$ and $\hat{S}_2(t_1,t_2)=\hat{S}_1(t_1\mid t_2)\hat{S}_2(t_2)$

where $\hat{S}_i(t_i\mid t_j)=\exp\{-\hat{\Lambda}_i(t_1,t_2)\}$; $i,j=1,2, i\neq j$, $\hat{S}_1(t_1)=\exp\{-\hat{\Lambda}_1(t_1,0)\}$

and $\hat{S}_2(t_2)=\exp\{-\hat{\Lambda}_2(0,t_2)\}$

$$\sqrt{n}\left(\hat{S}_1(t_1,t_2)-S(t_1,t_2)\right)=\sqrt{n}\left(\hat{S}_2(t_2\mid t_1)-S_2(t_2\mid t_1)\right)\hat{S}_1(t_1)+$$

$$\sqrt{n}\,S_2(t_2\mid t_1)\left(\hat{S}_1(t_1)-S_1(t_1)\right) \tag{2.46}$$

Now,

$$\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)=\sum_{j=1}^{\gamma_2}\int_0^{t_2}\frac{I(Y_1>t_1,Y_2>u)}{\hat{H}(t_1,u)}dM_{2j}(t_1,u)$$

where for fixed t_1, $M_{2j}(t_1,u)=\hat{H}_{2j}^{uc}(t_1,u)-\int_0^{t_2}\hat{H}(t_1,u)d\Lambda_j^{(2)}(t_1,u)$ is the martingale representation of $\Lambda_j^{(2)}(t_1,u)$ with $\hat{H}_{2j}^{uc}(t_1,u)$ is the estimator of $H_{2j}^{uc}(t_1,u)$ = $P(Y_1>t_1,Y_2\geq u,\delta_2=1,C_2=j)$. Then for fixed t_1, $\sqrt{n}\left(\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right)$ converges to a Gaussian process with mean zero. Further, for fixed t_1, $S_2(t_2\mid t_1)$ = $\prod_{s\leq t_2}\left(1-d\Lambda_2(t_1,s)\right)$ is the Hadamard differentiability of the product integral. Thus the asymptotic normality of $\hat{\Lambda}_2(t_1,t_2)$ carries over to the asymptotic normality of the corresponding estimator of $S_2(t_2\mid t_1)$. This shows that $\sqrt{n}\left(\hat{S}_2(t_2\mid t_1)-S_2(t_2\mid t_1)\right)$ converges to a Gaussian process with mean zero. Since $\hat{S}_1(t_1)$ is strongly consistent and $\sqrt{n}\left(\hat{S}_2(t_2\mid t_1)-S_2(t_2\mid t_1)\right)$converges weakly to a mean zero Gaussian process, the first factor of (2.46) weakly converges to a Gaussian process. Since $S_2(t_2\mid t_1)$ is bounded and $\sqrt{n}\left(\hat{S}_1(t_1)-S_1(t_1)\right)$ converges weakly, the second factor of (2.46) converges weakly to a Gaussian process. Thus, $\sqrt{n}\left(\hat{S}_1(t_1,t_2)-S(t_1,t_2)\right)$ weakly converges to a Gaussian process. On similar lines we can show that $\sqrt{n}\left(\hat{S}_2(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero. Thus, $\sqrt{n}\left(\hat{S}_a(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance of $\sqrt{n}\,\hat{S}_a(t_1,t_2)$ is

$$\sigma^2(t_1,t_2) = a^2(t_1,t_2)\sigma^2{}_1(t_1,t_2) + (1-a(t_1,t_2))^2\sigma^2{}_2(t_1,t_2)$$
$$+a(t_1,t_2)(1-a(t_1,t_2))\sigma_{12}(t_1,t_2) \qquad (2.47)$$

where $\sigma_k^2(t_1,t_2)$ is the asymptotic variance of $\sqrt{n}\ \hat{S}_k(t_1,t_2)$; k = 1,2 and $\sigma_{12}(t_1,t_2)$ is the asymptotic covariance between $\sqrt{n}\ \hat{S}_1(t_1,t_2)$ and $\sqrt{n}\ \hat{S}_2(t_1,t_2)$.From Appendix B of Lawless(2003, page 539), the asymptotic variance of $\sqrt{n}\ \hat{S}_1(t_1,t_2)$ is

$\sigma^2{}_1(t_1,t_2) = S_1^2(t_1,t_2)\ As\ \mathrm{var}\left(\log\hat{S}_1(t_1,t_2)\right)$ Now,

$$\sqrt{n}\left(\log\hat{S}_1(t_1,t_2) - \log S(t_1,t_2)\right) = \sqrt{n}\left(\hat{\Lambda}_1(t_1,0) - \Lambda_1(t_1,0) + \hat{\Lambda}_2(t_1,t_2) - \Lambda_2(t_1,t_2)\right).$$

Thus the asymptotic variance of $\sqrt{n}\left(\log\hat{S}_1(t_1,t_2) - \log S(t_1,t_2)\right)$ is

$$E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{H(u_1^-,0)}I(Y_{1u}\in du_1,\delta_{1u}=1,C_{1u}=i) + I(Y_{1u}\ge u_1)\frac{p_{1i}(du_1,0)}{H^2(u_1^-,0)}\right\}\right] +\right.$$
$$\left.\sum_{j=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{H(t_1,u_2^-)}I(Y_{1u}>t_1,Y_{2u}\in du_2,\delta_{2u}=1,C_{2u}=i) +\right.\right.\right.$$
$$\left.\left.\left. I(Y_{1u}>t_1,Y_{2u}\ge u_2)\frac{p_{2j}(t_1,du_2)}{H^2(t_1,u_2^-)}\right\}\right]\right)^2$$

Thus,

$$\sigma^2{}_1(t_1,t_2) = S_1^2(t_1,t_2)\ E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{H(u_1^-,0)}I(Y_{1u}\in du_1,\delta_{1u}=1,C_{1u}=i)\right.\right.\right.$$
$$\left.\left.\left. + I(Y_{1u}\ge u_1)\frac{p_{1i}(du_1,0)}{H^2(u_1^-,0)}\right\}\right] + \sum_{j=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{H(t_1,u_2^-)}I(Y_{1u}>t_1,Y_{2u}\in du_2,\delta_{2u}=1,C_{2u}=i)\right.\right.\right.$$
$$\left.\left.\left. + I(Y_{1u}>t_1,Y_{2u}\ge u_2)\frac{p_{2j}(t_1,du_2)}{H^2(t_1,u_2^-)}\right\}\right]\right)^2$$

Similarly the asymptotic variance of $\sqrt{n}\left(\log\hat{S}_2(t_1,t_2) - \log S(t_1,t_2)\right)$ is

$$\sigma^2{}_2(t_1,t_2) = S_2^2(t_1,t_2)\, E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{H(u_1^-,t_2)} I(Y_{1u}\in du_1, Y_{2u}>t_2, \delta_{1u}=1, C_{1u}=i)\right.\right.\right.$$

$$\left.\left.\left. + I(Y_{1u}\ge u_1, Y_{2u}>t_2)\frac{p_{1i}(du_1,t_2)}{H^2(u_1^-,t_2)}\right\}\right] + \sum_{i=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{H(0,u_2^-)} I(Y_{2u}\in du_2, \delta_{2u}=1, C_{2u}=j)\right.\right.\right.$$

$$\left.\left.\left. + I(Y_{2u}\ge u_2)\frac{p_{2i}(0,du_2)}{H^2(0,u_2^-)}\right\}\right]\right)^2$$

The asymptotic covariance between $\hat{S}_1(t_1,t_2)$ and $\hat{S}_2(t_1,t_2)$ is

$$\sigma_{12}(t_1,t_2) = S_1(t_1,t_2)\, S_2(t_1,t_2)\ As\,\mathrm{cov}\left(\log\hat{S}_1(t_1,t_2), \log\hat{S}_2(t_1,t_2)\right),$$

which is given by

$$\sigma_{12}(t_1,t_2) = n\, S_1(t_1,t_2)\, S_2(t_1,t_2)\left\{E\left[\left(\hat{\Lambda}_1(t_1,t_2)-\Lambda_1(t_1,t_2)\right)\left(\hat{\Lambda}_1(t_1,0)-\Lambda_1(t_1,0)\right)+\right.\right.$$

$$\left(\hat{\Lambda}_1(t_1,t_2)-\Lambda_1(t_1,t_2)\right)\left(\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right)+\left(\hat{\Lambda}_2(0,t_2)-\Lambda_2(0,t_2)\right)\left(\hat{\Lambda}_1(t_1,0)-\Lambda_1(t_1,0)\right)$$

$$\left.\left.+\left(\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right)\left(\hat{\Lambda}_2(0,t_2)-\Lambda_2(0,t_2)\right)\right]\right\}$$

Theorem 2.7

Under the assumptions of Theorem 2.2, for all $(t_1,t_2)\in D$ $\sqrt{n}\left(\hat{F}_{ij}^{(1)}(t_1,t_2)-F_{ij}(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance given by (2.52).

Proof

$$\sqrt{n}\left(\hat{F}_{ij}^{(1)}(t_1,t_2)-F_{ij}(t_1,t_2)\right) = \sqrt{n}\int_0^{t_1}\int_0^{t_2}\left(\frac{\hat{F}_{ij}^{*}(du_1,du_2)}{\hat{G}(u_1^-,u_2^-)}-\frac{F_{ij}^{*}(du_1,du_2)}{G(u_1^-,u_2^-)}\right)$$

$$= \sqrt{n}\int_0^{t_1}\int_0^{t_2}\left(\frac{\hat{F}_{ij}(du_1,du_2)}{\hat{G}(u_1^-,u_2^-)}-\frac{\hat{F}_{ij}(du_1,du_2)}{G(u_1^-,u_2^-)}\right.$$

$$\left.+\frac{\hat{F}_{ij}(du_1,du_2)}{G(u_1^-,u_2^-)}-\frac{F_{ij}^{*}(du_1,du_2)}{G(u_1^-,u_2^-)}\right). \qquad (2.48)$$

For large n, using (2.1), (2.48) can be approximated by

$$\sqrt{n}\int_0^{t_1}\int_0^{t_2}\frac{1}{G(u_1^-,u_2)}\left(\frac{1}{n}\sum_{l=1}^{n}I(T_{1l}\in du_1,T_{2l}\in du_2,\delta_{1l}=1,\delta_{2l}=1,C_{1l}=i,C_{2l}=j)-F_{ij}^{*}(du_1,du_2)\right)$$

$$-\sqrt{n}\int_0^{t_1}\int_0^{t_2}\frac{F_{ij}^{*}(du_1,du_2)}{G^2(u_1^-,u_2^-)}\left(\frac{1}{n}\sum_{l=1}^{n}I(Z_{1l}\geq u_1,Z_{2l}\geq u_2)-G(u_1,u_2)\right) \tag{2.49}$$

By multivariate central limit theorem, the terms in the simple brackets of the first and second integrals converges to a mean zero normal variate. Then the asymptotic normality of (2.49) follows from the Delta method using the maps $(x_1,x_2)\to(x_1+x_2)$ (van der Vaart and Wellner(1996)).

To find the asymptotic variance, let

$$A=\int_0^{t_1}\int_0^{t_2}\frac{1}{G(u_1^-,u_2^-)}I(T_{1l}\in du_1,T_{2l}\in du_2,\delta_{1l}=1,\delta_{2l}=1,C_{1l}=i,C_{2l}=j) \tag{2.50}$$

and

$$B=\int_0^{t_1}\int_0^{t_2}\frac{F_{ij}^{*}(du_1,du_2)}{G^2(u_1^-,u_2^-)}I(Z_{1l}\geq u_1,Z_{2l}\geq u_2). \tag{2.51}$$

Thus, the asymptotic variance of the process is

$$\sigma^{*2}(t_1,t_2)=E(A-B)^2 \tag{2.52}$$

Therefore, a consistent estimator of variance is given by

$$\hat{\sigma}^{*2}(t_1,t_2)=\frac{1}{n}\sum\left(\hat{A}-\hat{B}\right)^2$$

where $\hat{A}$ and $\hat{B}$ are obtained using (2.50) and (2.51) replacing the unknown quantities by their estimators.

Theorem 2.8

Under the assumptions of Theorem 2.2, for all $(t_1,t_2)\in D$ $\sqrt{n}\left(\hat{F}_{ij}^{(2)}(t_1,t_2)-F_{ij}(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and variance given by

$\sigma^{2}(t_1,t_2)=E(P-Q)^2$ where

$$P=\int_0^{t_1}\int_0^{t_2}\frac{S(u_1^-,u_2^-)}{H(u_1,u_2^-)}I(T_{1l}\in du_1,T_{2l}\in du_2,\delta_{1l}=\delta_{2l}=1,C_{1l}=i,C_{2l}=j) \tag{2.53}$$

and

$$Q=\int_0^{t_1}\int_0^{t_2}\frac{S(u_1^-,u_2)F_{il}^*(du_1,du_2)}{\left(H(u_1^-,u_2^-)\right)^2}I(Y_{1l}>u_1,Y_{2l}>u_2). \tag{2.54}$$

A consistent estimator of variance is

$$\hat{\sigma}^{'2}(t_1,t_2)=\frac{1}{n}\sum\left(\hat{P}-\hat{Q}\right)^2$$

where $\hat{P}$ and $\hat{Q}$ are obtained using (2.53) and (2.54) replacing the unknown quantities by their estimators.

The proof is similar to that of Theorem 2.7.

2.5 Data Analysis

We illustrate the estimation procedure given in Section 2.3 using a real data. Table 2.1 shows data concerning the times to tumour appearance or death for 50 pairs of mice from the same litter in a tumor genesis experiment (Mantel and Ciminera, 1979), as reported in Ying and Wei (1994). In this data, T_1 and T_2 represent failure times (in weeks) for a pair of mice, and C_j (j = 1,2) indicates whether the failure was the appearance of a tumour (C_j = 1) or the occurrence of death prior to tumour appearance (C_j = 2). The censored observations are denoted by C_j = 0. The experiment was terminated at 104 weeks, so there is a common censoring time across all animals of 104.

The estimators $\hat{S}_1(t_1,t_2)$ and $\hat{S}_2(t_1,t_2)$ can be obtained directly from the data using the approach in Section 2.3.1. To obtain $a(t_1,t_2)$, we use the extension of Efron (1981)'s bootstrap procedure for one-dimensional censored data. Given the data $(T_{1u},T_{2u},\Delta_{1u},\Delta_{2u},C_{1u},C_{2u})$, u =1,2...n, where Δ_{1u} and Δ_{2u} are the censoring indicator, we generate the bootstrap data $(T_{1u}^*,T_{2u}^*,\Delta_{1u}^*,\Delta_{2u}^*,C_{1u}^*,C_{2u}^*)$, u =1,2...n from the

distribution function $\frac{1}{n}\sum_{u=1}^{n} I\left(T_{1u}^{*} \le t_1, T_{2u} \le t_2, \Delta_{1u}^{*} = \delta_{1u}, \Delta_{2u}^{*} = \delta_{2u}, C_{1u}^{*} = i, C_{2u} = j\right)$

We take 1000 such re-samples and find $\hat{S}_1(t_1,t_2)$ and $\hat{S}_2(t_1,t_2)$ as explained in Section 2.3.1, which we denote by $S_{1j}^{*}(t_1,t_2)$ and $S_{2j}^{*}(t_1,t_2)$, j = 1,2...1000. Since the biases are negligible as shown in Section 2.6, we find the weight $a(t_1,t_2)$ as

$$a(t_1,t_2) = \frac{\mathrm{var}\left(S_2^{*}(t_1,t_2)\right) - \mathrm{cov}\left(S_1^{*}(t_1,t_2), S_2^{*}(t_1,t_2)\right)}{\mathrm{var}\left(S_1^{*}(t_1,t_2)\right) + \mathrm{var}\left(S_2^{*}(t_1,t_2)\right) - 2\,\mathrm{cov}\left(S_1^{*}(t_1,t_2), S_2^{*}(t_1,t_2)\right)}.$$

Then the estimator of the survivor function $\hat{S}_a(t_1,t_2)$ at different time points (55,90), (81,66), (94,91) and (73,74) is obtained using (2.38). The estimator of bivariate survivor function due to Dabrowska (1988), $\hat{S}(t_1,t_2)$, is also found at the same time points for comparison. The values of $a(t_1,t_2)$, $\hat{S}_a(t_1,t_2)$, $\hat{S}(t_1,t_2)$ and their variances, given in brackets, are presented in Table 2.2. It can be observed that the variance of $\hat{S}_a(t_1,t_2)$ is lesser than that of $\hat{S}(t_1,t_2)$ except at (55,90).

Table 2.3 gives the estimator $\hat{F}_{ij}(t_1,t_2)$ at all the observed points (t_1,t_2) and the plots of $\hat{F}_{ij}(t_1,t_2)$ for i,j = 1,2, is shown in Figures 2.1, 2.2, 2.3 and 2.4. It is to be noted that since all the units have a common censoring time, the two estimators in (2.39) and (2.40) are identical and simply equal to the observed fraction of pairs with $\left(Y_{1u} \le t_1, Y_{2u} \le t_2, C_{1u} = i, C_{2u} = j\right)$. From Table 2.3, it follows that the probability of failure due to different causes at different time points is uniform.

We estimate $\pi_{ij} = P\left(C_1 = i, C_2 = j\right)$ as $\hat{\pi}_{ij} = \hat{F}_{ij}(\infty,\infty)$. Since the last pair is a censored one, we normalize the estimate as $\hat{\pi}_{ij} = \dfrac{\hat{\pi}_{ij}}{\sum_k \sum_l \hat{\pi}_{kl}}$ We also estimated the marginal probabilities $\pi_i^{(1)} = P\left(C_1 = i\right)$ i, j = 1,2 using the estimate of the marginal sub distribution function. The estimates $\hat{\pi}_i^{(1)^{*}}$ and $\hat{\pi}_{ij}$ are given in Table 2.4 from which it follows that the two causes are not independent.

Table 2.1: Data concerning the times to tumour appearance or death for 50 pairs of mice

T_1	C_1	T_2	C_2	T_1	C_1	T_2	C_2
49	1	104*	0	104*	0	104*	0
102	2	104*	0	104*	0	104*	0
104*	0	104*	0	81	1	64	1
97	2	79	2	55	1	94	2
104*	0	104*	0	104*	0	54	1
96	1	104*	0	87	2	74	2
94	2	77	1	73	1	84	1
104*	0	104*	0	104*	0	83	2
77	2	104*	0	104*	0	73	2
104*	0	77	2	79	2	104*	0
91	2	90	2	104*	0	104*	0
70	2	92	2	104*	0	104*	0
45	2	50	1	101	1	94	2
69	2	91	2	84	1	78	1
104*	0	103	2	81	1	76	2
72	2	104*	0	95	2	104*	0
63	2	104*	0	104*	0	66	1
104*	0	74	2	104*	0	102	1
104*	0	69	2	98	2	73	2
104*	0	68	1	104*	0	104*	0
104*	0	104*	0	83	2	77	2
104*	0	104*	0	104*	0	104*	0
83	2	40	1	79	2	99	2
104*	0	104*	0	91	2	104*	0
104*	0	104*	0	104*	0	79	1

(* indicates censored time)

Table 2.2: Estimates of the survivor function $S(t_1,t_2)$

(t_1,t_2)	(55,90)	(81,66)	(94,91)	(73,74)
$a(t_1,t_2)$	.827674	.626292	.734062	1
$\hat{S}_a(t_1,t_2)$	.415157 (.00128)	.568151 (.01051)	.268659 (.000367)	.568186 (.005874)
$\hat{S}(t_1,t_2)$	.422768 (.00108)	.519525 (.01311)	.261610 (.000704)	.567488 (.00594)

Table 2.3: Estimates of the distribution function $F_{11}(t_1,t_2)$

T_1	C_1	T_2	C_2	$\hat{F}_{11}(t_1,t_2)$
45	2	50	1	.02
55	1	94	2	.04
69	2	91	2	.04
70	2	92	2	.06
73	1	84	1	.04
79	2	99	2	.12
81	1	64	1	.04
81	1	76	2	.06
83	2	40	1	.02
83	2	77	2	.10
84	1	78	1	.12
87	2	74	2	.08
91	2	90	2	.18
94	2	77	1	.14
97	2	79	2	.18
98	2	73	2	.08
101	1	99	2	.32

Table 2.4: Estimates of $\pi_i^{(j)}$ and π_{ij}.

$\hat{\pi}_1^{(1)*}$	$\hat{\pi}_2^{(1)*}$	$\hat{\pi}_1^{(2)*}$	$\hat{\pi}_2^{(2)*}$	$\hat{\pi}_{11}^{*}$	$\hat{\pi}_{12}^{*}$	$\hat{\pi}_{21}^{*}$	$\hat{\pi}_{22}^{*}$
$\frac{4}{13}$	$\frac{9}{13}$	$\frac{11}{28}$	$\frac{17}{28}$	$\frac{3}{17}$	$\frac{3}{17}$	$\frac{3}{17}$	$\frac{8}{17}$

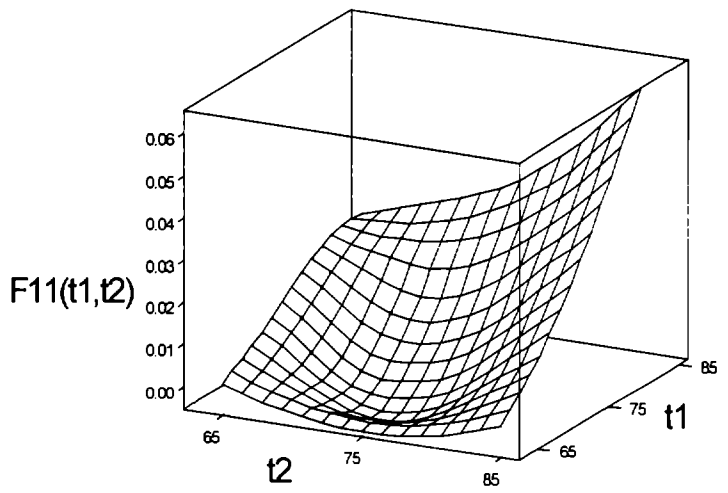

Fig 2.1: Estimator of cause-specific distribution function for the data on times to tumour appearance or death of 50 pair of mice corresponding to the pair of cause (1,1).

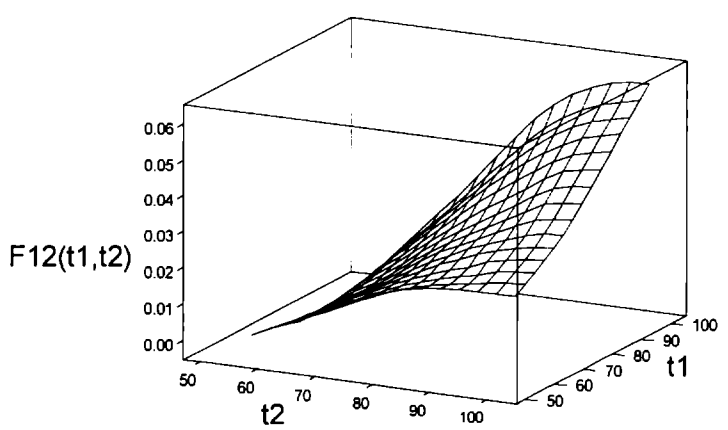

Fig 2.2: Estimator of cause-specific distribution function for the data on times to tumour appearance or death of 50 pair of mice corresponding to the pair of cause (1,2).

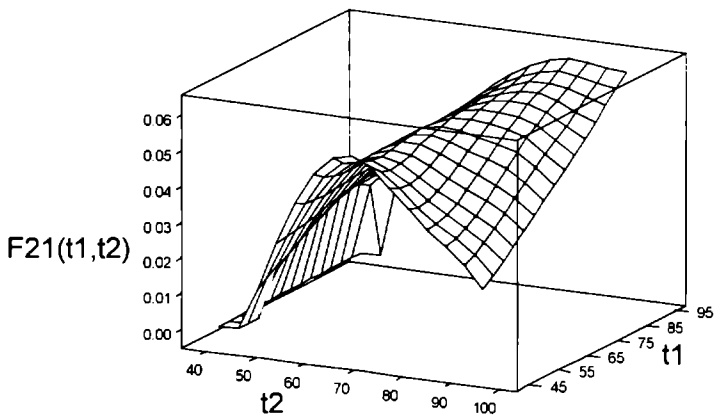

Fig 2.3: Estimator of cause-specific distribution function for the data on times to tumour appearance or death of 50 pair of mice corresponding to the pair of cause (2,1).

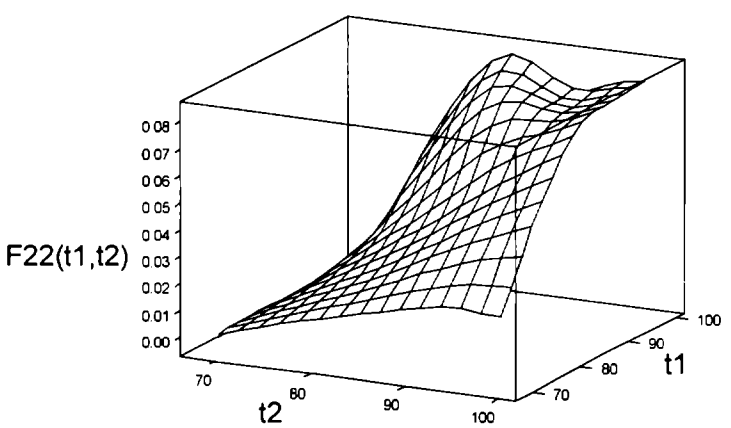

Fig 2.4: Estimator of cause-specific distribution function for the data on times to tumour appearance or death of 50 pair of mice corresponding to the pair of cause (2,2).

2.6 A Simulation Study

To assess the empirical performance of the estimator of $S(t_1,t_2)$, we carried out a series of 1000 simulations of size *n* from a bivariate Dirichlet Distribution with survivor function

$$S(t_1,t_2)=(1-t_1-t_2)^{\alpha+1} \quad 0<t_1,t_2<t_1+t_2<1,$$

with α = 2,3 and 4 for various sample sizes *n*= 20,50 and 100. We also generated censoring times (Z_1,Z_2) from the bivariate Dirichlet distribution with survivor function

$$G(t_1,t_2)=(1-t_1-t_2)^{\beta+1} \quad 0<t_1,t_2<t_1+t_2<1$$

with β = 4,5 and 10 so that the censoring times are random. We used the algorithm given in Gentle (1998, pages 94 and 111) for generating the observations. We considered two types of causes C_i, i = 1,2.The causes 1 and 2 are distributed randomly among the observed failure times with equal probability. We found the estimator of $S(t_1,t_2)$ using the equation (2.38) at five time points. The estimator of bivariate survivor function, $\hat{S}(t_1,t_2)$, given in Dabrowska (1988) is also calculated for the same time points. A careful examination of the tables reveals that biases and variances decrease with increasing sample size, as is expected. Further, the biases due to the estimator of Dabrowska (1988) have negative values in most of the cases and those due to the proposed one have positive values.

For empirical studies on the performance of the estimators of $F_{ij}(t_1,t_2)$, which is more important in the competing risk set up, we carried out a series of 1000 simulations of size *n* from

(*a*) a bivariate Dirichlet distribution with survivor function

$$S(t_1,t_2)=(1-t_1-t_2)^{\alpha+1} \quad 0<t_1,t_2<t_1+t_2<1$$

with α = 2,3 and 4 and for

(*b*) a bivariate Gumbel's (1960) exponential distribution with survivor function

$$S(t_1,t_2)=\exp\{-t_1-t_2-\lambda t_1 t_2\},\ 0<t_1,t_2<\infty,\ 0\le\lambda\le 1$$

with $\lambda = 0.5$ and 1 for various sample sizes $n = 20$, 50 and 100.

We generated censoring times (Z_1, Z_2) for case (a) from the bivariate Dirichlet distribution

$$G(t_1,t_2)=(1-t_1-t_2)^{\beta+1} \quad 0<t_1,t_2<t_1+t_2<1$$

with $\beta = 4$ .5 and 10 so that the censoring times are random. We used the algorithm given in Gentle (1998, pages 94 and 111) for generating the Dirichlet observations.

