MODELS FOR PROGNOSTIC VARIABLES IN MATCHED GROUPS WITH CENSORED DATA

Carol Jagger, B.Sc. (Hons), M.Sc. (Leeds)

Submitted for the degree of Doctor of Philosophy

University of Leicester 1984

CONTENTS

PAGE NUMBER

CHAPTER 8 - APPENDICES

PREFACE

This thesis is organised and presented in nine chapters.

The first chapter, the introduction, is in two broad sections and begins by discussing the origins of matched data and the reasons for matching. The general problems of censored data are mentioned and brief descriptions of the past attempts to analyse matched censored data are given, together with their shortcomings. The second section defines the notation used and presents the background to the failure time distributions and the types of censoring considered.

Chapter 2 is concerned with the analysis of data from the proportional hazards model. The two existing methods are reviewed then a new solution, the integrated method, is proposed and the theory developed. These methods are compared in the following chapter, Chapter 3.

Chapter 4 concentrates on data arising from the normal theory accelerated failure model. The previous solution is discussed and the results are derived for a new solution based upon the EM algorithm. This is extended to allow for right and interval censored data. The existing solution and the new solution are compared in Chapter 5.

Chapter 6 provides analyses of some data sets to compare the results arising from the new methods and the existing solutions, in a practical framework.

Chapter 7 discusses the relative merits of the new methods as compared with the previous solutions in the analysis of matched censored data and concludes with an outline of other areas in this field which require further research and the way in which the problems might be tackled.

Chapter 8 comprises four appendices whilst Chapter 9 lists the references cited in the text.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor P. R. Freeman for his advice and constructive criticism.

I am also grateful to my colleagues in the Department of Community Health for their encouragement and patience, in particular David Clayton for helpful discussion and John Woods for computing advice. Special thanks go to Illa Vora for typing the majority of this manuscript.

Part of this work was carried out whilst the author held an S.R.C. studentship.

CHAPTER 1

INTRODUCTION

1.1 The problem - past and present.

It has long been realised in clinical trials and epidemiological studies that variation in a variable between individuals is often of such a magnitude to obscure the effect of other differences that there might be such as treatment. In an attempt to overcome this, the groups to be compared were matched for factors which accounted for some of the variation. These factors were often age and sex. To tighten this up further, one-to-one matching was introduced in the form of the case-control study.

The special feature of studies involving twin samples is that they provide opportunities for the control of genetic and environmental factors to an extent impossible when using cases and controls. Monozygotic (MZ) twins are genetically identical and, even dizygotic (DZ) twins, although they are less concordant than MZ twins, are much more similar than the usually employed age- and sex- matched control. Twins share the same environment before birth and often very similar environments after birth . Hence MZ twins may be used to investigate effects of environmental factors since the 'noise' produced in case-control studies by genetic variability and/or gene-environment interaction, is absent.

The likelihood for matched pair data will necessarily contain one parameter for every pair in the data set. These are known as nuisance parameters since they are not in themselves of interest and it is not required that they be estimated. By looking at the difference or ratio of the outcome measures within the pairs these parameters can usually be conditioned out of the likelihood, leaving only the parameters of interest to be estimated, as in the case of the paired t-test.

When the outcome measure is time to an event such as death or remission, difficulties arise since it is possible that for some individuals the event may not have occurred by the end of the study. These observations are said to be censored and only a lower bound is known for the time to the event. Censoring can also occur if the event took place prior to the start of the study, if follow-up is at stated intervals only and the event cannot be placed exactly but only to have happened within an interval, or if death occurs from some unrelated cause or the patient is lost to follow-up or withdraws. If matched pair studies use an outcome measure which is censored for some of the data then the nuisance parameters cannot be conditioned out of the likelihood by simply considering the difference in the outcome measures within pairs.

The history of twin studies to investigate the effects of factors goes back to the 1960 's when WHO sponsored a meeting (WHO,1966) which reviewed how twins might contribute to the study of chronic diseases. The discussion centred mainly on the use of twins who were discordant in the factor of interest (e.g. smoking in respect of coronary heart disease). It was mentioned that these sets of twins may not be typical of all twins. Indeed, if a high proportion of pairs in the sample are concordant with respect to the exposure factor it seems intuitively wrong to attach no weight to these by ignoring their outcomes. Since that time a number of countries have set up twin registers. Both the Swedish and U.S. Registers have been used to investigate the effect of smoking by observing morbidity in twins (Liljefors,1970; Cederlof et al,1969). The analyses used contingency table methods for the estimate of risk with the twins who smoked forming the rows and their partners the columns. Again only twins who were discordant in respect of smoking, were analysed. Discordant twins on the Swedish register have been used to compare the extent of Ischaemic Heart Disease (IHD) in twins whose co-twin had died of IHD as compared with those twins whose co-twin had died of causes other than IHD (de Faire,1974). The Danish Twin register and the Danish psychiatric register established at the Institute of Human Genetics have provided data for a study of schizophrenia (Fischer et al,1969). With no censoring present these data can be analysed using existing statistical procedures for matched data. Registers have, more recently, been set up in Toronto, Budapest and Australia. To emphasise the growth that has taken place in this field, regular annual symposia are held on twin studies. In the laboratory setting twin samples become litter matched experiments e.g. the data of Mantel et al (1977) to study the effect of a carcinogen.

Natural pairing may also occur in situations other than human or animal twin studies. Batchelor and Hackett (1970) investigated the length of survival of closely matched and poorly matched skin grafts on the same individual. Another example is the Diabetic Retinopathy study , begun in 1971, which was a controlled clinical trial designed to determine whether photocoagulation was of benefit in preserving the vision of patients with proliferative diabetic retinopathy. One eye of each patient was randomly assigned to the treatment and the other was followed up without treatment (Diabetic Retinopathy Study Research Group,1976). No attempt was made in this paper to use the matching and the data were examined as two separate groups :- treated and untreated. Pairs of kidneys from the same donor transplanted into different recipients may also be looked at in this light (Kuhback and Tiilikainen,1975). The problem of censored times was overcome by Kuhback and Tiilikainen by considering the binary outcome success/ failure as opposed to the time to failure. This generally results in much loss of information.

Over the last decade since the paper by Cox (1972) a large theory has been built up around survival time data analysis. In the unrelated sample field there now exists parametric, non-parametric and Bayesian methods of analysis for grouped and ungrouped survival times together with goodness-of-fit tests, and graphical techniques. However, the analysis of data in the form of survival times of matched pairs/n-tuplets has not yet been resolved satisfactorily. In fact there are two aspects which might be of interest:- (i) the association between pairs i.e familial resemblance (with respect to the outcome variable) or (ii) the effect of some treatment or other covariate on the outcome. The former problem has been looked at by Clayton (1978), Mak and Ng (1981) and Oakes (1982) and will not be pursued here. The purpose of this thesis is to propose a solution to the second problem not only in terms of testing the treatment effect but also on estimating its magnitude. This was first examined by Sampford and Taylor as early as 1959. However no further progress was made until the paper of Holt and Prentice (1974) which used Cox's 1972 paper to define the problem more rigidly and to propose two solutions - a non-parametric one and a solution based on the marginal likelihood. This paper will be discussed in detail in Chapter 2. Both Sampford and Taylor and Holt and Prentice used the observable part of the likelihood obtained when the nuisance parameters have been conditioned out. This involves omitting from the analysis those pairs in which the ordering of the times is indeterminate. Since 1974 rank tests have been devised by Mantel et al (1977), Mantel and Ciminera (1979), Wei (1980) and Woolson and Lachenbruch (1980) although these all test the hypothesis of no effect of the covariate rather than actually computing the magnitude of the effect. The methods of Mantel and Ciminera involves an extension of the logrank procedure. A logrank score is produced for each member irrespective of litter and then these are summed within litters to produce a score for each litter. Allowance can only be made for other covariates by stratification on these covariates and, since especially in the laboratory situation, trials rarely involve more than a hundred pairs and often much less, this can result in small numbers in each strata. The approach of Mantel et al involved a modified Mantel-Haenzel test and hence only the rank order of the times to failure are used. Both methods of Mantel assume weak intra-pair correlation. Woolson and Lachenbruch extend the usual rank test for matched pairs to include censored observations. The test allows only for matched pairs and since it is based on within pair differences, pairs which are both censored are omitted from the analysis as in the methods of Holt and Prentice and Sampford and Taylor. Woolson and Lachenbruch also make the assumption that the censoring time for each member of the pair is the same although these censoring times are allowed to differ between pairs. This is clearly a strong assumption to make.

Hence a method of analysis is required which, using all the data, can not only estimate the size of the effect of a factor but test its effect as well. It will be necessary to make some assumptions about the parameters which describe the matching but this can be offset against the relative ease of dealing with other covariates and with the lack of assumptions about the censoring times, other than the usual one of independence between the censoring mechanism and the factor of interest.

1.2 Notation.

1.2.1 Failure time distributions.

Let T be a nonnegative random variable representing the failure time of an individual. The distribution of T may be described by any one of the following functions:

(i) Density function, f(t)

$$f(t) = \mathop{Lt}_{\delta t \to 0} \frac{Pr(t < T < t + \ t)}{\delta t} \quad .$$

(ii) Distribution function, F(t)

$$F(t) = Pr(T \leq t) \quad .$$

(iii) Survivor function, S(t)

$$S(t) = Pr(T \geq t) \quad .$$

(iv) Hazard function (instantaneous failure rate, intensity)

$\lambda(t)$

$$\lambda(t) = \underset{\delta t \to 0}{\mathrm{Lt}} \frac{\Pr(t < T < t + \delta t \mid T > t)}{\delta t} .$$

(v) Integrated hazard, $\Lambda(t)$

$$\Lambda(t) = \int_0^t \lambda(u)du .$$

These functions are interrelated by the following expressions.

$$\lambda(t) = \frac{f(t)}{F(t)}$$

$$f(t) = \lambda(t)\exp[-\Lambda(t)]$$

$$S(t) = \exp[-\Lambda(t)]$$

$$\lambda(t) = \frac{\partial}{\partial t} \left\{ -\ln S(t) \right\} .$$

There are a number of parametric models which have been used for failure time distributions. Three of these will be considered here. The first two, the exponential and Weibull, are frequently used because these distributions give closed form expressions for survivor functions and hazard functions. The third, the log normal distribution, does not have this property but is still frequently used as it has other useful properties.

(a) The Exponential Distribution.

The density function is defined as

$$f(t) = \lambda \exp(-\lambda t) \qquad t \geqslant 0, \ \lambda > 0$$

for some constant λ. This gives the hazard rate $\lambda(t) = \lambda$, a constant, which implies that the probability of failure in a time interval of given length is independent of the length of time the individual has been on trial.

The survivor function, S(t) is given by

$$S(t) = \exp(-\lambda t) \ .$$

(b) The Weibull Distribution.

This distribution is a generalisation of the exponential distribution. The two parameter Weibull distribution (λ, $\eta > 0$) has density function, hazard rate and survivor function given by

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

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

$$S(t) = \exp[- (\lambda t)^{\eta}] \quad .$$

When $\eta = 1$ this reduces to the exponential distribution. The hazard function is monotone decreasing for $\eta < 1$ and monotone increasing for $\eta > 1$.

(c) <u>The Log-Normal Distribution.</u>

Let

$$Z(v) = (2\pi)^{-\frac{1}{2}} \exp(- v^2 /2)$$

be the ordinate of a standard normal variate,

$$Q(v) = \int_v^{\infty} Z(u)du \quad ,$$

and the Mills ratio be defined by

$$M(v) = \frac{Z(v)}{Q(v)} \quad .$$

Then if T has a log-normal distribution, ln T is normally distributed, giving for $\sigma > 0$

$$f(t) = \frac{1}{\sigma t} Z\left(\frac{\ln t - \mu}{\sigma}\right)$$

$$S(t) = Q\left(\frac{\ln t - \mu}{\sigma}\right)$$

and

$$\lambda(t) = \frac{1}{\sigma t} M\left(\frac{\ln t - \mu}{\sigma}\right) .$$

The hazard function is zero at t = 0, increasing to a maximum and then decreasing more slowly, approaching zero again as t tends to infinity.

A distribution which can be used to model failure time but which will be met here in the context of the general linear model representation is the extreme minimum value distribution.

(d) Extreme Minimum Value Distribution.

If t has an exponential (λ) distribution and w = $(1/\eta)\ln t$ then w is said to have an extreme minimum value distribution defined by

$$f(w) = \lambda\eta \exp[\eta w - \lambda \exp(\eta w)] \qquad -\infty < w < \infty$$

$$S(w) = \exp[-\lambda \exp(\eta w)]$$

$$\lambda(w) = \lambda\eta \exp(\eta w) .$$

The unit extreme minimum value distribution with $\lambda = \eta = 1$ has density

$$f(w) = \exp[\, w - \exp(w)] \,.$$

1.2.2 The General Linear Model.

The exponential, Weibull and log normal distribution are contained in the log linear model represented by

$$\ln t = \mu + \sigma\xi \qquad (1.1)$$

where μ and σ are constants and ξ has some distribution $f(\xi)$.

Table 1.1 below classifies the three distributions according to this representation.

Distribution of t.	μ	σ	$f(\xi)$
exponential	$-\ln\lambda$	1	s.e.v.
Weibull	$-(1/\eta)\ln\lambda$	$1/\eta$	s.e.v.
lognormal	μ	σ	s.n.d.

s.e.v standard (unit) extreme value

s.n.d standard normal distribution.

Table 1.1.

Form of the general linear model when the survival times have exponential, Weibull or lognormal distributions.

1.2.3 Effect of Covariates.

There are often explanatory variables (or covariates) upon which the failure time may or is known to depend. Usually certain covariates are of interest e.g presence or absence of treatment; others may be known to affect the fa-ilure time and therefore need to be allowed for e.g age, stage of a disease.

The two main forms in which covariates affect the fai-lure time, are known as the Proportional Hazards class and the Accelerated Failure Class. Let $\underline{z}^T = (z_1, \ldots, z_p)$ be the observed vector of covariates for an individual.

(a) The Proportional Hazards Class.

The covariates are assumed to act multiplicatively on the hazard rate. If $\lambda(t;\underline{z})$ is defined as the hazard rate at time t for an individual with observed covariate vector $\underline{z}$ then this can be written as

$$\lambda(t;\underline{z}) = \lambda_0(t)\ h(\underline{\beta}^T \underline{z})$$

where $\lambda_0(t)$ is some baseline hazard rate, $\beta^T = (\beta_1, \ldots, \beta_p)$ is a vector of regression parameters and $h(\cdot)$ is a specified function. Forms for $h(\underline{\beta}^T \underline{z})$ which have been assumed are $h(\underline{\beta}^T \underline{z}) = (1 + \underline{\beta}^T \underline{z})$; $h(\underline{\beta}^T \underline{z}) = (1 + \underline{\beta}^T \underline{z})^{-1}$ and $h(\underline{\beta}^T \underline{z}) = \exp(\underline{\beta}^T \underline{z})$.

The latter, employed by Cox (1972), is the form most used since $h(\underline{\beta}^T \underline{z}) > 0$ for all possible $\underline{z}$. This form will be used throughout.

This gives

$$\lambda(t;\underline{z}) = \lambda_0(t)\exp(\underline{\beta}^T \underline{z})$$

and the log linear model (1.1) becomes

$$\ln t = \mu - \underline{\beta}^T \underline{z} + \sigma\xi .$$

The function $\lambda_0(t)$ may be allowed to vary within subsets (strata) of the data. If the data are divided into s strata then the hazard function can be written as

$$\lambda_j(t;\underline{z}) = \lambda_{0j}(t)\exp(\underline{\beta}^T \underline{z}) \qquad (j=1,\ldots,s) \qquad (1.2)$$

(b) <u>The Accelerated Failure Class.</u>

A disadvantage with the proportional hazards model is that there is no direct relationship between t and $\underline{z}$. The relationship with the hazard is very difficult to communicate to clinicians. With this class of models the covariates act multiplicatively on the failure time, in the form

$$T = \exp(\underline{\beta}^T \underline{z})T'$$

where T' has baseline hazard $\lambda_0(t')$, independent of $\underline{\beta}$.

In terms of the hazard function this is

$$\lambda(t;\underline{z}) = \lambda_0(te^{-\underline{\beta}^T\underline{z}})e^{-\underline{\beta}^T\underline{z}}$$

and the general linear model

$$\ln t = \mu + \underline{\beta}^T\underline{z} + \sigma\xi \; .$$

The difference in sign of $\underline{\beta}^T\underline{z}$ between the two classes is due to the fact that the proportional hazards model is linking the covariates to the probability of failure in the next instant whilst the accelerated failure class links the covariates to the survival time.

1.2.4 Censoring.

If it is only known that a failure time, T, exceeds some value, t^* say, this is said to be a right censored observation. i.e. $T > t^*$. Similarly a left censored observation is one where $T < t^*$. Interval censoring is where T is known to lie in some interval (t_1^*, t_2^*). Hence right and left censoring are special cases of interval censoring with the intervals being (t^*, ∞) and $(-\infty, t^*)$ respectively. Left and interval censoring are not often found in practice and this work will concentrate on the problem of right censored observations although left censoring could be accommodated without much difficulty. One mechanism to produce right censored observations is the assumption of random censorship. If $T_1, \ldots, T_n$ and $C_1, \ldots, C_n$ are the failure and censoring times respectively for n individuals with T_i , C_i ,i=1,...,n being all stochastically independent then the observation for the i-th individual is

$$Y_i = \min(T_i \ , \ C_i \ )$$

Special cases of this are Type I, Type II and Progressive Type II censoring. For Type I censoring the C_i are fixed in advance, and this most frequently occurs when items enter on test at the start of, or randomly throughout, a study period of fixed length and observation ceases at this prearranged time. Type II censoring occurs when all the items enter together and the study continues until the r-th smallest failure time is observed. In Progressive Type II censoring, all items again enter together but d_1 are censored at the r_1-th smallest time, a further r_2 failures are observed, and d_2 are censored, etc.

A more general censoring mechanism, of which the above are three special cases, is that of independent censoring. The assumption in this case is that, for an individual with given covariate vector $\underline{z}$, at time t the censoring mechanism at t does not depend upon $\underline{z}$. Hence individuals are not withdrawn because they are at a high or low risk of failure. In some studies, especially clinical trials, this may not be a valid assumption, as a clinician may withdraw his patient from a trial if the patient's condition has deteriorated. A simple check of comparing the covariate values for those censored and those who failed, can be made on the data, before analysis. This will help to ascertain whether the assumption of a random censoring mechanism is reasonable although it is by no means infallible. If this seems a valid assumption then, since the censoring is not informative about $\underline{\beta}$ (the regression parameter), the contribution to the likelihood for a failure at time t is $f(t; \underline{\beta}, \underline{z})$ and for a censored observation $S(t; \underline{\beta}, \underline{z})$.

With regard to the problem considered in this thesis, of survival time analysis for matched n-tuplets it will be assumed that the likelihood factors out in the above way. It remains to be shown what restrictions this makes on the C_i for each member of the n-tuplet but it seems reasonable that they should be independent of the T_i but may themselves be correlated, i.e C_i for members of the same n-tuplet may be related. This is likely to happen in practice as matched pairs e.g twins are usually recruited to a study at the same time.

CHAPTER 2

PROPORTIONAL HAZARDS MODEL

2.1 The Model.

Let t_{ji} , $\underline{z}_{ji}$ be the observed times and covariate vector for the j-th individual, j=1,...,n, belonging to the i-th n-tuplet, i=1,...,N. A censoring indicator δ_{ji} is observed for each individual with

$$\delta_{ji} = \begin{cases} 1 & \text{if the individual fails at } t_{ji} \\ 0 & \text{if the individual is censored at } t_{ji} \end{cases}$$

The proportional hazards model stated, by Holt and Prentice took the form of (1.2) with the n-tuplets forming the strata. If the hazard rate at time t for the j-th individual in the i-th n-tuplet is $\lambda_{ji}(t;\underline{z}_{ji})$ this means

$$\lambda_{ji}(t;\underline{z}) = \lambda_{oi}(t)\exp[\underset{\sim}{\beta}^T \underline{z}] \qquad j=1,\ldots,n;\ i=1,\ldots,N.$$

A further assumption was made by Holt and Prentice that the matching properties themselves acted multiplicatively on the baseline hazard. This gives rise to

$$\lambda_{ji}(t;\ \underline{z}_{ji}) = \alpha_i \lambda_0(t)\exp[\underline{\beta}^T \underline{z}_{ji}] \tag{2.1}$$

i.e each n-tuplet shares an α_i $(i=1,..,N)$ which can be regarded as a 'matching' variable. From this, three models are considered. Firstly the non-parametric model of Holt and Prentice which is defined by (2.1). The two parametric models considered for $\lambda_0(t)$ are the exponential and the more general Weibull. These are given by

$$\lambda_{ji}(t;\underline{z}_{ji}) = \alpha_i \exp[\underline{\beta}^T \underline{z}_{ji}] \tag{2.2}$$

and

$$\lambda_{ji}(t;\underline{z}_{ji}) = \alpha_i\, t^{\eta-1} \exp[\underline{\beta}^T \underline{z}_{ji}] \tag{2.3}$$

For simplicity the theory will take the case of n = 2 i.e matched pairs and a single covariate with dim $(\underline{z}_{ji}) = 1$, the corresponding covariate values being z_{1i} and z_{2i} . Sections 2.2 and 2.3 are , in the main, reviews of the work of Holt and Prentice.