For case (b), censoring observations are generated from the Gumbel's (1960) exponential distribution with survivor function

$$G(t_1,t_2)=\exp\{-t_1-t_2-\theta t_1 t_2\} \quad 0<t_1,t_2<\infty,\ 0\le\theta\le 1$$

with $\theta = 0.6$ and 0.9 so that the censoring times are random. We used the algorithm given in Devroye (1986) for generating the observations from Gumbel's distribution. We considered two types of causes for C_i, $i = 1,2$. The causes 1 and 2 are distributed randomly among the observed failure times with equal probability. The estimator of bivariate survivor function, $\hat{S}(t_1,t_2)$, is calculated as explained in Section 2.3.1 and the estimator given in Burke (1988) is used to find $\hat{G}(t_1,t_2)$. The estimators of $F_{ij}(t_1,t_2)$ are computed based on the equations (2.39) and (2.40). The empirical biases and empirical variances of the estimators are given in Tables 2.5, 2.6, 2.7, 2.8 and 2.9. A careful examination of the tables reveals the following patterns.

a. The estimate $\hat{F}_{ij}^{(2)}$ has small biases and small variances compared to the estimate $\hat{F}_{ij}^{(1)}$ for all $i, j = 1,2$.

b. The estimate $\hat{F}_{ij}^{(1)}$ has positive biases for all values of (t_1,t_2) and n. However, the biases of $\hat{F}_{ij}^{(2)}$ are negative for most of the values of (t_1,t_2).

c. For both the estimators, the bias and variance decreases with increasing sample size.

Table 2.5:Biases ($\times 10^3$) and Variances ($\times 10^3$), given in brackets, of $\hat{S}_a(t_1,t_2)$ and $\hat{S}(t_1,t_2)$ at five (t_1,t_2) pairs for bivariate Dirichlet distribution for various α and β

n	α	β	(t_1,t_2)	Bias and variance of $\hat{S}_a(t_1,t_2)$	Bias and variance of $\hat{S}(t_1,t_2)$
20	2	5	(.102,.0890)	29(22)	-11(3)
			(.077,.0990)	29(21)	-12(3)
			(.054,.0897)	40(27)	-11(3)
			(.079,.1228)	31(21)	-9(3)
			(.077,.1050)	50(27)	-12(3)
50	2	5	(.102,.0890)	13(9)	-6(7)
			(.077,.0990)	11(9)	-5(7)
			(.054,.0897)	14(12)	.9(6)
			(.079,.1228)	14(9)	-9(7)
			(.077,.1050)	14(13)	-5(5)
100	2	5	(.102,.0890)	5(5)	-5(17)
			(.077,.0990)	2(4)	-4(16)
			(.054,.0897)	3(5)	-10(16)
			(.079,.1228)	7(6)	-8(16)
			(.077,.1050)	6(5)	-9(15)
20	3	10	(.102,.0890)	14(16)	-20(5)
			(.077,.0990)	24(25)	-22(5)
			(.054,.0897)	13(19)	-19(4)
			(.079,.1228)	17(20)	-21(5)
			(.077,.1050)	21(20)	-23(5)
50	3	10	(.102,.0890)	2(6)	-15(9)
			(.077,.0990)	9(11)	-15(9)
			(.054,.0897)	5(7)	-11(8)
			(.079,.1228)	5(9)	-18(10)
			(.077,.1050)	5(9)	17(10)
100	3	10	(.102,.0890)	1(3)	-.6(23)
			(.077,.0990)	4(5)	-3(22)
			(.054,.0897)	4(4)	-2(23)
			(.079,.1228)	1(4)	1(25)
			(.077,.1050)	5(4)	-1(23)
20	4	4	(.102,.0890)	18(15)	-44(3)
			(.077,.0990)	27(19)	-40(3)
			(.054,.0897)	24(18)	-33(3)
			(.079,.1228)	38(19)	-45(3)
			(.077,.1050)	16(15)	-43(32)
50	4	4	(.102,.0890)	12(7)	-37(6)
			(.077,.0990)	14(8)	-35(6)
			(.054,.0897)	13(7)	-28(6)
			(.079,.1228)	21(8)	-34(6)
			(.077,.1050)	13(6)	-35(6)
100	4	4	(.102,.0890)	3(3)	-20(13)
			(.077,.0990)	5(4)	-22(13)
			(.054,.0897)	6(4)	-21(14)
			(.079,.1228)	5(3)	-20(13)
			(.077,.1050)	1(3)	-21(13)

Table 2.6: Biases ($\times 10^4$) and Variances ($\times 10^4$), given in brackets, of $\widehat{F}_{ij}^{(1)}(t_1,t_2)$ at three (t_1,t_2) pairs for bivariate Dirichlet distribution

(t_1,t_2)	n	α	β	$\widehat{F}_{11}$	$\widehat{F}_{12}$	$\widehat{F}_{21}$	$\widehat{F}_{22}$
	20	2	5	85 (38)	84(46)	93(33)	92(38)
(.07, .08)	50	2	5	66(9)	77(11)	60(8)	76(10)
	100	2	5	55(3)	63(4)	53(4)	75(4)
	20	2	5	44(23)	51(25)	62(19)	54(16)
(.20, .20)	50	2	5	41(5)	39(4)	36(4)	50(5)
	100	2	5	37(2)	38(2)	29(1)	42(2)
	20	2	5	146(76)	119(65)	211(117)	173(89)
(.14, .12)	50	2	5	128(25)	115(32)	123(25)	124(24)
	100	2	5	113(9)	104(7)	105(4)	120(9)
	20	3	10	32(12)	50(18)	47(10)	44(18)
(.07, .08)	50	3	10	29(4)	41(5)	41(5)	36(4)
	100	3	10	23(2)	27(2)	31(2)	26(2)
	20	3	10	121(87)	128(92)	98(44)	123(80)
(.20, .20)	50	3	10	86(16)	107(21)	92(19)	101(24)
	100	3	10	83(6)	80(7)	78(7)	82(7)
	20	3	10	48(19)	75(42)	58(20)	62(26)
(.14, .12)	50	3	10	47(7)	64(9)	54(7)	55(8)
	100	3	10	43(4)	42(3)	45(4)	50(4)
	20	4	4	157(107)	152(128)	142(111)	159(110)
(.07, .08)	50	4	4	150(69)	149(65)	141(75)	157(72)
	100	4	4	141(25)	139(33)	133(23)	133(27)
	20	4	4	135(94)	142(110)	115(105)	133(88)
(.20, .20)	50	4	4	130(55)	130(56)	122(60)	125(59)
	100	4	4	126(21)	118(23)	111(18)	114(20)
	20	4	4	86(68)	87(76)	85(77)	78(66)
(.14, .12)	50	4	4	84(31)	81(35)	80(34)	74(37)
	100	4	4	83(11)	79(13)	73(10)	73(10)

Table 2.7: Biases ($\times 10^4$) and Variances ($\times 10^4$), given in brackets, of $\widehat{F}_{ij}^{(2)}(t_1,t_2)$ at three (t_1,t_2) pairs for bivariate Dirichlet distribution

(t_1,t_2)	n	α	β	$\widehat{F}_{11}$	$\widehat{F}_{12}$	$\widehat{F}_{21}$	$\widehat{F}_{22}$
	20	2	5	-15 (6)	-30(5)	-5(5)	-25(5)
(.07, .08)	50	2	5	-8(2)	-12(2)	-7(3)	-2(3)
	100	2	5	-2(1)	-8(1)	1(1)	1(1)
	20	2	5	-17(3)	-27(9)	-11(8)	-33(6)
(.20, .20)	50	2	5	-14(3)	-13(3)	-9(4)	-6(4)
	100	2	5	-2(2)	-2(2)	-1(2)	-1(2)
	20	2	5	-28(19)	-29(20)	-37(13)	-52(13)
(.14, .12)	50	2	5	-18(6)	-18(7)	-6(7)	-2(9)
	100	2	5	-5(4)	-5(4)	-2(3)	-5(4)
	20	3	10	1(4)	-8(3)	-5(4)	-3(4)
(.07, .08)	50	3	10	-1(1)	-2(2)	1(2)	7(2)
	100	3	10	-1(1)	2(1)	-1(1)	-3(1)
	20	3	10	7(4)	-5(3)	-8(3)	4(4)
(.20, .20)	50	3	10	-4(1)	-4(1)	-1(1)	4(2)
	100	3	10	-1(1)	-1(1)	-1(1)	-1(1)
	20	3	10	1(4)	-6(4)	-11(3)	-1(4)
(.14, .12)	50	3	10	1(2)	-1(2)	-3(2)	-1(2)
	100	3	10	1(1)	-1(1)	-7(1)	-1(1)
	20	4	4	-5(6)	-17(30)	-14(6)	-22(6)
(.07, .08)	50	4	4	-4(4)	-3(13)	-5(4)	-3(4)
	100	4	4	-1(3)	2(3)	-2(3)	-2(3)
	20	4	4	-12(14)	-21(14)	-30(12)	-25(15)
(.20, .20)	50	4	4	-8(7)	-7(9)	-10(9)	-18(6)
	100	4	4	-7(7)	-1(7)	-1(1)	-1(1)
	20	4	4	-1(1)	-3(1)	-1(1)	-1(1)
(.14, .12)	50	4	4	-1(1)	-1(1)	-1(1)	-1(1)
	100	4	4	1(1)	-1(1)	-1(1)	1(1)

Table 2.8: Biases ($\times 10^4$) and Variances ($\times 10^4$), given in brackets, of $\widehat{F}_{ij}^{(1)}(t_1,t_2)$ at three (t_1,t_2) pairs for Gumbel's (1960) bivariate exponential distribution

(t_1,t_2)	n	λ	θ	$\widehat{F}_{11}$	$\widehat{F}_{12}$	$\widehat{F}_{21}$	$\widehat{F}_{22}$
	20	.5	.6	-170(60)	110(71)	190(58)	210(59)
(.5, .5)	50	.5	.6	101(31)	98(39)	115(28)	155(29)
	100	.5	.6	77(16)	-30(21)	65(19)	66(17)
	20	.5	.6	160(54)	120(53)	145(61)	166(57)
(1, 2)	50	.5	.6	112(33)	90(34)	104(30)	121(33)
	100	.5	.6	68(15)	54(17)	61(18)	75(16)
	20	.5	.6	-140(49)	115(51)	151(48)	139(50)
(2, 1)	50	.5	.6	104(33)	99(37)	103(30)	98(27)
	100	.5	.6	68(16)	70(20)	81(17)	63(14)
	20	1	.9	144(48)	123(50)	139(41)	155(39)
(.5, .5)	50	1	.9	119(29)	106(27)	105(31)	116(25)
	100	1	.9	77(14)	68(13)	79(19)	63(14)
	20	1	.9	151(47)	128(44)	145(45)	158(48)
(1, 2)	50	1	.9	110(28)	99(31)	98(29)	111(26)
	100	1	.9	66(14)	58(13)	60(15)	73(11)
	20	1	.9	-120(51)	126(50)	138(49)	141(60)
(2, 1)	50	1	.9	100(39)	98(40)	101(41)	93(45)
	100	1	.9	60(26)	-40(22)	45(25)	-35(26)

Table 2.9: Biases ($\times 10^4$) and Variances ($\times 10^4$), given in brackets, of $\widehat{F}_{ij}^{(2)}(t_1,t_2)$ at three (t_1,t_2) pairs for Gumbel's (1960) bivariate exponential distribution

(t_1,t_2)	n	λ	θ	$\widehat{F}_{11}$	$\widehat{F}_{12}$	$\widehat{F}_{21}$	$\widehat{F}_{22}$
	20	.5	.6	-30(34)	-50(60)	-25(40)	-36(52)
(.5, .5)	50	.5	.6	-20(25)	-10(40)	20(31)	-15(38)
	100	.5	.6	-8(11)	-5(13)	3(10)	5(15)
	20	.5	.6	-45(37)	-60(38)	-47(41)	-41(44)
(1, 2)	50	.5	.6	-30(26)	-50(29)	-28(38)	-23(40)
	100	.5	.6	5(15)	10(19)	-1(11)	-5(13)
	20	.5	.6	-51(52)	-58(49)	-49(53)	-44(51)
(2, 1)	50	.5	.6	-25(41)	-32(44)	-10(39)	-11(37)
	100	.5	.6	8(11)	-9(15)	5(14)	10(11)
	20	1	.9	-25(52)	-35(45)	16(39)	20(55)
(.5, .5)	50	1	.9	-17(34)	-31(39)	14(39)	18(41)
	100	1	.9	10(11)	12(19)	9(13)	8(19)
	20	1	.9	-31(49)	-37(60)	-18(65)	-25(55)
(1, 2)	50	1	.9	-19(31)	-28(58)	18(47)	-17(52)
	100	1	.9	10(18)	-11(19)	14(14)	15(17)
	20	1	.9	-52(61)	-39(70)	-45(77)	-20(71)
(2, 1)	50	1	.9	-23(49)	-21(51)	-21(58)	16(53)
	100	1	.9	17(19)	14(22)	10(16)	6(18)

2.7 Conclusion

In literature, there are different non-parametric estimators for bivariate survivor function under censoring (see van der Laan (1997)). In the present work, we developed a non-parametric estimator of the bivariate survivor function for competing risk models using cause specific hazard function. We proved the consistency and weak convergence of the estimator. The proposed estimator is compared with the well-known estimator of bivariate survivor function due to Dabrowska (1988) using a simulation study and a real data. It is found that the new estimator is at par. Then we developed two simple non-parametric estimators of

cause-specific sub-distribution function for bivariate competing risk models. Both estimators are consistent. The weak convergence of the estimators is established. We illustrated the estimation procedure with a real data. Simulation study shows that the bias and variance of the estimators are less. The extension of the method to the multivariate set up is straightforward. The results presented in this chapter are summarized in Ansa (2004) and Sankaran et.al. (2005b).

Chapter 3

BIVARIATE SURVIVOR FUNCTION UNDER MASKING

3.1 Introduction

There are situations in the analysis of competing risk lifetime data where for some of the systems or study subjects being studied, the exact failure cause cannot be identified easily. Then we say the cause of failure is masked. This may be due to inadequacy in the diagnostic mechanism or some individuals may be reluctant to report any specific failure cause (disease). Miyakawa (1984), Racine-Poon and Hoel (1984), Lo (1991) and Mukerjee and Wang (1993) considered the problem of masking with two failure types. Goetghebuer and Ryan (1990, 1995) considered the regression problem under masking using partial likelihood for two types of failure with the assumption that the cause-specific hazards for the two failure types are proportional.

In all the works mentioned above, it was assumed that no information on failure cause is available at all. But test procedures can restrict the cause to some subset of the possible failure causes. If a failure type is not observed, we observe a set of possible types of causes containing a true type, along with failure time, which may be subject to censoring. When the set of possible failure types consists of more than one element, the cause of failure is masked. When it is a singleton set, then the failure type is exactly observed and when it contains all the possible failure types, the missingness is total. Flehinger et.al. (1998) estimated the probability of survival due to different types in two stages with the assumption that the hazards of various risks are proportional to each other. They assumed that definitive diagnosis for a small sample of the masked causes could be obtained in the second stage. Recently, Dewanji and Sengupta (2003) obtained nonparametric estimator of the different cause-specific hazards with out the assumption of proportional hazards using counting processes and the multiplicative intensity process of Aalen (1978).

The problem of masking may arise in multivariate lifetime data with competing causes as well. The estimation of survivor function in such situations is not carried out so far. Motivated by this, in the present chapter, our attempt is to obtain an estimator for the bivariate survivor function when the causes of failure corresponding to the component lifetimes are masked. The proposed method is an extension of Dewanji and Sengupta (2003) to the bivariate set up. The extension to multivariate set up is direct.

In Section 3.2, we consider a system with two study objects (components) whose causes of failure are masked. We, in Section 3.3, develop a nonparametric estimator for the bivariate survivor function of competing risk models under masked causes of failure based on the vector hazard rate. Asymptotic properties of the estimator are established in Section 3.4. We illustrate the method with a data set in Section 3.5.A simulation study is presented in Section 3.6 to assess the performance of the estimator. We conclude the chapter in Section 3.7

3.2 Survivor Function and Hazard Function

Let $T = (T_1, T_2)$ be a pair of non-negative random variables defined on a probability space $(\Omega, \mathbb{F}, \mathrm{P})$. Let $\delta = (\delta_1, \delta_2)$ denote the censoring indicator and $C = (C_1, C_2)$ be a set of causes corresponding to $T = (T_1, T_2)$. Let $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$ be the survivor function of T Suppose that there are γ_1 causes of failure for T_1 and γ_2 causes of failure for T_2. The cause of failures either C_1 or C_2 or both C_1 and C_2 may be missing. Let $G = (G_1, G_2)$, where G_i is the set of possible causes for the ith component; i =1,2. If there are γ_i causes of failure for T_i, then $G_i \subseteq \{1,2\ldots \gamma_i\}$. Assume that failure type j must be a unique element of $\{1,2\ldots \gamma_i\}$.

Now define

$$P_{g_i,j}(t_1, t_2) = P(G_i = g_i \mid T_i \in dt_i, T_k > t_k, C_i = j, \delta_i = 1) \tag{3.1}$$

with $P_{g_i,j}(t_1, t_2) = 0$ if $j \notin g_i$; $j = 1,2,\ldots \gamma_i$; $i,k = 1,2,\ i \neq k$

Therefore, for fixed j, $\sum_{g_i} P_{g_i,j}(t_1,t_2) = 1 \qquad j = 1,2,\ldots\ \gamma_i :\ i = 1,2.$ (3.2)

Assume that the missing mechanism is independent of the censoring mechanism. Then (3.1) becomes

$$P_{g_i,j}(t_1,t_2) = P\left(G_i = g_i \mid T_i \in dt_i, T_k > t_k, C_i = j\right);\ j = 1,2,\ldots\ \gamma_i;\ i,k = 1,2,\ i \neq k \quad (3.3)$$

Thus $P_{g_i,j}(t_1,t_2)$ gives the conditional probability of observing $g_i \ni j$ as the set of possible causes, given failure of the ith component at time t_i due to the cause j and survival of the kth component at time t_k; $i,k = 1,2,\ i \neq k$

The hazard function for failure of the ith component due to cause j at time t_i and with $g_i \ni j$ observed as the set of possible causes is given by

$$\Lambda_{g_i,j|t_k}(dt_i) = \lim_{\Delta t_i \to 0} \frac{P(T_i \le t_i + \Delta t_i, C_i = j, G_i = g_i \mid T_i \ge t_i, T_k > t_k)}{\Delta t_i};\ j = 1,2,\ldots\ \gamma_i;$$

$$i,k = 1,2;\ i \neq k$$

Denote the events $\{T_i \le t_i + \Delta t_i, C_i = j\}$, $\{G_i = g_i\}$ and $\{T_i \ge t_i, T_k > t_k\}$ as A, B and C respectively. Since $P(\mathrm{A} \cap \mathrm{B}|\mathrm{C}) = P(\mathrm{A}|\mathrm{C})P(\mathrm{B}|\mathrm{A} \cap \mathrm{C})$, we obtain $\Lambda_{g_i,j|t_k}(dt_i)$ as

$$\Lambda_{g_i,j|t_k}(dt_i) = P_{g_i,j}(t_1,t_2)\ h_{ij}(t_1,t_2);\ j = 1,2,\ldots\ \gamma_i;\ i,k = 1,2,\ i \neq k$$

where $h_{ij}(t_1,t_2)$ is the cause-specific hazard function given in (2.6) and (2.7).

Thus the hazard function for failure of ith component at time t_i with g_i observed as the set of possible causes is given by

$$\Lambda_{g_i,t_k}(dt_i) = \lim_{\Delta t_i \to 0} \frac{P(T_i \le t_i + \Delta t_i, G_i = g_i \mid T_i \ge t_i, T_k > t_k)}{\Delta t_i}$$

$$= \sum_{j \in g_i} P_{g_i,j}(t_1,t_2) h_{ij}(t_1,t_2); \qquad i,k = 1,2;\ i \neq k \quad (3.4)$$

Using (3.2), summing (3.4) over all non-empty subsets g_i of G_i, we get

$$h_i(t_1,t_2) = \sum_{g_i} \Lambda_{g_i,t_k}(dt_i) = \sum_{j=1}^{\gamma_i} h_{ij}(t_1,t_2); \qquad i,k = 1,2;\ i \neq k \quad (3.5)$$

Denote $\Lambda_{t_k}^{(i)}(t_i)$ as the $(2^{\gamma_i}-1)\times 1$ vector of $\Lambda_{g_i|t_k}(t_i)$'s and $\Lambda_{t_k}^{*}(t_i)$ as the $\gamma_i \times 1$ vector of the cumulative cause specific hazards corresponding to $h_{ij}(t_1,t_2)$ We assume that $P_{g_i j}(t_1,t_2)$'s are independent of t_1 and t_2, though it may depend on g_i and j. So, let us denote $P_{g_i j}(t_1,t_2)$ by $P_{g_i j}$ and let P_i denote the $\left(2^{\gamma_i}-1\right)\times\gamma_i$ matrix of the $P_{g_i j}$'s.

Using (3.4),

$$\Lambda_{t_k}^{(i)}(t_i) = P_i\,\Lambda_{t_k}^{*}(t_i); \qquad i,k = 1,2;\ i \neq k \tag{3.6}$$

Let $I_{1\times\gamma_i}$ be a $1\times\gamma_i$ vector of unity. Then, using (3.5), we obtain the cumulative hazard functions as

$$\Lambda_i(t_1,t_2) = I_{1\times\gamma_i}\,\Lambda_{t_k}^{*}(t_i); \qquad i,k = 1,2;\ i \neq k \tag{3.7}$$

3.3 Nonparametric Estimation

Let $Z = (Z_1,Z_2)$ be a pair of random censoring times. Under the bivariate right random censoring, the observable variables are given by $Y=(Y_1,Y_2)$ and $\delta = (\delta_1,\delta_2)$ where $Y_i = \min(T_i,Z_i)$ and $\delta_i = I(T_i = Y_i)$; i = 1, 2. Let $H(t_1,t_2)$ denote the survivor function of $Y=(Y_1,Y_2)$. Let the observed data be $(Y_{1u},Y_{2u},\delta_{1u},\delta_{2u},G_{1u},G_{2u})$; u =1,2...n where Y_{1u} and Y_{2u} are observed or censored lifetimes corresponding to the ith unit. Consider the $\left(2^{\gamma_i}-1\right)$ dimensional counting process $\left\{N_{g_i t_k}(t_i)\right\}_{g_i\in\mathfrak{M}_i}$ where $\mathfrak{M}_i$ consists of all non-empty subsets of $\{1,2\ldots\gamma_i\}$ and $N_{g_i|t_k}(t_i)$ represents the number of failures of ith component up to time t_i for which $T_k > t_k$ with g_i as the observed set of possible causes; i,k = 1,2; $i \neq k$ .The corresponding intensity process for fixed t_k, using the multiplicative intensity process of Aalen (1978), is given by

$$\alpha_{g_i t_k}(t_i) = Y_{t_k}^{*}(t_i)\,\Lambda_{g_i|t_k}(dt_i); \qquad i,k = 1,2;\ i \neq k$$

where $Y_{t_k}^{*}(t_i)$ is the number of units with $Y_i \geq t_i$ and $Y_k > t_k$.

For each non-empty subset g_i of $\{1,2\ldots \gamma_i\}$ and fixed t_k,

$$dN_{g_i|t_k}(t_i) = \alpha_{g_i|t_k}(t_i)dt_i + dM_{g_i|t_k}(t_i); \qquad i,k = 1,2;\ i \neq k$$

where $M_{g_i|t_k}(t_i)$'s are local square integrable martingales. Therefore the estimator of $\Lambda_{g_i|t_k}(t_i)$ is given by

$$\hat{\Lambda}_{g_i|t_k}(t_i) = \int_0^{t_i} \frac{I(Y_{t_k}^{*}(s) > 0)}{Y_{t_k}^{*}(s)} dN_{g_i|t_k}(s). \tag{3.8}$$

Using Theorem 3.1 given in Section 3. 4 and equation (3.6), we obtain

$$\hat{\Lambda}_{t_k}^{(i)}(t_i) = P_i \Lambda_{t_k}^{*}(t_i) + \varepsilon_{t_k}(t_i); \qquad i,k = 1,2;\ i \neq k \tag{3.9}$$

where $\hat{\Lambda}_{t_k}^{(i)}(t_i)$ is the vector of $\hat{\Lambda}_{g_i|t_k}(t_i)$'s and for fixed t_k, $\varepsilon_{t_k}(t_i)$ is a vector process converging to a vector of Gaussian martingales whose variance function is consistently estimated by the matrix $diag\left(\hat{\tau}_{g_i|t_k}(t_i)\right)$ with $\hat{\tau}_{g_i|t_k}(t_i)$ is the variance function of $\hat{\Lambda}_{g_i|t_k}(t_i)$ given by equation (3.15). Equation (3.9) can be considered as a linear model with the design matrix P_i to be estimated. Let $\hat{P}_i$ denote a consistent estimator of P_i Then, using the principle of weighted least squares, a consistent estimator of $\Lambda_{t_k}^{*}(t_i)$ is

$$\hat{\Lambda}_{t_k}^{*}(t_i) = \left(\hat{P}_i^{T} W_{t_k}(t_i)\hat{P}_i\right)^{-1} \hat{P}_i^{T} W_{t_k}(t_i)\hat{\Lambda}_{t_k}^{(i)}(t_i); \qquad i,k = 1,2;\ i \neq k \tag{3.10}$$

where $W_{t_k}(t_i)$ is the inverse of the estimated $\left(2^{\gamma_i}-1\right)\times\left(2^{\gamma_i}-1\right)$ diagonal covariance matrix of $\hat{\Lambda}_{t_k}^{(i)}(t_i)$ as given by

$$W_{t_k}(t_i) = diag\left(\frac{1}{\hat{\tau}_{g_i|t_k}(t_i)}\right).$$

For $g_i \ni j$,

$$P_{g_i|j} = P(G_i = g_i \mid C_i = j)$$

$$= \frac{P(C_i = j \mid G_i = g_i)P(G_i = g_i)}{\sum_{g_i' \ni j} P(C_i = j \mid G_i = g_i')P(G_i = g_i')}; \qquad j = 1,2,\ldots \gamma_i;\ i = 1,2.$$

The estimator of $P_{g_i j}$ is given by

$$\hat{P}_{g_i j} = \frac{f_{g_i} q_{j g_i}}{\sum_{g_i \ni j} f_{g_i} q_{j g_i}}$$

where f_{g_i} denotes the number of failures with G_i observed as g_i and

$$q_{j g_i} = P(C_i = j \mid G_i = g_i); \qquad j = 1,2,\ldots \gamma_i; \quad i = 1,2.$$

From (3.7), we get the estimate of $\Lambda_i(t_1,t_2)$

$$\hat{\Lambda}_i^{M}(t_1,t_2) = I_{1\times\gamma_i}\, \hat{\Lambda}^{*}_{i_k}(t_i); \qquad i,k = 1,2;\ i \neq k \tag{3.11}$$

From (2.4) and (2.5), we obtain,

$$\hat{S}_1^{M}(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1^{M}(t_1,0) - \hat{\Lambda}_2^{M}(t_1,t_2)\right\} \tag{3.12}$$

and

$$\hat{S}_2^{M}(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1^{M}(t_1,t_2) - \hat{\Lambda}_2^{M}(0,t_2)\right\} \tag{3.13}$$

Thus we have two consistent estimators for $S(t_1,t_2)$, obtained in (3.12) and (3.13), which may be different. As in Chapter 2, to get a unique estimator, we follow the approach given in Akritas and van Keilegom (2003). The proposed estimator for $S(t_1,t_2)$ is a convex combination of two expressions (3.12) and (3.13). Thus the estimator for $S(t_1,t_2)$ is given by

$$\hat{S}_b^{M}(t_1,t_2) = b(t_1,t_2)\hat{S}_1^{M}(t_1,t_2) + (1 - b(t_1,t_2))\hat{S}_2^{M}(t_1,t_2) \tag{3.14}$$

As explained in Section 2.3.2, we choose the weight $b(t_1,t_2)$ in such a way that the mean squared error (MSE) of $\hat{S}_b^{M}(t_1,t_2)$ is minimum.

Remark 3.1 The extension to the multivariate set up is direct, as the survivor function $S(t_1,t_2,\ldots t_k)$ of $(T_1,T_2,\ldots T_k)$ can be uniquely represented as shown in Remark 2.1.

Remark 3.2 If both G_1 and G_2 are singleton sets, then the estimation reduces to the bivariate competing risk case, given in Ansa (2004).

Remark 3.3 When $t_k \to 0$, (3.10) reduces to the univariate case given in Dewanji and Sengupta (2003).

Remark 3.4 The estimation of $F_{ij}(t_1,t_2)$ under masking is a challenging problem in the bivariate competing risk set up.

3.4 Properties of the Estimators

In this section, we discuss various properties of the estimator. Let D = $[0,b_1]\times[0,b_2]$ with $H(b_1,b_2)>0$.

Theorem 3.1

For each fixed t_k, $\hat{\Lambda}_{g_i|t_k}(t_i)$ converges in distribution to a Gaussian process with mean $\Lambda_{g_i|t_k}(t_i)$ and a variance function, which can be consistently estimated by

$$\hat{\tau}_{g_i|t_k}(t_i) = \int_0 \frac{I(Y_{t_k}^{\bullet}(s)>0)}{Y_{t_k}^{\bullet 2}(s)} dN_{g_i|t_k}(s); \qquad i,k = 1,2;\ i \neq k \tag{3.15}$$

Proof

For each fixed t_k, $\hat{\Lambda}_{g_i|t_k}(t_i)$ is the estimator of the cumulative hazard function of *i*th component at time t_i with g_i observed as the set of possible causes conditioned on $T_k > t_k$ The weak convergence of the estimator of the hazard rate in the univariate set up with g_i observed as the set of possible causes is proved in Andersen and Borgan (1985). Thus, for fixed t_k, the proof follows directly from Andersen and Borgan (1985).

Theorem 3.2

Assume that the missing mechanism is independent of the failure mechanism and censoring mechanism. Then $\underset{D}{Sup}\left\|\hat{S}_h^{M} - S\right\| \to 0$ almost surely.

Proof

For fixed t_k, $\hat{\Lambda}_{ik}^{\bullet}(t_i)$ is uniformly strong consistent; $i,k = 1,2;\ i \neq k$ Therefore the proof follows from the proof of the Theorem 2.3 using (3.11), (3.12), (3.13) and (3.14).

Theorem 3.3

Under the assumptions of Theorem 3.2, for all $(t_1,t_2) \in D$ and for large n, $\sqrt{n}\left(\hat{S}_h^{M}(t_1,t_2) - S(t_1,t_2)\right)$ converges weakly to a mean zero Gaussian process with variance given by (3.18).

Proof

Using (3.14) we can write

$$\sqrt{n}\left(\hat{S}_h^{M}(t_1,t_2) - S(t_1,t_2)\right) = \sqrt{n}\, b(t_1,t_2)\left(\hat{S}_1^{M}(t_1,t_2) - S(t_1,t_2)\right) + \sqrt{n}\,(1-b(t_1,t_2))\left(\hat{S}_2^{M}(t_1,t_2) - S(t_1,t_2)\right). \quad (3.16)$$

Let $\hat{S}_i^{M}(t_i \mid t_k) = \exp\left\{-\hat{\Lambda}_i^{M}(t_1,t_2)\right\};\ i,k = 1,2;\ i \neq k$,

$\hat{S}_1^{M}(t_1) = \exp\left\{-\hat{\Lambda}_1^{M}(t_1,0)\right\}$ and $\hat{S}_2^{M}(t_2) = \exp\left\{-\hat{\Lambda}_2^{M}(0,t_2)\right\}$

Consider

$$\sqrt{n}\left(\hat{S}_1^{M}(t_1,t_2) - S(t_1,t_2)\right) = \sqrt{n}\left(\hat{S}_1^{M}(t_1) - S_1(t_1)\right)S_2(t_2 \mid t_1) + \sqrt{n}\,\hat{S}_1^{M}(t_1)\left(\hat{S}_2^{M}(t_2 \mid t_1) - S(t_2 \mid t_1)\right) \quad (3.17)$$

where $S_i(t_i)$ is the marginal distribution of T_i and $S_i(t_i \mid t_j)$ is the conditional distribution of T_i given $T_j > t_j$ and $\hat{S}_i^{M}(t_i)$ and $\hat{S}_i^{M}(t_i \mid t_j)$ are their nonparametric estimators; $i, j = 1,2;\ i \neq j$

The asymptotic normality of $\hat{\Lambda}_i^M(t_1,t_2)$, for i = 1,2, follows from Theorem 3.1 and equation (3.7). For fixed t_1, $S_2(t_2|t_1)=\prod_{s\le t_2}\left(1-d\Lambda_2(t_1,s)\right)$ is the Hadamard differentiability of the product integral. Thus the asymptotic normality of $\hat{\Lambda}_2^M(t_1,t_2)$ carries over to the asymptotic normality of $\hat{S}_2^M(t_2|t_1)$. This shows that $\sqrt{n}\left(\hat{S}_2^M(t_2|t_1)-S(t_2|t_1)\right)$ converges to a Gaussian process with mean zero. Since $\hat{S}_1^M(t_1)$ is strongly consistent and $\sqrt{n}\left(\hat{S}_2^M(t_2|t_1)-S(t_2|t_1)\right)$ converges weakly to a Gaussian process with mean zero, the second factor of (3.17) weakly converges to a Gaussian process. Since $S_2(t_2|t_1)$ is bounded and $\sqrt{n}\left(\hat{S}_1^M(t_1)-S_1(t_1)\right)$ converges weakly to a Gaussian process with mean zero, the first factor of (3.17) converges weakly to a Gaussian process with mean zero. Thus, $\sqrt{n}\left(\hat{S}_1^M(t_1,t_2)-S(t_1,t_2)\right)$ weakly converges to a Gaussian process with mean zero. On similar lines, we can prove that $\sqrt{n}\left(\hat{S}_2^M(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero. Thus, $\sqrt{n}\left(\hat{S}_b^M(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance of $\sqrt{n}\,\hat{S}_b^M(t_1,t_2)$ is

$$\begin{aligned}\sigma^2(t_1,t_2) = {} & b^2(t_1,t_2)\sigma_{11}(t_1,t_2)+\left(1-b(t_1,t_2)\right)^2\sigma_{22}(t_1,t_2)\\ & +b(t_1,t_2)\left(1-b(t_1,t_2)\right)\sigma_{12}(t_1,t_2) \qquad (3.18)\end{aligned}$$

where $\sigma_{ij}(t_1,t_2)$ is the asymptotic covariance between $\sqrt{n}\,\hat{S}_i^M(t_1,t_2)$ and $\sqrt{n}\,\hat{S}_j^M(t_1,t_2)$; i,j=1,2.

From Appendix B of Lawless (2003, page 539), we have

$$\sigma_{ii}(t_1,t_2)=\left(S_i(t_1,t_2)\right)^2 \; As\,\mathrm{var}\left(\log\hat{S}_i^M(t_1,t_2)\right) \qquad \text{for } i = 1,2.$$

From Theorem 3.1 we obtain

$$As\ \mathrm{var}\left(\log \hat{S}_i^{M}(t_1,t_2)\right) = diag\left(\hat{\tau}_{g_i t_i}(t_i)\right) \qquad \text{for } i = 1,2.$$

The asymptotic covariance between $\hat{S}_1^{M}(t_1,t_2)$ and $\hat{S}_2^{M}(t_1,t_2)$ is

$$\sigma_{12}(t_1,t_2) = S_1(t_1,t_2)S_2(t_1,t_2)\ As\ cov\left[\log \hat{S}_1^{M}(t_1,t_2), \log \hat{S}_2^{M}(t_1,t_2)\right].$$

Thus, from (3.17),

$$\begin{aligned}\sigma_{12}(t_1,t_2) = nS_1(t_1,t_2)S_2(t_1,t_2)\Big\{E\Big[&\left(\hat{\Lambda}_1^{M}(t_1,0)-\Lambda_1(t_1,0)\right)\left(\hat{\Lambda}_1^{M}(t_1,t_2)-\Lambda_1(t_1,t_2)\right)+\\ &\left(\hat{\Lambda}_1^{M}(t_1,t_2)-\Lambda_1(t_1,t_2)\right)\left(\hat{\Lambda}_2^{M}(t_1,t_2)-\Lambda_2(t_1,t_2)\right)+\\ &\left(\hat{\Lambda}_2^{M}(0,t_2)-\Lambda_2(0,t_2)\right)\left(\hat{\Lambda}_2^{M}(t_1,t_2)-\Lambda_2(t_1,t_2)\right)+\\ &\left(\hat{\Lambda}_1^{M}(t_1,0)-\Lambda_1(t_1,0)\right)\left(\hat{\Lambda}_2^{M}(0,t_2)-\Lambda_2(0,t_2)\right)\Big]\Big\}.\end{aligned}$$

3.5 Data Analysis

We could not find an appropriate real life masked data in the bivariate set up. However, to illustrate the estimation procedure given in Section 3.3, we use the data concerning the times to tumour appearance or death for 50 pairs of mice from the same litter in a tumor genesis experiment (Mantel and Ciminera, 1979), as reported in Ying and Wei (1994). We consider T_1 and T_2 as failure times (in weeks) for a pair of mice, and C_j (j = 1,2) indicates whether the failure was the appearance of a tumour (C_j = 1) or the occurrence of death prior to tumour appearance (C_j = 2). The censored observations are denoted by C_j = 0. The experiment was terminated at 104 weeks, so there is a common censoring time across all animals of 104. To introduce masking, we randomly allocated the masked set {1,2} among the observed lifetimes. The modified data is given in Table 3.1.

Table 3.1: Data concerning the times to tumour appearance or death for 50 pairs of mice

T_1	G_1	T_2	G_2	T_1	G_1	T_2	G_2
49	1	104*	0	104*	0	104*	0
102	2	104*	0	104*	0	104*	0
104*	0	104*	0	81	1	64	{1,2}
97	2	79	2	55	{1,2}	94	2
104*	0	104*	0	104*	0	54	1
96	{1,2}	104*	0	87	2	74	2
94	2	77	{1,2}	73	1	84	1
104*	0	104*	0	104*	0	83	{1,2}
77	{1,2}	104*	0	104*	0	73	2
104*	0	77	2	79	{1,2}	104*	0
91	2	90	2	104*	0	104*	0
70	2	92	{1,2}	104*	0	104*	0
45	2	50	1	101	1	94	{1,2}
69	{1,2}	91	2	84	1	78	1
104*	0	103	2	81	{1,2}	76	2
72	2	104*	0	95	2	104*	0
63	2	104*	0	104*	0	66	1
104*	0	74	2	104*	0	102	{1,2}
104*	0	69	2	98	{1,2}	73	2
104*	0	68	1	104*	0	104*	0
104*	0	104*	0	83	2	77	2
104*	0	104*	0	104*	0	104*	0
83	{1,2}	40	{1,2}	79	2	99	{1,2}
104*	0	104*	0	91	{1,2}	104*	0
104*	0	104*	0	104*	0	79	1

(* indicates censored time)

The estimators $\hat{S}_1^M(t_1,t_2)$ and $\hat{S}_2^M(t_1,t_2)$ can be obtained directly from the data using the approach in Section 3.3 for three cases of q_{jg_i}, given by (*i*) q_{1g_1} = 0.98 and q_{2g_1}=0.02 (*ii*) q_{1g_1} = 0.5 and q_{2g_1}=0.5 (*iii*) q_{1g_1} = 0.02 and q_{2g_1}=0.98. Since $\hat{S}_1^M(t_1,t_2)$ and $\hat{S}_2^M(t_1,t_2)$ are not very different, we take $b(t_1,t_2)$ = 0.5. Then the estimator of the survivor function $\hat{S}_b^M(t_1,t_2)$ at different time points (55,90), (97,79), (87,74) and (73,74) is obtained using (3.14). The value of $\hat{S}_b^M(t_1,t_2)$ for three cases of q_{jg_i} is given in Table 3.2 and the graphs are also plotted, which are shown in Figures 3.1, 3.2and 3.3. The value of $\hat{S}_b^M(t_1,t_2)$ for the cases (*i*), (*ii*) and (*iii*) are respectively denoted in graphs by $S(p_1)$, $S(p_2)$and $S(p_3)$.

Table 3.2: Estimates of the survivor function $S(t_1,t_2)$

(t_1,t_2)	Cases	$\hat{S}_1^M(t_1,t_2)$	$\hat{S}_2^M(t_1,t_2)$	$\hat{S}_b^M(t_1,t_2)$
(73,74)	(*i*)	.735941	.718968	.727904
	(*ii*)	.729264	.709157	.71921
	(*iii*)	.727244	.708576	.7179
(97,79)	(*i*)	.399942	.377815	.388878
	(*ii*)	.342390	.340210	.341300
	(*iii*)	.369583	.348013	.358798
(87,74)	(*i*)	.506083	.539301	.522692
	(*ii*)	.466971	.510490	.488731
	(*iii*)	.493082	.507350	.497408
(55,90)	(*i*)	.570129	.516792	.54346
	(*ii*)	.611161	.590559	.60086
	(*iii*)	.648197	.628784	.63849

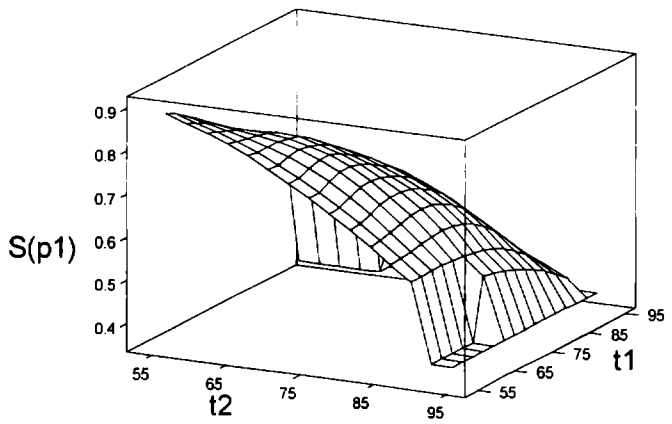

Fig 3.1: Estimated survivor function for the data on times to tumour appearance or death of 50 pair of mice for case (*i*)

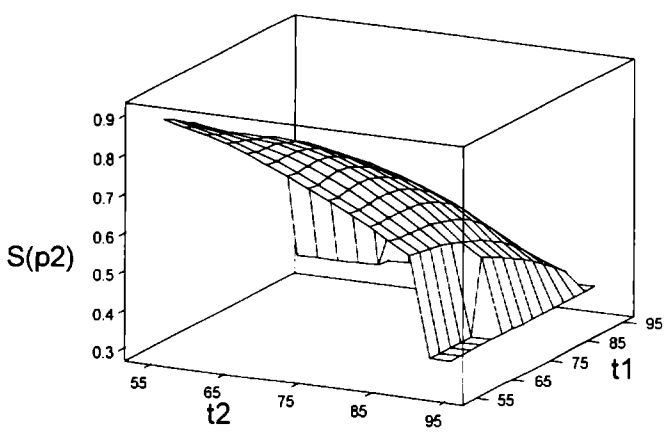

Fig 3.2: Estimated survivor function for the data on times to tumour appearance or death of 50 pair of mice for case (*ii*)

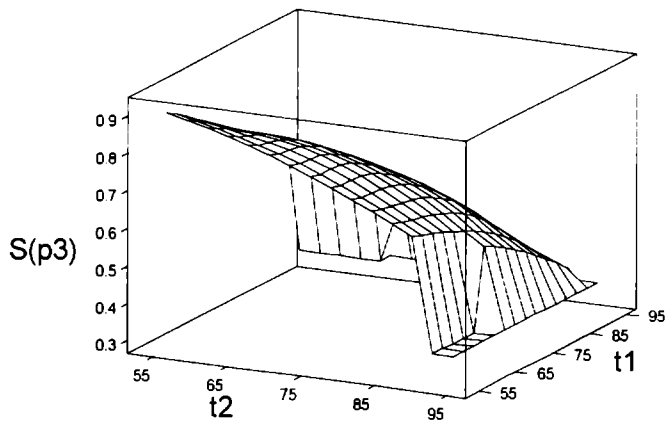

Fig 3.3: Estimated survivor function for the data on times to tumour appearance or death of 50 pair of mice for case (*iii*)

3.6 A Simulation Study

For empirical studies on the performance of the estimators, we carried out a series of 1000 simulations of size n from a Gumbel's (1960) bivariate exponential distribution with survivor function $S(t_1,t_2)=\exp\{-t_1-t_2-\lambda t_1 t_2\}$, $0<t_1,t_2<\infty$, $0\le\lambda\le 1$ with $\lambda=0.7$ and 0.8 for various sample sizes n = 50 and 100. We generated censoring times (Z_1,Z_2) from the Gumbel's (1960) exponential distribution with survivor function $G(t_1,t_2)=\exp\{-t_1-t_2-\theta t_1 t_2\}$, $0<t_1,t_2<\infty$, $0\le\theta\le 1$ with $\theta=0.7$ and 0.9 so that the censoring times are random. We used the algorithm given in Devroye (1986) for generating the observations. The observed times are given by $Y=(Y_1,Y_2)$ where $Y_i=\min(T_i,Z_i)$; i = 1, 2. Then we obtained $\delta=(\delta_1,\delta_2)$ where $\delta_i=I(T_i=Y_i)$; i = 1, 2. We considered two types of causes for C_i; i = 1, 2. The masked set $G=\{1,2\}$ was randomly allocated to the observed lifetimes so that the chance for an observed lifetime to be masked is 0.5. The causes 1 and 2 are distributed randomly among the observed failure times with unmasked

causes of failure. Then we computed $\hat{\Lambda}_i^M(t_1,t_2)$ and $\hat{S}_i^M(t_1,t_2)$; i =1,2 using (3.11), (3.12) and (3.13). Since $\hat{S}_1^M(t_1,t_2)$ and $\hat{S}_2^M(t_1,t_2)$ are not very different, we take $b(t_1,t_2)$ = 0.5 and found the estimator of $\hat{S}_b^M(t_1,t_2)$ using the equation (3.14) at four time points, namely (1) (0.1,0.2), (2) (0.1,0.3), (3) (0.2,0.2) and (4) (0.2,0.1), and for three cases of q_{jg_i}, by giving the values (*i*) q_{1g} = 0.98 and q_{2g} = 0.02 (*ii*) q_{1g} = 0.5 and q_{2g} = 0.5 (*iii*) q_{1g} = 0.02 and q_{2g} = 0.98. The value of $\hat{S}_b^M(t_1,t_2)$ for the cases (*i*), (*ii*) and (*iii*) are respectively denoted by $S(p_1)$, $S(p_2)$ and $S(p_3)$ The empirical biases and empirical variances of the estimators are given in Table 3.3. From Table 3.3, we observe that variance is small when q_{1g} = 0.5 and q_{2g} =0.5.

Table 3.3: Biases ($\times 10^3$) and variances ($\times 10^3$) of $\hat{S}_b(t_1,t_2)$ at five (t_1,t_2) pairs of Gumbel's (1960) exponential distribution for various values of λ and θ

	t	n = 50						n = 100					
		$S(p_1)$		$S(p_2)$		$S(p_3)$		$S(p_1)$		$S(p_2)$		$S(p_3)$	
		Bias	Var	Bias	Var	Bias	Var	Bias	Var	Bias	Var	Bias	Var
	1	8.53	42.7	13.1	3.79	31.6	31.4	3.33	29.3	4.14	2.78	5.65	28.0
λ = 0.8	2	27.73	40.1	28.3	6.45	-37.0	40.4	25.5	27.2	4.54	3.09	6.42	21.4
θ = 0.9	3	1.136	48.6	44.1	6.25	3.32	34.6	.458	21.0	5.56	2.95	4.05	.334
	4	-19.1	20.5	28.3	4.84	-11.7	29.7	13.4	8.01	4.92	1.90	9.49	24.5
	1	1.005	36.9	31.5	4.03	40.7	49.7	1.27	30.1	24.8	2.95	-4.22	30.2
λ = 0.7	2	-3.68	52.3	31.9	6.40	-17.2	33.3	2.84	50.6	25.0	2.95	-13.9	29.5
θ = 0.9	3	38.57	25.6	48.3	7.15	-39.3	61.8	31.4	24.1	40.4	4.70	4.41	40.2
	4	18.28	29.9	42.7	3.64	32.1	40.7	-17.4	23.2	40.4	3.24	-21.0	34.1
	1	-2.87	49.8	30.2	2.93	-14.2	19.7	5E-4	44.1	30.5	2.74	-7.84	13.9
λ = 0.8	2	33.5	39.1	17.4	6.79	-33.9	35.8	30.5	34.8	16.4	2.58	-15.5	28.0
θ = 0.7	3	23.8	30.7	49.3	6.14	-31.7	51.7	15.7	9.99	46.8	3.84	4.48	30.6
	4	-39.2	49.9	23.8	4.34	-43.5	45.1	-16.2	30.4	13.5	3.55	-2.71	29.8
	1	-.496	44.4	31.9	6.09	-38.2	50.1	-.123	41.1	30.3	3.15	-37.0	44.2
λ = 0.7	2	13.02	43.5	34.2	5.77	-15.3	26.5	11.2	42.3	20.4	4.62	-13.9	22.4
θ = 0.7	3	15.31	45.0	42.7	7.40	-28.3	49.7	14.3	40.0	36.3	3.59	23.8	22.4
	4	-30.3	40.8	35.4	5.00	-40.5	53.4	-11.5	37.8	25.0	4.08	-11.5	35.8

3.7 Conclusion

In this chapter, we developed a nonparametric estimator for the bivariate survivor function of competing risk models under masked causes of failure based on the vector hazard rate. Asymptotic properties of the estimator are established. A simulation study is carried out to assess the performance of the estimator. We also illustrated the method with a data set. The procedure can be directly extended to the multivariate set up. The results in this chapter are presented in Ansa and Sankaran (2005).

Chapter 4
COMPETING RISK MODELS WHEN CENSORING TIMES ARE MISSING

4.1 Introduction

As mentioned in the Chapter 1, there are situations in the analysis of failure time or lifetime data where the censoring times of unfailed units are missing. For example, suppose that T_i is the time to failure for product unit i in a population of M manufactured units. In some applications T_i is measured in calendar time from the date of sale of the unit. For many types of products the manufacturers do not know the date of sale for most units and therefore the censoring time, which is the elapsed time between the sales of the item and when the data are assembled, for most unfailed items is unknown. The non-parametric estimator of the lifetime distribution for such data is available in literature (see Hu et.al. (1998)).

The problem of missing censoring time can arise in the competing risk set up as well. The analysis of such lifetime in competing risk set up is not discussed in literature. In the present study, we discuss nonparametric estimation of the distribution of lifetime in the presence of competing risks when censoring times are missing. Maximum likelihood estimator and simple moment estimator of cause-specific sub-density for such univariate competing risk data are obtained in Section 4.2. We, then in Section 4.3, consider a bivariate situation where (T_1, T_2) represent lifetimes of components associated with the systems in a population of manufactured systems where each component is exposed to competing causes of failure. The censoring times for unfailed components are missing, but we assume that the censoring time distribution is known. We present the maximum likelihood and simple moment estimators of bivariate distribution of (T_1, T_2) in this set up. The proposed method is an extension of Hu et.al. (1998) to competing risk set up. We

also conducted a simulation study to observe the asymptotic behaviour of the estimators, which is presented in Section 4.4. Finally, a brief conclusion is given in Section 4.5 at the end of the chapter.

4.2 Univariate Competing Risk Set up

Let T be a non-negative random variable representing the lifetime of a product. Let $F(t)$ be the distribution function and $S(t)=1 - F(t)$ be the survivor function of T Suppose that the population of M units has independent lifetimes t_1, t_2 t_M generated from the distribution $F(t)$. Let there be k competing causes of failure and let $F_j(t) = P(T \le t, C = j)$, where C denote the cause of failure, be the cause-specific sub-distribution function; $j = 1,2...k$. Let $S_j(t) = P(T > t, C = j)$. There are also censoring times Z_1, Z_2 Z_M associated with the units and we assume that Z_i's are independent of each other with common distribution function $G(\tau) = P(Z_i \le \tau)$. The observed data is as follows.

If $t_i \le Z_i$, we observe t_i and C_i where C_i is the cause of failure of ith unit and if $t_i > Z_i$, we know only that fact and not the value of Z_i or t_i; $i=1,2...M$.

We assume that $G(\tau)$ is known and that the lifetime T and censoring time Z are independent. Assume that lifetime T and censoring time Z take discrete values 1,2.... Let $f(t)=P(T_i=t), g(\tau) = P(Z_i=\tau)$ and $\bar{G}(\tau)=P(Z_i \ge \tau)$. Let $n_j(t)$ denote the number of observed failures at time t due to cause j, m_j denote the number of failures with cause j, m denote the total number of observed lifetime and $\tau_{\max} = Sup\{\tau : \bar{G}(\tau) > 0\}$ Then

$$m_j = \sum_{t=1}^{\tau_{\max}} n_j(t)$$

and

$$m = \sum_{j=1}^{k} m_j = \sum_{j=1}^{k} \sum_{t=1}^{\tau_{\max}} n_j(t).$$

Now, we find the maximum likelihood and simple moment estimators of the cause-specific distribution $f_j(t) = P(T_i = t, C_i = j)$; $j = 1,2\ldots k$.

4.2.1 Maximum Likelihood Estimator

Based on the observed data, the likelihood may be written as

$$L(f_j) = \prod_{j=1}^{k}\prod_{t_i \le \tau}\left(f_j(t_i)\right)^{\delta_{ij}} \prod_{t_i > \tau_i} S(Z_i) \tag{4.1}$$

where δ_{ij} = 1 if the ith unit fails due to cause j.

= 0 otherwise; $j = 1,2\ldots k$; $i = 1,2\ldots M$.

Assume that the cause of failure C_i is a unique element of the set $\{1,2\ldots k\}$. The over-all survivor function $S(Z_i)$ will be equal to

$$\begin{aligned} S(Z_i) &= \prod_{j=1}^{k} S_j(Z_i) \\ &= \prod_{j=1}^{k}\sum_{\tau=1}^{\tau_{\max}} P(T > \tau, C = j) g(\tau) \\ &= \prod_{j=1}^{k}\sum_{\tau=1}^{\tau_{\max}}\left(1 - \sum_{t \le \tau} f_j(t)\right) g(\tau) \\ &= \prod_{j=1}^{k}\left(1 - \sum_{t=1}^{\tau_{\max}} f_j(t)\overline{G}(t)\right). \end{aligned} \tag{4.2}$$

Then, (4.1) can be written as

$$L(f_j) = \prod_{j=1}^{k}\prod_{t=1}^{\tau_{\max}}\left(f_j(t)\right)^{n_j(t)}\left(S(Z_i)\right)^{M-m} \tag{4.3}$$

Using (4.2), (4.3) becomes

$$L\left(f_{T|C=j}\right) = \prod_{j=1}^{k}\prod_{t=1}^{\tau_{\max}}\left(f_j(t)\right)^{n_j(t)}\left(1 - \sum_{t=1}^{\tau_{\max}} f_j(t)\overline{G}(t)\right)^{M-m} \tag{4.4}$$

Taking logarithm and differentiating (4.4) with respect to $f_j(t)$, and then equating to zero, we get

$$n_j(t)\left(1-\sum_s f_j(s)\bar{G}(s)\right) = (M-m)\bar{G}(t)f_j(t) \tag{4.5}$$

Summing both sides of (4.5) over $t = 1,2\ldots\tau_{\max}$ we get

$$\sum_s f_j(s)\bar{G}(s) = \frac{m_j}{M-m+m_j}; \qquad j = 1,2\ldots k. \tag{4.6}$$

Substituting (4.6) in (4.5), the estimator of the cause-specific sub-density $f_j(t)$ is given by

$$\hat{f}_j(t) = \frac{n_j(t)}{\left(M-m+m_j\right)\bar{G}(t)}; \qquad j = 1,2\ldots k\,;\ t = 1,2\ldots\tau_{\max}. \tag{4.7}$$

Hence, the estimators of the cause-specific sub-distribution functions $F_j(t)$ and overall survivor function $S(t)$ are respectively obtained as

$$\hat{F}_j(t) = \sum_{s=1}^{t}\hat{f}_j(s) \text{ and } \hat{S}(t) = \prod_{j=1}^{k}\hat{S}_j(t); \quad j = 1,2\ldots k\,;\ t = 1,2\ldots\tau_{\max}$$

where $\hat{S}_j(t) = 1-\hat{F}_j(t)$.

4.2.2 Simple Moment Estimator

We can write $n_j(t) = \sum_i I(t_i = t, Z_i \geq t, C_i = j); \quad j = 1,2\ldots k$

where $I(.)$ denotes the indicator function.