2.2 The Non-Parametric Approach of Holt and Prentice.

Consider the hazard rate given by (2.1). In the paired case and single covariate this reduces to

$$\lambda_{ji}(\underline{t},\ z_{ji}) = \alpha_i \lambda_0(t)\ \exp[\beta\, z_{ji}] \quad j=1,2;\ i=1,\ldots,N.$$

The main problem is the estimation of β in the presence of the 'nuisance pararameters' α_i. In the case of no censoring and a covariate independent of time the ranks of the t_{ji} within pairs are sufficient for β in the absence of knowledge of $\lambda_0(t)$, (Kalbfleisch and Prentice,1973). Hence the marginal likelihood for β , $L_N(\beta)$, is proportional to the product, over the pairs i=1,..,N, of

$$\mathrm{pr}\ (t_{2i} < t_{1i}) = [1+\exp(\beta d_i)]^{-1}$$

$$\mathrm{pr}\ (t_{1i} < t_{2i}) = [1+\exp(-\beta d_i)]^{-1}$$

with $d_i = z_{1i} - z_{2i}$. If $\theta = \exp(\beta)$ then the ratio of hazards is

$$\frac{\lambda_{1i}(t;z_{1i})}{\lambda_{2i}(t;z_{2i})} = \exp[\beta d_i] = \theta^{d_i}$$

and for the usual 'treatment' covariate with $z_{1i} = 1$ and $z_{2i} = 0$, θ measures the hazard ratio. Therefore

$$L_N(\beta) \propto \prod_{i=1}^{N} [1 + \exp(\epsilon_i \beta d_i)]^{-1} , \qquad (2.4)$$

where

$\epsilon_i = 1$ if $t_{2i} < t_{1i}$

and

$\epsilon_i = -1$ if $t_{1i} < t_{2i}$.

There is no contribution from tied pairs , i.e where $t_{1i} = t_{2i}$.

If censoring is included then the pair ranks are known for pairs with both times uncensored i.e $\delta_{1i} = \delta_{2i} = 1$, and for pairs with one of the pair censored, provided $\min(t_{1i}, t_{2i})$ is uncensored. The pair ranks are indeterminable for pairs with $\delta_{1i} = \delta_{2i} = 0$ and those pairs with $\min(t_{1i}, t_{2i})$ censored. Figure 2.1 below shows the possibile combinations for pairs (t_{1i}, t_{2i}), with $\epsilon_i = 0$ if the pair ranks are unknown.

			ϵ_i
(a)	t_{1i} $\bullet$	t_{2i} $\bullet$	1
(b)	t_{2i} $\bullet$	t_{1i} $\bullet$	-1
(c)	t_{1i} $\longrightarrow$	t_{2i} $\bullet$	0
(d)	t_{2i} $\longrightarrow$	t_{1i} $\bullet$	0
(e)	t_{1i} $\bullet$	t_{2i} $\longrightarrow$	-1
(f)	t_{2i} $\bullet$	t_{1i} $\longrightarrow$	1
(g)	t_{1i} $\longrightarrow$	t_{2i} $\longrightarrow$	0
(h)	t_{2i} $\longrightarrow$	t_{1i} $\longrightarrow$	0
(i)	t_{1i} $\longrightarrow$ t_{2i} $\longrightarrow$		0

0 $\longrightarrow$ Time

$\bullet$ denotes a failure

$\longrightarrow$ denotes a censored observation.

Figure 2.1

Holt and Prentice considered only the Type 1 censoring mechanism in which both members enter at the same time and are censored only if they survive the study period. In this case observations of type (c),(d),(g),(h) will not occur. Holt and Prentice argue that those pairs for whom the pair ranks are undefined provide no information about β and thus the likelihood for β is still given by (2.4), if ϵ_i is taken as zero for those pairs providing no information about β i.e (c),(d),(g),(h) and (i). The log likelihood for β is given by

$$\ln L_N(\beta) = -\sum_{i=1}^{N} \ln[1 + \exp(\epsilon_i \beta d_i)]$$

and therefore

$$\frac{\partial \ln L_N(\beta)}{\partial \beta} = -\sum_{i=1}^{N} \frac{\epsilon_i d_i \exp(\epsilon_i \beta d_i)}{1 + \exp(\epsilon_i \beta d_i)} \qquad (2.5)$$

and the maximum likelihood estimate of β, $\hat{\beta}$, is given by the solution of

$$\sum_{i=1}^{N} \frac{\epsilon_i d_i}{1+\exp(-\epsilon_i \beta d_i)} = 0 \qquad (2.6)$$

This shows that pairs which are concordant with respect to the covariate, i.e with $d_i = 0$, do not contribute to the likelihood.

For the case of a single covariate representing treatment, say, with $z_{1i} = 1$, $z_{2i} = 0$ and hence $d_i = 1$, $i=1,\ldots,N$, the work of Holt and Prentice can be extended further with (2.6) having an explicit solution. Assume no ties and let q pairs have $\epsilon_i = 1$, r-q pairs have $\epsilon_i = -1$ and N-r pairs have $\epsilon_i = 0$. Then $\hat{\beta}$ is given by the solution of

$$\frac{q}{1+\exp(-\beta)} = \frac{(r-q)}{1+\exp(\beta)}$$

and since $\exp(\beta) > 0$

$$\hat{\beta} = \ln \frac{(r-q)}{q} .$$

From the above equation for $\hat{\beta}$ can be seen a disadvantage of the non-parametric method, since it can produce a M.L. estimate of β, $\hat{\beta} = -\infty$ (i.e $\hat{\theta} = 0$) if $t_{1i} > t_{2i}$ for all $i=1,\ldots,N$ or $\hat{\beta} = \infty$ ($\hat{\theta} = \infty$) if $t_{1i} < t_{2i}$ for all $i=1,\ldots,N$. The probability of this occurrence in the case of no censoring and the single treatment indicator covariate is $(1+\theta)^{-N}+(1+\theta^{-1})^{-N}$, since this is the probability of 0 or N successes in N binomial trials each with probability $(1+\theta)^{-1}$ of success. Table 2.1 below gives selected values of this probability for varying values of θ (the true hazard ratio) and N. The effect of censoring is to increase this probability as ignoring pairs with $\epsilon_i = 0$ effectively reduces the sample size N.

N \ θ	0.1	0.25	0.5	0.75	1.0	1.5
5	0.621	0.328	0.127	0.075	0.063	0.088
10	0.386	0.107	0.017	0.004	0.002	0.006
20	0.149	0.012	0.000	0.000	0.000	0.000
30	0.057	0.001	0.000	0.000	0.000	0.000
60	0.000	0.000	0.000	0.000	0.000	0.000

Table 2.1.

Probabilities of obtaining $\hat{\theta} = 0$ or ∞ for N matched pairs with one covariate and no censoring when the true hazard ratio is θ.

The asymptotic variance of the M.L.E of β, $\hat{\beta}$, given by the solution of (2.6), can be estimated by $I^{-1}(\hat{\beta})$ where $I(\beta)$ is the observed information, i.e.

$$I(\beta) = \frac{-\partial^2 \ln L_N(\beta)}{\partial \beta^2} .$$

$I(\beta)$ is a consistent estimator of the Fisher information $\mathcal{I}(\beta)$, with

$$\mathcal{I}(\beta) = -E\left[\frac{\partial^2 \ln L_N(\beta)}{\partial \beta^2}\right].$$

Thus, differentation of (2.5) gives

$$I(\beta) = \sum_{i=1}^{N} \frac{(\epsilon_i d_i)^2 \exp(\epsilon_i \beta d_i)}{[1+\exp(\epsilon_i \beta d_i)]^2} \;.$$

If, as before, $d_i = z_{1i} - z_{2i} = 1$ for all $i=1,\ldots,N$ with q pairs having $\epsilon_i = 1$ (i.e $t_{2i} < t_{1i}$), r-q pairs having $\epsilon_i = -1$ and N-r pairs having $\epsilon_i = 0$ then this is simplified to give

$$I(\beta) = \frac{r \exp(\beta)}{[1+\exp(\beta)]^2}$$

and hence

$$S.E(\hat{\beta}) = \sqrt{I^{-1}(\hat{\beta})} = \frac{\exp(-0.5\hat{\beta})+\exp(0.5\hat{\beta})}{\sqrt{r}}$$

2.3 The Marginal Likelihood Approach of Holt and Prentice.

Rather than using only the within pair ranks of the twins, Holt and Prentice showed that marginal likelihoods can be derived for β for the models (2.2) and (2.3) by conditioning out the nuisance parameters α_i. In the case of all failures being observed, by conditioning on some statistic w_i, the likelihood is partitioned into two parts viz.

$$L(\text{data} \mid \beta, \alpha_i) = L(\text{data} \mid w_i, \beta, \alpha_i) L(w_i \mid \beta, \alpha_i) \;, \quad (2.7)$$

with the w_i being chosen so that $L(w_i | \beta, \alpha_i)$ is independent of α_i. This quantity is then the marginal likelihood for β since the w_i are marginally sufficient for β in the absence of knowledge of the α_i.

For the Weibull model defined by (2.3), and a single covariate z_{ji}, j=1,2, then

$$f(t_{ji}) = \alpha_i e^{\beta z_{ji}} t_{ji}^{\eta-1} \exp\left\{-\frac{\alpha_i}{\eta} e^{\beta z_{ji}} t_{ji}^{\eta}\right\} \qquad \begin{matrix} j=1,2 \\ i=1,\dots,N, \end{matrix}$$

and the linear model formulation is

$$\ln t_{ji} = \frac{-1}{\eta} \ln\left(\frac{\alpha_i}{\eta}\right) - \frac{\beta}{\eta} z + \frac{1}{\eta} \xi_{ji} \qquad \begin{matrix} j=1,2 \\ i=1,\dots,N, \end{matrix}$$

where ξ_{ji}, j=1,2 are independent with standard extreme value distributions.

Hence, taking $w_i = t_{1i}/t_{2i}$ and $d_i = z_{1i} - z_{2i}$ gives

$$\ln w_i = -\frac{\beta}{\eta} d_i + \frac{1}{\eta} \psi_i ,$$

with ψ_i having a logistic distribution i.e.

$$f(\psi_i) = \frac{\exp(\psi_i)}{[1+\exp(\psi_i)]^2} \qquad \begin{matrix} -\infty < \psi_i < \infty, \\ i=1,\dots,N. \end{matrix}$$

Thus the marginal density of w_i is

$$f(w_i|\beta,\alpha_i) = \frac{\eta \; e^{\beta d_i} w_i^{\eta-1}}{[1+e^{\beta d_i} \; w_i^{\eta} \;]^2} \tag{2.8}$$

and the density for the exponential model defined by (2.2) is found by putting η = 1 in the above equation. The marginal likelihood for the Weibull model is therefore given by $L_M(\beta,\eta)$ with

$$L_M(\beta,\eta) = \prod_{i=1}^{N} \frac{\eta \; e^{\beta d_i} w_i^{\eta-1}}{[1+e^{\beta d_i} \; w_i^{\eta} \;]^2} \quad . \tag{2.9}$$

Now consider the effect of censoring. Let W_i be a random variable with probability density function given by (2.8), and $w_i = t_{1i}/t_{2i}$. Then, returning to Figure 2.1, for cases (a),(b) there is a contribution to the likelihood of the form $\eta \, e^{\beta d_i} \, w_i^{\eta-1} \, [1+e^{\beta d_i} \, w_i^{\eta}]^{-2}$;in cases (c) and (f) all that is known of W_i is $W_i > w_i$, giving a contribution to the likehood of $[1+e^{\beta d_i} \, w_i^{\eta}]^{-1}$; cases (d) and (e) contribute $e^{\beta d_i} \, w_i^{\eta} \, [1+e^{\beta d_i} \, w_i^{\eta}]^{-1}$ since $W_i < w_i$. The doubly censored pairs (g),(h),(i) provide no contribution and Holt and Prentice omit these though they state that this omission has the tendency to bias the estimate of β away from the value β = 0. With these pairs omitted the likelihood is no longer a true likelihood.

The marginal likelihood (2.9) can be amended for censoring as follows

$$L_M(\beta,\eta) \propto \prod_{A_1} \frac{\eta\, e^{\beta d_i} w_i^{\eta-1}}{[1+e^{\beta d_i} w_i^{\eta}]^2} \prod_{A_2} \frac{1}{[1+e^{\beta d_i} w_i^{\eta}]} \prod_{A_3} \frac{e^{\beta d_i} w_i^{\eta}}{[1+e^{\beta d_i} w_i^{\eta}]}, \qquad (2.10)$$

where $A_1 = \{$ pairs of observations of type (a),(b) i.e both members of pair uncensored$\}$; $A_2 = \{$ pairs of type (c),(f) i.e pairs with t_{1i} censored and t_{2i} failed$\}$; $A_3 = \{$ pairs of type (d),(e) i.e t_{1i} failed and t_{2i} censored$\}$.

The log likelihood is given by

$$\ln L_M(\beta,\eta) = \beta \sum_{A_1 \cup A_3} d_i + \eta \sum_{A_1 \cup A_3} \ln w_i + N_1 \ln \eta - \sum_{i=1}^{N} \epsilon_i \ln[1+e^{\beta d_i} w_i^{\eta}] + \text{const} \qquad (2.11)$$

where N_1 = number of pairs in set A_1 and ϵ_i is the number of uncensored times in (t_{1i}, t_{2i}), i.e. $\epsilon_i = \delta_{1i} + \delta_{2i}$.

The maximum likelihood estimate of β and η are the solutions of

$$\sum_{A_1 \cup A_3} d_i = \sum_{i=1}^{N} \frac{\epsilon_i\, d_i\, e^{\beta d_i} w_i^{\eta}}{[1+e^{\beta d_i} w_i^{\eta}]} \qquad (2.12)$$

and

$$\frac{N_1}{\eta} + \sum_{A_1 \cup A_3} \ln w_i = \sum_{i=1}^{N} \frac{\epsilon_i\, e^{\beta d_i} w_i^{\eta} \ln w_i}{[1+e^{\beta d_i} w_i^{\eta}]}. \qquad (2.13)$$

The maximum likelihood estimate of β for the exponential model (2.2) is the solution of the equation

$$\sum_{A_1 \cup A_3} d_i = \sum_{i=1}^{N} \frac{\epsilon_i d_i e^{\beta d_i} w_i}{[1+e^{\beta d_i} w_i]} . \tag{2.14}$$

From (2.14) it can be seen, as in the non-parametric method, that pairs which are concordant with respect to the covariate, do not contribute to the estimation of β in the exponential model.

Differentiating (2.11) twice gives the terms of the observed information matrix. In the uncensored case the Fisher information matrix can be shown to be

$$\frac{1}{c}\begin{pmatrix} \frac{N}{9}(3+\pi^2) + \frac{\beta^2 \sum d_i^2}{3\eta^2} & \frac{\beta \sum d_i^2}{3\eta} \\ \frac{\beta \sum d_i^2}{3\eta} & \frac{1}{3}\sum d_i^2 \end{pmatrix}$$

with $c = \frac{N}{27\eta^2}(3+\pi^2)\sum d_i^2$.

Thus the asymptotic variances of $\hat{\beta}$ and $\hat{\eta}$ are $\frac{3}{\sum d_i^2} + \frac{9\hat{\beta}^2}{N(3+\pi^2)}$ and $\frac{9\hat{\eta}^2}{N(3+\pi^2)}$ respectively.

For the exponential model the asymptotic variance of $\hat{\beta}$ is $3/\sum d_i^2$, i.e for a single covariate with $z_{1i} = 1$, $z_{2i} = 0$, i=1,...,N the asymptotic variance of $\hat{\beta}$ is 3/N.

2.4 Fitting in GLIM.

Holt and Prentice used the usual function maximisation routines to estimate β by the non-parametric method and β and η by the marginal method. The likelihoods of both of these methods can be written in the general linear model form and fitted using the GLIM package (Baker and Nelder,1978). Consider first the non-parametric likelihood. Since

$$L_N(\beta) = \prod_{i=1}^{N} \pi_i^{e_i} (1 - \pi_i)^{1-e_i}$$

with $e_i = 1$ if $t_{1i} < t_{2i}$ and zero otherwise and $\pi_i = [1+\exp(-\beta d_i)]^{-1}$, $L_N(\beta)$ is the likelihood for n binomial variates e_i with $pr(e_i = 1) = \pi_i$. Hence

$$\ln \frac{\pi_i}{1 - \pi_i} = \beta d_i$$

and in the GLIM notation this becomes \$YVAR $\{e_i\}$; \$LINK LOG; \$ERROR B N; \$WEIGHT W and \$FIT $\underline{d}$ where $\underline{d} = \{d_i\}$; N $= \{1\}$, a vector of 1's and the weight vector W is 0 for the pairs of type (c),(d),(g),(h) and (i) in Figure 2.1 and 1 otherwise.

The marginal likelihood is less simple. If, this time, $\pi_i = 1/[1 + w_i^{\eta} \exp(\beta d_i)]$, $\delta_i = \delta_{1i} + \delta_{2i}$ and $\epsilon_i = \delta_{1i}\delta_{2i}$ then (2.10) can be written as

$$L_M(\beta, \eta) \propto \prod_{i=1}^{N} \eta^{\epsilon_i} \pi_i^{\delta_{1i}} (1 - \pi_i)^{\delta_i - \delta_{1i}}$$

and $\ln[\pi_i/(1 - \pi_i)] = \beta d_i + \eta \ln w_i$.

If the exponential model is to be fitted, i.e. $\eta = 1$, or there are no uncensored pairs, then the likelihood is of the binomial form again and this is fitted by the following GLIM declarations :- \$YVAR $\{\delta_{1i}\}$; \$ERROR B $\{\delta_i\}$; \$LINK LOG; \$OFFSET $\{\ln w_i\}$; \$FIT $\{d_i\}$. When η is to be estimated and the data includes pairs whose members have both failed then estimation of β can be found using the method just mentioned. Estimation of η , however, requires either an iterative procedure or the inclusion of a dummy binomial observation, with observed proportion ϵ_i out of ϵ_i. A more general form of GLIM macros (of which these two are special cases) can be found in Bennett and Whitehead (1981) and Roger and Peacock (1983) respectively.

2.5 The Integrated Likelihood Approach.

One alternative to the two previously mentioned methods is to consider the α_i as fixed effects and to simply estimate them along with β and η . This approach has not been considered since as each pair contributes an α_i , the set of observations are only partially consistent with respect to the α_i. Even if the maximum likelihood estimate of β from the full likelihood is consistent it need not possess asymptotic efficiency (Neyman and Scott,1948). The problems of estimating many parameters are well known. For completeness it should be added that the full likelihood and marginal likelihood approaches give the same maximum likelihood estimate of β under the exponential model, although the asymptotic variances differ. It may also be assumed further, that $\alpha_i = \alpha$ for all $i=1,\ldots,N$, which reduces the number of parameters to be estimated but at the expense of the very strong assumption that the pairing variable is constant over all the pairs.

The solution to the problem which will be considered further here is to make some assumptions about the distribution of the α_i. From a frequentist viewpoint the α_i might be considered as random effects, coming from some distribution with known form though unknown parameters. If covariates are also being included then the model will be a mixed effects model. This is in contrast to the full likelihood method which treats the α_i as fixed effects to be estimated. Justified by the exchangeability of the pairs, a Bayes approach would be to feed in information about the pairing variables in the form of a prior distribution. The likelihood approach uses the resulting 'marginal' likelihood when the α_i have been integrated out. An Empirical Bayes procedure considers that if the pairs are selected randomly then it is reasonable to assume that the α_i are independently and identically distributed. In this thesis the distribution of the α_i will be referred to as a prior distribution, for ease, and this can be interpreted as above by the other statistical approaches, where necessary.

The full likelihood for the Weibull model is given by

$$L(\underline{t}_1, \underline{t}_2 | \beta, \eta, \underline{\alpha}) = \prod_{i=1}^{N} f(t_{1i}, t_{2i} | \beta, \eta, \alpha_i)$$
$$= \prod_{i=1}^{N} \alpha_i^{(\delta_{1i}+\delta_{2i})} e^{\beta(\delta_{1i}z_{1i}+\delta_{2i}z_{2i})} (t_{1i}^{\delta_{1i}} t_{2i}^{\delta_{2i}})^{\eta-1} \ x$$
$$\exp\left\{-\frac{\alpha_i}{\eta}(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}})\right\} , \quad (2.15)$$

where $\underline{t}_1 = (t_{11}, \ldots, t_{1N})$, $\underline{t}_2 = (t_{21}, \ldots, t_{2N})$, $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ and

$$\delta_{ji} = \begin{cases} 1 & \text{if } t_{ji} \text{ fails ,} \\ 0 & \text{if } t_{ji} \text{ is censored ,} \end{cases}$$

for $j=1,2$, $i=1,\ldots,N$.

A convenient form for the distribution of the α_i is to assume the α_i have a gamma distribution with parameters a and b, a, b > 0 ($\alpha_i \sim \gamma(a,b)$) such that

$$f(\alpha_i) = \frac{a^b}{\Gamma(b)} \alpha_i^{b-1} \exp(-a\,\alpha_i) \quad , \quad 0 < \alpha_i < \infty \quad (2.16)$$
$$i=1,\ldots,N .$$

As b tends towards zero, the prior distribution reduces to the case mentioned previously, where all the α_i are equal. The form with a = b so that the α_i have unit mean, has been taken by Clayton (1978) although the main concern in this paper was with the estimation of the association parameter 1/a which represents the variance of the α_i .

Holt and Prentice considered using prior distributions for α_i of the form α_i^{-1} and α_i^{-b} , i.e $\gamma(0,0)$ and $\gamma(0,1-b)$. The former gives an improper integral to evaluate for doubly censored pairs. They found that the second type of prior, for fixed b, gave a bias on the estimate of β exceeding that for the marginal and non-parametric methods and this bias was found to increase with b.

From (2.15) the unconditional likelihood of β and η is given by

$$L(\underline{t}_1, \underline{t}_2 | \beta, \eta, a, b) = \int_{\alpha_1=0}^{\infty} \cdots \int_{\alpha_N=0}^{\infty} \prod_{i=1}^{N} [f(t_{1i}, t_{2i} | \beta, \eta, \alpha_i) f(\alpha_i) d\alpha_i]$$

and changing the order of integration this gives

$$L(\underline{t}_1, \underline{t}_2 | \beta, \eta, a, b) = \prod_{i=1}^{N} \int_{\alpha_i=0}^{\infty} f(t_{1i}, t_{2i} | \beta, \eta, \alpha_i) f(\alpha_i) d\alpha_i$$

$$= \prod_{i=1}^{N} \frac{a^b (t_{1i}^{\delta_{1i}} t_{2i}^{\delta_{2i}})^{\eta-1} e^{\beta(\delta_{1i} z_{1i} + \delta_{2i} z_{2i})}}{\Gamma(b)} \times$$

$$\int_{\alpha_i=0}^{\infty} \alpha_i^{b+\delta_{1i}+\delta_{2i}-1} \exp\left\{ -\frac{\alpha_i}{\eta}(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta) \right\} d\alpha_i$$

$$= \prod_{i=1}^{N} a^{b} \eta^{b+\delta_{1i}+\delta_{2i}} \frac{\Gamma(b+\delta_{1i}+\delta_{2i}) (t_{1i}^{\delta_{1i}} t_{2i}^{\delta_{2i}})^{\eta-1} e^{\beta(\delta_{1i}z_{1i}+\delta_{2i}z_{2i})}}{\Gamma(b) \quad (t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta)^{b+\delta_{1i}+\delta_{2i}}} .$$

Thus, in contrast to the non-parametric and marginal likelihood approaches, doubly censored pairs contribute a term $(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta)^{-b}$ to the likelihood. Pairs concordant in the covariate also contribute. A paper by Wild (1983), published whilst this thesis was being written, used a similar solution under the heading of Empirical Bayes. The form of the baseline hazard $\lambda_0(t)$, however, was taken to be $\eta t^{\eta-1}$ for the Weibull model. This is equivalent to a reparametrisation of the above likelihood in terms of β, η, a' and b where $a' = a\eta$.

The log likelihood is

$$\begin{aligned}
\ln L_I(\underline{t}_1, \underline{t}_2 \mid \beta, \eta, a, b) = {} & \sum [\ln\Gamma(b+\delta_{1i}+\delta_{2i}) - \ln\Gamma(b)] \\
& + \beta \sum (\delta_{1i} z_{1i} + \delta_{2i} z_{2i}) + \ln\eta \sum (b+\delta_{1i}+\delta_{2i}) \\
& + \eta \sum (\delta_{1i} \ln t_{1i} + \delta_{2i} \ln t_{2i}) + Nb\ln a \qquad (2.17) \\
& - \sum (b+\delta_{1i}+\delta_{2i}) \ln[t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta] + \text{const}
\end{aligned}$$

the summation being taken for i=1,...,N.

Differentiating (2.17) gives

$$\frac{\partial \ln L_I}{\partial \beta} = \sum_i (\delta_{1i} z_{1i} + \delta_{2i} z_{2i}) - \sum_i \frac{(b+\delta_{1i}+\delta_{2i})(t_{1i}^{\eta} z_{1i} e^{\beta z_{1i}} + t_{2i}^{\eta} z_{2i} e^{\beta z_{2i}})}{(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta)} \tag{2.18}$$

$$\frac{\partial \ln L_I}{\partial \eta} = \frac{1}{\eta} \sum_i (b+\delta_{1i}+\delta_{2i}) + \sum_i (\delta_{1i} \ln t_{1i} + \delta_{2i} \ln t_{2i})$$

$$- \sum_i \frac{(b+\delta_{1i}+\delta_{2i})(t_{1i}^{\eta} e^{\beta z_{1i}} \ln t_{1i} + t_{2i}^{\eta} e^{\beta z_{2i}} \ln t_{2i} + a)}{(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta)} \tag{2.19}$$

$$\frac{\partial \ln L_I}{\partial a} = \frac{Nb}{a} - \sum_i \frac{(b+\delta_{1i}+\delta_{2i})\eta}{(t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta)} \tag{2.20}$$

$$\frac{\partial \ln L_I}{\partial b} = \sum_i \sum_{j=1}^{\delta_{1i}+\delta_{2i}} \frac{1}{b+j-1} + N\ln(a\eta)$$

$$- \sum_i \ln[t_{1i}^{\eta} e^{\beta z_{1i}} + t_{2i}^{\eta} e^{\beta z_{2i}} + a\eta] . \tag{2.21}$$

Maximum likelihood estimates for β , η , a, b are found from solutions of

$$\frac{\partial \ln L_I}{\partial \beta} = 0 \quad ; \quad \frac{\partial \ln L_I}{\partial \eta} = 0 \quad ;$$

$$\frac{\partial \ln L_I}{\partial a} = 0 \quad ; \quad \frac{\partial \ln L_I}{\partial b} = 0 \quad .$$

The maximum likelihood estimates for the exponential model are given by the solutions to the equations

$$\sum_i (\delta_{1i} z_{1i} + \delta_{2i} z_{2i}) = \sum_i \frac{(b+\delta_{1i}+\delta_{2i})(t_{1i} z_{1i} e^{\beta z_{1i}} + t_{2i} z_{2i} e^{\beta z_{2i}})}{(t_{1i} e^{\beta z_{1i}} + t_{2i} e^{\beta z_{2i}} + a)}$$

$$\frac{Nb}{a} = \sum_i \frac{(b+\delta_{1i}+\delta_{2i})}{(t_{1i} e^{\beta z_{1i}} + t_{2i} e^{\beta z_{2i}} + a)}$$

$$N \ln a + \sum_i \left[\frac{\delta_{1i}\delta_{2i}}{b+1} + \frac{(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})}{b} \right] = \sum_i \ln(t_{1i} e^{\beta z_{1i}} + t_{2i} e^{\beta z_{2i}} + a) \;.$$

As in the marginal likelihood case, no closed form of solution exists to the equations for the exponential or Weibull model and some iterative solution needs to be employed.

The asymptotic variances can be found as in the marginal method by inverting the matrix of second derivatives. On differentiating (2.18) to (2.21) again, the matrix of second derivatives is, with k,l = 1,...,4 , $(i_{k\ell}) =$

$$\begin{pmatrix} -\frac{\partial^2 \ln L_I}{\partial \beta^2} & -\frac{\partial^2 \ln L_I}{\partial \beta \partial \eta} & -\frac{\partial^2 \ln L_I}{\partial \beta \partial a} & -\frac{\partial^2 \ln L_I}{\partial \beta \partial b} \\ & -\frac{\partial^2 \ln L_I}{\partial \eta^2} & -\frac{\partial^2 \ln L_I}{\partial \eta \partial a} & -\frac{\partial^2 \ln L_I}{\partial \eta \partial b} \\ & & -\frac{\partial^2 \ln L_I}{\partial a^2} & -\frac{\partial^2 \ln L_I}{\partial a \partial b} \\ & & & -\frac{\partial^2 \ln L_I}{\partial b^2} \end{pmatrix}$$

and

$$i_{11} = \sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} \times$$
$$[(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)(\sum_j t_{ji}^{\eta} z_{ji} e^{\beta z_{ji}}) - (\sum_j t_{ji}^{\eta} z_{ji} e^{\beta z_{ji}})^2]$$

$$i_{12} = \sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} \times$$
$$[(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)(\sum_j t_{ji}^{\eta} z_{ji} e^{\beta z_{ji}} \ln t_{ji})$$
$$-(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} \ln t_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} z_{ji})]$$

$$i_{13} = -\sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} (\sum_j t_{ji}^{\eta} z_{ji} e^{\beta z_{ji}})$$

$$i_{14} = \sum_i (\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-1} (\sum_j t_{ji}^{\eta} z_{ji} e^{\beta z_{ji}})$$

$$i_{22} = \eta^{-2} \sum_i (b + \sum_j \delta_{ji}) + \sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} \times$$
$$\{[\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta][\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} (\ln t_{ji})^2]$$
$$- (\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} \ln t_{ji} + a)^2\}$$

$$i_{23} = \sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} \times$$
$$(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} - \eta \sum_j t_{ji}^{\eta} e^{\beta z_{ji}} \ln t_{ji})$$

$$i_{24} = -N\eta^{-1} + \sum_i (\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2} (\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} \ln t_{ji} + a)$$

$$i_{33} = Nba^{-2} - \eta^2 \sum_i (b + \sum_j \delta_{ji})(\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-2}$$

$$i_{34} = \eta \sum_i (\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta)^{-1} - Na^{-1}$$

$$i_{44} = \sum_i [\delta_{1i}\delta_{2i}(b+1)^{-2} + (\delta_{1i} + \delta_{2i} - \delta_{1i}\delta_{2i})b^{-2}]$$

for j=1,2 ; i=1,...,N.

It is interesting to note that i_{44} is independent of $\underline{t}_{ji}$ and z_{ji} .

For the exponential model and the case of no censoring the lower triangle of the Fisher information matrix is found to be

$$\begin{pmatrix} \dfrac{\sum_i (z_{1i} - z_{2i})^2 + b\sum_i (z_{1i}^2 + z_{2i}^2)}{(b+3)} & & \\ & & \\ \dfrac{-b\sum_i (z_{1i} + z_{2i})}{a(b+3)} & \dfrac{2Nb}{a^2(b+3)} & \\ & & \\ \dfrac{\sum_i (z_{1i} + z_{2i})}{(b+2)} & \dfrac{-2N}{a(b+2)} & \dfrac{N(2b\ +2b+1)}{b^2(b+1)} \end{pmatrix} .$$

The inverse of this matrix gives the asymptotic variance - covariance matrix of $(\hat{\beta}, a, b)$ and the asymptotic variance of $\hat{\beta}$ is

$$\frac{2N(b+3)}{2N\sum_i (z_{1i} - z_{2i})^2 + 2Nb\sum_i (z_{1i}^2 + z_{2i}^2) - b[\sum_i (z_{1i} + z_{2i})]^2}, \qquad (2.22)$$

which is independent of β . With a single covariate such that $z_{1i} = 1$, $z_{2i} = 0$ this reduces to

$$\frac{2}{N}\left[1 + \frac{1}{(b+2)}\right] . \qquad (2.23)$$

2.6 Relative efficiencies of the three methods for the exponential model.

Fisher (1958,p.147) suggests that the relative efficiency of two estimation procedures can be found from the ratio of the informations in the two statistics. If efficient methods of estimation are used this ratio is asymptotically equal to the ratio of the variances of the estimators.

For the exponential model the informations about β contained in the statistics derived from the non-parametric, marginal and integrated methods are respectively

$$I_N(\beta) = \sum_{i=1}^{N} \frac{(\epsilon_i d_i)^2 \exp(\epsilon_i \beta d_i)}{[1+\exp(\epsilon_i \beta d_i)]^2} \quad ; \quad I_M = \frac{\sum_{i=1}^{N} d_i^2}{3}$$

and

$$I_I = \frac{2N\sum_i d_i^2 + 2Nb\sum_i (z_{1i}^2 + z_{2i}^2) - b[\sum_i (z_{1i} + z_{2i})]^2}{2N(b+3)}$$

Hence the relative efficiency of the non-parametric method to the integrated method is given by

$$R_1(0) = \frac{I_N(0)}{I_I}$$

$$= \frac{N(b+3)\sum_i d_i^2}{2\left\{2N\sum_i d_i^2 + 2Nb\sum_i (z_{1i}^2 + z_{2i}^2) - b[\sum_i (z_{1i} + z_{2i})]^2\right\}}.$$

For a fixed scalar covariate with $z_{1i} = 1$, $z_{2i} = 0$

$$R_1(0) = \frac{b+3}{2(b+2)} = \frac{1}{2} + \frac{1}{2(b+2)}.$$

Thus, as b tends to infinity, the efficiency of the non-parametric method decreases to 0.5, from its maximum of 0.75 (approached as b tends to zero).

The relative efficiency of the marginal method to the integrated approach, R_2, does not depend on the value of β and is given by

$$R_2 = \frac{2N(b+3)\sum_i d_i^2}{3\left\{2N\sum_i d_i^2 + 2Nb\sum_i (z_{1i}^2 + z_{2i}^2) - b[\sum_i (z_{1i} + z_{2i})]^2\right\}}.$$

For a fixed scalar covariate with $z_{1i} = 1$, $z_{2i} = 0$ this becomes

$$R_2 = \frac{2}{3}\left[1 + \frac{1}{(b+2)}\right]$$

which approaches 2/3 as b approaches infinity. As b tends to zero, the relative efficiency of the marginal method to the integrated method approaches 1. When b = 0 this reduces to the case where the pairing variables, α_i, are assumed to be the same for all pairs.

It should be added that the integrated method can be expected to be more efficient than the other two methods when b > 0 as extra information about the α_i is being included. If the number of doubly censored pairs is large, the number of pairs analysed in the marginal and non-parametric methods is reduced and it might be expected that the relative efficiencies R_1 and R_2 would also be reduced. In choosing between the methods, this reduction has to be offset against making the extra assumptions necessary for the integrated likelihood. If the integrated method is fairly robust to the true distribution of the α_i , the extra assumptions should be worth including so that the full data set is analysed.

2.7 Extension to n-tuplets.

The density function for the j-th member, j=1,...,n, of the i-th n-tuplet, i=1,...,N is

$$f_{ji}(t) = \alpha_i \lambda_0(t) e^{\beta z_{ji}} \exp[-\alpha_i \Lambda_0(t) e^{\beta z_{ji}}] \qquad 0 \leq t < \infty$$

In the non-parametric model the marginal likelihood for ranks, if no ties within n-tuplets are present, can be shown to be, from Kalbfleisch and Prentice (1973),

$$L_N(\beta) \propto \prod_{i=1}^{N} \prod_{u} \frac{\alpha_i \exp(\beta z_{(j)i})}{\sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki})} = \prod_{i=1}^{N} \prod_{u} \frac{\exp(\beta z_{(j)i})}{\sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki})}$$

where $z_{(j)i}$ is the covariate for the j-th ordered time, $t_{(j)i}$, within the i-th n-tuplet and the product $\prod_{u}$ is taken over all the uncensored times in the i-th n-tuplet. This is, in fact, the partial likelihood of Cox (1972) which is obtained by arguing conditionally on the set of times, within n-tuplets, at which failures occur.

Hence the log likelihood is

$$\ln L_N(\beta) = \sum_{i=1}^{N} \sum_{j=1}^{n} \delta_{(j)i} [\beta z_{(j)i} - \ln\{ \sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki})]$$

$$+ \text{const} \qquad (2.24)$$

with $\delta_{(j)i}$ the usual censoring indicator for the survival time $t_{(j)i}$ of the j-th ordered member of the i-th n-tuplet.

Thus $\hat{\beta}$ is given by the solution of the equation

$$\frac{\partial \ln L_N(\beta)}{\partial \beta} = \sum_{i=1}^{N} \sum_{j=1}^{n} \delta_{(j)i} \left[z_{(j)i} - \frac{\sum_{t_{ki} \geq t_{(j)i}} z_{ki} \exp(\beta z_{ki})}{\sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki})} \right]$$

$$= 0 \qquad (2.25)$$

The asymptotic variance of $\hat{\beta}$ can be found from the second derivative of the log likelihood

$$\frac{\partial^2 \ln L_N(\beta)}{\partial \beta^2} = -\sum_{i=1}^{N} \sum_{j=1}^{n} [\sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki})]^{-2} [\sum_{t_{ki} \geq t_{(j)i}} \exp(\beta z_{ki}) \times$$

$$\sum_{t_{ki} \geq t_{(j)i}} z_{ki}^2 \exp(\beta z_{ki}) - \{ \sum_{t_{ki} \geq t_{(j)i}} z_{ki} \exp(\beta z_{ki}) \}^2] \qquad (2.26)$$

Obviously any of the n-tuplets which have all the times censored, do not contribute to $L_N(\beta)$.

The marginal likelihood is less easily extended to the case of n-tuplets. Let $w_{1i} = t_{1i}$, $w_{ji} = t_{ji}/t_{1i}$ for j=2,...,n. Then, it can be shown that the log likelihood ln $L_M(\beta, \eta)$ is given by

$$\ln L_M(\beta,\eta) = \sum_{i=1}^{N} [\sum_{\substack{\epsilon_i \geq 1 \\ \delta_{1i}=1}} \{\beta \sum_{j=2}^{n} \delta_{ji} d_{ji} + \ln\eta \sum_{j=2}^{n} \delta_{ji} + \eta \sum_{j=2}^{n} \delta_{ji} \ln w_{ji}\}$$

$$- \epsilon_i \ln\{1 + \sum_{j=2}^{n} \exp(\beta d_{ji}) w_{ji}^{\eta}\} \quad]$$

$$+ \sum_{\delta_{1i}=0} \ln\left\{ 1 + \sum_{k=1}^{n-1} (-1)^k \sum_{r=1}^{\binom{n-1}{k}} \left[\frac{\prod_{u \in B_{r,k}} \delta_{ui}}{1 + \sum_{u \in B_{r,k}} w_{ui}^{\eta} \exp(\beta d_{ui})} \right] \right. \qquad (2.27)$$

with $\epsilon_i = \sum_{j=1}^{n} \delta_{ji}$ and $B_{r,k}$ being the r-th subset of size k from the integers 2,...,n. The final term is only included for those n-tuplets in which t_{1i} is censored. Any n-tuplets which have all the w_{ji}, j=2,...,n, indeterminate i.e t_{ji}, j=1,...,n all censored, are omitted completely from the analysis .

The integrated likelihood method extends very easily to the problem of matched n-tuplets. The log likelihood given by (2.17) becomes

$$\ln L_I(\underline{t}_1, \underline{t}_2, \ldots, \underline{t}_n | \beta, \eta, a, b)$$

$$=\sum_i [\ln\Gamma(b+\sum_j \delta_{ji}) - \ln\Gamma(b)] + \beta\sum_i\sum_j \delta_{ji} z_{ji}$$

$$+ \ln\eta \sum_i (b+\sum_j \delta_{ji}) \quad + \eta\sum_i\sum_j \delta_{ji} \ln t_{ji} \quad + N b \ln a$$

$$-\sum_i (b+\sum_j \delta_{ji}) \ln[\sum_j t_{ji}^{\eta} e^{\beta z_{ji}} + a\eta] \quad + \text{const.} \qquad (2.28)$$

with $j=1,\ldots,n$; $i=1,\ldots,N$. The terms in (2.18) to (2.21), the first derivatives of the log likelihood, extend similarly.