Let $u_{ij}(t) = \dfrac{I(t_i = t, Z_i \geq t, C_i = j)}{\bar{G}(t)}$ $\quad t$=1,2....; i =1,2...M; j = 1,2...k.

Assume that the cause of failure and censoring mechanism is independent. Then

$$E\left(n_j(t)\right) = M\bar{G}(t)\,f_j(t). \tag{4.8}$$

Then, a simple moment estimator of $f_j(t)$ is given by

$$\hat{f}_j^{SM}(t) = \frac{1}{M}\sum_i u_{ij}(t)$$

$$= \frac{n_j(t)}{M\bar{G}(t)}; \quad j = 1,2\ldots k\,;\ t = 1,2\ldots\tau_{\max} \tag{4.9}$$

Hence, the simple moment estimators of the cause-specific sub-distribution function $F_j(t)$ and overall survivor function $\bar{F}(t)$ are respectively obtained as

$$\hat{F}_j^{SM}(t) = \sum_{s=1}^{t} \hat{f}_j^{SM}(s) \qquad j = 1,2\ldots k.$$

and

$$\hat{S}^{SM}(t) = \prod_{j=1}^{k} \hat{S}_j^{SM}(t); \qquad t = 1,2\ldots \tau_{\max}$$

Under the assumption that the lifetime T and censoring time Z are independent, $n_j(t)$ is binomial$\left(M, f_j(t)\bar{G}(t)\right)$ and hence

$$Var\left(\hat{f}_j^{SM}(t)\right) = \frac{f_j(t)\left(1 - f_j(t)\bar{G}(t)\right)}{M\bar{G}(t)} \qquad t = 1,2\ldots \tau_{\max}, j = 1,2\ldots k.$$

The estimator of the variance is

$$\widehat{Var}\left(\hat{f}_j^{SM}(t)\right) = \frac{\sum_i \left(u_{ij}(t) - \bar{u}_j(t)\right)^2}{M^2}$$

$$= \frac{n_j(t)\left(M - n_j(t)\right)}{M^3 \bar{G}^2(t)} \text{ with } \bar{u}_j(t) = \frac{\sum_i u_{ij}(t)}{M};$$

$$j = 1,2\ldots k.$$

A consistent estimate for the variance of $\hat{F}^{SM}(t)$ is

$$\widehat{Var}\left(\hat{F}_j^{SM}(t)\right) = \frac{\sum_i \left\{\sum_{s=1}^{t}\left(u_{ij}(s) - \bar{u}_j(s)\right)\right\}^2}{M^2}$$

$$= \frac{1}{M^3}\sum_{s=1}^{t} \frac{n_j(s)\left(M - n_j(s)\right)}{\bar{G}^2(s)}$$

$$- \sum_{\substack{s_1, s_2 = 1 \\ s_1 \neq s_2}}^{t} \frac{n_j(s_1) n_j(s_2)}{M^3 \bar{G}(s_1)\bar{G}(s_2)}; \qquad j = 1,2\ldots k.$$

Remark 4.1 When k =1, the maximum likelihood and simple moment estimators reduces to the corresponding estimators given in Hu et.al. (1998).

4.3 Bivariate Competing Risk Set up

Consider a two-component system with lifetime vector $T = (T_1, T_2)$. Let there be k_1 causes of failure for T_1 and k_2 causes of failure for T_2 Let $C = (C_1, C_2)$ denote the cause of failure of the two components. Let $Z = (Z_1, Z_2)$ denote the bivariate censoring time and $\delta = (\delta_1, \delta_2)$ be the censoring indicator. Then, the observed data from a population of M units is as follows.

If $T_{1i} \leq Z_{1i}$ and $T_{2i} \leq Z_{2i}$, we observe $(t_{1i}, t_{2i}, \delta_{1i} = \delta_{2i} = 1, C_{1i}, C_{2i})$,

if $T_{1i} \leq Z_{1i}$ and $T_{2i} > Z_{2i}$, we observe $(t_{1i}, \delta_{1i} = 1, \delta_{2i} = 0, C_{1i})$,

if $T_{1i} > Z_{1i}$ and $T_{2i} \leq Z_{2i}$, we observe $(t_{2i}, \delta_{1i} = 0, \delta_{2i} = 1, C_{2i})$,

if $T_{1i} > Z_{1i}$ and $T_{2i} > Z_{2i}$, we observe $(\delta_{1i} = 0, \delta_{2i} = 0)$; $i = 1,2 \ldots M$.

Note that neither Z_{1i} nor Z_{2i} is observable in any case. We use the following notations in this section.

$f(t_1, t_2)$ joint p.d.f of (T_1, T_2).

$F(t_1, t_2)$ joint distribution function of (T_1, T_2).

(j_1, j_2) observed causes of failure.

$f_{j_1 j_2}(t_1, t_2)$ bivariate cause-specific joint p.d.f of lifetime.

$F_{j_1 j_2}(t_1, t_2)$ bivariate cause-specific sub-distribution function.

$$f(t_1, t_2) = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1, t_2).$$

$$F(t_1, t_2) = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} F_{j_1 j_2}(t_1, t_2)$$

$G(\tau_1, \tau_2)$ bivariate distribution function of censoring time.

$\overline{G}(\tau_1,\tau_2) = P(Z_1 > \tau_1, Z_2 > \tau_2)$.

$g(\tau_1,\tau_2)$ bivariate p.d.f of censoring time.

$G_{Z_1}(\tau_1)$ marginal distribution function of Z_1

$g_{Z_1}(\tau_1)$ marginal p.d.f of Z_1

$G_{Z_2}(\tau_2)$ marginal distribution function of Z_2.

$g_{Z_2}(\tau_2)$ marginal p.d.f of Z_2.

$\overline{G}_{Z_i}(\tau)$ $1-G_{Z_i}(\tau)$.

$\tau_{i\max} = Sup\{\tau : \overline{G}_{Z_i}(\tau) > 0\}$; $i = 1,2$.

$f_{j_1}^{T_1}(t_1) = P(T_1 = t_1, C_1 = j_1) = \sum_{j_2=1}^{k_2}\sum_{t_2=1}^{\tau_{2\max}} f_{j_1 j_2}(t_1,t_2)$ cause-specific marginal of T_1.

$f_{j_2}^{T_2}(t_2) = P(T_2 = t_2, C_2 = j_2) = \sum_{j_1=1}^{k_1}\sum_{t_1=1}^{\tau_{1\max}} f_{j_1 j_2}(t_1,t_2)$ cause-specific marginal of T_2

$f_{j_1}(t_1,t_2) = P(T_1 = t_1, T_2 = t_2, C_1 = j_1) = \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1,t_2)$; $j_1 = 1,2\ldots k_1$

$f_{j_2}(t_1,t_2) = P(T_1 = t_1, T_2 = t_2, C_2 = j_2) = \sum_{j_1=1}^{k_1} f_{j_1 j_2}(t_1,t_2)$; $j_2 = 1,2\ldots k_2$.

We assume that the censoring time distribution function $G(\tau_1,\tau_2)$ and the marginal distribution functions $G_{Z_1}(\tau_1)$ and $G_{Z_2}(\tau_2)$ are known or at least estimated from other sources. Further, we assume (T_1,T_2,C_1,C_2) and (Z_1,Z_2) are independent. As carried out by Hu et.al. (1998) in the univariate case, we assume that each of T_1,T_2, Z_1 and Z_2 take discrete values 1,2....

4.3.1 Maximum Likelihood Estimator

Define $\delta_{ij_1} = I(T_{1i} = t_{1i}, C_{1i} = j_1)$ and $\delta_{ij_2} = I(T_{2i} = t_{2i}, C_{2i} = j_2)$,

$i = 1,2\ldots M$; $j_1 = 1,2\ldots k_1$; $j_2 = 1,2\ldots k_2$.

where $I(.)$ denotes the usual indicator function.

Then, the likelihood can be written as

$$L\left(f_{i_1 i_2}\right) = \prod_{\substack{i_{1r} \le Z_{1r} \\ i_{2r} \le Z_{2r}}} \prod_{i_1=1}^{k_1} \prod_{i_2=1}^{k_2} f_{i_1 i_2}^{\delta_{i_1} \delta_{i_2}} \left(t_{1r}, t_{2r}\right) \prod_{\substack{i_{1r} \le Z_{1r} \\ i_{2r} > Z_{2r}}} \prod_{i_1=1}^{k_1} \left(P\left(T_1 = t_{1r}, T_2 > Z_{2r}, C_{1r} = j_1\right)\right)^{\delta_{i_1}}$$

$$\prod_{\substack{i_{1r} > Z_{1r} \\ i_{2r} \le Z_{2r}}} \prod_{i_2=1}^{k_2} \left(P\left(T_1 > Z_{1r}, T_2 = t_{2r}, C_{2r} = j_2\right)\right)^{\delta_{i_2}} \prod_{\substack{i_{1r} > Z_{1r} \\ i_{2r} > Z_{2r}}} P\left(T_1 > Z_{1r}, T_2 > Z_{2r}\right). \qquad (4.10)$$

We consider

$$P\left(T_1 = t_{1r}, T_2 > Z_{2r}, C_{1r} = j_1\right) = \sum_{\tau_2=1}^{\tau_{2\max}} P\left(T_1 = t_{1r}, T_2 > \tau_2, C_{1r} = j_1\right) g_{Z_2}\left(\tau_2\right)$$

$$= \sum_{\tau_2=1}^{\tau_{2\max}} P\left(T_1 = t_{1r}, C_{1r} = j_1\right) P\left(T_2 > \tau_2 \mid T_1 = t_{1r}, C_{1r} = j_1\right) g_{Z_2}\left(\tau_2\right)$$

$$= \sum_{\tau_2=1}^{\tau_{2\max}} P\left(T_1 = t_{1r}, C_{1r} = j_1\right) \left[1 - \sum_{t_2 \le \tau_2} f_{T_2 \mid T_1 = t_{1r}, C_1 = j_1}\left(t_2\right)\right] g_{Z_2}\left(\tau_2\right)$$

$$= P\left(T_1 = t_{1r}, C_{1r} = j_1\right) \sum_{\tau_2=1}^{\tau_{2\max}} \left[g_{Z_2}\left(\tau_2\right) - \sum_{t_2 \le \tau_2} f_{T_2 \mid T_1 = t_{1r}, C_1 = j_1}\left(t_2\right) g_{Z_2}\left(\tau_2\right)\right]$$

$$= P\left(T_1 = t_{1r}, C_{1r} = j_1\right) \left[\sum_{\tau_2=1}^{\tau_{2\max}} g_{Z_2}\left(\tau_2\right) - \sum_{t_2=1}^{\tau_{2\max}} f_{T_2 \mid T_1 = t_{1r}, C_1 = j_1}\left(t_2\right) \sum_{\tau_2=t_2}^{\tau_{2\max}} g_{Z_2}\left(\tau_2\right)\right]$$

$$= P\left(T_1 = t_{1r}, C_{1r} = j_1\right) \left[1 - \sum_{t_2=1}^{\tau_{2\max}} f_{T_2 \mid T_1 = t_{1r}, C_1 = j_1}\left(t_2\right) \overline{G}_{Z_2}\left(t_2\right)\right]$$

$$= f_{i_1}^{T_1}\left(t_{1r}\right) - \sum_{t_2=1}^{\tau_{2\max}} \sum_{i_2=1}^{k_2} f_{i_1 i_2}\left(t_{1r}, t_2\right) \overline{G}_{Z_2}\left(t_2\right)$$

$$= \sum_{i_2=1}^{k_2} \sum_{t_2=1}^{\tau_{2\max}} f \quad \left(t_1, t_2\right) - \sum_{i_2=1}^{\tau_{2\max}} \sum_{t_2=1}^{k_2} f_{i_1 i_2}\left(t_{1r}, t_2\right) \overline{G}_{Z_2}\left(t_2\right)$$

$$= L_{i_1}^{(1)}\left(t_{1r}\right), \text{ say.} \qquad (4.11)$$

Similarly we obtain

$$P\left(T_1 > Z_{1r}, T_2 = t_{2r}, C_{2r} = j_2\right) = \sum_{i_1=1}^{k_1} \sum_{t_1=1}^{\tau_{1\max}} f_{i_1 i_2}\left(t_1, t_2\right) - \sum_{t_1=1}^{\tau_{1\max}} \sum_{i_1=1}^{k_1} f_{i_1 i_2}\left(t_1, t_{2r}\right) \overline{G}_{Z_1}\left(t_1\right)$$

$$= L_{i_2}^{(2)}\left(t_{2r}\right), \text{ say.} \qquad (4.12)$$

Now we consider

$$P(T_1 > Z_{1'}, T_2 > Z_{2'}) = \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_1 > \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2)$$

$$= \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} \left(1 - P(T_2 \le \tau_2) - P(T_1 \le \tau_1, T_2 > \tau_2)\right) g(\tau_1, \tau_2)$$

$$= 1 - \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_2 \le \tau_2) g(\tau_1, \tau_2) - \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_1 \le \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2)$$

$$= 1 - \sum_{\tau_2=1}^{\tau_{2\max}} P(T_2 \le \tau_2) g_{Z_2}(\tau_2) - \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_1 \le \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2)$$

$$= 1 - \sum_{t_2=1}^{\tau_{2\max}} f_{T_2}(t_2) \overline{G}_{Z_2}(t_2)$$

$$- \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_1 \le \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2). \qquad (4.13)$$

We can write

$$\sum_{\tau_1=1}^{\tau_{1\max}} P(T_1 \le \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2) = \sum_{\tau_1=1}^{\tau_{1\max}} \sum_{t_1 \le \tau_1} P(T_1 = t_1, T_2 > \tau_2) g(\tau_1, \tau_2)$$

$$= \sum_{t_1=1}^{\tau_{1\max}} \sum_{\tau_1=t_1}^{\tau_{1\max}} P(T_1 = t_1, T_2 > \tau_2) g(\tau_1, \tau_2)$$

$$= \sum_{t_1=1}^{\tau_{1\max}} P(T_1 = t_1, T_2 > \tau_2) P(Z_1 \ge t_1, Z_2 = \tau_2).$$

$$\sum_{\tau_1=1}^{\tau_{1\max}} \sum_{\tau_2=1}^{\tau_{2\max}} P(T_1 \le \tau_1, T_2 > \tau_2) g(\tau_1, \tau_2) = \sum_{\tau_2=1}^{\tau_{2\max}} \sum_{t_1=1}^{\tau_{1\max}} P(T_1 = t_1, T_2 > \tau_2) P(Z_1 \ge t_1, Z_2 = \tau_2)$$

$$= \sum_{\tau_2=1}^{\tau_{2\max}} \sum_{t_1=1}^{\tau_{1\max}} P(T_1 = t_1) P(T_2 > \tau_2 \mid T_1 = t_1) P(Z_1 \ge t_1, Z_2 = \tau_2)$$

$$= \sum_{t_1=1}^{\tau_{1\max}} P(T_1 = t_1) \sum_{\tau_2=1}^{\tau_{2\max}} \left[1 - \sum_{t_2 \le \tau_2} f_{T_2 \mid T_1 = t_1}(t_2)\right] P(Z_1 \ge t_1, Z_2 = \tau_2)$$

$$= \sum_{t_1=1}^{\tau_{1\max}} P(T_1 = t_1) \left[P(Z_1 > t_1) - \sum_{t_2=1}^{\tau_{2\max}} f_{T_2 \mid T_1 = t_1}(t_2) \overline{G}(t_1, t_2)\right]$$

$$= \sum_{t_1=1}^{\tau_{1\max}} f_{T_1}(t_1) \overline{G}_{Z_1}(t_1) - \sum_{t_1=1}^{\tau_{1\max}} \sum_{t_2=1}^{\tau_{2\max}} f(t_1, t_2) \overline{G}(t_1, t_2). \qquad (4.14)$$

Substituting (4.14) in (4.13), we obtain

$$
\begin{aligned}
P(T_1 > Z_{1i}, T_2 > Z_{2i}) &= 1 - \sum_{t_2=1}^{\tau_{2\max}} f_{j_2}(t_2)\overline{G}_{Z_2}(t_2) - \sum_{t_1=1}^{\tau_{1\max}} f_{j_1}(t_1)\overline{G}_{Z_1}(t_1) \\
&\quad + \sum_{t_1=1}^{\tau_{1\max}} \sum_{t_2=1}^{\tau_{2\max}} \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1, t_2)\overline{G}(t_1, t_2) \\
&= 1 - \sum_{t_1=1}^{\tau_{1\max}} \sum_{t_2=1}^{\tau_{2\max}} \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1, t_2)\overline{G}_{Z_2}(t_2) \\
&\quad - \sum_{t_1=1}^{\tau_{1\max}} \sum_{t_2=1}^{\tau_{2\max}} \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1, t_2)\overline{G}_{Z_1}(t_1) \\
&\quad + \sum_{t_1=1}^{\tau_{1\max}} \sum_{t_2=1}^{\tau_{2\max}} \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} f_{j_1 j_2}(t_1, t_2)\overline{G}(t_1, t_2) \\
&= L^{(3)}, \text{ say.}
\end{aligned} \tag{4.15}
$$

Denote

$$n_{j_1 j_2}(t_1, t_2) = \sum_i I(t_{1i} \le Z_{1i}, t_{2i} \le Z_{2i}, t_{1i} = t_1, t_{2i} = t_2, C_{1i} = j_1, C_{2i} = j_2),$$

$$m_{j_1}(t_1) = \sum I(t_{1i} \le Z_{1i}, t_{2i} > Z_{2i}, t_{1i} = t_1, C_{1i} = j_1),$$

$$m_{j_2}(t_2) = \sum_i I(t_{1i} > Z_{1i}, t_{2i} \le Z_{2i}, t_{2i} = t_2, C_{2i} = j_2),$$

and

$$m_3 = \sum_i I(t_{1i} > Z_{1i}, t_{2i} > Z_{2i}).$$

Then, the likelihood (4.10) can be written as

$$
\begin{aligned}
L(f_{j_1 j_2}) &= \prod_{t_1=1}^{\tau_{1\max}} \prod_{t_2=1}^{\tau_{2\max}} \prod_{j_1=1}^{k_1} \prod_{j_2=1}^{k_2} \left(f_{j_1 j_2}(t_1, t_2)\right)^{n_{j_1 j_2}(t_1, t_2)} \prod_{t_1=1}^{\tau_{1\max}} \prod_{j_1=1}^{k_1} \left(L_{j_1}^{(1)}(t_1)\right)^{m_{j_1}(t_1)} \\
&\quad \prod_{t_2=1}^{\tau_{2\max}} \prod_{j_2=1}^{k_2} \left(L_{j_2}^{(2)}(t_2)\right)^{m_{j_2}(t_2)} \left(L^{(3)}\right)^{m_3}
\end{aligned} \tag{4.16}
$$

Therefore the only unknown quantity in the likelihood function expression (4.16) is $f_{j_1 j_2}(t_1, t_2)$. Taking logarithm and differentiating (4.16) with respect to $f_{j_1 j_2}(t_1, t_2)$ and equating to zero, we get

$$\hat{f}_{j_1 j_2}(t_1,t_2) = n \quad (t_1,t_2)\left[\frac{m_{j_1}(t_1)G_{Z_2}(t_2)}{L^{(1)}_{j_1}(t_1)}+\frac{m_{j_2}(t_2)G_{Z_1}(t_1)}{L^{(2)}_{j_2}(t_2)}\right.$$

$$\left.-\frac{m_3\left(G_{Z_1}(t_1)+G_{Z_2}(t_2)-\overline{G}(t_1,t_2)\right)}{L^{(3)}}\right]^{-1} \qquad (4.17)$$

$$j_1 = 1,2\ldots k_1;\ j_2 = 1,2\ldots k_2.$$

Finally, $f_{j_1 j_2}(t_1,t_2)$ can be obtained from (4.17) by numerical iteration.

Remark 4.2 When $k_1 = k_2 = 1$, this reduces to the bivariate non-competing risk case. In this situation (4.17) becomes

$$\hat{f}(t_1,t_2) = n(t_1,t_2)\left[\frac{m_1(t_1)G_{Z_2}(t_2)}{L_1(t_1)}+\frac{m_2(t_2)G_{Z_1}(t_1)}{L_2(t_2)}\right.$$

$$\left.-\frac{m_3\left(G_{Z_1}(t_1)+G_{Z_2}(t_2)-\overline{G}(t_1,t_2)\right)}{L_3}\right]^{-1} \qquad (4.18)$$

where

$$n(t_1,t_2) = \sum_i I(t_{1i} \le Z_{1i}, t_{2i} \le Z_{2i}, t_{1i} = t_1, t_{2i} = t_2),$$

$$m_1(t_1) = \sum_i I(t_{1i} \le Z_{1i}, t_{2i} > Z_{2i}, t_{1i} = t_1),$$

$$m_2(t_2) = \sum_i I(t_{1i} > Z_{1i}, t_{2i} \le Z_{2i}, t_{2i} = t_2),$$

$$m_3 = \sum_i I(t_{1i} > Z_{1i}, t_{2i} > Z_{2i})$$

$$L_1(t_1,t_2) = \sum_{t_2=1}^{\tau_{2\max}} f(t_1,t_2) - \sum_{t_2=1}^{\tau_{2\max}} f(t_1,t_2)\overline{G}_{Z_2}(t_2),$$

$$L_2(t_1,t_2) = \sum_{t_1=1}^{\tau_{1\max}} f(t_1,t_2) - \sum_{t_1=1}^{\tau_{1\max}} f(t_1,t_2)\overline{G}_{Z_1}(t_1)$$

and

$$L_3(t_1,t_2) = 1 - \sum_{t_2=1}^{\tau_{2\max}}\sum_{t_1=1}^{\tau_{1\max}} f(t_1,t_2)\overline{G}_{Z_2}(t_2) - \sum_{t_1=1}^{\tau_{1\max}}\sum_{t_2=1}^{\tau_{2\max}} f(t_1,t_2)\overline{G}_{Z_1}(t_1) + \sum_{t_1=1}^{\tau_{1\max}}\sum_{t_2=1}^{\tau_{2\max}} f(t_1,t_2)\overline{G}(t_1,t_2).$$

Then, $\hat{f}(t_1,t_2)$ can be obtained from (4.18) by numerical iteration procedure.

Remark 4.3 The cause-specific sub-distribution function given by

$$F_{j_1j_2}(t_1,t_2) = P(T_1 \le t_1, T_2 \le t_2, C_1 = j_1, C_2 = j_2)$$

can be written as

$$F_{j_1j_2}(t_1,t_2) = \sum_{s_1 \le t_1} \sum_{s_2 \le t_2} f \quad (s_1,s_2),$$

and its estimator is given by

$$\hat{F}_{j_1j_2}(t_1,t_2) = \sum_{s_1 \le t_1} \sum_{s_2 \le t_2} \hat{f}_{j_1j_2}(s_1,s_2); \qquad j_1 = 1,2\ldots k_1;\ j_2 = 1,2\ldots k_2. \tag{4.19}$$

Remark 4.4 The joint distribution function of (T_1,T_2) is given by

$$F(t_1,t_2) = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} F_{j_1j_2}(t_1,t_2)$$

and its estimator is given by

$$\hat{F}(t_1,t_2) = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \hat{F}_{j_1j_2}(t_1,t_2).$$

Thus the estimator of bivariate survivor function is obtained as

$$\hat{S}(t_1,t_2) = 1 - \hat{F}(t_1,\infty) - \hat{F}(\infty,t_2) + \hat{F}(t_1,t_2).$$

4.3.2 Simple Moment Estimator

Denote by $m_{j_1j_2}$, the number of failures with cause (j_1,j_2) and let m denote the total number of units for which lifetime is observed for both the components. We can write

$$n_{j_1j_2}(t_1,t_2) = \sum_i I(t_{1i} \le Z_{1i}, t_{2i} \le Z_{2i}, t_{1i} = t_1, t_{2i} = t_2, C_{1i} = j_1, C_{2i} = j_2),$$

$$j_1 = 1,2\ldots k_1;\ j_2 = 1,2\ldots k_2.$$

where $I(.)$ denotes the indicator function.

Let $u'_{j_1 j_2}(t_1,t_2) = \dfrac{I\left(t_{1i} \le Z_{1i}, t_{2i} \le Z_{2i}, t_{1i} = t_1, t_{2i} = t_2, C_{1i} = j_1, C_{2i} = j_2\right)}{\overline{G}(t_1,t_2)}$

$i = 1,2\ldots M;\ j_1 = 1,2\ldots k_1;\ j_2 = 1,2\ldots k_2;\ t_1 = 1,2\ldots;t_2 = 1,2\ldots$

Then a simple moment estimator of $f_{j_1 j_2}(t_1,t_2)$ is obtained as

$$\hat{f}^{SM}_{j_1 j_2}(t_1,t_2) = \frac{1}{M}\sum_i u'_{j_1 j_2}(t_1,t_2)$$

$$= \frac{n_{j_1 j_2}(t_1,t_2)}{M\overline{G}(t_1,t_2)};\ t_1 = 1,2\ldots\tau_{1\max}\quad t_2 = 1,2\ldots\tau_{2\max} \tag{4.20}$$

which is obtained from the fact that

$$E\left(n_{j_1 j_2}(t_1,t_2)\right) = M\overline{G}(t_1,t_2)f_{j_1 j_2}(t_1,t_2).$$

Under the assumption that (T_1,T_2,C_1,C_2) and (Z_1,Z_2) are independent, $n_{j_1 j_2}(t_1,t_2)$ is binomial$\left(M, f_{j_1 j_2}(t_1,t_2)\overline{G}(t_1,t_2)\right)$ and hence the variance of the estimator is obtained as

$$Var\left(\hat{f}^{SM}_{j_1 j_2}(t_1,t_2)\right) = \frac{f_{j_1 j_2}(t_1,t_2)\left(1 - f_{j_1 j_2}(t_1,t_2)\overline{G}(t_1,t_2)\right)}{M\overline{G}(t_1,t_2)}$$

$$t_1 = 1,2\ldots\tau_{1\max};t_2 = 1,2\ldots\tau_{2\max}.$$

The estimator of the variance is

$$\widehat{Var}\left(\hat{f}^{SM}_{j_1 j_2}(t_1,t_2)\right) = \frac{\sum_i\left(u'_{j_1 j_2}(t_1,t_2) - \bar{u}_{j_1 j_2}(t_1,t_2)\right)^2}{M^2}$$

where $\bar{u}_{j_1 j_2}(t_1,t_2) = \dfrac{\sum_i u'_{j_1 j_2}(t_1,t_2)}{M}$

and thus

$$\widehat{Var}\left(\hat{f}^{SM}_{j_1 j_2}(t_1,t_2)\right) = \frac{n_{j_1 j_2}(t_1,t_2)\left(M - n_{j_1 j_2}(t_1,t_2)\right)}{M^3\overline{G}^2(t_1,t_2)}$$

Remark 4.5 When there is only one possible cause of failure for both the components $(k_1 = k_2 = 1)$, (4.10) becomes

$$\hat{f}^{SM}(t_1,t_2) = \frac{1}{M}\sum u_i(t_1,t_2) = \frac{n(t_1,t_2)}{M\bar{G}(t_1,t_2)} \qquad t_1 = 1,2\ldots\tau_{1\max};\ t_2 = 1,2\ldots\tau_{2\max}$$

where $u_i(t_1,t_2) = \dfrac{I(t_{1i} \le Z_{1i}, t_{2i} \le Z_{2i}, t_{1i} = t_1, t_{2i} = t_2)}{\bar{G}(t_1,t_2)}$

A consistent estimator for the variance of $\hat{f}^{SM}(t_1,t_2)$ is

$$\widehat{Var}\left(\hat{f}^{SM}(t_1,t_2)\right) = \frac{\sum_{i=1}^{M}\left(u_i(t_1,t_2) - \bar{u}(t_1,t_2)\right)^2}{M^2} \text{ where } \bar{u}(t_1,t_2) = \frac{\sum_{i=1}^{M} u_i(t_1,t_2)}{M}$$

$$= \frac{n(t_1,t_2)\left[M - n(t_1,t_2)\right]}{M^3\bar{G}(t_1,t_2)}.$$

Finally, the estimator of distribution function is

$$\hat{F}^{SM}(t_1,t_2) = \sum_{s_1 \le t_1}\sum_{s_2 \le t_2}\hat{f}^{SM}(s_1,s_2)$$

with an estimator of variance as

$$\widehat{Var}\left(\hat{F}^{SM}(t_1,t_2)\right) = \frac{\sum_{i=1}^{M}\left(\sum_{s_1=1}^{t_1}\sum_{s_2=1}^{t_2}\left(u_i(s_1,s_2) - \bar{u}(s_1,s_2)\right)\right)^2}{M^2}$$

$$= \frac{1}{M^3}\sum_{s_1=1}^{t_1}\sum_{s_2=1}^{t_2}\frac{n(s_1,s_2)\left(M - n(s_1,s_2)\right)}{\bar{G}^2(s_1,s_2)}$$

$$- \frac{1}{M^3}\sum_{\substack{s_1,r_1=1 \\ s_1 \ne r_1}}^{t_1}\sum_{\substack{s_2,r_2=1 \\ s_2 \ne r_2}}^{t_2}\frac{n(s_1,s_2)n(r_1,r_2)}{\bar{G}(s_1,s_2)\bar{G}(r_1,r_2)}$$

4.4 A Simulation Study

For empirical studies on the performance of the estimators, we carried out a series of 1000 simulations of size M from a bivariate Dirichlet Distribution with survivor function $S(t_1,t_2) = (1-t_1-t_2)^{\alpha+1}$ $0 < t_1, t_2 < t_1+t_2 < 1$ with $\alpha = 2,3$ and 5 for M = 1000 and 4000. We used the algorithm from Gentle (1998, pages 94

and 111) for generating the observations. We assumed that both the components T_1 and T_2 are censored by a common censoring variable Z, which is a uniform random variable in the interval (0,1). Accordingly, we generated M values of Z, Z_i, i = 1,2...M. Then we discretized both the lifetimes and censoring times. We considered two types of causes C i = 1,2. The causes 1 and 2 are distributed randomly among the observed lifetimes with equal probability. We found both the simple moment estimator and maximum likelihood estimator of $F_{j_1 j_2}(t_1, t_2)$; j_1, j_2 = 1,2, as explained in Section 4.3, at five time points. The empirical biases and empirical variances of the estimators are given in the Tables 4.1, 4.2 and 4.3. We observe that the biases and variances of the estimators are small and that as sample size increases the variance and bias of the estimators' decreases.