The non-parametric method involves sorting the times within n-tuplets to obtain the risk sets. This should not pose too great a problem as the number within n-tuplets will usually be fairly small. The marginal method is difficult to program if the n-tuplets are of different sizes and especially if there are n-tuplets with t_{1i} censored, as the final term in (2.27) is then required. The integrated method, on the other hand, requires very little amendment to extend to either cases with n > 2 and/or n-tuplets of differing sizes.

2.8 Summary.

In this chapter the non-parametric and marginal methods of Holt and Prentice have been reviewed. Solutions were found for the non-parametric and marginal likelihoods using the generalised linear models program GLIM. The theory was developed for a proposed improved analysis of matched pair data subject to censoring, referred to as the integrated method. For this method, the equations giving the maximum likelihood estimates of the parameters and the observed information matrix, from which the asymptotic variances of the paramters can be obtained, were found. The efficiencies of the three methods in their estimation of β were compared in the case of no censoring. All three methods were extended to deal with matched n-tuplets.

CHAPTER 3

COMPARISONS OF THE THREE METHODS

It has already been stated that the integrated method can be expected to perform better than the other two methods since extra information is being included. To investigate the amount of improvement, various measures such as bias and mean-squared error can be calculated. Since in both the marginal and the integrated methods, closed forms for the estimates do not exist, the bias cannot be calculated algebraically. Thus various sets of simulations have been performed in an attempt to determine the relative merits of the three methods.

3.1 Method.

The data for the simulations was generated under two models for $\lambda_o(t)$ viz. the Weibull with $\eta = 1.5$ and the exponential (which is equivalent to the Weibull with $\eta = 1$). The marginal and the integrated methods treated the data generated under the exponential model both as exponential (estimating β in the marginal method and β, a, b in the integrated method) as well as Weibull (estimating β and η in the marginal method and β, η, a, b in the integrated method). As in the Holt and Prentice (1974) paper, a single covariate was used with $z_{1i} = 1$, $z_{2i} = 0$ for $i=1,\ldots,N$ and $\exp(\beta) = \theta = 0.5$.

The first set of simulations took a common censoring time, T , for all pairs with T taking initially a large value which resulted in no censoring, and then four successively smaller values which resulted in increasing numbers of censored observations. This method was used to simulate the situation of a study where all items enter at the start and censoring occurs only if the item has not failed by the end of the study period at time T. Pairs of observations of types (c), (d), (g) and (h) in Figure 2.1 will not be produced. For the exponential model three values for N, the number of pairs, were taken:- N = 30, 60 and 120, to provide a reasonable range of size of study. For most studies 30 - 60 pairs would be the likely size used (see data analysed in Chapter 6), whilst 120 is an attempt to see what might happen in the larger studies. For the Weibull model N was taken as 60.

3.2 Computation.

With the above constraints on $\underline{z}$ and θ , the failure times t_{1i} and t_{2i} for the i-th pair have distributions, for the Weibull model, of

$$f(t_{1i} \mid \alpha_i) = \frac{\alpha_i}{2} t_{1i}^{\eta-1} \exp\left\{- \frac{\alpha_i}{2\eta} t_{1i}^{\eta}\right\} \qquad (3.1)$$

and

$$f(t_{2i}|\alpha_i) = \alpha_i \, t_{2i}^{\eta-1} \exp\left\{-\frac{\alpha_i}{\eta} \, t_{2i}^{\eta}\right\} \qquad (3.2)$$

$$0 \leqslant t_{1i}, t_{2i} < \infty$$

Hence, if y_{1i} and y_{2i} are random variables from a uniform (0,1) distribution then

$$t_{1i} = \left[-\frac{2\eta}{\alpha_i} \ln y_{1i}\right]^{1/\eta} \quad \text{and} \quad t_{2i} = \left[-\frac{\eta}{\alpha_i} \ln y_{2i}\right]^{1/\eta}$$

have the distributions in (3.1) and (3.2). The failure times were thus produced by generating 2N random values from a uniform (0,1) distribution, using the NAG library routine G05CBF, and assigning the first N to $\underline{y}_1$ and the second N to $\underline{y}_2$. To generate the times an α_i for each pair is required. Various prior distributions for the α_i were tried. The integrated method makes the assumption that the α_i are a random sample from a gamma distribution. Three gamma distributions were used to generate sets of α_i for the simulations :- $\gamma(40,3)$, $\gamma(6,3)$ and $\gamma(20,5)$ where a $\gamma(a,b)$ distribution is defined in (2.16). The integrated method ought to behave best in these circumstances. Two other non-gamma prior distributions were used to determine the robustness of the integrated method. The first non-gamma prior to be used was that taken by Holt and Prentice with α_i^{-1}, i=1,...,N, as cycles of the sequence 5.5,6.0,...,35.0. A uniform (0,1) prior for α_i was also investigated. The differences between these five distributions for the α_i can be summarised in Table 3.1 below.

Distribution of α_i	Mean	Median	S.D	Range
$\alpha_i^{-1} = 5.5,\ldots,35.$	0.06	0.05	0.04	(.03 , .18)
$\alpha_i \sim$ uniform (0,1)	0.5	0.5	0.29	(0 , 1)
$\alpha_i \sim \gamma(40,3)$	0.08	0.07	0.04	(0 , ∞)
$\alpha_i \sim \gamma(20,5)$	0.25	0.23	0.11	(0 , ∞)
$\alpha_i \sim \gamma(6,3)$	0.5	0.45	0.29	(0 , ∞)

Table 3.1.

Summary measures of the prior distributions for α_i considered.

The maximisation routine used for the integrated and marginal methods was the NAG library routine E04KAF. For this reason the likelihood was maximised in terms of $\theta = \exp(\beta)$ rather than β, in order that the lower bounds on all the parameters should be zero. In each set of simulations the average percentage of singly censored, doubly censored and uncensored pairs was recorded.

For each set of simulated data, an estimate of θ was produced by each of the three methods and the mean and standard deviation of the estimates of θ over the 100 simulations were calculated, for each method. Since the same data were used for each method within the 100 simulations, direct comparisons could be made of the closeness of the estimates to the true value, produced by the integrated method with that of the marginal and non-parametric methods. The results are shown in Tables 3.2 to 3.14. The subscripts I,M and N on the estimates of θ and η denote the method producing the estimate viz integrated, marginal and non-parametric respectively.

Table 3.2 gives the means and standard deviations for $\hat{\theta}_N$, $\hat{\theta}_M$ and $\hat{\theta}_I$ when the α_i^{-1} were cycles of the sequence 5.5,...,35. Both the integrated and non-parametric methods give similar results. If the standard deviations multiplied by 0.1 are taken as estimates of the standard errors of the means then the means for $\hat{\theta}_I$ and $\hat{\theta}_N$ are within two standard errors of the true values of $\theta = 0.5$ for the vast majority of values of N, the number of pairs, and T the censoring time. The standard errors for the non-parametric method were found to be greater than those of the integrated method. The marginal method fared worst of the three overall, resulting in estimates below the true value of $\theta = 0.5$ and hence biased away from $\theta = 1$ (i.e. $\beta = 0$) with the heaviest censoring. It should be remembered that both the non-parametric and marginal methods discard those pairs which are doubly censored and in the heaviest censoring case (T = 5) this resulted in less than half the pairs being used in these analyses. Similar conclusions can be drawn from the other choices of α_i in Tables 3.3 to 3.6.

Table 3.7 shows the direct comparison within simulations for the three methods. For each distribution of α_i , the number of simulations in which $|\hat{\theta}_I - 0.5| \leqslant |\hat{\theta}_N - 0.5|$ and $|\hat{\theta}_I - 0.5| \leqslant |\hat{\theta}_M - 0.5|$ are shown .

Columns 3 and 4 of the table show the intra-simulation comparisons when α_i^{-1} were cycles of the sequence 5.5,...,35. For column 3 the comparison of the integrated and non-parametric methods, within censoring times shows that the number of simulations in which the integrated method is closer than the non-parametric method increases as N, the number of pairs, increases. This improvement is not as marked within each set of pairs, N, for decreasing censoring time (i.e. increasing number of doubly censored pairs). Column 4 compares the integrated and marginal methods and here the improvement achieved by the integrated method varies little with the number of pairs but quite dramatically as the number of doubly censored pairs increases, reflecting the heavily biased estimates that the marginal method produces with heavy censoring.

The same pattern was found when the α_i followed a uniform distribution the results being shown in Table 3.3 with the intra-simulation comparisons in columns 5 and 6 of Table 3.7. In this case the marginal likelihood method produced less biased estimates and the intra-simulation comparisons showed that the integrated method was closer than the other two methods on slightly fewer simulations than previously. Results for the gamma priors on α_i are given in Tables 3.4 to 3.6 and the final six columns of Table 3.7. As expected the integrated method performed better than previously when non-gamma priors were used for α_i . When the gamma distribution had parameters a = 40 and b = 3 the marginal method gave more biased estimates as N, the number of pairs, increased. This held for the other two gamma distributions although the bias was less marked. A similar pattern in the intra-simulation comparisons was found, to that when α_i^{-1} = 5.5,...,35 with the integrated method being closer on slightly more of the 100 simulations.

The results for the Weibull model are given in Tables 3.8 to 3.14 similarly. With the assumption of this model the means of $\hat{\theta}_N$, $\hat{\theta}_M$ and $\hat{\theta}_I$ from the three methods are very similar. The marginal method does not seem to give estimates heavily biased away from $\theta = 1$ as in the exponential case. However, heavy censoring does seem to produce an underestimate of η for the marginal method and since the estimated ratio of survival times for this method is exp($\hat{\beta}$ / $\hat{\eta}$) this will be overestimated. This gradient of bias with decreasing T was not as marked with the α_i distributed uniformly. The integrated method fared worst with estimation of η when the α_i were uniformly distributed , producing estimates with a bias directed away from η = 0. Comparison betweem the integrated and non-parametric and integrated and marginal methods, for each simulation, are shown in Tables 3.13 and 3.14 with an extra column for each set of α_i comparing the estimates of η achieved by the integrated and marginal methods to the true value of η . As regards the integrated and non-parametric methods, similar results were found to the exponential model previously discussed. For the integrated and marginal methods the number of simulations in which the integrated method produced estimates closer to the true value of θ was slightly less than in the earlier exponential model. The number of simulations in which the integrated method resulted in a closer estimate of η increased rapidly as the censoring time decreased. Over all the models for the α_i , the integrated method produced closer estimates of θ than the non-parametric in around 60% of simulations. This percentage was slightly less comparing the integrated and marginal methods on their estimation of θ , though much higher for η with the integrated method being closer in all the 100 simulations for $\alpha_i^{-1} = 5.5,\ldots,35$ with the heaviest censoring. When the α_i were gamma distributed the same pattern was found with respect to η as was found in respect of θ in the exponential model ,with $\alpha_i \sim \gamma(40,3)$ giving the most biased estimates for the marginal method.

3.3 Random censoring.

To examine the effect of pairs of type (c) and (d) with min (t_{1i}, t_{2i}) censored the above simulations were repeated with a random censoring mechanism, the t_{ji} , $j=1,2$ being generated as previously. Potential censoring times c_{1i} , c_{2i} were produced from an exponential distribution with parameters λ_1 and λ_2 . If $t_{ji} > c_{ji}$ ($j=1,2,\ i=1,\ldots,N$) then t_{ji} was said to be censored at c_{ji} . Five pairs of values of λ_1 and λ_2 were taken to produce various combinations of proportions of observations with min (t_{1i}, t_{2i}) censored, singly censored and doubly censored. These were (.075,.075), (.1,.1), (.075,.015), (.25,.25) and (.5,.01). The first set produced mostly uncensored pairs. The fourth and fifth set produced pairs with the majority having min (t_{1i}, t_{2i}) censored and the majority doubly censored respectively for $\alpha_i^{-1} = 5.5,\ldots,35$ and the two gamma priors $\gamma(40,3)$ and $\gamma(20,5)$ whilst these were reversed for the uniform prior on α_i and the $\gamma(6,3)$ prior.

The results for the exponential model are given in Tables 3.15 to 3.19. The non-parametric method omits pairs of type (c), (d), (g), (h) and (i) hence the analysis is based on the percentage of pairs obtained by summing the uncensored and singly censored columns. The marginal method omits only the doubly censored pairs of type (g), (h) and (i) and thus omits the percentage of pairs given in the doubly censored column. The same distribution for the pairing variables, as in the fixed censoring mechanism, were taken. Table 3.15 gives the results for $\alpha_i^{-1} = 5.5,\ldots,35$. The integrated method again appeared the more stable of the three methods although with 30 pairs and 63% of observations doubly censored , the mean of the estimates of θ produced by this method was more than two standard errors above the true value of $\theta = 0.5$ indicating that the estimates were somewhat biased away from $\theta = 0$. In Tables 3.15, 3.17 and 3.19, a number of the entries in the columns of mean and standard deviation of the estimates produced by the non-parametric method are asterisked. The mean and standard deviations in these entries are based on fewer than 100 simulations. In the simulations omitted the non-parametric method produced estimates of θ of 0 or ∞ (i.e. $\beta = -\infty$ or ∞) as mentioned in Chapter 2. The intra-simulation comparisons are, however, based on the whole set of 100 simulations as it is valid to say, in the simulations where the non-parametric method produced an estimate of θ of 0 or ∞, that the integrated method produced an estimate closer to the true value of θ. With random censoring the integrated method performed much better with a uniform prior for the α_i than for the fixed censoring (Table 3.16). The results for the gamma distribution with a = 40 and b = 3 followed closely those for α_i^{-1} = 5.5,...,35.

The last two simulations in each set, with values of (λ_1, λ_2) of (.25,.25) and (.5,.01), have a majority of observations doubly censored or with the minimum time in the pair censored. The non-parametric method seemed to give slightly better results when the majority of observations was doubly censored rather than when the majority had min (t_{1i}, t_{2i}) censored. The integrated method again was much more stable and this stability increased with increasing N as for the fixed censoring. The marginal method, as before, came out worst of the three overall with very biased estimates of θ even with few doubly censored pairs. As with the non-parametric method, the amount of bias was greatest when there was a high number of pairs with min (t_{1i}, t_{2i}) censored rather than a high number of doubly censored pairs.

Intra-simulation comparisons for the exponential model are shown in Tables 3.20 to 3.24 and similar results were yielded to those obtained for the fixed censoring mechanism although the integrated method gave improved estimates over the other two methods on slightly more of the simulations.

The results for the Weibull model and the random censoring mechanism begin in Table 3.25. As in the fixed censoring case, two values of η were used to generate the data, $\eta = 1.0$ and $\eta = 1.5$. In Tables 3.25 and 3.27 the asterisked entries again infer that the non-parametric method produced estimates of $\theta = 0$ or ∞ in some simulations and hence these are not included in the calculation of the means and standard deviations. The means and standard deviations of the estimates of θ for the non-parametric method and θ and η for the integrated method appear much the same as for the fixed censoring case, with the integrated method giving the closer estimates on average. This was not true of the marginal method. In the fixed censoring case it was found that the joint estimation of θ and η produced very biased estimates of η with heavy censoring but that the estimation of θ was much less biased than the exponential model. In the random censoring case, however, the estimates of θ were still heavily biased especially when a high percentage of pairs had the minimum time within a pair censored. Tables 3.26 to 3.29 give the mean and standard deviations of the estimates of θ and η for the different prior distributions of α_i. The five tables, Tables 3.30 to 3.34, show the intra-simulation comparisons. For the estimation of θ, the integrated method is closer on between 52 out of 100 and 69 out of 100 simulations as compared to the non-parametric method and between 44 out of 100 and 88 out of 100 simulations as compared to the marginal method. With regard to the estimation of η , the integrated method gave improved estimates over the marginal method in between 47 out of 100 and 90 out of 100 simulations. As with the fixed censoring, over the five prior distributions taken for α_i , the integrated method seemed to perform worst when the α_i had a uniform distribution.

3.4 Discussion.

The integrated method has both advantages and disadvantages over the other two methods considered. The main disadvantage is the increased computer time needed for estimation using the integrated likelihood. With fixed censoring and few doubly censored pairs there seems little to choose between the three methods. Although, on the whole, the non-parametric method performed well overall, with heavy censoring it may produce estimates of θ of zero or infinity. The integrated method appears to remain fairly stable with heavy censoring whilst the marginal method, which performed worst overall of the three methods, has a tendency to result in very biased estimates of θ when the exponential model for $\lambda_0(t)$ is assumed and for η when the Weibull model for $\lambda_0(t)$ is assumed. When a random censoring mechanism is used the non-parametric method ceases to be as reliable and with a large number of the pairs having the minimum time within pairs censored, the estimates of θ are biased towards $\theta = 1$ ($\beta = 0$). In this case the marginal method also performs worse and produces heavily biased estimates for both θ and η , assuming the Weibull model for

$\lambda_0(t)$. For both censoring mechanisms, the integrated method was fairly robust with regard to the prior distribution on the α_i. The uniform prior resulted in the worst estimates of θ over all the α_i distributions. This may be due to the fact that this prior is not unimodal and this is confirmed by the $\gamma(6,3)$ prior, which has the same mean and variance, producing much better results.

The main parameter of interest here has been β, the effect of the covariate. However the integrated likelihood method, with a little modification, could be used to look at the association within pairs, as defined by Clayton (1978). This involves putting a = b and reparametrising the likelihood in terms of β, η and γ, where γ = 1/a measures the association within pairs. It can be easily shown that

$$\frac{\lambda_{1i}(t|t_{2i} = T_2)}{\lambda_{1i}(t|t_{2i} > T_2)} = \frac{\lambda_{2i}(t|t_{1i} = T_1)}{\lambda_{2i}(t|t_{1i} > T_1)} = \gamma$$

or, if the observations are ages at which members succumb to a disease, then the ratio of age-specific incidence rates for one member of the pair with and without the disease, should be constant for all t.

Number of Pairs N	Censoring Time T	Average % of Pairs Un-censored	Average % of Pairs Singly Censored	Average % of Pairs Doubly Censored	Non-parametric Method $\hat{\theta}_N$(S.D)	Marginal Method $\hat{\theta}_M$(S.D)	Integrated Method $\hat{\theta}_I$(S.D)
30	Uncens.	100.0	0.0	0.0	.499(.22)	.498(.18)	.495(.14)
30	75	93.1	6.8	0.1	.498(.22)	.484(.18)	.491(.14)
30	50	83.5	15.9	0.6	.499(.22)	.469(.18)	.495(.14)
30	25	56.3	37.6	6.1	.499(.23)	.421(.20)	.489(.16)
30	5	7.9	39.4	52.7	.529(.34)	.409(.35)	.518(.30)
60	Uncens.	100.0	0.0	0.0	.510(.14)	.508(.11)	.517(.11)
60	75	81.3	17.6	1.1	.509(.14)	.474(.11)	.511(.11)
60	50	66.9	28.9	4.2	.504(.14)	.447(.12)	.508(.11)
60	25	39.4	44.2	16.4	.510(.16)	.412(.12)	.513(.14)
60	5	4.4	31.1	64.5	.506(.21)	.359(.23)	.510(.20)
120	Uncens.	100.0	0.0	0.0	.506(.11)	.502(.09)	.502(.07)
120	75	81.3	17.7	1.0	.507(.10)	.468(.09)	.499(.07)
120	50	66.7	29.6	3.7	.503(.11)	.439(.09)	.496(.08)
120	25	39.4	44.9	15.7	.509(.11)	.406(.12)	.507(.10)
120	5	4.6	31.3	64.1	.517(.18)	.358(.21)	.504(.17)

Table 3.2.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i^{-1} generated from cycles of the sequence 5.5,...,35.

Number of Pairs N	Censoring Time T	Average % of Pairs Un-censored	Singly Censored	Doubly Censored	Non-parametric Method $\hat{\theta}_N$(S.D)	Marginal Method $\hat{\theta}_M$(S.D)	Integrated Method $\hat{\theta}_I$(S.D)
30	Uncens.	100.0	0.0	0.0	.499(.22)	.498(.18)	.497(.16)
30	75	95.6	2.4	2.0	.499(.22)	.497(.18)	.501(.16)
30	50	94.0	3.6	2.4	.499(.22)	.496(.18)	.500(.16)
30	25	89.7	6.3	4.0	.498(.22)	.489(.18)	.501(.16)
30	5	55.7	29.7	14.6	.508(.24)	.437(.22)	.490(.18)
60	Uncens.	100.0	0.0	0.0	.510(.14)	.508(.11)	.515(.11)
60	75	95.3	3.5	1.2	.507(.13)	.504(.11)	.517(.12)
60	50	92.8	5.2	2.0	.507(.13)	.503(.11)	.514(.12)
60	25	86.7	8.7	4.6	.506(.13)	.497(.11)	.514(.11)
60	5	52.0	30.9	17.1	.510(.14)	.445(.12)	.509(.12)
120	Uncens.	100.0	0.0	0.0	.506(.11)	.501(.09)	.497(.08)
120	75	95.9	3.1	1.0	.508(.10)	.497(.09)	.498(.08)
120	50	94.0	4.4	1.6	.508(.11)	.496(.09)	.499(.08)
120	25	89.3	7.4	3.3	.508(.11)	.490(.09)	.502(.08)
120	5	55.3	30.7	14.0	.510(.11)	.441(.11)	.499(.08)

Table 3.3.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i generated from a uniform (0,1) distribution.

Number of Pairs N	Censoring Time T	Average % of Pairs: Un-censored	Singly Censored	Doubly Censored	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	Uncens.	100.0	0.0	0.0	.499(.22)	.498(.18)	.505(.14)
30	75	86.5	11.5	2.0	.497(.22)	.468(.18)	.500(.15)
30	50	75.7	20.9	3.4	.494(.22)	.446(.19)	.497(.15)
30	25	49.1	39.7	11.2	.494(.23)	.403(.21)	.488(.17)
30	5	7.0	35.4	57.6	.523(.41)	.430(.47)	.509(.35)
60	Uncens.	100.0	0.0	0.0	.510(.14)	.508(.11)	.511(.09)
60	75	83.7	14.5	1.8	.508(.14)	.480(.12)	.509(.10)
60	50	72.7	23.0	4.3	.507(.14)	.459(.12)	.507(.10)
60	25	45.8	39.9	14.3	.500(.14)	.411(.13)	.500(.11)
60	5	5.8	33.9	60.3	.524(.22)	.381(.25)	.530(.20)
120	Uncens.	100.0	0.0	0.0	.506(.11)	.502(.09)	.501(.07)
120	75	85.8	12.9	1.3	.505(.10)	.475(.09)	.499(.07)
120	50	75.0	21.8	3.2	.504(.11)	.454(.09)	.498(.07)
120	25	48.8	39.5	11.7	.508(.11)	.411(.10)	.494(.08)
120	5	6.7	35.4	57.9	.513(.17)	.358(.18)	.503(.15)

Table 3.4.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i generated from a gamma (40,3) distribution.

Number of Pairs N	Censoring Time T	Average % of Pairs Un-censored	Singly Censored	Doubly Censored	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	Uncens.	100.0	0.0	0.0	.499(.22)	.498(.18)	.505(.14)
30	75	100.0	0.0	0.0	.499(.22)	.498(.18)	.505(.14)
30	50	99.9	0.1	0.0	.499(.22)	.498(.18)	.505(.14)
30	25	97.5	2.5	0.0	.499(.22)	.493(.18)	.503(.15)
30	5	54.1	38.0	7.9	.492(.22)	.425(.21)	.498(.16)
60	Uncens.	100.0	0.0	0.0	.510(.14)	.508(.11)	.511(.09)
60	75	99.8	0.2	0.0	.510(.14)	.508(.11)	.511(.09)
60	50	99.4	0.6	0.0	.510(.14)	.507(.11)	.510(.09)
60	25	96.5	3.3	0.2	.509(.14)	.501(.11)	.509(.09)
60	5	57.5	33.8	8.7	.506(.14)	.432(.12)	.504(.10)
120	Uncens.	100.0	0.0	0.0	.506(.11)	.502(.09)	.501(.07)
120	75	99.8	0.2	0.0	.506(.11)	.501(.09)	.501(.07)
120	50	99.5	0.5	0.0	.506(.11)	.501(.09)	.500(.07)
120	25	96.8	3.1	0.1	.506(.11)	.495(.09)	.500(.07)
120	5	58.8	33.2	8.0	.508(.11)	.432(.10)	.498(.08)

Table 3.5.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i generated from a gamma (6,3) distribution.

Number of Pairs N	Censoring Time T	Average % of Pairs Un-censored	Singly Censored	Doubly Censored	Non-parametric Method $\hat{\vartheta}_N$ (S.D)	Marginal Method $\hat{\vartheta}_M$ (S.D)	Integrated Method $\hat{\vartheta}_I$ (S.D)
30	Uncens.	100.0	0.0	0.0	.499(.22)	.498(.18)	.488(.13)
30	75	99.5	0.5	0.0	.499(.22)	.497(.18)	.487(.13)
30	50	98.3	1.7	0.0	.499(.22)	.496(.18)	.487(.13)
30	25	88.5	11.1	0.4	.499(.22)	.475(.18)	.483(.13)
30	5	29.0	49.5	21.5	.517(.28)	.411(.28)	.507(.21)
60	Uncens.	100.0	0.0	0.0	.510(.14)	.508(.11)	.508(.10)
60	75	99.4	0.6	0.0	.510(.14)	.507(.11)	.507(.10)
60	50	98.3	1.7	0.0	.510(.14)	.504(.11)	.508(.10)
60	25	90.0	9.4	0.6	.511(.14)	.489(.11)	.507(.10)
60	5	32.6	47.5	19.9	.510(.15)	.398(.14)	.508(.13)
120	Uncens.	100.0	0.0	0.0	.506(.11)	.502(.09)	.503(.07)
120	75	99.5	0.5	0.0	.506(.11)	.500(.09)	.502(.07)
120	50	98.4	1.5	0.1	.506(.11)	.499(.09)	.503(.07)
120	25	90.7	8.8	0.5	.506(.11)	.484(.09)	.501(.07)
120	5	33.5	47.3	19.2	.512(.12)	.389(.11)	.499(.09)

Table 3.6.

Means and standard deviations over 100 simulations of estimates of ϑ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i generated from a gamma (20,5) distribution.

N	Cens Time T	$\alpha_i^{-1}=5.5,\ldots,35$ % with $\lvert\hat\theta_I-0.5\rvert \leqslant$ $\lvert\hat\theta_N-0.5\rvert$	$\lvert\hat\theta_M-0.5\rvert$	α_i uniform % with $\lvert\hat\theta_I-0.5\rvert \leqslant$ $\lvert\hat\theta_N-0.5\rvert$	$\lvert\hat\theta_M-0.5\rvert$	$\alpha_i \sim \gamma(40,3)$ % with $\lvert\hat\theta_I-0.5\rvert \leqslant$ $\lvert\hat\theta_N-0.5\rvert$	$\lvert\hat\theta_M-0.5\rvert$	$\alpha_i \sim \gamma(6,3)$ % with $\lvert\hat\theta_I-0.5\rvert \leqslant$ $\lvert\hat\theta_N-0.5\rvert$	$\lvert\hat\theta_M-0.5\rvert$	$\alpha_i \sim \gamma(20,5)$ % with $\lvert\hat\theta_I-0.5\rvert \leqslant$ $\lvert\hat\theta_N-0.5\rvert$	$\lvert\hat\theta_M-0.5\rvert$
30	Uncens.	55	56	56	57	63	61	63	61	59	63
30	75	54	59	58	60	58	67	63	61	59	63
30	50	54	65	60	59	54	74	63	61	58	62
30	25	56	66	60	63	59	76	60	60	58	65
30	5	51	77	61	71	53	74	61	73	59	71
60	Uncens.	61	56	67	54	68	60	68	60	67	56
60	75	58	55	63	56	72	67	68	60	67	57
60	50	61	65	64	53	66	70	68	61	66	59
60	25	58	68	64	55	66	73	67	58	67	64
60	5	53	77	59	64	58	76	69	71	62	72
120	Uncens.	65	56	56	49	63	56	63	56	66	57
120	75	64	64	58	59	67	68	63	56	66	56
120	50	66	76	61	57	64	78	63	58	65	53
120	25	60	78	65	59	63	80	61	62	67	59
120	5	61	77	57	77	50	74	60	79	64	80

Table 3.7.

Percentage of simulations with $\lvert\hat\theta_I - 0.5\rvert \leqslant \lvert\hat\theta_N - 0.5\rvert$ and $\lvert\hat\theta_I - 0.5\rvert \leqslant \lvert\hat\theta_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$.

η	Censoring Time T	Average % of Pairs: Un-cens.	Singly Cens.	Doubly Cens.	Non-Par. Method: $\hat{\theta}_N$ (S.D)	Marginal Method: $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method: $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.02(.11)	.519(.11)	0.99(.09)
1.0	75	81.7	17.2	1.1	.509(.14)	.503(.11)	0.91(.11)	.511(.12)	0.99(.09)
1.0	50	67.7	28.3	4.0	.505(.14)	.498(.11)	0.83(.11)	.506(.12)	1.00(.09)
1.0	25	40.6	43.6	15.8	.512(.16)	.504(.14)	0.66(.11)	.509(.14)	1.02(.11)
1.0	5	5.0	31.5	63.5	.509(.21)	.505(.19)	0.28(.16)	.507(.19)	1.01(.14)
1.5	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.519(.11)	1.48(.14)
1.5	75	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.519(.11)	1.48(.14)
1.5	50	99.4	0.6	0.0	.510(.14)	.501(.12)	1.53(.16)	.518(.11)	1.48(.14)
1.5	25	85.1	14.3	0.6	.511(.14)	.503(.11)	1.40(.16)	.513(.12)	1.49(.14)
1.5	5	9.6	39.0	51.4	.509(.19)	.501(.16)	0.55(.23)	.506(.16)	1.59(.30)

Table 3.8.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_o(t)$, with α_i^{-1} generated from cycles of the sequence 5.5,...,35 and with 60 pairs.

η	Censoring Time T	Average % of Pairs Un-cens.	Singly Cens.	Doubly Cens.	Non-Par. Method $\hat{\theta}_N$(S.D)	Marginal Method $\hat{\theta}_M$(S.D)	$\hat{\eta}_M$(S.D)	Integrated Method $\hat{\theta}_I$(S.D)	$\hat{\eta}_I$(S.D)
1.0	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.02(.11)	.499(.12)	1.05(.10)
1.0	75	95.3	3.5	1.2	.508(.13)	.503(.12)	1.00(.11)	.499(.12)	1.05(.10)
1.0	50	92.8	5.2	2.0	.507(.13)	.500(.12)	0.99(.11)	.497(.12)	1.05(.10)
1.0	25	86.7	8.7	4.6	.506(.13)	.500(.12)	0.97(.11)	.496(.12)	1.06(.10)
1.0	5	52.0	30.9	17.1	.502(.14)	.495(.12)	0.77(.11)	.499(.12)	1.04(.11)
1.5	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.499(.12)	1.58(.14)
1.5	75	99.4	0.6	0.0	.509(.14)	.501(.12)	1.53(.16)	.498(.12)	1.58(.14)
1.5	50	98.8	1.0	0.2	.509(.14)	.501(.12)	1.53(.16)	.499(.12)	1.58(.14)
1.5	25	95.8	3.2	1.0	.509(.13)	.503(.12)	1.51(.16)	.499(.12)	1.58(.14)
1.5	5	64.9	22.8	12.3	.504(.13)	.501(.12)	1.29(.17)	.504(.12)	1.58(.15)

Table 3.9.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a uniform (0,1) distribution and with 60 pairs.

η	Censoring Time T	Average % of Pairs: Un-cens.	Singly Cens.	Doubly Cens.	Non-Par. Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.02(.11)	.509(.10)	1.01(.10)
1.0	75	84.4	13.5	2.1	.508(.14)	.501(.12)	0.94(.11)	.505(.10)	1.02(.10)
1.0	50	72.3	23.4	4.3	.507(.14)	.500(.11)	0.86(.11)	.505(.10)	1.01(.10)
1.0	25	46.1	39.9	14.0	.500(.14)	.489(.12)	0.71(.12)	.497(.12)	1.02(.11)
1.0	5	6.4	34.0	59.6	.524(.22)	.523(.20)	0.30(.14)	.529(.20)	1.03(.20)
1.5	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.509(.10)	1.52(.15)
1.5	75	99.7	0.3	0.0	.510(.14)	.501(.12)	1.54(.16)	.508(.10)	1.52(.15)
1.5	50	98.5	1.4	0.1	.510(.14)	.502(.12)	1.52(.16)	.508(.10)	1.52(.15)
1.5	25	86.6	11.8	1.6	.509(.14)	.502(.12)	1.42(.16)	.506(.10)	1.52(.14)
1.5	5	11.8	40.9	47.4	.522(.19)	.518(.18)	0.58(.18)	.528(.18)	1.54(.22)

Table 3.10.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a gamma (40,3) distribution and with 60 pairs.

	Censoring	Average % of Pairs			Non-Par. Method	Marginal Method		Integrated Method	
η	Time T	Un-cens.	Singly Cens.	Doubly Cens.	$\hat{\theta}_N$ (S.D)	$\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	$\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.02(.11)	.509(.10)	1.01(.10)
1.0	75	99.8	0.2	0.0	.510(.14)	.501(.12)	1.02(.11)	.509(.10)	1.01(.10)
1.0	50	99.4	0.6	0.0	.510(.14)	.501(.12)	1.02(.11)	.508(.10)	1.01(.10)
1.0	25	96.5	3.3	0.2	.509(.14)	.501(.12)	1.00(.11)	.506(.10)	1.01(.10)
1.0	5	57.5	33.7	8.7	.506(.15)	.494(.12)	0.78(.11)	.501(.11)	1.01(.10)
1.5	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.509(.10)	1.52(.15)
1.5	75	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.509(.10)	1.52(.15)
1.5	50	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.509(.10)	1.52(.15)
1.5	25	99.8	0.2	0.0	.510(.14)	.501(.12)	1.54(.16)	.509(.10)	1.52(.15)
1.5	5	72.1	23.5	4.3	.507(.14)	.500(.11)	1.29(.17)	.504(.10)	1.52(.15)

Table 3.11.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a gamma (6,3) distribution and with 60 pairs.

η	Censoring Time T	Average % of Pairs Un-cens.	Singly Cens.	Doubly Cens.	Non-Par. Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.02(.11)	.504(.11)	1.02(.10)
1.0	75	99.4	0.6	0.0	.510(.14)	.501(.12)	1.02(.11)	.504(.11)	1.02(.10)
1.0	50	98.2	1.7	0.1	.510(.14)	.501(.12)	1.01(.11)	.504(.11)	1.02(.10)
1.0	25	90.0	9.4	0.6	.510(.14)	.501(.12)	0.96(.11)	.503(.11)	1.02(.10)
1.0	5	32.6	47.5	19.9	.510(.15)	.498(.12)	0.59(.11)	.504(.13)	1.02(.12)
1.5	Uncens.	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.504(.11)	1.53(.14)
1.5	75	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.504(.11)	1.53(.14)
1.5	50	100.0	0.0	0.0	.510(.14)	.501(.12)	1.54(.16)	.504(.11)	1.53(.14)
1.5	25	99.6	0.4	0.0	.510(.14)	.501(.12)	1.53(.16)	.504(.11)	1.53(.14)
1.5	5	48.2	40.8	11.0	.508(.14)	.499(.12)	1.06(.16)	.503(.12)	1.53(.16)

Table 3.12.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a gamma (20,5) distribution and with 60 pairs.

η	Censoring Time T	$\alpha_i^{-1}=5.5,\ldots,35$	% with		α_i Uniform	% with	
		$\lvert\hat{\theta}_I-0.5\rvert \leq \lvert\hat{\theta}_N-0.5\rvert$	$\lvert\hat{\theta}_I-0.5\rvert \leq \lvert\hat{\theta}_M-0.5\rvert$	$\lvert\hat{\eta}_I-\eta\rvert \leq \lvert\hat{\eta}_M-\eta\rvert$	$\lvert\hat{\theta}_I-0.5\rvert \leq \lvert\hat{\theta}_N-0.5\rvert$	$\lvert\hat{\theta}_I-0.5\rvert \leq \lvert\hat{\theta}_M-0.5\rvert$	$\lvert\hat{\eta}_I-\eta\rvert \leq \lvert\hat{\eta}_M-\eta\rvert$
1.0	Uncens.	59	51	29	62	54	59
1.0	75	56	44	93	65	53	52
1.0	50	58	45	100	60	51	50
1.0	25	56	47	100	59	52	51
1.0	5	51	54	100	55	45	85
1.5	Uncens.	59	51	66	62	54	58
1.5	75	59	51	66	62	55	58
1.5	50	59	53	64	63	53	52
1.5	25	55	48	68	61	53	51
1.5	5	55	46	96	53	45	66

Table 3.13.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, with 60 pairs and with α_i^{-1} generated from cycles of the sequence 5.5,...,35 and α_i having a uniform (0,1) distribution.

		$\alpha_i \sim \gamma(40,3)$			$\alpha_i \sim \gamma(6,3)$			$\alpha_i \sim \gamma(20,5)$		
		% with			% with			% with		
η	Cens. Time T	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$
1.0	Uncens.	66	61	58	66	61	58	60	62	63
1.0	75	69	59	64	66	61	58	60	61	67
1.0	50	69	63	75	66	60	59	60	61	62
1.0	25	66	57	89	65	64	67	57	59	66
1.0	5	58	47	96	72	64	81	59	54	94
1.5	Uncens.	66	61	60	66	61	60	60	62	61
1.5	75	66	61	59	66	61	60	60	62	61
1.5	50	67	63	66	66	61	60	60	62	61
1.5	25	71	57	70	66	61	60	59	61	63
1.5	5	54	48	98	69	63	75	60	54	89

Table 3.14.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, with 60 pairs and with α_i generated from the gamma distributions with parameters (40,3), (6,3) and (20,5).

Number of Pairs N	Average % of Pairs Un-cens	Singly cens	Min cens	Doubly cens	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	70.2	15.1	12.5	2.2	.536(.21)	.498(.17)	.525(.16)
30	34.5	28.4	23.1	14.0	.506(.24)	.399(.19)	.487(.15)
30	30.4	27.7	31.2	10.7	.542(.30)	.374(.23)	.517(.23)
30	7.0	13.0	69.9	10.1	.805(.71)*	.215(.19)	.503(.34)
30	4.2	16.4	15.6	63.0	.777(.67)*	.605(.52)	.595(.40)
60	61.8	18.3	15.1	4.8	.519(.16)	.481(.13)	.515(.11)
60	27.5	25.4	26.8	20.3	.556(.21)	.402(.14)	.531(.14)
60	23.2	25.9	34.0	16.9	.531(.23)	.352(.25)	.524(.17)
60	5.2	10.1	69.6	15.1	.586(.36)*	.187(.14)	.529(.26)
60	2.8	13.1	12.0	72.1	.573(.40)*	.453(.36)	.515(.26)
120	61.9	18.1	15.3	4.7	.531(.12)	.495(.10)	.518(.08)
120	26.6	26.6	25.8	21.0	.505(.13)	.379(.11)	.518(.10)
120	22.8	26.0	34.4	16.8	.511(.14)	.338(.22)	.521(.12)
120	5.3	10.5	69.4	14.8	.519(.31)	.162(.09)	.521(.18)
120	2.9	12.9	12.8	71.4	.591(.34)	.453(.29)	.522(.19)

Table 3.15.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$ and with α_i^{-1} generated from cycles of the sequence 5.5,...,35 with random censoring. Asterisked entries based on 89, 91, 99 and 98 simulations respectively (reading down the column).

Number of Pairs N	Un- cens	Average % of Pairs Singly cens	Min cens	Doubly cens	Non-parametric Method $\hat{\theta}_N$(S.D)	Marginal Method $\hat{\theta}_M$(S.D)	Integrated Method $\hat{\theta}_I$(S.D)
30	86.9	6.3	4.6	2.2	.528(.19)	.506(.16)	.517(.15)
30	70.6	12.6	11.7	5.1	.503(.21)	.453(.15)	.491(.13)
30	66.3	15.5	13.1	5.1	.539(.25)	.475(.19)	.532(.17)
30	29.8	23.3	20.6	26.3	.597(.28)	.517(.28)	.535(.21)
30	29.4	23.0	42.6	5.0	.565(.31)	.324(.15)	.499(.20)
60	86.3	6.5	5.4	1.9	.504(.14)	.497(.12)	.507(.10)
60	68.2	14.1	11.8	5.9	.535(.15)	.481(.12)	.524(.11)
60	63.2	16.8	15.3	4.7	.535(.16)	.457(.16)	.524(.14)
60	28.3	22.4	44.3	5.0	.519(.19)	.310(.12)	.510(.15)
60	27.3	22.4	20.1	30.2	.529(.19)	.460(.17)	.516(.15)
120	88.1	5.6	4.8	1.5	.530(.11)	.515(.09)	.519(.08)
120	70.4	13.7	10.9	5.0	.514(.11)	.471(.10)	.516(.09)
120	65.6	15.5	14.7	4.2	.521(.11)	.456(.15)	.520(.09)
120	38.0	23.6	19.8	26.6	.531(.16)	.455(.14)	.519(.12)
120	29.4	24.0	42.4	4.2	.489(.12)	.291(.08)	.475(.10)

Table 3.16.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$, with α_i generated from a uniform (0,1) distribution and with random censoring.