Table 4.1: Biases ($\times 10^3$) and Variances ($\times 10^5$) of the simple moment and maximum likelihood estimators of $F\ (t_1,t_2)$ at various time points when α=2.

(t_1,t_2)	(C_1,C_2)	M = 1000				M = 4000			
		Simple Moment		Maximum Likelihood		Simple Moment		Maximum Likelihood	
		Bias	Var	Bias	Var	Bias	Var	Bias	Var
(0.1,0.1)	(1,1)	-1.57	1.29	-9.86	.178	1.10	.736	-4.34	.148
	(1,2)	-0.81	1.27	-9.52	.189	0.15	.699	-4.38	.129
	(2,1)	-1.27	9.58	-9.78	.148	1.06	.626	-4.45	.108
	(2,2)	-1.68	1.03	-9.79	.162	1.42	.637	-4.20	.133
(0.1,0.2)	(1,1)	-5.00	2.46	-4.91	.392	3.18	1.38	-3.06	.308
	(1,2)	-4.24	2.32	-4.89	.359	3.60	1.47	-3.04	.301
	(2,1)	-4.82	1.85	-4.92	.321	3.24	1.12	-3.06	.254
	(2,2)	-5.32	1.95	-4.92	.309	3.20	1.01	-1.00	.237
(0.2,0.1)	(1,1)	-5.15	2.41	-1.93	.331	1.55	1.26	-1.04	.297
	(1,2)	-3.87	2.68	-1.87	.401	1.31	1.34	-1.05	.397
	(2,1)	-4.65	1.78	-1.92	.320	1.48	1.31	-1.04	.315
	(2,2)	-4.73	1.90	-1.90	.274	1.67	1.56	-1.03	.205
(0.3,0.3)	(1,1)	-2.59	12.3	-2.20	2.36	1.55	5.47	-2.13	1.63
	(1,2)	-2.39	9.64	-2.13	1.81	1.62	6.61	-2.09	1.05
	(2,1)	-2.62	10.07	-2.22	1.84	1.57	6.73	-2.13	0.99
	(2,2)	-2.52	9.77	-2.19	1.65	1.72	6.37	-2.11	1.00
(0.5,0.4)	(1,1)	-4.75	24.9	-1.26	5.21	1.75	11.1	-1.11	4.16
	(1,2)	-4.66	18.5	-1.26	4.00	1.88	12.6	-1.12	3.80
	(2,1)	-4.90	21 7	-1.27	3.52	1.75	13.6	-1.01	2.84
	(2,2)	-4.67	22.7	-1.26	4.70	2.04	12.1	-1.00	4.20
(0.4,0.5)	(1,1)	-4.81	25.2	-1.26	5.16	1.79	10.9	-1.18	4.42
	(1,2)	-4.61	18.7	-1.26	3.39	1.82	13.7	-1.18	3.33
	(2,1)	-4.90	21.2	-1.27	4.41	1.75	15.4	-1.11	4.39
	(2,2)	-4.72	23.0	-1.26	4.29	2.02	11.6	-.143	4.17

Table 4.2: Biases ($\times 10^3$) and Variances ($\times 10^5$) of the simple moment and maximum likelihood estimators of $F_{t_1,t_2}(t_1,t_2)$ at various time points when α=3.

(t_1,t_2)	(C_1,C_2)	M = 1000				M = 4000			
		Simple Moment		Maximum Likelihood		Simple Moment		Maximum Likelihood	
		Bias	Var	Bias	Var	Bias	Var	Bias	Var
(0.1,0.1)	(1,1)	6.50	3.64	-1.21	.954	4.79	2.08	1.09	.454
	(1,2)	6.00	3.93	-1.23	.826	4.84	2.56	1.14	.636
	(2,1)	4.42	3.55	-1.30	.903	3.87	2.39	1.04	.623
	(2,2)	5.09	2.98	-1.28	.828	4.80	2.04	1.22	.583
(0.1,0.2)	(1,1)	6.16	5.54	-2.37	1.78	4.05	4.41	2.25	1.68
	(1,2)	4.74	8.31	-2.45	1.86	4.10	4.26	2.36	1.50
	(2,1)	2.66	6.33	-2.53	1.61	2.16	4.12	2.19	1.24
	(2,2)	3.15	5.31	-2.56	1.26	3.07	4.20	2.30	1.17
(0.2,0.1)	(1,1)	6.17	6.29	-2.39	1.71	5.00	4.08	2.12	1.34
	(1,2)	4.42	6.55	-2.46	1.38	4.12	5.17	2.35	1.27
	(2,1)	3.08	6.91	-2.52	1.81	2.16	4.58	2.19	1.74
	(2,2)	4.63	5.54	-2.47	1.72	4.03	3.95	2.46	1.47
(0.3,0.3)	(1,1)	-4.65	27.5	-4.30	8.78	4.45	16.5	-2.50	6.77
	(1,2)	-7.83	28.8	-6.42	7.51	4.54	13.1	-4.27	5.34
	(2,1)	-12.4	22.4	-8.61	6.10	4.74	12.4	-4.56	4.91
	(2,2)	-9.06	23.0	-8.55	5.87	4.48	14.5	-4.21	5.79
(0.5,0.4)	(1,1)	-1.61	46.7	-1.27	14.3	0.93	23.9	-1.17	13.6
	(1,2)	-1.97	41.6	-1.28	11.4	1.00	21.3	-1.08	10.8
	(2,1)	-2.62	39.7	-1.31	10.8	1.02	20.6	-1.22	10.8
	(2,2)	-2.02	30.9	-1.29	8.58	1.88	21.1	-1.15	7.72
(0.4,0.5)	(1,1)	-1.66	43.2	-1.27	12.8	0.96	23.3	-1.12	11.5
	(1,2)	-1.96	42.8	-1.28	11.7	1.00	18.6	-1.16	10.4
	(2,1)	-2.63	37.9	-1.31	11.3	1.02	22.1	-1.19	10.7
	(2,2)	-2.10	31.1	-1.29	8.67	1.93	22.4	-1.13	7.5

Table 4.3: Biases ($\times 10^3$) and Variances ($\times 10^5$) of the simple moment and maximum likelihood estimators of $F_{t_1 t_2}(t_1, t_2)$ at various time points when α=5.

(t_1, t_2)	(C_1, C_2)	M = 1000				M = 4000			
		Simple Moment		Maximum Likelihood		Simple Moment		Maximum Likelihood	
		Bias	Var	Bias	Var	Bias	Var	Bias	Var
(0.1,0.1)	(1,1)	11.2	16.0	0.39	9.24	10.4	8.04	0.38	8.25
	(1,2)	10.7	17.6	0.20	9.88	10.5	8.25	0.185	7.14
	(2,1)	19.4	18.5	-0.61	9.97	10.6	10.6	0.567	7.57
	(2,2)	10.6	22.5	-0.17	11.1	10.5	9.51	0.158	10.7
(0.1,0.2)	(1,1)	18.2	26.0	-9.17	15.1	14.3	15.0	8.08	12.8
	(1,2)	18.2	24.9	-8.52	14.8	14.6	13.3	7.11	13.0
	(2,1)	16.7	23.3	-10.0	12.9	14.6	14.3	6.11	11.1
	(2,2)	17.9	32.6	-9.28	16.3	14.6	13.1	8.11	13.6
(0.2,0.1)	(1,1)	10.1	22.4	-7.51	14.9	9.55	14.6	6.09	10.0
	(1,2)	19.1	28.6	-8.43	16.1	14.5	13.2	6.11	11.4
	(2,1)	18.9	26.5	-8.06	15.2	14.7	17.0	7.11	13.4
	(2,2)	10.1	32.4	-7.73	17.2	9.67	14.7	7.11	14.3
(0.3,0.3)	(1,1)	27.2	44.5	-14.3	29.6	24.8	36.1	11.5	25.0
	(1,2)	28.8	59.5	-14.4	33.1	25.0	33.8	12.9	24.6
	(2,1)	26.8	46.4	-15.9	28.5	25.1	34.7	13.7	26.7
	(2,2)	29.4	63.7	-13.9	36.6	25.0	33.8	13.8	30.7
(0.5,0.4)	(1,1)	24.2	49.3	-16.2	33.2	13.7	41.3	12.9	26.6
	(1,2)	25.3	69.3	-16.9	41.1	14.8	36.4	12.5	32.9
	(2,1)	23.1	53.0	-18.2	32.5	16.8	41.8	12.4	28.5
	(2,2)	26.3	77.0	-15.6	44.8	16.1	35.9	13.3	38.6
(0.4,0.5)	(1,1)	24.0	52.9	-16.1	34.1	13.7	41.0	12.8	28.7
	(1,2)	25.0	72.7	-17.0	41.0	14.6	37.8	12.3	36.0
	(2,1)	23.2	55.5	-18.2	34.0	16.4	42.2	11.8	28.3
	(2,2)	25.8	77.3	-16.0	46.0	16.0	35.6	13.1	38.3

4.5 Conclusion

We developed non-parametric estimatior of lifetime distribution for univariate and bivariate competing risk models when the censoring times are missing but the censoring time distribution is known. The maximum likelihood and simple moment estimators of cause-specific sub-distribution function in univariate and bivariate set up are obtained. The estimators are generalization of the estimators given in Hu et.al. (1998). We also conducted a simulation study to assess the performance of the estimators. We observe that the biases and variances of the estimators are small and that as sample size increases the variance and bias of the estimators' decreases. The work done in this chapter is presented in Sankaran and Ansa (2005a).

Chapter 5

BIVARIATE COMPETING RISK MODELS UNDER RANDOM LEFT TRUNCATION AND RIGHT CENSORING

5.1 Introduction

In survival or reliability studies, it is common to have truncated data due to the limited time span of the study or dropouts of the subjects for various reasons. The estimator of survivor function under left truncation was first discussed by Kaplan and Meier by extending the well known product-limit estimator of the survivor function. Later, Lynden-Bell (1971) discussed the nonparametric estimation of survivor function under right truncation. Nonparametric estimation of survivor function for truncated data was also considered by Peto (1973) and Turnbull (1976). Efron and Petrosian (1999) discussed the analysis of lifetime data that are simultaneously left and right truncated. Examples of lifetime data that are left truncated and right censored is provided in Andersen et.al. (1993). The existing literature on the study of bivariate lifetime models under truncation is very limited. Most studies in bivariate set up refer to the case when only one component of the bivariate vector is subject to truncation (see Gurler, 1996, 1997). But in practical situations, we come across bivariate lifetime data where both the component lifetimes are under the possibility of truncation. For example, in the case of Swedish twin lifetime data, both twins had to be alive in a certain year in order to be included in the sample (Cederlof and Lorich, 1978). Accordingly, van der Laan (1996) developed a nonparametric maximum likelihood estimator for the bivariate distribution function where both components are randomly truncated. Recently, Gurler (2004) considered the diverse hazard vector, which is an extension of Dabrowska's (1988) hazard vector for censored observations, for the estimation of truncated data where both components are randomly truncated.

The estimation of the bivariate distribution function for truncated data under competing risk set up is not considered so far. Motivated by this, in this chapter, we address the problem of random left truncation in the presence of multiple causes of failure and right censoring, which is the generalization of Gurler (2004). The focus of this chapter is on the nonparametric estimation of the survivor function and the cause-specific sub-distribution functions in bivariate competing risk set up, when the observations are subject to random left truncation and right censoring. In Section 5.2, we consider the bivariate truncation model where both the components of the lifetime vector are subject to random left truncation and they are exposed to multiple causes of failure. Section 5.3 gives nonparametric estimators for the bivariate survivor function and the cause-specific sub-distribution functions. In Section 5.4, we discuss various asymptotic properties of the estimators. Section 5.5 presents a simulation study discussing the empirical behaviour of the estimator. The chapter ends with a brief summary in Section 5.6.

5.2 Bivariate Competing Risk Model for Random Left Truncated and Right Censored Data

Let $T = (T_1, T_2)$ be a pair of non-negative random variables defined on a probability space $(\Omega, \mathbb{F}, \mathrm{P})$. Let $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$ be the survivor function of T. We assume that each component of (T_1, T_2) is exposed to more than one cause of failure. Let $C = (C_1, C_2)$ represents the cause of failure corresponding to $T = (T_1, T_2)$. If there are γ_i causes of failure corresponding to T_i, then C_i is a unique element of $\{1, 2 \ldots \gamma_i\}$; $i = 1, 2$. Let $L = (L_1, L_2)$ be a non-negative random vector representing the truncating variables corresponding to $T = (T_1, T_2)$. Let $Z = (Z_1, Z_2)$ be a pair of random censoring time corresponding to (T_1, T_2). Assume that the truncating vector L and censoring vector Z has a joint density such that $(L_1 < Z_1, L_2 < Z_2)$ with probability one and denote $G(t_1, t_2) = P(L_1 < t_1 < Z_1, L_2 < t_2 < Z_2)$. We also assume that the truncating vector L and censoring vector Z are independent of $T = (T_1, T_2)$ and that there is a positive probability that $(T_1 > L_1, T_2 > L_2)$. Now, one observes

$(Y_1, Y_2, \delta_1, \delta_2, C_1, C_2, L_1, L_2)$ only if $(Y_1 > L_1, Y_2 > L_2)$ where $Y_i = \min(T_i, Z_i)$ is the observed lifetime or censoring time and $\delta_i = I(T_i = Y_i)$ is the censoring indicator for $i = 1,2$. Since L and Z are independent of T, we obtain

$$P(L_1 < t_1 < Y_1, L_2 < t_2 < Y_2) = G(t_1, t_2) S(t_1, t_2). \tag{5.1}$$

Define $$S_i^{(1)}(t_1, t_2) = P(T_1 \geq t_1, T_2 > t_2, C_1 = i), \tag{5.2}$$

$$S_j^{(2)}(t_1, t_2) = P(T_1 > t_1, T_2 \geq t_2, C_2 = j), \tag{5.3}$$

and

$$S_{ij}^{(12)}(t_1, t_2) = P(T_1 \geq t_1, T_2 \geq t_2, C_1 = i, C_2 = j); \quad i = 1,2 \ldots \gamma_1 \; ; \; j = 1,2 \ldots \gamma_2. \tag{5.4}$$

Combining equations (5.2), (5.3) and (5.4) with (2.6), (2.7), (2.8), (2.9) and (2.25), we get

$$\Lambda_i^{(1)}(dt_1, t_2) = \frac{-S_i^{(1)}(dt_1, t_2)}{S(t_1^-, t_2)}, \tag{5.5}$$

$$\Lambda_j^{(2)}(t_1, dt_2) = \frac{-S_j^{(2)}(t_1, dt_2)}{S(t_1, t_2)}, \tag{5.6}$$

and

$$\Lambda_{ij}(dt_1, dt_2) = \frac{S_{ij}^{(12)}(dt_1, dt_2)}{S(t_1^-, t_2)}; \quad i = 1,2 \ldots \gamma_1 \; ; \; j = 1,2 \ldots \gamma_2. \tag{5.7}$$

Let $\alpha = P(Y_1 > L_1, Y_2 > L_2)$. Let $(\Omega^*, \mathbb{F}^*, P^*)$ be the probability space conditional on $(Y_1 > L_1, Y_2 > L_2)$

Define the functions

$$\begin{aligned} S^*(t_1, t_2) &= P^*(L_1 < t_1 < Y_1, L_2 < t_2 < Y_2) \\ &= P(L_1 < t_1 < Y_1, L_2 < t_2 < Y_2 \mid Y_1 > L_1, Y_2 > L_2), \end{aligned} \tag{5.8}$$

$$\begin{aligned} S_i^{(1)*}(t_1, t_2) &= P^*(L_1 < t_1 \leq Y_1, L_2 < t_2 < Y_2, \delta_1 = 1, C_1 = i) \\ &= P(L_1 < t_1 \leq Y_1, L_2 < t_2 < Y_2, \delta_1 = 1, C_1 = i \mid Y_1 > L_1, Y_2 > L_2), \end{aligned} \tag{5.9}$$

$$S_j^{(2)^*}(t_1,t_2) = P^*(L_1 < t_1 < Y_1, L_2 < t_2 \le Y_2, \delta_2 = 1, C_2 = j)$$

$$= P(L_1 < t_1 < Y_1, L_2 < t_2 < Y_2, \delta_2 = 1, C_2 = j \mid Y_1 > L_1, Y_2 > L_2), \quad (5.10)$$

and

$$S_{ij}^{(12)^*}(t_1,t_2) = P^*(L_1 < t_1 \le Y_1, L_2 < t_2 \le Y_2, \delta_1 = \delta_2 = 1, C_1 = i, C_2 = j)$$

$$= P(L_1 < t_1 \le Y_1, L_2 < t_2 \le Y_2, \delta_1 = \delta_2 = 1, C_1 = i, C_2 = j \mid Y_1 > L_1, Y_2 > L_2) \quad (5.11)$$

$$i = 1,2\ldots\gamma_1 \; ; \; j = 1,2\ldots\gamma_2 .$$

Since L and Z are independent of T and C, from (5.1), (5.8), (5.9), (5.10) and (5.11), we obtain,

$$S^*(t_1,t_2) = \alpha^{-1}G(t_1,t_2)S(t_1,t_2), \quad (5.12)$$

$$S_i^{(1)^*}(dt_1,t_2) = \alpha^{-1}G(t_1^-,t_2)S_i^{(1)}(dt_1,t_2), \quad (5.13)$$

$$S_j^{(2)^*}(t_1,dt_2) = \alpha^{-1}G(t_1,t_2^-)S_j^{(2)}(t_1,dt_2), \quad (5.14)$$

and

$$S_{ij}^{(12)^*}(dt_1,dt_2) = \alpha^{-1}G(t_1^-,t_2^-)S_{ij}^{(12)}(dt_1,dt_2); \qquad i = 1,2\ldots\gamma_1 \; ; \; j = 1,2\ldots\gamma_2 \quad (5.15)$$

where

$$G(t_1^-,t_2) = P(L_1 < t_1 \le Z_1, L_2 < t_2 < Z_2),$$

$$G(t_1,t_2^-) = P(L_1 < t_1 < Z_1, L_2 < t_2 \le Z_2),$$

and

$$G(t_1^-,t_2^-) = P(L_1 < t_1 \le Z_1, L_2 < t_2 \le Z_2).$$

From (5.5), (5.12) and (5.13), we get

$$\Lambda_i^{(1)}(dt_1,t_2) = \frac{-S_i^{(1)^*}(dt_1,t_2)}{S^*(t_1^-,t_2)}; \qquad i = 1,2\ldots\gamma_1 \quad (5.16)$$

From (5.6), (5.12) and (5.14), we get

$$\Lambda_j^{(2)}(t_1,dt_2) = \frac{-S_j^{(2)^*}(t_1,dt_2)}{S^*(t_1,t_2^-)}; \qquad j = 1,2\ldots\gamma_2 . \quad (5.17)$$

From (5.7), (5.12) and (5.15), we get

$$\Lambda_{ij}(dt_1,dt_2)=\frac{S_{ij}^{(12)^*}(dt_1,dt_2)}{S^*(t_1^-,t_2^-)};\qquad i=1,2\ldots\gamma_1\ ;\ j=1,2\ldots\gamma_2. \tag{5.18}$$

Let $\mathfrak{A}=[0,b_1]\times[0,b_2]$ with $S^*(b_1^-,b_2^-)>0$

From (2.8), (2.10) and (5.16), for all $(t_1,t_2)\in\mathfrak{A}$, we get

$$\Lambda_1(t_1,t_2)=\sum_{i=1}^{\gamma_1}\int_0^{t_1}\Lambda_i^{(1)}(du,t_2)=-\sum_{i=1}^{\gamma_1}\int_0^{t_1}\frac{S_i^{(1)^*}(du,t_2)}{S^*(u^-,t_2)} \tag{5.19}$$

Similarly, from (2.9), (2.11) and (5.17), for all $(t_1,t_2)\in\mathfrak{A}$, we get

$$\Lambda_2(t_1,t_2)=\sum_{j=1}^{\gamma_2}\int_0^{t_2}\Lambda_j^{(2)}(t_1,dv)=-\sum_{j=1}^{\gamma_2}\int_0^{t_2}\frac{S_j^{(2)^*}(t_1,dv)}{S^*(t_1,v^-)}. \tag{5.20}$$

and from (5.18), for all $(t_1,t_2)\in\mathfrak{A}$, we obtain

$$\Lambda_{ij}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{S_{ij}^{(12)^*}(du,dv)}{S^*(u^-,v^-)};\qquad i=1,2\ldots\gamma_1\ ;\ j=1,2\ldots\gamma_2. \tag{5.21}$$

From (2.28) and (5.21), we get

$$F_{ij}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}S(u^-,v^-)\Lambda_{ij}(du,dv);\ i=1,2\ldots\gamma_1\quad j=1,2\ldots\gamma_2. \tag{5.22}$$

5.3 Nonparametric Estimation

Now suppose that $(Y_{1l},Y_{2l},\delta_{1l},\delta_{2l},C_{1l},C_{2l},L_{1l},L_{2l})$, $l=1,2\ldots n$ be an i.i.d sample from the conditional probability space $(\Omega^*,\mathbb{F}^*,\mathrm{P}^*)$. To estimate the bivariate survivor function $S(t_1,t_2)$ and the cause-specific distribution functions $F_{ij}(t_1,t_2)$; $i=1,2\ldots\gamma_1$; $j=1,2\ldots\gamma_2$, define

$$\hat{S}^*(t_1,t_2)=\frac{1}{n}\sum_l I(L_{1l}<t_1<Y_{1l},L_{2l}<t_2<Y_{2l}),$$

$$\hat{S}_i^{(1)^*}(t_1,t_2)=\frac{1}{n}\sum_l I(L_{1l}<t_1\le Y_{1l},L_{2l}<t_2<Y_{2l},\delta_{1l}=1,C_{1l}=i),$$

$$\hat{S}_I^{(2)^*}(t_1,t_2) = \frac{1}{n}\sum_I I\left(L_{1I} < t_1 < Y_{1I}, L_{2I} < t_2 \le Y_{2I}, \delta_{2I} = 1, C_{2I} = j\right).$$

and

$$\hat{S}_{II}^{(12)^*}(t_1,t_2) = \frac{1}{n}\sum_I I\left(L_{1I} < t_1 \le Y_{1I}, L_{2I} < t_2 \le Y_{2I}, \delta_{1I} = \delta_{2I} = 1, C_{1I} = i, C_{2I} = j\right);$$

$$i = 1,2\ldots\gamma_1 \;;\; j = 1,2\ldots\gamma_2,$$

with $I(.)$ as the usual indicator function.

Then from (5.19) and (5.20), the estimators of $\Lambda_1(t_1,t_2)$ and $\Lambda_2(t_1,t_2)$ are obtained as

$$\hat{\Lambda}_1^{I.}(t_1,t_2) = -\sum_{i=1}^{\gamma_1}\int_0^{t_1}\frac{\hat{S}_I^{(1)^*}(du,t_2)}{\hat{S}^*(u\ ,t_2)} \tag{5.23}$$

for all (t_1,t_2) such that $\hat{S}^*(t_1^-,t_2) > 0$ and 0 otherwise, and

$$\hat{\Lambda}_2^{I.}(t_1,t_2) = -\sum_{I=1}^{\gamma_2}\int_0^{t_2}\frac{\hat{S}_I^{(2)^*}(t_1,dv)}{\hat{S}^*(t_1,v^-)} \tag{5.24}$$

for all (t_1,t_2) such that $\hat{S}^*(t_1,t_2^-) > 0$ and 0 otherwise.

Thus from (2.4), (5.23) and (5.24), the estimator of $S(t_1,t_2)$ is obtained as

$$\hat{S}_1^{I.}(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1^{I.}(t_1,0) - \hat{\Lambda}_2^{I.}(t_1,t_2)\right\} \tag{5.25}$$

and

$$\hat{S}_2^{I.}(t_1,t_2) = \exp\left\{-\hat{\Lambda}_1^{I.}(t_1,t_2) - \hat{\Lambda}_2^{I.}(0,t_2)\right\} \tag{5.26}$$

The estimator of $S(t_1,t_2)$ obtained by (5.25) and (5.26) may be different. As mentioned in Section 2.3.1, the proposed estimator for $S(t_1,t_2)$ is a convex combination of two expressions (5.25) and (5.26). Thus the estimator for $S(t_1,t_2)$ is given by

$$\hat{S}_c^{I.}(t_1,t_2) = c(t_1,t_2)\hat{S}_1^{I.}(t_1,t_2) + (1 - c(t_1,t_2))\hat{S}_2^{I.}(t_1,t_2). \tag{5.27}$$

As given in Section 2.3.1, $c(t_1,t_2)$ is chosen such that the mean squared error (MSE) of $\hat{S}_c^{L}(t_1,t_2)$ is minimum.

From (5.21) and (5.22), the estimator of $F_{ij}(t_1,t_2)$ is

$$\hat{F}_{ij}^{L}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\hat{S}_c^{L}(u^-,v^-)\hat{\Lambda}_{ij}^{L}(du,dv), \tag{5.28}$$

where

$$\Lambda_{ij}^{L}(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}\frac{\hat{S}_{ij}^{(12)^*}(du,dv)}{\hat{S}^{*}(u^-,v^-)}; \qquad i=1,2\ldots\gamma_1 \quad j=1,2\ldots\gamma_2, \tag{5.29}$$

for all (t_1,t_2) such that $\hat{S}^{*}(t_1^-,t_2^-)>0$ and 0 otherwise.

Remark 5.1 The proposed method is a generalization of the method given in Gurler (2004) to the competing risk set up where both the components are subject to random left truncation and right censoring.

5.4 Properties of Estimators

In this section we prove the consistency and weak convergence of the estimators of $S(t_1,t_2)$ and $F_{ij}(t_1,t_2)$

Theorem 5.1

Suppose that both L and Z are independent of T and C. Then $\underset{\mathfrak{A}}{Sup}\left\|\hat{\Lambda}_k^{L}-\Lambda_k\right\|\to 0$ almost surely; k =1,2.

Proof

When k =1, from (2.10) and (5.23),

$$\underset{\mathfrak{A}}{Sup}\left\|\hat{\Lambda}_1^{L}-\Lambda_1\right\| = \underset{\mathfrak{A}}{Sup}\left\|\sum_{i=1}^{\gamma_1}\left(\hat{\Lambda}_i^{L(1)}-\Lambda_i^{(1)}\right)\right\|$$

$$\le\sum_{i=1}^{\gamma_1}\underset{\mathfrak{A}}{Sup}\left\|\hat{\Lambda}_i^{L(1)}-\Lambda_i^{(1)}\right\| \tag{5.30}$$

where $\hat{\Lambda}_i^{L(1)}(t_1,t_2) = \int_0^{t_1} \frac{\hat{S}_i^{(1)^*}(du,t_2)}{\hat{S}^*(u^-,t_2)}$; $\quad i = 1,2\ldots\gamma_1$

For fixed i, consistency of $\hat{\Lambda}_i^{L(1)}$ follows from the Glivenko-Cantelli theorem. Then, from (5.30), $\underset{\mathfrak{A}}{Sup}\left\|\hat{\Lambda}_i^{L}-\Lambda_i\right\| \to 0$. Thus the strong consistency of $\hat{\Lambda}_1^{L}$ is proved. The proof for $\hat{\Lambda}_2^{L}$ is similar.

Theorem 5.2

Under the assumptions of Theorem 5.1, $\underset{\mathfrak{A}}{Sup}\left\|\hat{S}_c^{L}-S\right\| \to 0$ almost surely.

The proof is similar to that of Theorem 2.3.

Theorem 5.3

Under the assumptions of Theorem 5.1, $\underset{\mathfrak{A}}{Sup}\left\|\hat{F}_{ij}^{L}-F_{ij}\right\| \to 0$ almost surely for every i =1,2... γ_1 and j =1,2... γ_2.

Proof

Using (5.28),

$$\hat{F}_{ij}^{L}(t_1,t_2)-F_{ij}(t_1,t_2) = \int_0^{t_1}\int_0^{t_2}\left\{\hat{S}_c^{L}(u_1^-,u_2^-)\left[\hat{\Lambda}_{ij}^{L}(du_1,du_2)-\Lambda_{ij}(du_1,du_2)\right]\right.$$

$$\left.+\left[\hat{S}_c^{L}(u_1^-,u_2^-)-S(u_1^-,u_2^-)\right]\Lambda_{ij}(du_1,du_2)\right\} \qquad (5.31)$$

For fixed i and j, the strong consistency of $\hat{\Lambda}_{ij}^{L}$ follows from the Glivenko-Cantelli theorem. From (5.31) and Theorem 5.2, it follows that $\underset{\mathfrak{A}}{Sup}\left\|\hat{F}_{ij}^{L}-F_{ij}\right\| \to 0$ almost surely; i =1,2... γ_1 and j =1,2... γ_2.

Theorem 5.4

Under the assumptions of Theorem 5.1, for all $(t_1,t_2)\in\mathfrak{A}$, $\sqrt{n}\left(\hat{S}_c^{L}(t_1,t_2)-S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance given by (5.33).

Proof

From (5.25) and (5.26), we can have

$$\hat{S}_1^{I.}(t_1,t_2) = \hat{S}_2^{I.}(t_2 \mid t_1)\hat{S}_1^{I.}(t_1) \text{ and } \hat{S}_2^{I.}(t_1,t_2) = \hat{S}_1^{I.}(t_1 \mid t_2)\hat{S}_2^{I.}(t_2)$$

where

$$\hat{S}_k^{I.}(t_k \mid t_m) = \exp\left\{-\hat{\Lambda}_k\ (t_1,t_2)\right\};\ k,m = 1,2;\ k \neq m$$

$$\hat{S}_1^{I.}(t_1) = \exp\left\{-\hat{\Lambda}_1^{I.}(t_1,0)\right\} \text{ and } \hat{S}_2^{I.}(t_2) = \exp\left\{-\hat{\Lambda}_2^{I.}(0,t_2)\right\}$$

Since $S(t_1,t_2) = S_k(t_k \mid t_m)\, S_k(t_k)$; k, m=1,2; $k \neq m$, $S_1(t_1) = \exp\{-\Lambda_1(t_1,0)\}$ and $S_2(t_2) = \exp\{-\Lambda_2(0,t_2)\}$, we have

$$\sqrt{n}\left(\hat{S}_1^{I.}(t_1,t_2) - S(t_1,t_2)\right) = \sqrt{n}\left(\hat{S}_2^{I.}(t_2 \mid t_1) - S_2(t_2 \mid t_1)\right)\hat{S}_1^{I.}(t_1) +$$

$$\sqrt{n}\, S_2(t_2 \mid t_1)\left(\hat{S}_1^{I.}(t_1) - S_1(t_1)\right). \qquad (5.32)$$

Now, $\hat{\Lambda}_2^{I.}(t_1,t_2) - \Lambda_2(t_1,t_2) = \sum_{l=1}^{\gamma_2} \int_0^{t_2} \frac{I(L_1 \le t_1 < Y_1, L_2 \le u \le Y_2)}{\hat{S}^{\bullet}(t_1,u^-)} dM_{2l}^{I.}(t_1,u)$

where for fixed t_1, $M_{2l}^{I.}(t_1,u) = \hat{H}_{2l}^{ucI.}(t_1,u) - \int_0^{t_2} \hat{S}^{\bullet}(t_1,u^-)\, d\Lambda_l^{(2)}(t_1,u)$ is the martingale representation of $\Lambda_l^{(2)}(t_1,u)$ with $\hat{H}_{2l}^{ucI.}(t_1,u)$ is the estimator of $H_{2l}^{ucI.}(t_1,u) = P(L_1 < t_1 < Y_1, L_2 < u \le Y_2, \delta_2 = 1, C_2 = j \mid Y_1 > L_1, Y_2 > L_2)$. Then for fixed t_1, $\sqrt{n}\left(\hat{\Lambda}_2^{I.}(t_1,t_2) - \Lambda_2(t_1,t_2)\right)$ converges to a Gaussian process with mean zero. Further, for fixed t_1, $S_2(t_2 \mid t_1) = \prod_{s \le t_2}(1 - d\Lambda_2(t_1,s))$ is the Hadamard differentiability of the product integral. Thus the asymptotic normality of $\hat{\Lambda}_2^{I.}(t_1,t_2)$ carries over to the asymptotic normality of the corresponding estimators of $S_2(t_2 \mid t_1)$. This shows that $\sqrt{n}\left(\hat{S}_2^{I.}(t_2 \mid t_1) - S_2(t_2 \mid t_1)\right)$ converges to a Gaussian process with mean zero. Since $\hat{S}_1^{I.}(t_1)$ is strongly consistent and $\sqrt{n}\left(\hat{S}_2^{I.}(t_2 \mid t_1) - S_2(t_2 \mid t_1)\right)$ converges

weakly, the first factor of (5.32) weakly converges to a Gaussian process. Since $S_2(t_2 \mid t_1)$ is bounded and $\sqrt{n}\left(\hat{S}_1^{I.}(t_1) - S_1(t_1)\right)$ converges weakly, the second factor converges weakly to a Gaussian process. Thus, $\sqrt{n}\left(\hat{S}_1^{I.}(t_1,t_2) - S(t_1,t_2)\right)$ weakly converges to a Gaussian process. On similar lines we can show that $\sqrt{n}\left(\hat{S}_2^{I.}(t_1,t_2) - S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero. Thus, $\sqrt{n}\left(\hat{S}_c^{I.}(t_1,t_2) - S(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and the asymptotic variance of $\sqrt{n}\,\hat{S}_c^{I.}(t_1,t_2)$ is

$$\begin{aligned}\sigma^{(I.)2}(t_1,t_2) &= c^2(t_1,t_2)\sigma_1^{(I.)2}(t_1,t_2) + \left(1-c(t_1,t_2)\right)^2 \sigma_2^{(I.)2}(t_1,t_2) \\ &\quad + c(t_1,t_2)\left(1-c(t_1,t_2)\right)\sigma_{12}^{I.}(t_1,t_2), \end{aligned} \tag{5.33}$$

where $\sigma_k^{(I.)2}(t_1,t_2)$ is the asymptotic variance of $\sqrt{n}\,\hat{S}_k^{I.}(t_1,t_2)$ $k = 1,2$ and $\sigma_{12}^{I.}(t_1,t_2)$ is the asymptotic covariance between $\sqrt{n}\,\hat{S}_1^{I.}(t_1,t_2)$ and $\sqrt{n}\,\hat{S}_2^{I.}(t_1,t_2)$. From Appendix B of Lawless(2003), the asymptotic variance of $\sqrt{n}\,\hat{S}_1^{I.}(t_1,t_2)$ is

$$\sigma_1^{(I.)2}(t_1,t_2) = \left(S_1^{I.}(t_1,t_2)\right)^2 \; As\,\mathrm{var}\left(\log \hat{S}_1^{I.}(t_1,t_2)\right).$$

Now,

$$\sqrt{n}\left(\log \hat{S}_1^{I.}(t_1,t_2) - \log S(t_1,t_2)\right) = \sqrt{n}\left(\hat{\Lambda}_1^{I.}(t_1,0) - \Lambda_1(t_1,0) + \hat{\Lambda}_2^{I.}(t_1,t_2) - \Lambda_2(t_1,t_2)\right).$$

Thus the asymptotic variance of $\sqrt{n}\left(\log \hat{S}_1^{I.}(t_1,t_2) - \log S(t_1,t_2)\right)$ is given by

$$E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{S^*(u_1^-,0)} I\left(L_{1i} < Y_{1i}, Y_{1i} \in du_1, \delta_{1i} = 1, C_{1i} = i\right) + \right.\right.$$

$$\left.\left. I\left(L_{1i} < u_1 \le Y_{1i}\right)\frac{S_i^{(1)*}(du_1,0)}{\left(S^*(u_1^-,0)\right)^2}\right\}\right] +$$

$$\sum_{j=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{S^*\left(t_1,v_2^-\right)}I\left(L_{1l}<t_1<Y_{1l},L_{1l}<Y_{2l},Y_{2l}\in dv_2,\delta_{2l}=1,C_{2l}=j\right)\right.\right.$$

$$\left.\left.+I\left(L_{1l}<t_1<Y_{1l},L_{2l}<v_2\le Y_{2l}\right)\frac{S_j^{(2)*}\left(t_1,dv_2\right)}{\left(S^*\left(t_1,v_2^-\right)\right)^2}\right\}\right]\Bigg)^2$$

Thus,

$$\sigma_1^{(L)2}\left(t_1,t_2\right)=\left(S_1^{L.}\left(t_1,t_2\right)\right)^2 E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{S^*\left(u_1^-,0\right)}I\left(L_{1l}<Y_{1l},Y_{1l}\in du_1,\delta_{1l}=1,C_{1l}=i\right)+\right.\right.\right.$$

$$\left.\left.\left.I\left(L_{1l}<u_1\le Y_{1l}\right)\frac{S_i^{(1)*}\left(du_1,0\right)}{\left(S^*\left(u_1^-,0\right)\right)^2}\right\}\right]+$$

$$\sum_{j=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{S^*\left(t_1,v_2^-\right)}I\left(L_{1l}<t_1<Y_{1l},L_{1l}<Y_{2l},Y_{2l}\in dv_2,\delta_{2l}=1,C_{2l}=j\right)\right.\right.$$

$$\left.\left.\left.+I\left(L_{1l}<t_1<Y_{1l},L_{2l}<v_2\le Y_{2l}\right)\frac{S_j^{(2)*}\left(t_1,dv_2\right)}{\left(S^*\left(t_1,v_2^-\right)\right)^2}\right\}\right]\right)^2$$

Similarly the asymptotic variance of $\sqrt{n}\left(\log\hat{S}_2^{L.}\left(t_1,t_2\right)-\log S\left(t_1,t_2\right)\right)$ is given by

$\sigma_2^{(L)2}\left(t_1,t_2\right)=$

$$\left(S_2^{L.}\left(t_1,t_2\right)\right)^2 E\left(\sum_{i=1}^{\gamma_1}\left[\int_0^{t_1}\left\{\frac{1}{S^*\left(u_1^-,t_2\right)}I\left(L_{1l}<Y_{1l},Y_{1l}\in du_1,L_{2l}<t_2<Y_{2l},\delta_{1l}=1,C_{1l}=i\right)\right.\right.$$

$$\left.\left.+\ I\left(L_{1l}<u_1\le Y_{1l},L_{2l}<t_2<Y_{2l}\right)\frac{S_i^{(1)*}\left(du_1,t_2\right)}{\left(S^*\left(u_1^-,t_2\right)\right)^2}\right\}\right]$$

$$+\sum_{l=1}^{\gamma_2}\left[\int_0^{t_2}\left\{\frac{1}{S^*\left(0,v_2^-\right)}I\left(L_{2l}<Y_{2l},Y_{2l}\in dv_2,\delta_{2l}=1,C_{2l}=j\right)+\right.\right.$$

$$\left.\left.I\left(L_{2l}\le v_2\le Y_{2l}\right)\frac{S_l^{(2)*}\left(0,dv_2\right)}{\left(S^*\left(0,v_2^-\right)\right)^2}\right\}\right]\Bigg)^2$$

The asymptotic covariance between $\hat{S}_1^{\,l}\left(t_1,t_2\right)$ and $\hat{S}_2^{\,l}\left(t_1,t_2\right)$ is

$$\sigma_{12}^{\,l}\left(t_1,t_2\right)=S_1^{\,l}\left(t_1,t_2\right)S_2^{\,l}\left(t_1,t_2\right)\ As\operatorname{cov}\left(\log\hat{S}_1^{\,l}\left(t_1,t_2\right),\log\hat{S}_2^{\,l}\left(t_1,t_2\right)\right).$$

which is given by

$$\sigma_{12}^{\,l}\left(t_1,t_2\right)=nS_1^{\,l}\left(t_1,t_2\right)S_2^{\,l}\left(t_1,t_2\right)\left\{E\left[\left(\hat{\Lambda}_1^{\,l}\left(t_1,t_2\right)-\Lambda_1\left(t_1,t_2\right)\right)\left(\hat{\Lambda}_1^{\,l}\left(t_1,0\right)-\Lambda_1\left(t_1,0\right)\right)+\right.\right.$$

$$\left(\hat{\Lambda}_1^{\,l}\left(t_1,t_2\right)-\Lambda_1\left(t_1,t_2\right)\right)\left(\hat{\Lambda}_2^{\,l}\left(t_1,t_2\right)-\Lambda_2\left(t_1,t_2\right)\right)+$$

$$\left(\hat{\Lambda}_2^{\,l}\left(0,t_2\right)-\Lambda_2\left(0,t_2\right)\right)\left(\hat{\Lambda}_1^{\,l}\left(t_1,0\right)-\Lambda_1\left(t_1,0\right)\right)+$$

$$\left.\left.\left(\hat{\Lambda}_2^{\,l}\left(t_1,t_2\right)-\Lambda_2\left(t_1,t_2\right)\right)\left(\hat{\Lambda}_2^{\,l}\left(0,t_2\right)-\Lambda_2\left(0,t_2\right)\right)\right]\right\}$$

Theorem 5.5

Under the assumptions of Theorem 5.1, for all $\left(t_1,t_2\right)\in\mathfrak{A}$, $\sqrt{n}\left(\hat{F}_{ij}^{\,l}\left(t_1,t_2\right)-F_{ij}\left(t_1,t_2\right)\right)$ converges weakly to a Gaussian process with mean zero and variance given by (5.38); $i=1,2\ldots\ \gamma_1$; $j=1,2\ldots\ \gamma_2$.

Proof

$$\sqrt{n}\left(\hat{F}_{ij}^{\,l}\left(t_1,t_2\right)-F_{ij}\left(t_1,t_2\right)\right)=\sqrt{n}\int_0^{t_1}\int_0^{t_2}\left(\frac{\hat{S}_c^{\,l}\left(u_1^-,u_2\right)\hat{S}_{ij}^{(12)*}\left(du_1,du_2\right)}{\hat{S}^*\left(u_1^-,u_2^-\right)}\right.$$

$$\left.-\frac{S\left(u_1^-,u_2\right)S_{ij}^{(12)}\left(du_1,du_2\right)}{S^*\left(u_1^-,u_2^-\right)}\right)$$

$$= \sqrt{n} \int_0^{t_1}\int_0^{t_2}\left(\frac{\hat{S}_c^{I.}\left(u_1^-, u_2^-\right)\hat{S}_{ij}^{(12)^*}\left(du_1, du_2\right)}{\hat{S}^*\left(u_1^-, u_2^-\right)} - \right.$$

$$\frac{\hat{S}_c^{I.}\left(u_1^-, u_2^-\right)\hat{S}_{ij}^{(12)^*}\left(du_1, du_2\right)}{S^*\left(u_1^-, u_2^-\right)} + \frac{\hat{S}_c^{I.}\left(u_1^-, u_2^-\right)\hat{S}_{ij}^{(12)^*}\left(du_1, du_2\right)}{S^*\left(u_1^-, u_2^-\right)}$$

$$\left. - \frac{S\left(u_1^-, u_2^-\right)S_{ij}^{(12)^*}\left(du_1, du_2\right)}{S^*\left(u_1^-, u_2^-\right)} \right)$$

$$= \sqrt{n} \int_0^{t_1}\int_0^{t_2}\left(-\left(\hat{S}^*\left(u_1^-, u_2^-\right) - S^*\left(u_1^-, u_2^-\right)\right)\frac{\hat{S}_c^{I.}\left(u_1^-, u_2^-\right)\hat{S}_{ij}^{(12)^*}\left(du_1, du_2\right)}{\hat{S}^*\left(u_1^-, u_2^-\right)S^*\left(u_1^-, u_2^-\right)} \right.$$

$$\left. + \frac{1}{S^*\left(u_1^-, u_2^-\right)}\left(\hat{S}_c^{I.}\left(u_1^-, u_2^-\right)\hat{S}_{ij}^{(12)^*}\left(du_1, du_2\right) - S_c^{I.}\left(u_1^-, u_2^-\right)S_{ij}^{(12)^*}\left(du_1, du_2\right)\right)\right). \quad (5.34)$$

Since $\underset{\Re}{Sup}\left\|\hat{S}_c^{I.} - S\right\| \to 0$ and $\underset{\Re}{Sup}\left\|\hat{F}_{ij}^{I.} - F_{ij}\right\| \to 0$, for large n, (5.34) becomes

$$\sqrt{n} \int_0^{t_1}\int_0^{t_2}\left(\frac{S\left(u_1^-, u_2^-\right)}{S^*\left(u_1^-, u_2^-\right)}\left(\frac{1}{n}\sum_{l=1}^{n} I\left(L_{1l} < Y_{1l}, Y_{1l} \in du_1, L_{2l} < Y_{2l}, Y_{2l} \in du_2, \delta_{1l} = \delta_{2l} = 1, C_{1l} = i, C_{2l} = j\right)\right.\right.$$

$$\left. - S_{ij}^{(12)^*}\left(du_1, du_2\right)\right) - \left(\frac{1}{n}\sum_{l=1}^{n} I\left(L_{1l} < u_1 < Y_{1l}, L_{2l} < u_2 < Y_{2l}\right) - S^*\left(u_1^-, u_2^-\right)\right)$$

$$\left. \frac{S\left(u_1^-, u_2^-\right)S_{ij}^{(12)^*}\left(du_1, du_2\right)}{\left(S^*\left(u_1^-, u_2^-\right)\right)^2} \right). \quad (5.35)$$

By multivariate central limit theorem, the terms in the simple brackets of the first and second integrals converges to a mean zero normal variate. Then the asymptotic normality of (5.35) follows from the delta method using the maps $\left(x_1, x_2\right) \to \left(x_1 + x_2\right)$ (van der Vaart and Wellner (1996)).

To find the asymptotic variance, let

$$A^{I.}=\int_0^{t_1}\int_0^{t_2}\frac{S\left(u_1^-,u_2\right)}{S^*\left(u_1^-,u_2^-\right)}I\left(L_{1l}<Y_{1l},Y_{1l}\in du_1,L_{2l}<Y_{2l},Y_{2l}\in du_2,\delta_{1l}=\delta_{2l}=1,C_{1l}=i,C_{2l}=j\right)$$

(5.36)

and

$$B^{I.}=\int_0^{t_1}\int_0^{t_2}I\left(L_{1l}<u_1\le Y_{1l},L_{2l}<u_2\le Y_{2l}\right)\frac{S\left(u_1,u_2\right)S_{ij}^{(12)*}\left(du_1,du_2\right)}{\left(S^*\left(u_1^-,u_2^-\right)\right)^2}. \tag{5.37}$$

Thus, the asymptotic variance of the process is

$$\sigma^{*2}\left(t_1,t_2\right)=E\left(A^{I.}-B^{I.}\right)^2 \tag{5.38}$$

and a consistent estimator of variance is

$$\hat{\sigma}^{*2}\left(t_1,t_2\right)=\frac{1}{n}\sum\left(\hat{A}^{I.}-\hat{B}^{I.}\right)^2$$

where $\hat{A}^{I.}$ and $\hat{B}^{I.}$ are obtained using (5.36) and (5.37) replacing the unknown quantities by their estimators.

5.5 A Simulation Study

To assess the empirical performance of the estimators of $S(t_1,t_2)$, we carried out a series of 1000 simulations of size n from a Gumbel's (1960) bivariate exponential distribution with survivor function

$$S\left(t_1,t_2\right)=\exp\left\{-t_1-t_2-\lambda t_1t_2\right\},0<t_1,t_2<\infty,0\le\lambda\le1 \tag{5.39}$$

with $\lambda=3$ and 5 with various sample sizes n = 20, 50 and 100.

We generated censoring times $\left(Z_1,Z_2\right)$ from the Gumbel's (1960) bivariate exponential distribution with survivor function

$$G\left(t_1,t_2\right)=\exp\left\{-t_1-t_2-\theta t_1t_2\right\},0<t_1,t_2<\infty,\ 0\le\theta\le1 \tag{5.40}$$

with $\theta=0.1$.

We consider two bivariate distributions for $\left(L_1,L_2\right)$, namely

(*a*) a Gumbel(1960)'s bivariate exponential distribution with survivor function

$$S(t_1,t_2)=\exp\{-t_1-t_2-\eta t_1 t_2\},\ 0<t_1,t_2<\infty,\ 0\le\eta\le 1 \tag{5.41}$$

with $\eta = 0.1$ and 0.5 and

(*b*) an independent uniform distributions in the intervals $(0,1)\times(0,0.7)$. (5.42)

We used the algorithm given in Devroye (1986) for generating the observations from Gumbel's (1960) bivariate exponential distribution. We considered two types of causes C_i, i = 1,2. The causes 1 and 2 are distributed randomly among the observed failure times with equal probability. We found the estimator of $S(t_1,t_2)$ using the equation (5.27) at five time points. The empirical biases and empirical variances of the estimators are given in Tables 5.1and 5.2. A careful examination of the tables reveals that the biases and variances of $\hat{S}^{L}(t_1,t_2)$ decreases with increasing sample size and that the biases and variances are small.

To study the performance of the estimators of $F_{ij}(t_1,t_2)$, we generated lifetimes (T_1,T_2) from the Gumbel's (1960) bivariate exponential distribution given in (5.39) for λ = 1and 3 with sample sizes n = 20, 50 and 100. The censoring times (Z_1,Z_2) and truncating times (L_1,L_2) were generated from the same distributions given in (5.40), (5.41) and (5.42) for the same values of the parameters. The estimators of $F_{ij}(t_1,t_2)$ i =1,2; j =1,2; are computed based on the equations (5.28) and (5.29). The empirical biases and empirical variances of the estimators are given in Table 5.3 and Table 5.4. We observe that the biases and variances of the estimators are small. As sample size increases, the biases and variances of the estimators decreases.

Table 5.1: Biases ($\times 10^3$) and Variances ($\times 10^6$), given in brackets, of $\hat{S}_c^{L}(t_1,t_2)$ at five (t_1,t_2) pairs for θ=0.1 and various values of λ and η when (L_1,L_2) has Gumbel's (1960) bivariate exponential distribution

n	(t_1,t_2)	λ	η	Bias	Var	λ	η	Bias	Var	λ	η	Bias	Var	λ	η	Bias	Var
20	(1,2)	3	0.1	41.1	375	3	0.5	34.5	91.6	5	0.1	34.9	118	5	0.5	32.4	44.4
	(1,3)			41.4	259			39.8	261			32.8	99.1			32.7	77.4
	(1.5,2)			38.4	204			37.1	149			32.9	57.9			32.0	45.1
	(2,1)			38.1	250			38.2	176			31.7	40.7			31.3	22.2
	(2,2)			37.4	171			41.0	222			33.7	59.9			32.9	58.9
50	(1,2)			18.2	68.9			17.9	74.6			14.1	11.9			13.4	13.1
	(1,3)			19.0	121			16.1	22.2			13.9	11.7			13.9	11.9
	(1.5,2)			16.7	34.8			16.8	55.5			13.8	8.22			13.7	13.8
	(2,1)			18.6	91.3			17.6	10.3			13.8	18.6			13.2	4.63
	(2,2)			15.7	24.5			15.6	18.6			13.2	5.91			13.7	10.2
100	(1,2)			10.7	38.4			9.9	19.1			7.3	4.57			7.68	8.01
	(1,3)			9.1	11.7			8.6	10.6			7.8	6.92			6.99	2.78
	(1.5,2)			9.0	8.98			9.4	12.5			6.98	2.65			7.08	3.27
	(2,1)			11.5	47.2			8.9	16.1			7.02	2.18			6.99	3.20
	(2,2)			8.2	9.28			8.3	7.7			7.09	4.62			7.03	2.12

Table 5.2: Biases ($\times 10^3$) and Variances ($\times 10^6$), given in brackets, of $\hat{S}_c^L(t_1,t_2)$ at five (t_1,t_2) pairs for θ=0.1 and various values of λ when (L_1,L_2) has independent uniform distributions

n	(t_1,t_2)	λ	Bias	Var	λ	Bias	Var
20	(1,2)	3	37.5	16.8	5	36.7	121
	(1,3)		32.9	81.9		36.6	99.9
	(1.5,2)		36.2	98.9		38.1	210
	(2,1)		33.9	55.9		32.4	48.2
	(2,2)		33.8	88.4		35.4	82.2
50	(1,2)		14.1	12.0		15.8	22.3
	(1,3)		13.9	10.4		16.3	36.4
	(1.5,2)		13.3	5.15		15.7	34.5
	(2,1)		13.6	9.05		16.1	52.4
	(2,2)		13.4	6.26		15.6	23.0
100	(1,2)		6.97	2.91		8.68	19.5
	(1,3)		7.3	3.57		8.51	11.4
	(1.5,2)		6.9	2.51		8.28	8.67
	(2,1)		6.89	1.78		7.78	6.79
	(2,2)		6.96	2.79		7.87	6.75

Table 5.3: Biases ($\times 10^3$) and Variances ($\times 10^6$), given in brackets, of $\widehat{F}_{ij}^{L}(t_1,t_2)$ at three (t_1,t_2) pairs for θ=0.1 and λ=1 when(L_1,L_2) has Gumbel's (1960) bivariate exponential distribution

n	(t_1,t_2)	η	$\widehat{F}_{11}^{L}(t_1,t_2)$		$\widehat{F}_{12}^{L}(t_1,t_2)$		$\widehat{F}_{21}^{L}(t_1,t_2)$		$\widehat{F}_{22}^{L}(t_1,t_2)$	
			Bias	Var	Bias	Var	Bias	Var	Bias	Var
20	(0.1,0.6)	0.1	-3.5	8	-3.5	10	-3.5	14	-3.0	24.9
	(0.2,1)		-13.9	58.9	-13	93.8	-13.4	83.1	-13	91.5
	(1.7,0.5)		-49.4	920	-51	635	-53.1	558	-49	858
50	(0.1,0.6)		-3.29	4.35	-3.3	4.17	-3.29	4.35	-2.5	0.1
	(0.2,1)		-13.3	33.4	-15	25.9	-13.3	35.7	-11	18.6
	(1.7,0.5)		-44.5	180	-49	411	-49.7	302	-42	224
100	(0.1,0.6)		-3.12	0.1	-3.1	0.1	-3.18	0.1	-2.0	0.1
	(0.2,1)		-13.2	4.72	-14	10.2	-13.0	4.15	-10	4.79
	(1.7,0.5)		-38.9	101	-48	401	-44.7	291	-41	200
20	(0.1,0.6)	0.5	-3.00	25.0	-3.5	11	-3.5	15	-2.9	27.7
	(0.2,1)		-12.3	131	-13	84.2	-13.5	68.4	-14	22.5
	(1.7,0.5)		-48.7	865	-46	1124	-44.1	978	-44	774
50	(0.1,0.6)		-2.5	10	-3.3	4.17	-2.8	12.6	-2.6	17.7
	(0.2,1)		-11.9	51.7	-12	47	-12.5	56.7	-12	17.2
	(1.7,0.5)		-42.3	482	-40	358	-38.7	649	-42	525
100	(0.1,0.6)		-1.41	4.00	-3.2	2.65	-2.1	0.1	-2.4	0.1
	(0.2,1)		-10.3	4.10	-11	10.2	-12.3	41.2	-9.7	0.1
	(1.7,0.5)		-41.0	213	-38	142	-36.1	471	-40	213

Table 5.4: Biases ($\times 10^3$) and Variances ($\times 10^6$), given in brackets, of $\widehat{F}_{ij}^{L}(t_1,t_2)$ at three (t_1,t_2) pairs for θ=0.1 and λ=3 when (L_1,L_2) has Gumbel's (1960) bivariate exponential distribution

n	(t_1,t_2)	η	$\widehat{F}_{11}^{L}(t_1,t_2)$		$\widehat{F}_{12}^{L}(t_1,t_2)$		$\widehat{F}_{21}^{L}(t_1,t_2)$		$\widehat{F}_{22}^{L}(t_1,t_2)$	
			Bias	Var	Bias	Var	Bias	Var	Bias	Var
20	(0.1,0.6)	0.1	19.1	241	15.4	366	15.8	279	16.1	327
	(0.2,1)		38.9	1424	27.8	996	28.6	1221	27.1	975
	(1.7,0.5)		48.9	4135	41.3	4755	46.5	3327	45.8	4433
50	(0.1,0.6)		18.1	160	13.1	60.7	12.6	50.6	13.1	79.9
	(0.2,1)		32.9	546	22.7	459	26.5	246	22.7	429
	(1.7,0.5)		46.5	2306	39.2	1942	38.6	2187	39.6	3888
100	(0.1,0.6)		16.2	99.7	11.2	41.3	10.3	40.1	10.2	66.7
	(0.2,1)		27.5	355	21.3	263	25.2	204	16.2	378
	(1.7,0.5)		45.3	984	36.7	677	34.9	426	33.5	2263
20	(0.1,0.6)	0.5	17.6	489	16.2	287	15.5	277	15.1	157
	(0.2,1)		64.7	600	6.47	821	69.4	626	69.4	725
	(1.7,0.5)		43.5	6855	63.9	5464	48.4	6145	55.4	5516
50	(0.1,0.6)		14.2	91.7	15.2	113	15.3	150	14.1	113
	(0.2,1)		35.0	549	3.75	704	32.7	509	36.7	698
	(1.7,0.5)		39.6	2012	55.1	1570	47.6	1837	51.6	2270
100	(0.1,0.6)		13.3	10.7	14.3	98.6	15.1	113	13.1	89.9
	(0.2,1)		29.7	368	3.61	312	36.7	418	30.0	579
	(1.7,0.5)		38.9	812	49.7	897	41.2	913	48.7	1261

5.6 Conclusion

In this chapter, we developed a nonparametric estimator of the bivariate survivor function for competing risk models using cause-specific hazard function under random left truncation and right censoring. This method is a generalization of the procedure for the bivariate lifetime data in non-competing risk set up given in Gurler (2004). We proved the consistency and weak convergence of the estimators. Then we developed nonparametric estimators of cause-specific sub-distribution functions for bivariate competing risk models. The weak convergence and strong consistency of the estimators are established. Simulation study shows that the bias and variance of the estimators are less. The extension of the method to the multivariate set up is straightforward. The estimation procedure and results presented in this chapter are summarized in Sankaran and Ansa (2005b).