Number of Pairs N	Average % of Pairs Un-cens	Singly cens	Min cens	Doubly cens	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	62.1	18.4	15.0	4.5	.536(.22)	.492(.18)	.518(.16)
30	27.8	27.5	24.4	20.3	.514(.27)	.398(.21)	.501(.17)
30	23.8	26.2	34.1	15.9	.563(.38)	.349(.23)	.509(.22)
30	5.1	10.9	69.7	14.3	.819(.77)*	.196(.20)	.486(.37)
30	3.2	13.6	12.4	70.8	.892(.88)*	.621(.56)	.625(.47)
60	64.0	16.6	14.7	4.7	.514(.15)	.481(.13)	.520(.11)
60	30.3	25.7	24.8	19.2	.536(.18)	.402(.15)	.517(.14)
60	25.4	26.8	32.8	15.0	.536(.22)	.361(.25)	.510(.15)
60	6.6	11.2	67.8	14.4	.611(.41)*	.209(.14)	.558(.28)
60	3.2	14.4	13.5	68.9	.553(.43)*	.434(.36)	.504(.23)
120	65.4	16.4	13.7	4.5	.532(.12)	.492(.10)	.514(.08)
120	30.7	26.3	24.1	18.9	.503(.14)	.390(.11)	.508(.10)
120	26.6	25.0	32.9	15.5	.520(.13)	.349(.23)	.505(.10)
120	6.5	11.9	68.3	13.3	.541(.33)	.179(.09)	.516(.19)
120	3.7	14.9	13.9	67.5	.567(.27)	.456(.26)	.535(.18)

Table 3.17.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$, with α_i generated from a gamma (40,3) distribution and with random censoring. Asterisked entries are based on 82, 78, 99 and 97 simulations respectively (reading down the column).

Number of Pairs N	Average % of Pairs: Un-cens	Singly cens	Min cens	Doubly cens	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	91.5	5.0	3.2	0.3	.528(.19)	.506(.16)	.516(.15)
30	74.1	14.6	9.6	1.7	.496(.20)	.456(.16)	.488(.12)
30	69.7	16.2	12.6	1.5	.538(.25)	.473(.19)	.525(.17)
30	29.2	24.7	22.2	23.9	.622(.42)	.520(.29)	.547(.23)
30	28.4	25.8	43.9	1.9	.514(.29)	.305(.17)	.530(.22)
60	91.7	4.2	3.9	0.2	.509(.13)	.505(.11)	.519(.10)
60	75.0	13.1	10.0	1.9	.522(.14)	.477(.12)	.519(.11)
60	70.2	15.8	12.6	1.4	.525(.16)	.454(.13)	.510(.11)
60	30.8	24.2	22.7	22.3	.512(.18)	.452(.16)	.521(.13)
60	30.7	24.8	42.4	2.1	.526(.18)	.311(.11)	.520(.14)
120	92.7	3.8	3.3	0.2	.529(.11)	.515(.08)	.516(.07)
120	75.4	13.2	9.8	1.6	.503(.11)	.471(.09)	.510(.08)
120	70.6	15.7	12.3	1.4	.517(.11)	.452(.09)	.510(.08)
120	32.2	24.9	21.8	21.1	.543(.16)	.459(.13)	.514(.11)
120	31.4	25.0	41.5	2.1	.497(.12)	.292(.07)	.493(.09)

Table 3.18.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$, with α_i generated from a gamma (6,3) distribution and with random censoring.

Number of Pairs N	Average % of Pairs: Un-cens	Singly cens	Min cens	Doubly cens	Non-parametric Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)
30	86.1	8.1	5.3	0.5	.528(.20)	.505(.16)	.516(.15)
30	59.0	20.7	16.1	4.2	.489(.22)	.432(.17)	.495(.13)
30	54.2	22.6	19.6	3.6	.526(.25)	.436(.21)	.521(.19)
30	17.1	21.4	57.3	4.2	.578(.43)*	.261(.18)	.508(.25)
30	15.4	23.7	22.6	38.3	.723(.81)	.530(.35)	.544(.25)
60	87.1	6.7	5.7	0.5	.508(.13)	.499(.11)	.516(.09)
60	62.8	18.3	15.2	3.7	.525(.15)	.458(.12)	.516(.11)
60	56.5	21.7	18.7	3.1	.526(.17)	.429(.16)	.509(.12)
60	19.2	22.1	54.9	3.8	.546(.23)	.266(.12)	.528(.18)
60	16.5	24.5	22.4	36.6	.526(.23)	.443(.22)	.516(.12)
120	87.5	6.7	5.4	0.4	.529(.11)	.510(.09)	.516(.07)
120	63.5	18.5	14.2	3.8	.509(.11)	.454(.10)	.506(.08)
120	56.8	21.1	18.8	3.3	.510(.11)	.426(.14)	.512(.09)
120	19.6	22.8	53.9	3.7	.501(.16)	.245(.07)	.505(.10)
120	17.2	25.2	21.9	35.7	.535(.16)	.443(.14)	.522(.12)

Table 3.19.

Means and standard deviations over 100 simulations of estimates of θ by the non-parametric, marginal and integrated methods assuming the exponential model for $\lambda_0(t)$, with α_i generated from a gamma (20,5) distribution and with random censoring. The asterisked entry is based on 98 simulations.

Number of Pairs N	Average % of Pairs Un-cens	Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$
30	70.2	15.1	12.5	2.2	63	57
30	34.5	28.4	23.1	14.0	60	74
30	30.4	27.7	31.2	10.7	64	78
30	7.0	13.0	69.9	10.1	65	82
30	4.2	16.4	15.6	63.0	63	89
60	61.8	18.3	15.1	4.8	65	68
60	27.5	25.4	26.8	20.3	65	71
60	23.2	25.9	34.0	16.9	58	76
60	5.2	10.1	69.6	15.1	62	79
60	2.8	13.1	12.0	72.1	58	77
120	61.9	18.1	15.3	4.7	66	67
120	26.6	26.6	25.8	21.0	63	71
120	22.8	26.0	34.4	16.8	58	76
120	5.3	10.5	69.4	14.8	67	81
120	2.9	12.9	12.8	71.4	69	87

Table 3.20.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$ and $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$, α_i^{-1} generated from cycles of the sequence 5.5,...,35 and random censoring.

Number of Pairs N	Average % of Pairs: Un-cens	Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$
30	86.9	6.3	4.6	2.2	63	59
30	70.6	12.6	11.7	5.1	72	72
30	66.3	15.5	13.1	5.1	69	66
30	29.8	23.3	20.6	26.3	56	70
30	29.4	23.0	42.6	5.0	65	81
60	86.3	6.5	5.4	1.9	66	63
60	68.2	14.1	11.8	5.9	61	64
60	63.2	16.8	15.3	4.7	54	66
60	28.3	22.4	44.3	5.0	64	74
60	27.3	22.4	20.1	30.2	61	70
120	88.1	5.6	4.8	1.5	71	64
120	70.4	13.7	10.9	5.0	56	60
120	65.6	15.5	14.7	4.2	56	57
120	30.0	23.6	19.8	26.6	59	77
120	29.4	24.0	42.4	4.2	56	81

Table 3.21.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$ and $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$, α_i generated from a uniform (0,1) distribution and random censoring.

Number of Pairs N	Un-cens	Average % of Pairs Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$
30	62.1	18.4	15.0	4.5	65	65
30	27.8	27.5	24.4	20.3	67	80
30	23.8	26.2	34.1	15.9	64	82
30	5.1	10.9	69.7	14.3	49	86
30	3.2	13.6	12.4	70.8	41	76
60	64.0	16.6	14.7	4.7	66	66
60	30.3	25.7	24.8	19.2	61	76
60	25.4	26.8	32.8	15.0	59	75
60	6.6	11.2	67.8	14.4	49	70
60	3.2	14.4	13.5	68.9	57	78
120	65.4	16.4	13.7	4.5	72	64
120	30.7	26.3	24.1	18.9	65	83
120	26.6	25.0	32.9	15.5	57	75
120	6.5	11.9	68.3	13.3	74	82
120	3.7	14.9	13.9	67.5	63	78

Table 3.22.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$ and $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$, α_i generated from a gamma (40,3) distribution and random censoring.

Number of Pairs N	Average % of Pairs: Un-cens	Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$
30	91.5	5.0	3.2	0.3	61	58
30	74.1	14.6	9.6	1.7	67	66
30	69.7	16.2	12.6	1.5	60	69
30	29.2	24.7	22.2	23.9	63	73
30	28.4	25.8	43.9	1.9	61	73
60	91.7	4.2	3.9	0.2	66	67
60	75.0	13.1	10.0	1.9	56	57
60	70.2	15.8	12.6	1.4	72	68
60	30.8	24.2	22.7	22.3	63	70
60	30.7	24.8	42.4	2.1	63	77
120	92.7	3.8	3.3	0.2	70	59
120	75.4	13.2	9.8	1.6	61	70
120	70.6	15.7	12.3	1.4	64	62
120	32.2	24.9	21.8	21.1	77	77
120	31.4	25.0	41.5	2.1	62	79

Table 3.23.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$ and $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$, α_i generated from a gamma (6,3) distribution and random censoring.

Number of Pairs N	Average % of Pairs Un-cens	Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$
30	86.1	8.1	5.3	0.5	60	61
30	59.0	20.7	16.1	4.2	71	73
30	54.2	22.6	19.6	3.6	63	73
30	17.1	21.4	57.3	4.2	64	81
30	15.4	23.7	22.6	38.3	61	78
60	87.1	6.7	5.7	0.5	68	65
60	62.8	18.3	15.2	3.7	62	66
60	56.5	21.7	18.7	3.1	71	73
60	19.2	22.1	54.9	3.8	60	77
60	16.5	24.5	22.4	36.6	78	76
120	87.5	6.7	5.4	0.4	72	60
120	63.5	18.5	14.2	3.8	70	71
120	56.8	21.1	18.8	3.3	68	65
120	19.6	22.8	53.9	3.7	69	77
120	17.2	25.2	21.9	35.7	69	76

Table 3.24.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$ and $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$ assuming the exponential model for $\lambda_0(t)$, α_i generated from a gamma (20,5) distribution and random censoring.

η	Average % of pairs Un-cens	Singly cens	Min cens	Doubly cens	Non-par. Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	61.7	18.3	15.1	4.9	.519(.16)	.518(.13)	0.88(.13)	.514(.11)	1.00(.10)
	27.5	25.4	26.8	20.3	.556(.21)	.456(.12)	0.70(.15)	.530(.15)	1.03(.12)
	23.2	25.9	34.0	16.9	.531(.23)	.384(.24)	0.71(.15)	.521(.17)	1.03(.13)
	5.2	10.1	69.6	15.1	.586(.37)*	.155(.09)	0.61(.27)	.541(.31)	1.01(.15)
	2.8	13.1	12.0	72.1	.573(.40)*	.534(.31)	0.43(.32)	.506(.29)	1.08(.30)
1.5	80.8	7.4	10.8	1.0	.511(.14)	.511(.13)	1.46(.18)	.516(.10)	1.49(.14)
	48.4	17.7	26.4	7.5	.536(.18)	.486(.13)	1.29(.21)	.518(.13)	1.51(.16)
	41.2	20.3	31.9	6.6	.531(.21)	.431(.17)	1.27(.24)	.512(.14)	1.52(.18)
	6.4	9.7	76.9	7.0	.626(.47)	.178(.10)	1.07(.45)	.540(.28)	1.50(.21)
	4.6	12.4	19.8	63.2	.534(.31)*	.547(.29)	0.91(.40)	.508(.25)	1.60(.36)

Table 3.25.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i^{-1} generated from cycles of the sequence 5.5,...35, with random censoring and 60 pairs. Asterisked entries are based on 99, 98 and 97 simulations respectively (reading down the column).

	Average % of pairs				Non-par. Method	Marginal Method		Integrated Method	
η	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\theta}_N$ (S.D)	$\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	$\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	86.2	6.5	5.4	1.9	.504(.14)	.504(.12)	0.98(.12)	.490(.11)	1.05(.10)
	68.2	14.1	11.8	5.9	.535(.15)	.507(.12)	0.92(.13)	.511(.12)	1.05(.11)
	63.2	16.8	15.3	4.7	.545(.18)	.481(.16)	0.89(.14)	.510(.14)	1.06(.11)
	28.3	22.4	44.3	5.0	.519(.19)	.332(.10)	0.76(.14)	.507(.15)	1.05(.14)
	27.3	22.4	20.1	30.2	.529(.19)	.540(.16)	0.70(.12)	.510(.14)	1.03(.12)
1.5	91.5	3.1	4.8	0.6	.508(.13)	.506(.12)	1.51(.17)	.493(.11)	1.58(.15)
	75.3	9.4	12.6	2.7	.536(.16)	.511(.12)	1.44(.19)	.506(.11)	1.58(.16)
	70.2	11.5	16.1	2.2	.544(.18)	.481(.13)	1.40(.22)	.506(.14)	1.58(.17)
	28.1	16.9	27.4	27.5	.522(.19)	.533(.16)	1.16(.20)	.500(.15)	1.55(.19)
	27.2	18.8	51.2	2.8	.522(.19)	.336(.11)	1.19(.23)	.505(.14)	1.58(.21)

Table 3.26.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a uniform (0,1) distribution, with random censoring and 60 pairs.

η	Average % of pairs Un-cens	Singly cens	Min cens	Doubly cens	Non-par. Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	64.0	16.6	14.7	4.7	.514(.15)	.515(.13)	0.89(.12)	.518(.11)	1.01(.10)
	30.3	25.7	24.8	19.2	.536(.18)	.458(.13)	0.71(.14)	.516(.15)	1.02(.13)
	25.5	26.7	32.8	15.0	.536(.22)	.394(.24)	0.72(.15)	.506(.15)	1.03(.13)
	6.6	11.2	67.8	14.4	.611(.41)*	.182(.11)	0.61(.23)	.579(.33)	1.02(.17)
	3.2	14.4	13.5	68.9	.553(.43)*	.525(.27)	0.48(.69)	.497(.24)	1.09(.38)
1.5	81.8	6.8	10.2	1.2	.512(.15)	.510(.13)	1.47(.17)	.519(.11)	1.52(.14)
	50.4	17.4	24.9	7.3	.521(.17)	.483(.12)	1.30(.20)	.516(.13)	1.52(.17)
	43.5	20.3	29.9	6.3	.538(.20)	.442(.16)	1.28(.24)	.507(.13)	1.52(.16)
	7.7	10.6	74.5	7.2	.647(.55)*	.207(.11)	1.08(.46)	.569(.28)	1.52(.24)
	5.4	13.3	21.1	60.2	.533(.41)*	.533(.27)	0.93(.40)	.505(.23)	1.58(.33)

Table 3.27.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_o(t)$, with α_i generated from a gamma (40,3) distribution, with random censoring and 60 pairs. Asterisked entries are based on 99, 97, 99 and 96 simulations respectively (reading down the column).

η	Average % of pairs Un-cens	Singly cens	Min cens	Doubly cens	Non-par. Method $\hat{\theta}_N$ (S.D)	Marginal Method $\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	Integrated Method $\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	91.7	4.2	3.9	0.2	.509(.13)	.508(.12)	0.99(.11)	.517(.10)	1.01(.09)
	75.0	13.1	10.0	1.9	.522(.14)	.500(.11)	0.93(.12)	.516(.12)	1.01(.10)
	70.2	15.8	12.6	1.4	.535(.18)	.475(.13)	0.91(.14)	.506(.12)	1.02(.10)
	30.7	24.2	22.7	22.4	.512(.18)	.531(.15)	0.71(.14)	.515(.14)	1.03(.12)
	30.7	24.8	42.4	2.1	.526(.18)	.335(.10)	0.75(.15)	.523(.15)	1.00(.13)
1.5	94.0	1.9	3.9	0.2	.511(.13)	.509(.12)	1.52(.18)	.518(.10)	1.52(.14)
	80.8	7.5	10.5	1.2	.523(.14)	.505(.11)	1.45(.18)	.517(.11)	1.52(.15)
	75.7	10.3	13.1	0.9	.533(.18)	.481(.13)	1.43(.22)	.502(.11)	1.52(.14)
	31.8	18.1	29.3	20.8	.514(.19)	.535(.16)	1.18(.21)	.515(.14)	1.53(.19)
	29.9	20.5	48.0	1.6	.539(.23)	.346(.12)	1.18(.24)	.531(.18)	1.49(.19)

Table 3.28.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a gamma (6,3) distribution, with random censoring and 60 pairs.

	Average % of pairs				Non-par. Method	Marginal Method		Integrated Method	
η	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\theta}_N$ (S.D)	$\hat{\theta}_M$ (S.D)	$\hat{\eta}_M$ (S.D)	$\hat{\theta}_I$ (S.D)	$\hat{\eta}_I$ (S.D)
1.0	76.9	11.7	9.9	1.5	.518(.15)	.512(.13)	.938(.12)	.513(.11)	1.02(.10)
	45.4	23.5	21.6	9.5	.527(.16)	.473(.12)	.799(.13)	.511(.12)	1.03(.11)
	38.7	26.5	26.9	7.9	.522(.18)	.419(.16)	.783(.13)	.498(.14)	1.03(.11)
	10.0	15.9	66.4	7.7	.583(.42)	.204(.10)	.628(.20)	.521(.25)	1.05(.19)
	6.7	19.3	18.3	55.7	.510(.30)	.519(.21)	.485(.20)	.490(.20)	1.03(.16)
1.5	87.8	4.7	7.0	0.5	.509(.14)	.508(.12)	1.49(.17)	.511(.10)	1.52(.14)
	61.8	14.0	20.5	3.7	.518(.15)	.489(.11)	1.37(.20)	.510(.11)	1.54(.17)
	54.7	17.4	24.7	3.2	.533(.20)	.458(.16)	1.33(.23)	.501(.13)	1.54(.16)
	11.6	14.2	69.7	4.5	.567(.41)	.233(.11)	1.09(.35)	.530(.23)	1.55(.24)
	10.3	16.3	26.0	47.4	.510(.28)	.538(.24)	0.98(.31)	.499(.19)	1.57(.25)

Table 3.29.

Means and standard deviations over 100 simulations of estimates of θ and η by the non-parametric, marginal and integrated methods assuming the Weibull model for $\lambda_0(t)$, with α_i generated from a gamma (20,5) distribution, with random censoring and 60 pairs.

η	Un-cens	Average % of pairs Singly cens	Min cens	Doubly cens	$\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leqslant \lvert\hat{\eta}_M - \eta\rvert$
1.0	61.7	18.3	15.1	4.9	64	61	73
	27.5	25.4	26.8	20.3	63	50	85
	23.2	25.9	34.0	16.9	58	72	81
	5.2	10.1	69.6	15.1	62	86	89
	2.8	13.1	12.0	72.1	55	57	85
1.5	80.8	7.4	10.8	1.0	67	68	62
	48.4	17.7	26.4	7.5	68	58	75
	41.2	20.3	31.9	6.6	63	63	77
	6.4	9.7	76.9	7.0	63	88	87
	4.6	12.4	19.8	63.2	63	58	78

Table 3.30.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leqslant \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, α_i^{-1} generated from cycles of the sequence 5.5,...35 , with random censoring and 60 pairs.

	Average % of pairs				% with		
η	Un-cens	Singly cens	Min cens	Doubly cens	$\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leqslant \lvert\hat{\eta}_M - \eta\rvert$
1.0	86.2	6.5	5.4	1.9	56	55	47
	68.2	14.1	11.8	5.9	66	54	61
	63.2	16.8	15.3	4.7	57	55	70
	28.3	22.4	44.3	5.0	62	77	81
	27.3	22.4	20.1	30.2	67	55	84
1.5	91.5	3.1	4.8	0.6	57	60	52
	75.3	9.4	12.6	2.7	64	55	62
	70.2	11.5	16.1	2.2	61	49	63
	28.1	16.9	27.4	27.5	58	47	78
	27.2	18.8	51.2	2.8	67	77	52

Table 3.31.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leqslant \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leqslant \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, α_i generated from a uniform (0,1) distribution, with random censoring and 60 pairs.