Chapter 6
A LOCAL DEPENDENCE MEASURE FOR BIVARIATE LIFETIME DATA

6.1 Introduction

Let (T_1,T_2) be a nonnegative random vector representing lifetime of a two-component system. For example, (T_1,T_2) could be cohort study ages at diagnosis of breast cancer of mother and daughter or lifetimes of components of two components in an industrial system. In the analysis of lifetime data, it is often of interest to estimate the degree of dependence between lifetimes. Accordingly, Clayton (1978) introduced a measure of dependence as $\theta(t_1,t_2) = \dfrac{S(t_1,t_2)D_1D_2(S(t_1,t_2))}{D_1(S(t_1,t_2))D_2(S(t_1,t_2))}$ where $S(t_1,t_2) = P(T_1 > t_1, T_2 > t_2)$ is the survivor function of (T_1,T_2) and D_j denotes the operator $-\dfrac{\partial}{\partial t_j}$ j = 1,2. This function may be interpreted as, ratio of the hazard function of conditional distribution of T_1 given $T_2 = t_2$ to that of T_1 given $T_2 > t_2$. The value of $\theta(t_1,t_2)$ is greater than 1, 1 or less than 1 according to whether there is a positive association, no association or negative association between the two failure times at (t_1,t_2). Later Oakes (1989) defined a measure of dependence as $g(t_1,t_2) = \dfrac{\theta(t_1,t_2)-1}{\theta(t_1,t_2)+1}$ which is the conditional version of Kendall's concordance measure (Kendall, 1938). Later, Fan et.al. (1998) have proposed two nonparametric summary measures of dependence that are weighted averages of the above two local dependence measures. Recently, Fan et.al. (2000) have proposed a class of weighted dependence measures, which includes the measure in Fan et.al. (1998), as special cases.

Basically, the local dependence measures discussed above are the generalizations of the Kendall's concordance measure. The measures given by Fan et.al. (1998) and Fan et.al. (2000) largely depend on the choice of weight function. Further, the computations involved in these measures are little complex. Motivated by this, in the present chapter, we present a simple measure of local dependence using covariance residual lives, introduced by Nair et.al. (2005). The proposed measure is the generalization of the well-known covariance function. Generally, zero covariance (correlation) does not imply independence. However, in Section 6.2, we prove that zero covariance residual life implies independence between the variables. In Section 6.3, we propose a nonparametric estimator for the dependence measure and its asymptotic properties are studied. The simulation study showing the empirical behaviour of the estimator is presented in Section 6.4.

Testing independence for bivariate lifetime data is important in many biomedical studies. For example, leukemia patients after bone marrow transplantation are at risk of acute graft versus host disease and cytomegalovirus. The times to these two diseases may be correlated. For patients with retinopathy in both eyes, the times to blindness of the two eyes may be associated due to natural pairing. Several researchers have developed tests of independence. Oakes (1982) developed a concordance test for independence in the presence of censoring. Cuzick (1982), Clayton and Cuzick (1985), Prentice and Self (1985) and Dabrowska (1986) have focused on semi parametric approaches. Shih and Louis (1996) present two alternative test statistics that are based on the covariance process of the martingale residuals for the marginal distributions. In Section 6.5, we develop a simple statistic for testing independence among the variables using the covariance residual life function. Finally, in Section 6.6, we illustrate the test procedure for a real data set. Section 6.7 gives a conclusion for the chapter.

6.2 Covariance Residual Life Function

Let (T_1,T_2) be a random vector defined on the positive octant $R_2^+ = \{(t_1,t_2)\,|\,t_1>0,t_2>0\}$ of the two dimensional space with absolutely continuous distribution function $F(t_1,t_2)$. Let $S(t_1,t_2) = P(T_1>t_1,T_2>t_2)$ be the survivor

function of (T_1,T_2). Assume that $S(t_1,t_2)$ is differentiable and the joint probability density function $f(t_1,t_2)$ exists. Further we assume that $E(T_1T_2)$, $E(T_1)$ and $E(T_2)$ are finite. Then the product moment residual life function of (T_1,T_2) is defined as

$$M(t_1,t_2) = E\left[(T_1-t_1)(T_2-t_2)\,|\,T_1>t_1,T_2>t_2\right].$$

Then the covariance residual life function (CVRL) is given by

$$C(t_1,t_2) = M(t_1,t_2) - r_1(t_1,t_2)\, r_2(t_1,t_2)$$

where

$$r_1(t_1,t_2) = E\left[T_1-t_1\,|\,T_1>t_1,T_2>t_2\right]$$

and

$$r_2(t_1,t_2) = E\left[T_2-t_2\,|\,T_1>t_1,T_2>t_2\right]$$

are the mean residual life functions in the bivariate case defined by Arnold and Zahedi (1988).

Note that $C(0,0) = E(T_1T_2) - E(T_1)\,E(T_2)$ is the covariance between T_1 and T_2 It is easy to see that $M(t_1,t_2)$ can be written as

$$M(t_1,t_2) = \frac{1}{S(t_1,t_2)}\int_{t_1}^{\infty}\int_{t_2}^{\infty} S(u_1,u_2)\,du_2\,du_1.$$

Theorem 6.1

The variables T_1 and T_2 are independent if and only if $C(t_1,t_2)=0$ for all $(t_1,t_2)\in R_2^+$

Proof

When T_1 and T_2 are independent, we have

$$r_1(t_1,t_2) = r_1(t_1,0) \tag{6.1}$$

and

$$r_2(t_1,t_2) = r_2(0,t_2) \tag{6.2}$$

Further, $S(t_1,t_2) = S(t_1,0)S(0,t_2)$,

which leads to

$$M(t_1,t_2) = r_1(t_1,0)r_2(0,t_2). \tag{6.3}$$

From (6.1), (6.2) and (6.3), we get

$$C(t_1,t_2) = 0.$$

Conversely,

$$C(t_1,t_2) = 0$$

implies

$$M(t_1,t_2) = r_1(t_1,t_2)r_2(t_1,t_2)$$

or

$$\int_{t_1}^{\infty}\int_{t_2}^{\infty} S(u_1,u_2)du_2du_1 = S(t_1,t_2)r_1(t_1,t_2)r_2(t_1,t_2) \tag{6.4}$$

Differentiating (6.4) with rsespect to t_1, we obtain

$$-\int_{t_2}^{\infty} S(t_1,u_2)du_2 = r_1(t_1,t_2)r_2(t_1,t_2)\frac{\partial S(t_1,t_2)}{\partial t_1} + S(t_1,t_2)r_2(t_1,t_2)\frac{\partial r_1(t_1,t_2)}{\partial t_1}$$

$$+ S(t_1,t_2)r_1(t_1,t_2)\frac{\partial r_2(t_1,t_2)}{\partial t_1}$$

or

$$-r_2(t_1,t_2) = r_1(t_1,t_2)r_2(t_1,t_2)\frac{1}{S(t_1,t_2)}\frac{\partial S(t_1,t_2)}{\partial t_1} + r_2(t_1,t_2)\frac{\partial r_1(t_1,t_2)}{\partial t_1}$$

$$+ r_1(t_1,t_2)\frac{\partial r_2(t_1,t_2)}{\partial t_1} \tag{6.5}$$

From the relationship

$$r_1(t_1,t_2)\frac{\partial S(t_1,t_2)}{\partial t_1} = -\left(1+\frac{\partial r_1(t_1,t_2)}{\partial t_1}\right)S(t_1,t_2),$$

(6.5) becomes

$$-r_2(t_1,t_2) = -r_2(t_1,t_2)\left(1+\frac{\partial r_1(t_1,t_2)}{\partial t_1}\right) + r_2(t_1,t_2)\frac{\partial r_1(t_1,t_2)}{\partial t_1}$$

$$+ r_1(t_1,t_2)\frac{\partial r_2(t_1,t_2)}{\partial t_1}$$

or

$$r_1(t_1,t_2)\frac{\partial r_2(t_1,t_2)}{\partial t_1} = 0.$$

Thus $r_2(t_1,t_2)$ is a function of t_2 alone and hence $r_2(t_1,t_2) = r_2(0,t_2)$. Similarly, (6.4) leads to $r_1(t_1,t_2) = r_1(t_1,0)$. From Nair and Nair(1989) it follows that T_1 and T_2 are independent.

Now, we propose $C(t_1,t_2)$ as a measure of dependence (association) between two variables. The value of $C(t_1,t_2)$ is greater that zero, zero or less than zero according to whether there is a positive association, no association or negative association between the two failure times at (t_1,t_2). For example, consider bivariate Lomax distribution with survivor function

$$S(t_1,t_2) = (1+a_1t_1+a_2t_2)^{-b} \quad t_1,t_2>0;\ b>2.$$

Then

$$C(t_1,t_2) = \frac{(1+a_1t_1+a_2t_2)^2}{a_1a_2(b-1)^2(b-2)}$$

which is positive. Thus T_1 and T_2 are positively associated.

For bivariate Dirichlet distribution with survivor function

$$S(t_1,t_2) = (1-p_1t_1-p_2t_2)^q \quad 0<t_1<\frac{1}{p_1};\ 0<t_2<\frac{1-p_1t_1}{p_2},$$

$C(t_1,t_2)$ is obtained as

$$C(t_1,t_2) = \frac{(1-p_1t_1-p_2t_2)^2}{p_1p_2(q+1)^2(q+2)}.$$

In this case $C(t_1,t_2)$ is negative and thus T_1 and T_2 are negatively associated. For more properties and characterizations, one could refer to Nair et.al. (2005).

6.3 Estimation of Covariance Residual Life Function

In this section, we propose a nonparametric estimator of $C(t_1,t_2)$ and study its properties.

In practical situations, lifetimes may not be fully observed because of loss to follow-up or the finite duration of the study. Under the bivariate right censoring, the observable variables are given by $Y = (Y_1,Y_2)$ and $\delta = (\delta_1,\delta_2)$ where $Y_i = \min(T_i,Z_i)$; $i = 1,2$ with $Z = (Z_1,Z_2)$ is a pair of random censoring times and $\delta_1 = I(T_1 = Y_1)$ and $\delta_2 = I(T_2 = Y_2)$. Let $H(t_1,t_2)$ and $G(t_1,t_2)$ denote the bivariate survivor function of Y and Z respectively. We further assume that (T_1,T_2) and (Z_1,Z_2) are independent, which gives

$$H(t_1,t_2) = S(t_1,t_2)\,G(t_1,t_2).$$

Suppose now that (Y_{1i},Y_{2i}) and $(\delta_{1i},\delta_{2i})$; $i = 1,2\ldots n$ are i.i.d samples, each having the same distribution as (Y,δ).

Then, the estimator of $C(t_1,t_2)$ is then obtained as

$$\hat{C}(t_1,t_2) = \frac{1}{\hat{S}(t_1,t_2)}\int_{t_1}^{b_1}\int_{t_2}^{b_2}\hat{S}(u_1,u_2)\,du_2\,du_1$$

$$\frac{1}{\hat{S}(t_1,t_2)}\int_{t_1}^{b_1}\hat{S}(u_1,t_2)\,du_1 \cdot \frac{1}{\hat{S}(t_1,t_2)}\int_{t_2}^{b_2}\hat{S}(t_1,u_2)\,du_2$$

where $\hat{S}(t_1,t_2)$ is any non-parametric estimator of $S(t_1,t_2)$ and (b_1,b_2) is such that $b_i = \sup\{t_i \mid S(t_1,t_2) > 0\}$; $i = 1,2$. In the complete sample setup, $\hat{C}(t_1,t_2)$ could be obtained by replacing $\hat{S}(t_1,t_2)$ with the usual empirical survivor function

$$S_n(t_1,t_2) = \frac{1}{n}\sum_{i=1}^{n} I(T_{1i} > t_1, T_{2i} > t_2).$$

When the observations are censored, in the present study, we consider $\hat{S}(t_1,t_2)$ as the estimator given in Dabrowska (1988), which is the extension of the well-known Kaplan-Meier (1958) estimator to the bivariate set up.

Now we derive the asymptotic properties of $\hat{C}(t_1,t_2)$.

Let $D = [0,b_1]\times[0,b_2]$ with $H(b_1,b_2) > 0$.

Theorem 6.2

$\underset{D}{Sup}\left\|\hat{C}-C\right\| \to 0$ almost surely. In other words, the estimator $\hat{C}(t_1,t_2)$ is uniformly strong consistent.

Proof

Denote $\widehat{M}(t_1,t_2) = \dfrac{1}{\hat{S}(t_1,t_2)}\int_{t_1}^{b_1}\int_{t_2}^{b_2}\hat{S}(u_1,u_2)du_2du_1$

Then

$$\underset{D}{Sup}\left\|\hat{C}-C\right\| \le \underset{D}{Sup}\left\|\widehat{M}-M\right\| + \underset{D}{Sup}\left\|\hat{r}_1\hat{r}_2 - r_1r_2\right\| \tag{6.6}$$

where $\hat{r}_1(t_1,t_2)$ and $\hat{r}_2(t_1,t_2)$ are the estimators of $r_1(t_1,t_2)$ and $r_2(t_1,t_2)$ obtained by replacing $S(t_1,t_2)$ by $\hat{S}(t_1,t_2)$

Thus, (6.6) can also be written as

$$\underset{D}{Sup}\left\|\hat{C}-C\right\| \le \underset{D}{Sup}\left\|\widehat{M}-M\right\| + \underset{D}{Sup}\left\|\hat{r}_1\hat{r}_2 - r_1r_2\right\|$$

which gives

$$\underset{D}{Sup}\left\|\hat{C}-C\right\| \le \underset{D}{Sup}\left\|\widehat{M}-M\right\| + \underset{D}{Sup}\left\|\hat{r}_1\left(\hat{r}_2 - r_2\right)\right\| + \underset{D}{Sup}\left\|r_2\left(\hat{r}_1 - r_1\right)\right\| \tag{6.7}$$

Now we prove that $\underset{D}{Sup}\left\|\widehat{M}-M\right\| \to 0$ almost surely.

$$\underset{D}{Sup}\left\|\widehat{M}-M\right\| = \underset{D}{Sup}\left\|\left(\hat{S}(t_1,t_2)\right)^{-1}\int_{t_1}^{b_1}\int_{t_2}^{b_2}\hat{S}(u_1,u_2)du_2du_1 - \left(S(t_1,t_2)\right)^{-1}\int_{t_1}^{b_1}\int_{t_2}^{b_2}S(u_1,u_2)du_2du_1\right\|$$

$$\leq \left[\hat{S}(b_1,b_2)S(b_1,b_2)\right]^{-1}\left[\left\|S(t_1,t_2)\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(\hat{S}(u_1,u_2)\right.\right.\right.$$

$$\left.\left.\left.-S(u_1,u_2)\right)du_1du_2\right\|+\left\|\left[\hat{S}(t_1,t_2)-S(t_1,t_2)\right]\int_{t_1}^{b_1}\int_{t_2}^{b_2}S(u_1,u_2)du_2du_1\right\|\right].$$

From Dabrowska (1988), it follows that $\underset{D}{Sup}\left\|\hat{S}-S\right\|\to 0$ almost surely which implies $\underset{D}{Sup}\left\|\widehat{M}-M\right\|\to 0$ almost surely, (van der Vaart and Wellner (1996)). Similarly, one can prove that $\underset{D}{Sup}\left\|\hat{r}_i-r_i\right\|\to 0$ almost surely for i = 1,2. Therefore, since $r_1(t_1,t_2)$ and $r_2(t_1,t_2)$ are bounded, the second and third factors in the right side of (6.7) approaches to zero. This completes the proof.

Theorem 6.3

For $(t_1,t_2)\in D$, $\sqrt{n}\left(\hat{C}(t_1,t_2)-C(t_1,t_2)\right)$ converges weakly to a Gaussian process with mean zero and asymptotic variance given in (6.10).

Proof

$$\begin{aligned}\sqrt{n}\left(\hat{C}(t_1,t_2)-C(t_1,t_2)\right) &= \sqrt{n}\left[\left(\hat{S}(t_1,t_2)\right)^{-1}\int_{t_1}^{b_1}\int_{t_2}^{b_2}\hat{S}(u_1,u_2)du_2du_1\right.\\
&\quad -\left(S(t_1,t_2)\right)^{-1}\int_{t_1}^{b_1}\int_{t_2}^{b_2}S(u_1,u_2)du_2du_1\\
&\quad -\left(\hat{S}(t_1,t_2)^2\right)^{-1}\int_{t_1}^{b_1}\hat{S}(u_1,t_2)du_1\int_{t_2}^{b_2}\hat{S}(t_1,u_2)du_2\\
&\quad \left.+\left(S(t_1,t_2)^2\right)^{-1}\int_{t_1}^{b_1}S(u_1,t_2)du_1\int_{t_2}^{b_2}S(t_1,u_2)du_2\right]\\
&= \sqrt{n}\left[\left(\hat{S}(t_1,t_2)S(t_1,t_2)\right)^{-1}\left[S(t_1,t_2)\int_{t_1}^{b_1}\int_{t_2}^{b_2}\hat{S}(u_1,u_2)du_2du_1\right.\right.\\
&\quad \left.\left.-\hat{S}(t_1,t_2)\int_{t_1}^{b_1}\int_{t_2}^{b_2}S(u_1,u_2)du_2du_1\right]\right.\end{aligned}$$

$$-\left(\hat{S}^2(t_1,t_2)S^2(t_1,t_2)\right)^{-1}\left[S^2(t_1,t_2)\int_{t_1}^{b_1}\hat{S}(u_1,t_2)\,du_1\int_{t_2}^{b_2}\hat{S}(t_1,u_2)\,du_2\right.$$

$$\left.-\hat{S}^2(t_1,t_2)\int_{t_1}^{b_1}S(u_1,t_2)\,du_1\int_{t_2}^{b_2}S(t_1,u_2)\,du_2\right] \qquad (6.8)$$

To prove the theorem we follow the approach for the proof of asymptotic normality given in Fan et.al. (2000, p.189). Since $H(t_1,t_2) = S(t_1,t_2)\,G(t_1,t_2)$, $\underset{D}{Sup}\left\|\hat{H}-H\right\|\to 0$, $\underset{D}{Sup}\left\|\hat{G}-G\right\|\to 0$ and $\underset{D}{Sup}\left\|\hat{S}-S\right\|\to 0$, (6.8) is asymptotically equal to

$$\sqrt{n}\left(\hat{C}(t_1,t_2)-C(t_1,t_2)\right)=$$

$$\sqrt{n}\left[\frac{1}{H(t_1,t_2)}\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(\frac{1}{n}\sum_{k=1}^{n}I(Y_{1k}\ge u_1,Y_{2k}\ge u_2)-H(u_1,u_2)\right)du_2du_1\right.$$

$$-\frac{\int_{t_1}^{b_1}\int_{t_2}^{b_2}H(u_1,u_2)\,du_2du_1}{H^2(t_1,t_2)}\left(\frac{1}{n}\sum_{k=1}^{n}I(Y_{1k}\ge t_1,Y_{2k}\ge t_2)-H(t_1,t_2)\right)$$

$$-\frac{\int_{t_1}^{b_1}H(u_1,t_2)\,du_1}{H^2(t_1,t_2)}\left(\int_{t_2}^{b_2}\frac{1}{n}\sum_{k=1}^{n}I(Y_{1k}\ge t_1,Y_{2k}\ge u_2)\,du_2-\int_{t_2}^{b_2}H(t_1,u_2)\,du_2\right)$$

$$-\frac{\int_{t_2}^{b_2}H(t_1,u_2)\,du_2}{H^2(t_1,t_2)}\left(\int_{t_1}^{b_1}\frac{1}{n}\sum_{k=1}^{n}I(Y_{1k}\ge u_1,Y_{2k}\ge t_2)\,du_1-\int_{t_1}^{b_1}H(u_1,t_2)\,du_1\right)$$

$$\left.+\frac{2}{H^3(t_1,t_2)}\int_{t_1}^{b_1}H(u_1,t_2)\,du_1\int_{t_2}^{b_2}H(t_1,u_2)\,du_2\left(\frac{1}{n}\sum_{k=1}^{n}I(Y_{1k}\ge t_1,Y_{2k}\ge t_2)-H(t_1,t_2)\right)\right]. \qquad (6.9)$$

By multivariate central limit theorem, each term in the simple braces of (6.9) converges to a mean zero normal variate. Thus the asymptotic normality of $\hat{C}(t_1,t_2)$ follows from the delta method using the maps $(a_1,a_2,a_3,a_4,a_5)\to(a_1+a_2+a_3+a_4+a_5)$ (see van der Vaart and Wellner (1996)). Then the asymptotic variance will be obtained as

$$\sigma^2(t_1,t_2) = E\Bigg[\frac{1}{H(t_1,t_2)}\int_{t_1}^{h_1}\int_{t_2}^{h_2} I(Y_{1k}\geq u_1, Y_{2k}\geq u_2)du_2du_1$$

$$-\frac{\int_{t_1}^{h_1}\int_{t_2}^{h_2} H(u_1,u_2)\,du_2du_1}{H^2(t_1,t_2)} I(Y_{1k}\geq t_1, Y_{2k}\geq t_2)$$

$$-\frac{\int_{t_1}^{h_1} H(u_1,t_2)\,du_1}{H^2(t_1,t_2)}\int_{t_2}^{h_2} I(Y_{1k}\geq t_1, Y_{2k}\geq u_2)\,du_2$$

$$-\frac{\int_{t_2}^{h_2} H(t_1,u_2)\,du_2}{H^2(t_1,t_2)}\int_{t_1}^{h_1} I(Y_{1k}\geq u_1, Y_{2k}\geq t_2)\,du_1$$

$$+\frac{2}{H^3(t_1,t_2)} I(Y_{1k}\geq t_1, Y_{2k}\geq t_2)\int_{t_1}^{h_1} H(u_1,t_2)\,du_1\int_{t_2}^{h_2} H(t_1,u_2)\,du_2\Bigg]^2 \qquad (6.10)$$

This completes the proof.

Remark 6.1 A consistent estimator of $\sigma^2(t_1,t_2)$ is given by

$$\hat{\sigma}^2(t_1,t_2) = \frac{1}{n}\sum_{k=1}^{n}\Bigg[\frac{1}{\hat{H}(t_1,t_2)}\sum_{u_1=t_1}^{h_1}\sum_{u_2=t_2}^{h_2} I(Y_{1k}\geq u_1, Y_{2k}\geq u_2)$$

$$-\frac{\sum_{u_1=t_1}^{h_1}\sum_{u_2=t_2}^{h_2}\hat{H}(u_1,u_2)}{\hat{H}^2(t_1,t_2)} I(Y_{1k}\geq t_1, Y_{2k}\geq t_2)$$

$$-\frac{\sum_{u_1=t_1}^{h_1}\hat{H}(u_1,t_2)}{\hat{H}^2(t_1,t_2)}\sum_{u_2=t_2}^{h_2} I(Y_{1k}\geq t_1, Y_{2k}\geq u_2)$$

$$-\frac{\sum_{u_2=t_2}^{h_2}\hat{H}(t_1,u_2)\,du_2}{\hat{H}^2(t_1,t_2)}\sum_{u_1=t_1}^{h_1} I(Y_{1k}\geq u_1, Y_{2k}\geq t_2)$$

$$+\frac{2}{\hat{H}^3(t_1,t_2)} I(Y_{1k}\geq t_1, Y_{2k}\geq t_2)\sum_{u_1=t_1}^{h_1}\hat{H}(u_1,t_2)\sum_{u_2=t_2}^{h_2}\hat{H}(t_1,u_2)\Bigg]^2 \qquad (6.11)$$

where $\widehat{H}(t_1,t_2)$ is the usual empirical estimator of $H(t_1,t_2)$

6.4 A Simulation Study

To study the performance of the estimator, we carried out a series of 1000 simulations each of size n from a bivariate Dirichlet distribution with survivor function

$$S(t_1,t_2) = (1-t_1-t_2)^q,\ 0<t_1,t_2<1,\ 0<t_1+t_2<1$$

for q = 3,5 and 10 for different sample sizes n = 20, 50 and 100. We generated the observations using the algorithm given in Gentle (1998). Simulation results are given in Table 6.1. In the table we give the empirical bias and the empirical variance of the estimate of $C(t_1,t_2)$. Here bias, $B(t_1,t_2) = C_1(t_1,t_2) - C(t_1,t_2)$ where C_1 is the average of estimators of $C(t_1,t_2)$ obtained from 1000 simulations.

From Table 6. 1 it follows that

(*i*) the empirical variance decreases with increasing sample size, as is expected.

(*ii*) the empirical bias is increasing in $t_1\,(t_2)$ for every fixed $t_2\,(t_1)$.

(*iii*) for q = 3 and 5 the empirical variance decreases in t_1 and t_2

6.5 Test for Independence

In this section, we give a simple method for test of independence between two variables using covariance residual life.

It is well known that the zero covariance (correlation) between two variables does not imply independence. From Theorem 6.1, it follows that zero covariance residual life implies the independence among variables, which can be used for developing test of independence. The null hypothesis is $C(t_1,t_2)$ 0 for all $t_1,t_2 \geq 0$.