η	Un-cens	Average % of pairs Singly cens	Min cens	Doubly cens	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$
1.0	64.1	16.6	14.7	4.7	61	62	70
	30.3	25.7	24.8	19.2	60	59	83
	25.5	26.7	32.8	15.0	58	70	83
	6.6	11.2	67.8	14.4	52	75	87
	3.2	14.4	13.5	68.9	58	53	87
1.5	81.8	6.8	10.2	1.2	67	62	59
	50.4	17.4	24.9	7.3	59	58	73
	43.5	20.3	29.9	6.3	57	66	79
	7.7	10.6	74.5	7.2	59	77	87
	5.4	13.3	21.1	60.2	60	50	81

Table 3.32.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, α_i generated from a gamma (40,3) distribution, with random censoring and 60 pairs.

η	Un-cens	Average % of pairs Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$
1.0	91.7	4.2	3.9	0.2	61	69	61
	75.0	13.1	10.0	1.9	55	44	64
	70.2	15.8	12.6	1.4	67	57	75
	30.7	24.2	22.7	22.4	59	53	81
	30.7	24.8	42.4	2.1	61	76	84
1.5	94.0	1.9	3.9	0.2	62	67	62
	80.8	7.5	10.5	1.2	60	49	66
	75.7	10.3	13.1	0.9	64	52	71
	31.8	18.1	29.3	20.8	62	58	80
	29.9	20.5	48.0	1.6	62	73	84

Table 3.33.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, α_i generated from a gamma (6,3) distribution, with random censoring and 60 pairs.

	Average % of pairs				% with		
η	Un-cens	Singly cens	Min cens	Doubly cens	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$	$\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$	$\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$
1.0	76.9	11.7	9.9	1.5	68	69	65
	45.4	23.5	21.6	9.5	64	58	80
	38.7	26.5	26.9	7.9	66	71	84
	10.0	15.9	66.4	7.7	61	80	86
	6.7	19.3	18.3	55.7	62	57	90
1.5	87.8	4.7	7.0	0.5	69	67	60
	61.8	14.0	20.5	3.7	56	60	73
	57.4	17.4	24.7	3.2	67	60	72
	11.6	14.2	69.7	4.5	58	82	84
	10.3	16.3	26.0	47.4	64	53	85

Table 3.34.

Percentage of simulations with $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_N - 0.5\rvert$, $\lvert\hat{\theta}_I - 0.5\rvert \leq \lvert\hat{\theta}_M - 0.5\rvert$ and $\lvert\hat{\eta}_I - \eta\rvert \leq \lvert\hat{\eta}_M - \eta\rvert$ assuming the Weibull model for $\lambda_0(t)$, α_i generated from a gamma (20,5) distribution, with random censoring and 60 pairs.

CHAPTER 4

NORMAL THEORY ACCELERATED FAILURE MODEL

4.1 The Model.

The accelerated failure class of models differs from the proportional hazards class in that the covariates act directly on the survival time rather than the hazard rate. This class of models has the advantage of the results being much more easily conveyed to clinicians as the effect of a covariate may be described as, say, 'doubling the survival time'. The Weibull and exponential models described so far belong to both the accelerated failure class and the proportional hazards class. An obvious model to consider, which belongs only to the accelerated failure class, is the normal model. This is in effect, an extension of the paired t-test to allow for censored data.

Let the survival times and covariate vector be T_{ji}, $\underline{z}_{ji}$, i=1,...,N, j=1,2. Then the normal theory accelerated failure model for the expected uncensored survival times is

$$T_{ji} = \alpha_i \exp(\underset{\sim}{\beta}^T \underline{z}_{ji}) \xi_{ji}$$

with β a vector of parameters as previously mentioned, α_i the 'matching' variable, $\xi_{ji} \sim N(0,\tau^2)$ independently j=1,2 and τ a scale factor. For simplification consider a single covariate, the usual 'treatment' effect, with $z_{1i} = 1$, $z_{2i} = 0$, i=1,...,N. Then the linear model formulation for the observed uncensored times becomes

$$t_{ji} = \ln T_{ji} = a_i + \beta^T z_{ji} + \tau \epsilon_{ji} \tag{4.1}$$

with $a_i = \ln \alpha_i$, $\epsilon_{ji} \sim N(0,1)$ independently j=1,2.

The data observed is

$$Y = \begin{pmatrix} Y_{11} & \dots & Y_{1N} \\ \delta_{11} & \dots & \delta_{1N} \\ Y_{21} & \dots & Y_{2N} \\ \delta_{21} & \dots & \delta_{2N} \end{pmatrix} \tag{4.2}$$

with $Y_{ji} = \min(t_{ji}, c_{ji})$,

$$\delta_{ji} = \begin{cases} 1 & \text{if } Y_{ji} = t_{ji} \\ 0 & \text{if } Y_{ji} = c_{ji} \end{cases}$$

and c_{ji} is the logarithm of the potential censoring time (independent of β).

4.2 Previous Solutions.

To eliminate the a_i , Sampford and Taylor (1959) considered the pairwise differences $w_i = t_{1i} - t_{2i}$, i=1,...,N.

From (4.1)

$$w_i = \beta d_i + \tau \epsilon_i \tag{4.3}$$

with $d_i = z_{1i} - z_{2i}$ and $\epsilon_i \sim N(0,1)$. As in the proportional hazards marginal likelihood method, no information on the distribution of w_i is provided, in this way, for those pairs doubly censored or with min (t_{1i}, t_{2i}) censored. These pairs were omitted from the analysis. Sampford and Taylor used a different parametrization to (4.3) and used an iterative technique to obtain maximum likelihood estimates of a and b where in (4.1) $a = -\beta/\tau$ and $b = 1/\tau$.

Wolynetz (1979b) extended Sampford and Taylor's methods to allow for observations which were left censored (i.e only the upper bound known) or confined between finite limits, using the parametrization in (4.3). He also used the EM algorithm (Dempster et al,1977) to find the maximum likelihood estimates and this procedure will be explained in the section following.

For simplification consider the `treatment' covariate again with $z_{1i} = 1$, $z_{2i} = 0$ and $d_i = 1$ for all i=1,...,N. Let Z(x), Q(x) and M(x) be defined as in section (1.2.1) and also

$$T(x) = \partial M(x)/\partial x = M(x)[\, M(x) - x \,].$$

Then the log likelihood for β and τ is, from Wolynetz

(1979a)

$$\ln L_w(\beta,\tau) = -N_1 \ln\tau - \frac{1}{2}\sum_{A_1}\frac{(w_i-\beta)^2}{\tau^2} + \sum_{A_2}\ln Q\left(\frac{w_i^*-\beta}{\tau}\right) + \sum_{A_3}\ln Q\left(\frac{-w_i^*+\beta}{\tau}\right), \quad (4.4)$$

where w_i , A_1 , A_2 , A_3 are as defined in Chapter 2, section 2.3, and $N_1 = \| A_1 \|$.

The maximum likelihood estimates of β and τ are the solutions of

$$\frac{\partial \ln L_w}{\partial\beta} = \frac{1}{\tau^2}\sum_{A_1}(w_i-\beta) + \frac{1}{\tau}\sum_{A_2} M\left(\frac{w_i^*-\beta}{\tau}\right) - \frac{1}{\tau}\sum_{A_3} M\left(\frac{-w_i^*+\beta}{\tau}\right) = 0 \quad (4.5)$$

and

$$\frac{\partial \ln L_w}{\partial\tau} = \frac{-r}{\tau} + \frac{1}{\tau^3}\sum_{A_1}(w_i-\beta)^2 + \frac{1}{\tau^2}\sum_{A_2}(w_i^*-\beta)M\left(\frac{w_i^*-\beta}{\tau}\right) - \frac{1}{\tau^2}\sum_{A_3}(w_i^*-\beta)M\left(\frac{-w_i^*+\beta}{\tau}\right) = 0 , \quad (4.6)$$

since $M(x) = \partial Q(x)/\partial x$.

Let u_i be defined as follows

$$\begin{array}{ll} \text{if } t_{1i}, t_{2i} \in A_1 & u_i = w_i \\ \text{if } t_{1i}, t_{2i} \in A_2 & u_i = E(w_i \mid w_i > w_i^*) = \beta + \tau M(w_i^*) \\ \text{if } t_{1i}, t_{2i} \in A_3 & u_i = E(w_i \mid w_i < w_i^*) = \beta - \tau M(-w_i^*). \quad (4.7) \end{array}$$

Substituting (4.7) into (4.5) and (4.6) gives

$$\hat{\beta} = \frac{\sum_{i=1}^{M} \hat{u}_i}{M} \qquad (4.8)$$

and

$$\hat{\tau}^2 = \frac{\sum_{i=1}^{M} (\hat{u}_i - \hat{\beta})}{\left[r + \sum_{A_2} T\left(\frac{w_i^* - \beta}{\hat{\tau}}\right) + \sum_{A_3} T\left(\frac{-w_i^* + \beta}{\hat{\tau}}\right)\right]}, \qquad (4.9)$$

with $M = \| A_1 \cup A_2 \cup A_3 \|$ i.e the number of pairs neither doubly censored nor with min (t_{1i}, t_{2i}) censored.

The EM procedure consists of alternately estimating $\hat{\beta}$ and $\hat{\tau}$ from (4.8) and (4.9) and $\{u_i\}$ from (4.7) until some convergence criteria are satisfied.

Wolynetz (1979a) gives an expression for an estimate of the variance - covariance matrix, $V(\hat{\beta}, \hat{\tau})$ based on the observed information matrix where $V(\hat{\beta}, \hat{\tau}) = I^{-1}(\hat{\beta}, \hat{\tau})$ and $I(\hat{\beta}, \hat{\tau})$ is the matrix of the negatives of the second derivatives of the log likelihood evaluated at $(\hat{\beta}, \hat{\tau})$.

The elements of $I(\hat{\beta}, \hat{\tau})$ are

$$-\left.\frac{\partial^2 \ln L_W}{\partial \beta^2}\right|_{\substack{\beta=\hat{\beta} \\ \tau=\hat{\tau}}} = \frac{1}{\hat{\tau}^2}\left[r + \sum_{A_2} T\left(\frac{w_i^* - \hat{\beta}}{\hat{\tau}}\right) + \sum_{A_3} T\left(\frac{-w_i^* + \hat{\beta}}{\hat{\tau}}\right)\right]$$

$$-\left.\frac{\partial^2 \ln L_W}{\partial \beta \partial \tau}\right|_{\substack{\beta=\hat{\beta} \\ \tau=\hat{\tau}}} = \frac{1}{\hat{\tau}^2}\left[\sum_{A_1}\left(\frac{w_i - \hat{\beta}}{\hat{\tau}}\right) + \sum_{A_2}\left(\frac{w_i^* - \hat{\beta}}{\hat{\tau}}\right) T\left(\frac{w_i^* - \hat{\beta}}{\hat{\tau}}\right) + \sum_{A_3}\left(\frac{w_i^* - \hat{\beta}}{\hat{\tau}}\right) T\left(\frac{-w_i^* + \beta}{\hat{\tau}}\right)\right]$$

$$-\frac{\partial^2 \ln L_W}{\partial \tau^2}\bigg|_{\substack{\beta=\hat{\beta}\\ \tau=\hat{\tau}}} = \frac{1}{\hat{\tau}^2}\left[r + \sum_{A_1}\left(\frac{w_i-\hat{\beta}}{\hat{\tau}}\right)^2 + \sum_{A_2}\left(\frac{w_i^*-\hat{\beta}}{\hat{\tau}}\right)^2 T\left(\frac{w_i^*-\hat{\beta}}{\hat{\tau}}\right) + \sum_{A_3}\left(\frac{w_i^*-\hat{\beta}}{\hat{\tau}}\right)^2 T\left(\frac{-w_i^*+\hat{\beta}}{\hat{\tau}}\right)\right]. \tag{4.10}$$

Wolynetz (1979b) extends the single covariate case to multiple covariates.

4.3 The EM Algorithm.

The EM method is an iterative procedure for computing maximum likelihood estimates in incomplete data problems. Let $\underline{X}$ be the complete data, $\underline{Y}$ the incomplete (observed) data, $f(\underline{x} \mid \phi)$ a family of sampling densities dependent upon parameters ϕ and a corresponding $g(\underline{y} \mid \phi)$ for the incomplete observations. These latter two are related by

$$g(\underline{y} \mid \phi) = \int_{\mathcal{R}} f(\underline{x} \mid \phi)\,d\underline{x}$$

where $\mathcal{R}$ is the region in which $\underline{x}$ is known to lie i.e satisfied by $\underline{y} = \underline{y}(\underline{x})$. As an example, in the case of a right censored observation $t > t^*$ say $\mathcal{R}$ is the region (t^*, ∞).

Basically, the EM algorithm consists of two steps :- the E-step and the M-step. If $\phi^{(p)}$ is the current estimate of ϕ, these are defined as follows

E-step Compute $\lambda(\phi \mid \phi^{(p)}) = E[\ln f(\underline{x} \mid \phi^{(p)}) \mid \underline{Y}, \phi]$.

M-step Choose $\underset{\sim}{\phi}^{(p+1)}$ to maximise $\lambda(\underset{\sim}{\phi} \mid \underset{\sim}{\phi}^{(p)})$. (4.11)

These two steps are repeated until convergence is reached. Theorem 3 of Dempster et al (1977) guarantees that this iterative procedure converges to the maximum likelihood estimate of $\underset{\sim}{\phi}$.

Louis (1982) derives a procedure to find the observed information matrix when using the EM algorithm and develops, as a consequence of this, a method of speeding up convergence of the EM procedure. The calculation of the observed information matrix requires the calculation of first and second derivatives of the log likelihood of the complete data only not the incomplete data. If $I_Y(\hat{\underset{\sim}{\phi}})$ is the observed information matrix evaluated at the maximum likelihood estimate of $\underset{\sim}{\phi} = \hat{\underset{\sim}{\phi}}$ then

$$\begin{aligned} I_Y(\hat{\underset{\sim}{\phi}}) = \; & E[B(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}] \\ & - E[G(\underline{X},\underset{\sim}{\phi})G^T(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}] , \end{aligned} \qquad (4.12)$$

where $\underline{X}$ and $\mathcal{R}$ are as previously defined and $G(\underline{X},\hat{\underset{\sim}{\phi}})$ and $B(\underline{X},\hat{\underset{\sim}{\phi}})$ are the gradient vector and the matrix of the negatives of the second derivatives of the log likelihood, respectively.

In general

$$\begin{aligned} I_Y(\underset{\sim}{\phi}) = \; & E[B(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}] - E[G(\underline{X},\underset{\sim}{\phi})G^T(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}] \\ & + E[G(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}]E[G^T(\underline{X},\underset{\sim}{\phi}) \mid \underline{X} \in \mathcal{R}], \end{aligned} \qquad (4.13)$$

with the final term, $E[G(\underline{X}, \underline{\Phi}) | \underline{X} \in \mathcal{R}]$ equating to zero when $\underline{\Phi} = \hat{\underline{\Phi}}$, the M.L.E. The first term is the observed information based on the complete data. The last two comprise the expected information conditional on $\underline{X} \in \mathcal{R}$.

Hence (4.13) can be written

$$I_Y = I_X - I_{X|Y} .$$

Louis (1982) suggests that the convergence of the EM procedure can be speeded up by the use of a further step after the E-and M-steps. This step can be formulated as

Step 3 Compute

$$\underline{\Phi}_1^{(p+1)} = \underline{\Phi}^{(p)} + (1-\hat{J})^{-1} (\underline{\Phi}^{(p+1)} - \underline{\Phi}^{(p)}) \qquad (4.14)$$

and use $\underline{\Phi}_1^{(p+1)}$ in place of $\underline{\Phi}^{(p+1)}$ in the next E-step, where if $\underline{\Phi}$ has dimension n then 1 is the nxn identity matrix and $(1-\hat{J})^{-1} = I_X I_Y^{-1}$.

If many parameters are to be estimated the cost of inverting I_Y must be compared to the cost of running through another iteration of the EM algorithm. In any case the approximations which produce step 3 are valid only around the maximum likelihood estimate and therefore step 3 should not be used until some iterations have been made.

4.4 Proposed EM solution.

4.4.1 Theory

Returning to the ideas of complete and incomplete data, if, in (4.1) the a_i and t_{ji} had all been observed the solution would be to simply fit the regression model of t_{ji} on z_{ji} and a_i to estimate β and τ. Hence this problem is incomplete in two aspects :- the a_i are unobserved and some of the t_{ji} are unobserved through censoring. The complete data $\underline{X}$ could be thought of as

$$\underline{X} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ c_{11} & c_{12} & \cdots & c_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ c_{21} & c_{22} & \cdots & c_{2N} \\ a_1 & a_2 & \cdots & a_N \end{pmatrix} \tag{4.15}$$

and the observed data are the incomplete data $\underline{Y}$ given in (4.2).

To use the EM algorithm some assumptions must be made about the distribution of the a_i in a similar way to the integrated method in the proportional hazards model. It will be assumed that the $a_i \sim N(\mu, \sigma^2)$ independently of ϵ_{ji} and let $\underline{\phi} = (\beta, \mu, \sigma, \tau)$.

Now

$$\ln\,[f(\underline{x} \mid \underline{\phi})] = -2N\ln\tau - N\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(a_i-\mu)^2$$
$$- \frac{1}{2\tau^2}\sum_{i=1}^{N}[(t_{1i}-\beta-a_i)^2 + (t_{2i}-a_i)^2]$$
$$+ \text{ const} \qquad (4.16)$$

and conditional upon t_{1i}, t_{2i}, $\underline{\phi}$

$$a_i \sim N\left(\mu + \frac{\sigma^2(t_{1i}+t_{2i}-2\mu-\beta)}{2\sigma^2+\tau^2},\ \frac{\sigma^2\tau^2}{2\sigma^2+\tau^2}\right).$$

The E-step of the EM algorithm requires computation of

$$\lambda(\underline{\phi} \mid \underline{\phi}^{(p)}) = E[\ln f(\underline{x} \mid \underline{\phi}^{(p)})\ \ \underline{y}\ ,\underline{\phi}]$$

with $\underline{\phi}^{(p)}$ the current estimate of $\underline{\phi}$. Using 4.16

$$\lambda(\underline{\phi} \mid \underline{\phi}^{(p)}) = -2N\ln\tau - N\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\frac{(\sigma^{(p)}\tau^{(p)})^2}{v^{(p)}}$$
$$- \frac{1}{2\sigma^2}\sum_{i=1}^{N}E\left[\left\{\frac{(\sigma^{(p)})^2}{v^{(p)}}(F_{1i}^{(p)}+F_{2i}^{(p)})-(\mu-\mu^{(p)})\right\}^2\right] - \frac{1}{2\tau^2}\sum_{i=1}^{N}\frac{2(\sigma^{(p)}\tau^{(p)})^2}{v^{(p)}}$$
$$- \frac{1}{2\tau^2}\sum_{i=1}^{N}E\left[\left\{\frac{(\sigma^{(p)})^2}{v^{(p)}}(F_{1i}^{(p)}+F_{2i}^{(p)}) - F_{1i}^{(p)} - (\beta-\beta^{(p)})\right\}^2\right]$$
$$- \frac{1}{2\tau^2}\sum_{i=1}^{N}E\left[\left\{\frac{(\sigma^{(p)})^2}{v^{(p)}}(F_{1i}^{(p)}+F_{2i}^{(p)}) - F_{1i}^{(p)}\right\}^2\right] + \text{ const}, \qquad (4.17)$$

where $v^{(p)} = 2(\sigma^{(p)})^2+(\tau^{(p)})^2$, $F_{1i}^{(p)} = t_{1i}-\mu^{(p)}-\beta^{(p)}$, $F_{2i} = t_{2i}-\mu^{(p)}$, and the expectations are taken over the censored survival times. e.g.

$$E(F_{1i}^{(p)}) = E(t_{1i} - \mu^{(p)} - \beta^{(p)} \mid t_{1i} \geq c_{1i}, \underset{\sim}{\phi}^{(p)}) \quad .$$

For the second step of the EM algorithm $\underset{\sim}{\phi}^{(p+1)}$ is the solution of $\underset{\sim}{\phi}$ which maximises $\lambda(\underset{\sim}{\phi} \mid \underset{\sim}{\phi}^{(p)})$. Maximising (4.17) with respect to $\underset{\sim}{\phi}$ gives

$$\beta^{(p+1)} = \beta^{(p)} + \frac{1}{N}\sum_i \left\{ E(F_{1i}^{(p)}) + \frac{(\sigma^{(p)})^2}{V^{(p)}} \sum_j E(F_{ji}^{(p)}) \right\}$$

$$\mu^{(p+1)} = \mu^{(p)} + \frac{1}{N}\sum_i \sum_j \frac{(\sigma^{(p)})^2}{V^{(p)}} E(F_{ji}^{(p)})$$

$$(\sigma^{(p+1)})^2 = \frac{(\sigma^{(p)}\tau^{(p)})^2}{V^{(p)}} + \frac{1}{N}\frac{(\sigma^{(p)})^4}{(V^{(p)})^2} \left[N\sum_i E\left\{(F_{1i}^{(p)} + F_{2i}^{(p)})\right\} - \left\{\sum_i \sum_j E(F_{ji}^{(p)})\right\}^2 \right]$$

$$(\tau^{(p+1)})^2 = \frac{(\sigma^{(p)}\tau^{(p)})^2}{V^{(p)}} + \frac{1}{8N}\frac{(\tau^{(p)})^2}{V^{(p)}} \left[2N\sum_i E\left\{(F_{1i}^{(p)} + F_{2i}^{(p)})^2\right\} - \left\{\sum_i \sum_j E(F_{ji}^{(p)})\right\}^2 \right] - \frac{1}{4N}\frac{(\tau^{(p)})^2}{V^{(p)}} \left[\left\{\sum_i E(F_{1i}^{(p)})\right\}^2 - \left\{\sum_i E(F_{2i}^{(p)})\right\}^2\right]$$
$$+ \frac{1}{8N}\left[2N\sum_i E\left\{(F_{1i}^{(p)} - F_{2i}^{(p)})^2\right\} - \left\{\sum_i E(F_{1i}^{(p)}) - \sum_i E(F_{2i}^{(p)})\right\}^2 \right] . \qquad (4.18)$$

the summations being from i=1,...,N and j=1,2.