Therefore one could use $D_n = \underset{\substack{0<t_1<t_0\\ 0<t_2<t_0}}{Sup} \dfrac{\sqrt{n}\left|\widehat{C}(t_1,t_2)\right|}{\widehat{\sigma}(t_1,t_2)}$ as a test of independence where

$\widehat{\sigma}(t_1,t_2)$ is the estimator of the standard deviation obtained from (6.11) and t_0 is

total time of the study period. From Theorem 6.3. it follows that we reject the null hypothesis when D_n exceeds $\underline{\phi}^{-1}(1-\alpha/2)$ at α significance level, where $\underline{\phi}$ is the standard normal distribution function.

Table 6.1: Empirical Bias ($\times 10^3$) and empirical variance ($\times 10^3$), given in brackets, for the estimator of $C(t_1,t_2)$ for q = 3,5 and 10

q	(t_1,t_2)	Sample Size		
		20	50	100
3	(0.240, 0.225)	-0.164402(1.26)	-0.005657(.046)	-0.005279(.042)
	(0.283, 0.275)	-0.095208(.587)	-0.003535(.021)	-0.003191(.018)
	(0.325, 0.326)	-0.044492(.226)	-0.001814(.009)	-0.001096(.007)
	(0.368, 0.377)	-0.012462(.051)	-0.000745(.003)	-0.000690(.002)
	(0.411, 0.427)	-0.001184(.004)	-0.000009(.0004)	-0.000006(.0004)
	(0.453, 0.478)	0.000067(.00007)	0.000055(.000004)	0.000043(.000003)
5	(0.150, 0.172)	-0.097042(.819)	-0.003906(.030)	-0.002317(.026)
	(0.185, 0.206)	-0.062710(.566)	-0.002849(.207)	-0.002192(.016)
	(0.220, 0.240)	-0.042367(.430)	-0.002644(.015)	-0.002341(.011)
	(0.255, 0.275)	-0.020960(.196)	-0.001293(.012)	-0.001163(.008)
	(0.290, 0.310)	-0.008305(.063)	-0.000590(.061)	-0.000142(.005)
	(0.326, 0.344)	-0.002044(.016)	-0.000043(.002)	-0.000038(.002)
10	(0.075, 0.068)	-0.017525(.10)	-0.000513(.002)	-0.000501(.001)
	(0.081, 0.082)	-0.014537(.092)	-0.000419(.002)	-0.000404(.001)
	(0.087, 0.095)	-0.011989(.008)	-0.000288(.002)	-0.000117(.001)
	(0.094, 0.109)	-0.011073(.081)	-0.000269(.002)	-0.000160(.001)
	(0.100, 0.123)	-0.008246(.063)	-0.000218(.002)	-0.000102(.001)
	(0.106, 0.136)	0.006504(.044)	-0.000096(.002)	-0.000033(.001)

6.6 Data Analysis

In this section, we obtain the estimator of $C(t_1,t_2)$ for the censored data given in Table 4 of Shih and Louis (1996, page 1446). Let (T_1,T_2) represents the days of survival of closely matched and poorly matched skin grafts on the same person. Table 6.2 provides the estimator of $C(t_1,t_2)$ for the data set and it is easy to observe that $\hat{C}(t_1,t_2)$ is non-positive.

Now we apply the test of independence using the statistic D_n to the data of Shih and Louis (1996). Although the sample size is small to draw conclusions based on asymptotic approaches, the data do illustrate the advantage of the new test. The value of D_n is 2.20282 indicating strong evidence against the hypothesis of independence. Thus the result suggests that early failures are strongly associated with pairing, which supports the arguments of Shih and Louis (1996).

Table 6.2. Estimate of $C(t_1,t_2)$ for the data set on days of survival of closely and poorly matched skins grafts on the same person (* denotes censored observations).

t_1	t_2	$\hat{C}(t_1,t_2)$
16	11	-8.07003
18	21	-8.12632
19	13	-5.05563
20	26	-14.9402
22	17	-7.05954
29	15	-8.41495
37	29	-5.72155
*57	15	-5.88678
*60	40	0
63	43	0
93	26	0

6.7 Conclusion

In this chapter, we proposed a local dependence measure between a pair of failure time variables based on covariance residual function. It is shown that zero covariance (correlation) residual life implies independence between the variables. We, then give a nonparametric estimator for the local dependence measure. Asymptotic properties of the estimator are discussed and a simulation study assessing the performance of the estimator is also carried out. We further developed a procedure for testing independence between two variables using covariance residual life function. Finally, the procedure is illustrated with a real data related to skin grafts. The results in this chapter are presented in Sankaran et.al. (2005a).

Chapter 7
CONCLUSION

7.1 Introduction

The pervasive importance of competing causes of failure in life testing experiments is established in literature. We come across lifetime data with more than one cause of failure in practical situations such as engineering applications, medical follow up studies, economics, demography, public health, actuarial science and demography. A glance through medical literature also reveals that medical investigators are often interested in multiple response variables. For example, times to death or times to initial contraction of disease may be of interest for littermate pairs of rats or for twin studies in humans. The time to deterioration level or the time to reaction of a treatment may be of interest in pairs of lungs, kidneys, eyes or ears of humans. In reliability analysis, one often has systems with more than one component and many systems, subsystems and components have more than one cause of failure. For the purpose of improving reliability of the system, it is necessary to identify the cause of failure down to the component level. Also, in the analysis of mortality data on married couples, one would be interested to compare the hazards for the same cause of death as well as to check whether death due to one cause is more important for the partners' risk of death from other causes. Accordingly, it is worthwhile to develop estimators of the multivariate survivor function and cause-specific sub-distribution function for competing risk lifetime data. Since parametric assumptions for lifetime data, especially in medical studies, is not realistic, semi parametric and nonparametric methods are popular in survival analysis. However, the non-parametric and semi parametric modelling of multivariate competing risk lifetime data under censoring has received less attention in literature. Under the assumption that the censoring time and lifetime are independent, we developed nonparametric estimators of cause-specific sub-distribution functions. We also derived a

nonparametric estimator for survivor function in bivariate (multivariate) competing risk set up using bivariate vector hazard rate. It is proved that the estimators satisfy large sample properties such as strong consistency and asymptotic normality. The procedure was illustrated with a real data set.

Apart from censoring, there are other features of competing risk lifetime data that makes the estimation of survivor function and cause-specific distribution function problematic. They are masking, missing censoring time and random left truncation. The nonparametric estimation of bivariate (multivariate) survivor function for masked competing risk lifetime data was presented in Chapter3. Chapter 4 addressed the nonparametric estimation of cause-specific sub-distribution functions of multivariate competing risk model under missing censoring time. In Chapter 5, we obtained nonparametric estimators for bivariate (multivariate) survivor function and cause-specific sub-distribution functions for bivariate (multivariate) competing risk data subjected to random left truncation and right random censoring. All these estimators possess desirable properties such as strong consistency and asymptotic normality. Simulation studies were carried out to study the performance of the estimators. Finally, in Chapter 6, we developed a simple local dependence measure for the bivariate lifetime data. The proposed measure could be used for testing independence between the variables.

7.2 Future Works

In all the above-mentioned works, it was assumed that the failure mechanism and censoring mechanism are independent to ensure identifiability of the marginal survivor functions. There are many life-testing situations in which the censoring mechanism depends on the failure mechanism. More generally, in most duration models with right-censored data, the usual assumption of independence between failure and censoring process is often doubtful. For example, it is known that the development of AIDS evolves from HIV incubation period to the period of clinical AIDS. The joint behaviour, especially their association, of the two duration variables is of interest. The longer the first duration, the greater is the chance that the second duration time will be censored. Wang and Wells (1998) considered nonparametric estimation for duration times of two successive events in non-

competing risk set up. The estimation of survivor function and cause-specific sub-distribution function for bivariate (multivariate) competing risk models under dependent censoring is a problem that is yet to be extensively studied.

The use of explanatory variables or covariates in a regression model is an important way to represent heterogeneity in a population. For example, in a survival study for lung cancer patients' factors such as the type of tumour, the age and general condition of the patient might be of importance. In industrial experiments on the time to failure of electrical insulation an important factor is the voltage the insulation is subjected to while in use. In clinical trials in medicine, the treatment assigned to a patient may be considered as a covariate. These may be factors that are of intrinsic importance for the application, like treatment in a drug trial, or factors that are known or suspected to influence the hazard of event. In fact, when there are covariates, we are more interested in the effect of those than in how the hazard changes over time. Indeed the main objective in many studies is to understand and exploit the relationship between lifetime and covariates. Accordingly, regression models are used to study the dependence between lifetimes and covariates. The most widely used semi parametric regression for univariate lifetime data is the Cox proportional hazards model (see Lawless, 2003). Some general classes of regression models for univariate competing risks, which are adaptations of ones widely used in univariate survival analysis, are presented in Crowder (2001, Chapter 1). The analysis of multivariate competing risk data in the presence of covariates is an area of research that remains to be explored.

REFERENCES

1. Aalen, O.O. (1976) *Nonparametric inference in connection with multiple decrement models*, Scandinavian Journal of Statistics, **3**, 15-27

2. Aalen, O.O (1978) *Non-parametric inference for a family of counting processes*, The Annals of Statistics, **6**, 701-726.

3. Akritas, M.G and van Keilegom, I. (2003) *Estimation of bivariate and marginal distributions with censored data*, Journal of the Royal Statistical Society, Series B, **65**, 457-471.

4. Andersen, P.K. and Borgan, O. (1985) *Counting process models for life history data: A review*, Scandinavian Journal of Statistics, **12**, 97-158.

5. Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993) Statistical Models Based on Counting Processes, Springer-Verlag, New York.

6. Anderson, J.W and Ryan, L.M. (1998) *Analysis of competing risk data with missing failure type when missingness depends on covariates*, Invited Paper, Presented at the Joint Statistical Meetings, Dallas, TX.

7. Ansa, A.A. (2004) *Estimation of bivariate survivor function of competing risk models under censoring*, Presented at the 24th Annual Conference of Indian Society for Probability and Statistics (ISPS) for the Dr. U.S. Nair Young Statistician Award Competition and won First prize.

8. Ansa, A.A. and Sankaran, P.G. (2005) *Estimation of bivariate survivor function under masked causes of failure* (communicated).

9. Aras, G. and Deshpande, J.V (1992) *Statistical analysis of dependent competing risks*, Statistics and Decisions, **10**, 323-336.

10. Arnold, B.C. and Zahedi, H.(1988) *On multivariate mean residual life functions*, Journal of Multivariate Analysis, **25**, 1-9.

11. Bandeen-Roche, K. and Liang, K-Y (2002) *Modeling multivariate failure time associations in the presence of a competing risk*, Biometrika, **89**, 2, 299-314.

12. Basu, A.P. (1971) *Bivariate failure rate*, Journal of the American Statistical Association, **66**, 103-104.

13. Basu, A.P and Sun, K. (1997) *Multivariate exponential distributions with constant total failure rates*, Journal of Multivariate Analysis, **61**, 159-169.

14. Berkson, J. and Elveback, L. (1960) *Competing exponential risks with particular reference to the study of smoking and lung cancer*, Journal of the American Statistical Association, **55**, 415-428.

15. Boag, J.W (1949) *Maximum likelihood estimates of the proportion of patients cured by cancer therapy*, Journal of the Royal Statistical Society, Series B, **11**, 15-44.

16. Boardman, T.J. and Kendell, P.J. (1970) *Estimation in compound exponential failure models*, Technometrics, **12**, 891-900.

17. Burke, M.D (1988) *Estimation of a bivariate distribution function under random censorship*, Biometrika, **75**, 379-82.

18. Cai, J. and Prentice, R.L (1995) *Estimating equations for hazard rate parameters based on correlated failure time data*, Biometrika, **82**, 151-164.

19. Campbell, G. (1981) *Non-parametric bivariate estimation with randomly censored data*, Biometrika, **68**, 417-22.

20. Campbell, G. and Foldes, A. (1982) *Large sample properties of non-parametric bivariate estimators with censored data*, Colloquia Methemetica-Societatis, Janos Bolyai, **32**,103-121.

21. Cederlof, R. and Lorich, U. (1978) *The Swedish twin registry*, Twin Research: Biology and Epidemiology, (W.E.Nance Ed.) 189-195, A.R. Liss New York.

22. Cheng, S.C., Fine, J.P and Wei, L.J. (1998) *Predictions of cumulative incidence function under proportional hazards model*, Biometrics, **54**, 219-228.

23. Clayton, D.G. (1978) *A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incid*ence, Biometrika, **65**, 141-151.

24. Clayton, D.G. and Cuzick, J. (1985) *Multivariate generalizations of the proportional hazards model*, Journal of the Royal Statistical Society, Series A, **148**, 82-117.

25. Cox, D.R. (1959) *The analysis of exponentially distributed failure times with two types of failure*, Journal of the Royal Statistical Society, Series B, **21**, 411-421.

26. Cox, D.R. (1972) *Regression models and life tables*, Journal of the Royal Statistical Society, Series B, **34**, 187-220.

27. Cronin, K.A. and Feuer, E.J.(2000) *Cumulative cause-specific mortality for cancer patients in the presence of other causes: A crude analogue of relative survival*, Statistics in Medicine, **19**, 1729-1740.

28. Crowder, M (2001) Classical Competing Risks, Chapman and Hall, London.

29. Cuzick, J. (1982) *Rank tests for association with right censored data*, Biometrika, **69**, 351-364.

30. Dabrowska, D.M. (1986) *Rank tests for independence for bivariate censored data*, The Annals of Statistics, **14**, 250-264.

31. Dabrowska, D.M. (1988) *Kaplan-Meier estimate on the plane*, The Annals of Statistics, **16**, 1475-1489.

32. David, H.A. (1957) *Estimation of means of normal populations from observed minima*, Biometrika, **44**, 282-286.

33. David, H.A. and Moeschberger, M.L. (1978) The Theory of Competing Risks, Griffin, London.

34. DeMasi, R.A. (2000) *Statistical methods for multivariate failure time data and competing risks*, Handbook of Statistics, **18**, 749-781.

35. Devroye, L. (1986) Non-Uniform Random Variate Generation, Springer Verlag, New York, First Edition.

36. Dewanji, A and Sengupta, D. (2003) *Estimation of competing risks with general missing pattern in failure types*, Biometrics, **59**, 1063-1070.

37. Dinse, G.E. (1982) *Non-parametric estimation for partially complete time and type of failure data*, Biometrics, **38**, 417-431.

38. Dinse, G.E. (1986) *Non-parametric prevalence and mortality estimators for animal experiments with incomplete cause of death data*, Journal of the American Statistical Association, **81**, 328-336.

39. Efron, B. (1967) *The two sample problem with censored data*, Proceedings of the Fifth Berkely Symposium on Mathematical Statistics and Probability, Berkely, USA: University of California Press, 831-853.

40. Efron, B. (1981) *Censored data and the bootstrap*, Journal of the American Statistical Association, **76**, 312-319.

41. Efron, B. and Petrosian, V (1999) *Nonparametric methods for doubly truncated data*, Journal of the American Statistical Association, **94**, 824-834.

42. Farley, T.M., Ali, M.M. and Slaymaker, E. (2001) *Competing approaches to analysis of failure times with competing risks*, Statistics in Medicine, **120**, 3601-3610.

43. Fan, J.J., Hsu, L. and Prentice, R.L. (1998) *Dependence estimation over a finite bivariate failure time region*, Technical Report 336. Division of Statistics, University of California at Davis, U.S.A.

44. Fan, J.J., Prentice, R.L. and Hsu, L. (2000) *A class of weighted dependence measures for bivariate failure time data*, Journal of the Royal Statistical Society, Series B, **62**, 181-190.

45. Flehinger, B.J., Reiser, B., and Yashchin, E. (1996) *Inference about defects in the presence of masking*, Technometrics, **38**, 247-55.

46. Flehinger, B.J., Reiser, B., and Yashchin, E. (1998) *Survival with competing risks and masked causes of failures*, Biometrika, **85**, 151-164.

47 Fleming, T.R. and Harrington, D.P (1991) Counting Processes and Survival Analysis, Wiley, New York.

48. Flinn, C.J. and Heckmann, J.J. (1983) *Are unemployment and out of the labour force behaviourally distinct labour force states*?, Journal of Labour Economics, **1**, 28-42.

49. Galambos, J. and Kotz, S. (1978) Characterization of Probability Distributions, Springer-Verlag, New York.

50. Gentle, J.E (1998) Random Number Generation and Monte Carlo Methods, Springer- Verlag, New York.

51. Gill, R.D. (1992) *Multivariate survival analysis*, Theory of Probability and its Applications, **37**, 18-31.

52. Goetghebeur, E. and Ryan, L. (1990) *A modified log rank test for competing risks with missing failure type*, Biometrika, **77**, 207-211.

53. Goetghebeur, E. and Ryan, L. (1995) *Analysis of competing risk data when some failure types are missing*, Biometrika, **82**, 821-833.

54. Gooley, T.A., Leisenring, W Crowley, J., and Storer, B.E. (1999) *Estimation of failure probabilities in the presence of competing risks: New representations of old estim*ators, Statistics in Medicine, **18**, 695-706.

55. Gray, R.J.(1988) *A class of k-sample tests for comparing the cumulative incidence of a competing risk*, The Annals of Statistics, **16**, 1141-1154.

56. Gumbel, E.J. (1960) *Bivarite exponential distributions*, Journal of the American Statistical Association, **55**, 698-707

57. Gurler, U. (1996) *Bivariate estimation with right truncated data*, Journal of the American Statistical Association, **91**, 1152-1165.

58. Gurler, U. (1997) *Bivariate distributions and hazard functions when a component is randomly truncated*, Journal of Multivariate Analysis, **60**, 20-47.

59. Gurler, U. (2004) *Nonparametric bivariate estimation with randomly truncated observations*, Handbook of Statistics, **23**, 195-207

60. Hanley, J.A. and Parnes, M.N. (1983) *Non parametric estimation of a multivariate distribution in the presence of censoring*, Biometrics, **39**, 129-139.

61. Herman, R.J. and Patel, R.K.N. (1971) *Maximum likelihood estimation for multi-risk model*, Technometrics, **13**, 385-396.

62. Hinds, P.A. (1996) Reliability Analysis Techniques for Equipment Management, M.Sc Dissertation, City University, London.

63. Hoel, D.G. (1972) *A representation of mortality data by competing risks*, Biometrics, **28**, 475-488.

64. Horvath, L. (1983) *The rate of strong uniform consistency for the multivariate product-limit estimator*, Journal of Multivariate Analysis, **13**, 202-209.

65. Hougaard, P. (2000) Analysis of Multivariate Survival Data, Springer-Verlag, New York.

66. Hougaard, P., Harvald, B., and Holn, N.V (1992) *Measuring the similarities between the lifetimes of adult Daniah twins born between 1881-1930*, Journal of the American Statistical Association, **87**, 17-24.

67. Hu, X.J. and Lawless, J.F (1996) *Estimation of rate and mean functions from truncated recurrent event data*, Journal of the American Statistical Association, **91**, 300-310.

68. Hu, X.J. and Lawless, J.F. and Suzuki, K. (1998) *Non-parametric estimation of a lifetime distribution when censoring times are missing*, Technometrics, **30**, 365-388.

69. Ichida, J.M., Wassell, J.T and Keller, M.D. (1993) *Evaluation of protocol change in burn-care management using the Cox proportional hazards model with time-dependent covariates*, Statistics in Medicine, **12**, 301-310.

70. Johnson, N.L. and Kotz, S. (1975) *A vector multivariate hazard rate*, Journal of Multivariate Analysis, **5**, 53-66.

71. Kalbfleisch, J.D. and Lawless, J.F (1988) *Estimation of reliability from field performance studies (with discussion)*, Technometrics, **30**, 365-388.

72. Kalbfleisch, J.D. and Prentice, R.L. (2002) The Statistical Analysis of Failure Time Data, John Wiley and Sons, New York.

73. Kaplan, E.L. and Meier, P (1958) *Non-parametric estimation from incomplete observations*, Journal of the American Statistical Association, **53**, 457-481.

74. Kendall, M.G. (1938) *A new measure of rank correlation*, Biometrika, **30**, 81-93.

75. Klein, J.P and Moeschberger, M.L. (1997) Survival Analysis: Techniques for Censored and Truncated Data, Springer-Verlag, New York.

76. Kodel, R.K. and Chen, J.J. (1987) *Handling cause of death in equivocal cases using the EM algorithm*, Communications in Statistics- Theory and Methods, **16**, 2565-2603.

77. Kulkarni, H.V and Rattihalli, R.N (2002) *Non-parametric estimation of a bivariate mean residual life function*, Journal of the American Statistical Association, **97**, 907-917

78. Lapidus, G., Braddock, M., Schwartz, R., Banco, L. and Jacobs, L. (1994) *Accuracy of fatal motorcycle injury reporting on death certificates*, Accident Analysis and Prevention, **26**(4), 535-542.

79. Lawless, J.F. (2003) Statistical Models and Methods for Life Time Data, John Wiley and Sons, New York.

80. Lin, D.Y (1997) *Non-parametric inference for cumulative incidence functions in competing risk studies*, Statistics in Medicine, **16**, 901-910.

81. Lin, D.Y Sun, W and Ying, Z. (1999) *Nonparametric estimation of the gap time distributions for serial events with censored data*, Biometrika, **86**, 59-70.

82. Lo, S.H. (1991) *Estimating a survival function with incomplete cause-of-death data*, Journal of Multivariate Analysis, **39**, 217-235.

83. Lynden-Bell, D. (1971) *A method of allowing for known observational selection in small samples applied to 3CR quasars*, Monthly Notices; Royal Astronomical Society, **155**, 95-118.

84. Makeham, W.M. (1874) *On an application of the theory of the composition of decremental forces*, Journal of the Institute of Actuaries, **18**, 317-322.

85. Mantel, N. and Ciminera, J.L. (1979) *Use of log rank series in the analysis of litter-matched data on time to tumour appearance*, Cancer Research, **39**, 4308-4315.

86. Marshall, A.W (1975) *Some comments on the hazard gradient*, Stochastic Processes and their Applications, **3**, 293-300.

87. Martz, H.F. and Waller, R.A. (1982) Bayesian Reliability Analysis, Wiley, New York.

88. McLachlan, G.J. and Peel, D. (2000) Finite Mixture Models, John Wiley and Sons, New York.

89. Miyakawa, M. (1984) *Analysis of incomplete data in competing risk model*, IEEE Transactions in Reliability, **33**, 293-296.

90. Moeschberger, M.L. (1974) *Life tests under dependent competing causes of failure*, Technometrics, **16**, 39-47

91. Moeschberger, M.L. and David, H.A. (1971) *Life tests under competing causes of failure and the theory of competing risks*, Biometrics, **27**, 909-933.

92. Mukerjee, H and Wang, J.L. (1993) *Non-parametric maximum likelihood estimation of increasing hazard rate for uncertain cause-of-death data*, Scandinavian Journal of Statistics, **20**, 17-33.

93. Nair, K.R.M. and Nair, N.U. (1989) *Bivariate mean residual life*, IEEE Transactions in Reliability, **38**, 362-364.

94. Nair, N.U., Asha, G. and Sankaran, P.G. (2005) *On covariance residual lives*, Statistica (to appear).

95. Nelson, W.B. (1969) *Hazard plotting methods for incomplete failure data*, Journal of Quality and Technology, **1**, 27-52.

96. Oakes, D. (1982) *A concordance test for independence in the presence of censoring*, Biometrics, **28**, 451-455.

97. Oakes, D. (1989) *Bivariate survival models induced by frailties*, Journal of the American Statistical Association, **84**, 487-493.

98. Oakes, D. (2001) *Biometrika centenary: Survival analysis*, Biometrika, **88**, 99-142.

99. Pepe, M.S. and Mori, M. (1993) *Kaplan-Meier, marginal or conditional probability curves in summarizing competing risk failure time data*, Statistics in Medicine, **12**, 737-751.

100. Peto, R. (1973) *Experimental survival curves for interval censored data*, Applied Statistics, **22**, 86-91.

101. Prentice, R.L and Cai, J. (1992) *Covariance and survivor function estimation using censored multivariate failure time data*, Biometrika, **79**, 495-512.

102. Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V Flournoy, N., Farewell, V.T. and Breslow, N.E. (1978) *The analysis of failure times in the presence of competing risks*, Biometrics, **34**, 541-554.

103. Prentice, R.L. and Self, S. (1985) *Comments on "Multivariate generalization of the proportional hazards model"* Journal of the Royal Statistical Society, Series A, **148**, 82-108.

104. Pruitt, R.C. (1991) *Strong consistency of self-consistent estimators: general theory and an application to bivariate survival analysis*, Technical Report 543, University of Minnesota, Minneapolis, U.S.A.

105. Racine-Poon, A.H. and Hoel, D.G. (1984) *Non-parametric estimation of the survival function when causes of death is uncertain*, Biometrics, **40**, 1151-1158.

106. Reiser, B., Guttman, I., Lin, D.K.J., Usher, J.S., and Guess, F.M., (1995) *Bayesian Inference for masked system life time data*, Applied Statistics, **44**, 79-90.

107. Sampford, M.R. (1952) *The estimation of response-time distributions. II. Multi-stimulus distributions*, Biometrics, **8**, 307-369.

108. Sankaran, P.G. and Ansa, A.A. (2005a) *Nonparametric estimation of competing risk models when censoring times are missing* (communicated).

109. Sankaran, P.G. and Ansa, A.A. (2005b) *Nonparametric estimation of competing risk models under random left truncation and right censoring* (communicated).

110. Sankaran, P.G., Abraham, B. and Ansa, A.A. (2005a) *A measure of local dependence for bivariate failure time data* (communicated).

111. Sankaran, P.G., Lawless, J.F., Abraham, B. and Ansa, A.A. (2005b) *Estimation of distribution function in competing risk models* (submitted after revision).

112. Shanbag, D.N and Kotz, S. (1987) *Some new approaches to multivariate probability distributions*, Journal of Multivariate Analysis, **22**, 189-211.

113. Shih, J.H. and Louis, T.A. (1996) *Tests of independence for bivariate survival data*, Biometrics, **52**, 1440-1449.

114. Sinha, S.K. (1986) Reliability and Life Testing, Wiley Eastern Limited, New Delhi.

115. Suzuki, K. (1985) *Estimation of lifetime parameters from incomplete field data*, Technometrics, **27**, 263-271.

116. Tsai W.-Y Leurgan, S. and Crowley, J. (1986) *Non-parametric estimation of a bivariate survival function in the presence of censoring*, The Annals of Statistics, **14**, 1351-1365.

117. Tsiatis, A.A. (1975) *A nonidentifiability aspect of the problem of competing risks*, Proceedings of the National Academy of Sciences, U.S.A, **72**,20-22.

118. Turnbull, B.W (1976) *The empirical distribution function with arbitrarily grouped, censored and truncated data*, Journal of the Royal Statistical Society, Series B, **38**, 290-295.

119. van der Laan, M.J.(1996) *Efficient estimation in the bivariate censoring model and repairing NPMLE*, The Annals of Statistics, **24**, 596-627

120. van der Laan, M.J.(1997) *Non parametric estimator of the bivariate survival function under random censoring*, Statistica Neerlandica, **51**,178-200.

121. van der Laan, M.J., Hubbard, A.E. and James, M.R. (2002) *Locally efficient estimation of a multivariate survivor function in longitudinal studies*, Journal of the American Statistical Association, **97**, 494-507

122. van der Vaart, A.W and Wellner, J.A (1996) Weak Convergence and **Empirical Processes: With Applications to Statistics**, Springer-Verlag, New York.

123. Wang, W and Wells, M.T. (1997) *Non-parametric estimators of the bivariate survival function under simplified censoring conditions*, Biometrika, **84**, 863-880.

124. Wang, W and Wells, M.T. (1998) *Non-parametric estimation of successive duration times under dependent censoring*, Biometrika, **85**, 561-572.

125. Wang, W. and Wells, M.T (1999) *Semi-parametric estimation of the bivariate survival function*, Technical report, Institute of Statistics, Academia Sinica, Taipei.

126. Wei, L.J., Lin, D.Y and Weissfeld, L. (1989) *Regression analysis of multivariate incomplete failure time data by modeling marginal distributions*, Journal of the American Statistical Association, **84**, 1065-1073.

127. Ying, Z. and Wei, L.J. (1994) *The Kaplan-Meier estimate for dependent failure time observations*, Journal of Multivariate Analysis, **50**, 17-29.