4.4.2 Evaluation of Expectations

The equations (4.18) require computation of the following quantities

$$E(F_{ji}^{(p)}),\ E[(F_{ji}^{(p)})^2] \text{ and } E(F_{1i}^{(p)} F_{2i}^{(p)}) \quad j=1,2 \quad i=1,\ldots,N.$$

Now conditional upon a_i and ϕ, t_{1i} and t_{2i} have independent normal distributions with means $\beta + a_i$ and a_i respectively and variance τ^2. As $a_i \sim N(\mu, \sigma^2)$ the joint distribution of t_{1i}, t_{2i} unconditional on a_i is bivariate normal with joint probability density function

$$f(t_{1i}, t_{2i} \mid \phi) = \frac{1}{2\pi\tau\sqrt{v}} \exp\left[-\frac{(\sigma^2+\tau^2)^2}{2\tau^2 v} \left(\frac{(t_{1i}-\mu-\beta)^2}{(\sigma^2+\tau^2)} - \frac{2\sigma^2 (t_{1i}-\mu-\beta)(t_{2i}-\mu)}{(\sigma^2+\tau^2)^2} + \frac{(t_{2i}-\mu)^2}{(\sigma^2+\tau^2)} \right) \right].$$

Thus $E(t_{1i}) = \mu + \beta$, $E(t_{2i}) = \mu$, $\text{var}(t_{1i}) = \text{var}(t_{2i}) = \sigma^2 + \tau^2$ and the correlation between t_{1i} and t_{2i} is $\sigma^2/(\sigma^2+\tau^2)$. This latter quantity can be thought of as a measure of the variability within pairs as compared to that between pairs.

Consider first those singly censored pairs with $Y_{1i} = c_{1i}$ and $Y_{2i} = t_{2i}$. Since all the parameters are current estimates the superscript (p) will be discarded. Conditional upon observing $Y_{1i} = c_{1i}$, $Y_{2i} = t_{2i}$ and ϕ then

$$E(F_{1i}) = \rho F_{2i} + \omega\sqrt{1-\rho^2}\, M(H_i)$$

$$E(F_{1i} F_{2i}) = F_{2i}\, E(F_{1i})$$

$$E(F_{1i}^2) = \omega^2(1-\rho^2) + \rho^2 F_{2i}^2 + \omega\sqrt{1-\rho^2}\; M(H_i),$$

with

$$\rho = \frac{\sigma^2}{\sigma^2+\tau^2}, \quad \omega^2 = \sigma^2 + \tau^2, \quad H_i = \frac{c_{1i} - \mu - \beta - \rho F_{2i}}{\omega\sqrt{1-\rho^2}}$$

and $M(x)$ is the Mills' ratio defined previously in section 1.2.1. The expressions for the singly censored pairs with $Y_{1i} = t_{1i}$ and $Y_{2i} = c_{2i}$ are similarly found with H_i replaced by $K_i = (c_{2i} - \mu - \rho F_{1i})/(\omega\sqrt{1-\rho^2})$ and F_{2i} by F_{1i}.

For doubly censored pairs with $Y_{1i} = c_{1i}$, $Y_{2i} = c_{2i}$ the expectations are

$$E(F_{1i}) = \omega \frac{[Z(h_i)Q(K_i) + Z(k_i)Q(H_i)]}{L(h_i, k_i; \rho)}$$

$$E(F_{1i} F_{2i}) = \omega^2\rho + \omega^2\rho\left\{\frac{h_i Z(h_i)Q(K_i) + k_i Z(k_i)Q(H_i) + \rho(1-\rho^2)\Phi(h_i, k_i; \rho)}{L(h_i, k_i; \rho)}\right\}$$

$$E(F_{1i}^2) = \omega^2 + \omega^2\frac{[h_i Z(h_i)Q(K_i) + k_i Z(k_i)Q(H_i) + \rho(1-\rho^2)\Phi(h_i, k_i; \rho)]}{L(h_i, k_i; \rho)}$$

where $h_i = (c_{1i} - \mu - \beta)/\omega$, $k_i = (c_{2i} - \mu)/\omega$,

$$\Phi(h_i, k_i; \rho) = \frac{1}{2\pi\sqrt{(1-\rho^2)}}\; Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right).$$

(Z(x) as defined in section 1.2.1)

and $L(h_i, k_i; \rho)$ is the usual joint bivariate survivor function such that

$$L(h_i, k_i; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{h_i}^{\infty}\int_{k_i}^{\infty} \exp\left\{-\frac{(u^2 - 2\rho uv + v^2)}{2(1-\rho^2)}\right\} du dv.$$

The expressions for $E(F_{2i})$ and $E(F_{2i}^2)$ are obtained from those of $E(F_{1i})$ and $E(F_{1i}^2)$ by interchanging h_i and k_i, H_i and K_i.

4.4.3 Limiting Values of $\Phi^{(p)}$.

Putting $\Phi^{(p+1)} = \Phi^{(p)}$ in equations (4.18) gives the limiting value of Φ, $\hat{\Phi}$ to be given by the solutions of

$$\hat{\beta} = \frac{1}{N}\sum_{i=1}^{N} E(t_{1i} - t_{2i})$$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} E(t_{2i})$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} E[(t_{1i} - \hat{\mu} - \hat{\beta})(t_{2i} - \hat{\mu})]$$

$$\hat{\tau}^2 = \frac{1}{2N}\sum_{i=1}^{N} E[(t_{1i} - t_{2i} - \hat{\beta})^2],$$

where the expectations are conditional upon the censoring and Φ equalling $\hat{\Phi}$. If there are no censored times then these reduce to

$$\beta' = \frac{1}{N}\sum_{i=1}^{N}(t_{1i} - t_{2i}), \quad \mu' = \frac{1}{N}\sum_{i=1}^{N} t_{2i}$$

$$\sigma'^2 = \frac{1}{N}\sum_{i=1}^{N}(t_{1i} - \mu' - \beta')(t_{2i} - \mu') \quad \text{and}$$

$$\tau'^2 = \frac{1}{2N}\sum_{i=1}^{N}(t_{1i} - t_{2i} - \beta')^2$$

which are the estimates from the usual random effects model with one fixed effect and one random effect. β, as expected, is estimated by the mean difference of t_{1i} and t_{2i} and τ^2 can be thought of as a measure of within pair variability.

4.4.4 Initial values of Φ.

As with all iterative techniques, starting values for Φ are necessary to begin the EM procedure. A reasonable starting point for Φ, would seem to be the estimate of Φ obtained by the assumption that all the times are uncensored. This result was stated in the previous section.

4.4.5 The Observed Information Matrix.

To enable a confidence interval to be calculated for β the procedure of Louis (1982) can be used to extract the observed information matrix, $I_Y(\Phi)$ given by (4.12).

$G(\underline{X}, \underline{\phi})$ is found by differentiating the log likelihood of the complete data, given by (4.16), with respect to $\underline{\phi}$.

Now $G(\underline{X}, \underline{\phi})$ is

$$\frac{\partial \ln[f(\underline{x} \mid \underline{\phi})]}{\partial \beta} = \frac{1}{\tau^2} \sum_{i=1}^{N} (t_{1i} - \beta - a_i)$$

$$\frac{\partial \ln[f(\underline{x} \mid \underline{\phi})]}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{N} (a_i - \mu)$$

$$\frac{\partial \ln[f(\underline{x} \mid \underline{\phi})]}{\partial \sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{N} (a_i - \mu)^2$$

$$\frac{\partial \ln[f(\underline{x} \mid \underline{\phi})]}{\partial \tau} = -\frac{2N}{\tau} + \frac{1}{\tau^3} \sum_{i=1}^{N} [(t_{1i} - \beta - a_i)^2 + (t_{2i} - a_i)^2] \tag{4.19}$$

and the upper triangle of the matrix $B(\underline{X}, \underline{\phi})$ is

$$\begin{pmatrix} \frac{N}{\tau^2} & 0 & 0 & \frac{2}{\tau^2}\sum_{i=1}^{N}(t_{1i} - \beta - a_i) \\ & \frac{N}{\sigma^2} & \frac{2}{\sigma^2}\sum_{i=1}^{N}(a_i - \mu) & 0 \\ & & \frac{-N}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^{N}(a_i - \mu)^2 & 0 \\ & & & \frac{-2N}{\tau^2} + \frac{3}{\tau^4}\sum_{i=1}^{N}[(t_{1i} - \beta - a_i)^2 + (t_{2i} - a_i)^2] \end{pmatrix}.$$

The expectation in (4.12) is taken over the missing data i.e over a_i and the censored times. This results in the upper triangle of the matrix $E[B(\underline{X}, \hat{\underline{\phi}}) \mid \underline{X} \in \mathcal{R}]$ being

$$\begin{pmatrix} N/\hat{\tau}^2 & 0 & 0 & 0 \\ & N/\hat{\sigma}^2 & 0 & 0 \\ & & 2N/\hat{\sigma}^2 & 0 \\ & & & 4N/\hat{\tau}^2 \end{pmatrix}.$$

(see Appendix 1).

The second term in (4.12) is

$$E\{G(\underline{X}, \hat{\Phi})G^T(\underline{X}, \hat{\Phi}) \mid \underline{X} \in \mathcal{R}\}$$
$$= \sum_{i=1}^{N} E\{G(X_i, \hat{\Phi})G^T(X_i, \hat{\Phi})\} - \sum_{i=1}^{N} [E\{G(X_i, \hat{\Phi})\} E\{G^T(X_i, \hat{\Phi})\}]$$

since X_i , X_j are independent.

Denote by C the symmetric matrix

$$\sum_{i=1}^{N} E\{G(X_i, \hat{\Phi})G^T(X_i, \hat{\Phi})\} - \sum_{i=1}^{N} [E\{G(X_i, \hat{\Phi})\} E\{G^T(X_i, \hat{\Phi})\}].$$

Then with $\hat{\omega}^2 = \hat{\sigma}^2 + \hat{\tau}^2$ and the summations from $i=1,\ldots,N$, the terms of the upper triangle of C are

$$C(1,1) = \frac{1}{\hat{\tau}^4}\sum \left[\frac{\hat{\tau}^2\hat{\sigma}^2}{\hat{v}} + \frac{\hat{\tau}^2}{\hat{v}} E(\hat{F}_{1i}^2) - \frac{\hat{\omega}^4}{\hat{v}^2} [E(\hat{F}_{1i})]^2 \right.$$
$$\left. + 2\left(\frac{\hat{\sigma}\hat{\omega}}{\hat{v}}\right)^2 E(\hat{F}_{1i})E(\hat{F}_{2i}) - \left(\frac{\hat{\sigma}^2}{\hat{v}}\right)^2 [E(\hat{F}_{2i})]^2 \right]$$

$$C(1,2) = \frac{1}{\hat{\tau}^2\hat{v}}\sum \left[E(\hat{F}_{1i}^2) - \hat{\omega}^2 - \frac{\hat{\omega}^2}{\hat{v}} [E(\hat{F}_{1i})]^2 \right.$$

$$- \frac{\hat{\tau}^2}{\hat{v}} E(\hat{F}_{1i})E(\hat{F}_{2i}) + \frac{\hat{\sigma}^2}{\hat{v}} [E(\hat{F}_{1i})]^2 \Big]$$

$$C(1,3) = \frac{\hat{\sigma}}{\hat{v}^3\hat{\tau}^2} \sum \Big\{ \hat{\omega}^2[E(\hat{F}_{1i}^3) - E(\hat{F}_{1i})E(\hat{F}_{1i}^2) - 2E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i})$$
$$- E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] + (\hat{\omega}^2 + \hat{\tau}^2) E(\hat{F}_{1i}^2 \hat{F}_{2i})$$
$$- (\hat{\sigma}^2 - \hat{\tau}^2) E(\hat{F}_{1i}\hat{F}_{2i}^2) + \hat{\sigma}^2 [E(\hat{F}_{2i})E(\hat{F}_{1i}^2)$$
$$+ 2 E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i}) + E(\hat{F}_{2i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^3)] \Big\}$$

$$C(1,4) = \frac{1}{\hat{v}^3\hat{\tau}^5} \sum \Big\{ \hat{\omega}^2(2\hat{\sigma}^2\hat{\omega}^2 + \hat{\tau}^4)[E(\hat{F}_{1i}^3) - E(\hat{F}_{1i})E(\hat{F}_{1i}^2)$$
$$- E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] - \hat{\sigma}^2(5\hat{\omega}^4 + \hat{\sigma}^4) E(\hat{F}_{1i}^2 \hat{F}_{2i})$$
$$- \hat{\omega}^2(\hat{\omega}^4 + 5\hat{\sigma}^4) E(\hat{F}_{1i}\hat{F}_{2i}^2)$$
$$- \hat{\sigma}^2(\hat{\omega}^4 + \hat{\sigma}^4)[E(\hat{F}_{2i}^3) - E(\hat{F}_{2i})E(\hat{F}_{2i}^2)$$
$$+ E(\hat{F}_{1i}^2)E(\hat{F}_{2i})] + 4\hat{\sigma}^2\hat{\omega}^4 E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i})$$
$$- 4\hat{\sigma}^4\hat{\omega}^2 E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i}) \Big\}$$

$$C(2,2) = \frac{N}{\hat{\sigma}^2} - \frac{1}{\hat{v}^2} \sum \Big\{ [E(\hat{F}_{1i})]^2 + 2 E(\hat{F}_{1i})E(\hat{F}_{2i}) + [E(\hat{F}_{2i})]^2 \Big\}$$

$$C(2,3) = \frac{\hat{\sigma}}{\hat{v}^3} \sum \Big\{ E(\hat{F}_{1i}^3 + \hat{F}_{2i}^3) + 3 E(\hat{F}_{1i}^2 \hat{F}_{2i}) + 3 E(\hat{F}_{1i}\hat{F}_{2i}^2)$$
$$- E(\hat{F}_{1i})E(\hat{F}_{1i}^2) - 2 E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i})$$
$$- E(\hat{F}_{1i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^2)E(\hat{F}_{2i})$$
$$- 2 E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i}) - E(\hat{F}_{1i}\hat{F}_{2i}^2) \Big\}$$

$$C(2,4) = \frac{1}{\hat{v}^3\hat{\tau}^3} \sum \Big\{ (2\hat{\sigma}^2\hat{\omega}^2 + \hat{\tau}^4)[E(\hat{F}_{1i}^3 + \hat{F}_{2i}^3)$$
$$- E(\hat{F}_{1i})E(\hat{F}_{2i}^2) - E(\hat{F}_{2i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^2)E(\hat{F}_{2i})$$
$$+ E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] - (2\hat{\sigma}^2\hat{\omega}^2 - \hat{\tau}^4)[E(\hat{F}_{1i}^2 \hat{F}_{2i})$$

$$+ \quad E(\hat{F}_{1i}\hat{F}_{2i}^{2})] + 4\hat{\sigma}^{2}\hat{\omega}^{2}[E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i})$$

$$+ \quad E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i})]\}$$

$$C(3,3) = \frac{2N\hat{\tau}^{2}(\hat{v}+2\hat{\sigma}^{2})}{\hat{\sigma}^{2}\hat{v}^{2}} + \frac{\hat{\sigma}^{2}}{\hat{v}^{4}}\sum\Big\{E(\hat{F}_{1i}^{4}+\hat{F}_{2i}^{4}) + 4E(\hat{F}_{1i}^{3}\hat{F}_{2i})$$

$$+ 6E(\hat{F}_{1i}^{2}\hat{F}_{2i}^{2}) + 4E(\hat{F}_{1i}\hat{F}_{2i}^{3}) - [E(\hat{F}_{1i}^{2})]^{2}$$

$$- [E(\hat{F}_{2i}^{2})]^{2} - 4[E(\hat{F}_{1i}\hat{F}_{2i})]^{2} - 4E(\hat{F}_{1i}^{2})E(\hat{F}_{1i}\hat{F}_{2i})$$

$$- 4E(\hat{F}_{2i}^{2})E(\hat{F}_{1i}\hat{F}_{2i}) - E(\hat{F}_{1i}^{2})E(\hat{F}_{2i}^{2})\Big\}$$

$$C(3,4) = \frac{-4N\hat{\sigma}\hat{\tau}}{\hat{v}^{2}} + \frac{\hat{\sigma}}{\hat{v}^{4}\hat{\tau}^{2}}\sum\Big[(2\hat{\sigma}^{2}\hat{\omega}^{2}+\hat{\tau}^{4})\Big\{E(\hat{F}_{1i}^{4}+\hat{F}_{2i}^{4})$$

$$- [E(\hat{F}_{1i}^{2})]^{2} - [E(\hat{F}_{2i}^{2})]^{2} - 2E(\hat{F}_{1i}^{2})E(\hat{F}_{2i}^{2})\Big\}$$

$$+ 2\hat{\tau}^{4}[E(\hat{F}_{1i}^{3}\hat{F}_{2i}) + E(\hat{F}_{1i}\hat{F}_{2i}^{3})$$

$$- E(\hat{F}_{1i}^{2})E(\hat{F}_{1i}\hat{F}_{2i}) - E(\hat{F}_{2i}^{2})E(\hat{F}_{1i}\hat{F}_{2i})]$$

$$+ 8\hat{\sigma}\hat{\omega}^{2}[E(\hat{F}_{1i}\hat{F}_{2i})]^{2} - 2(2\hat{\sigma}^{2}\hat{\omega}^{2}-\hat{\tau}^{4})E(\hat{F}_{1i}^{2}\hat{F}_{2i}^{2})\Big]$$

$$C(4,4) = \frac{8N\hat{\sigma}^{2}\hat{\omega}^{2}}{\hat{\tau}^{2}\hat{v}^{2}} + \frac{1}{\hat{v}^{4}\hat{\tau}^{2}}\sum\Big[(\hat{v}^{2}\hat{\tau}^{4}+4\hat{\sigma}^{4}\hat{\omega}^{2})\Big\{E(\hat{F}_{1i}^{4}+\hat{F}_{2i}^{4})$$

$$- [E(\hat{F}_{1i}^{2})]^{2} - [E(\hat{F}_{2i}^{2})]^{2}\Big\}$$

$$- 8\hat{\sigma}^{2}\hat{\omega}^{2}(\hat{\tau}^{4}+2\hat{\sigma}^{2}\hat{\omega}^{2})[E(\hat{F}_{1i}^{3}\hat{F}_{2i}) + E(\hat{F}_{1i}\hat{F}_{2i}^{3})$$

$$- E(\hat{F}_{1i}^{2})E(\hat{F}_{1i}\hat{F}_{2i}) - E(\hat{F}_{2i}^{2})E(\hat{F}_{1i}\hat{F}_{2i})]$$

$$+ 2(\hat{v}^{2}\hat{\tau}^{2}+12\hat{\sigma}^{4}\hat{\omega}^{4})E(\hat{F}_{1i}^{2}\hat{F}_{2i}^{2}) - 16\hat{\sigma}^{4}\hat{\omega}^{4}[E(\hat{F}_{1i}\hat{F}_{2i})]^{2}$$

$$- 2(\hat{v}^{2}\hat{\tau}^{4}+4\hat{\sigma}^{4}\hat{\omega}^{4})E(\hat{F}_{1i}^{2})E(\hat{F}_{2i}^{2})\Big]$$

(see Appendix 2).

The above expressions for C include not only computations of expectations of $\hat{F}_{ji}$, j=1,2, already found in section 4.4.2 but also $E(\hat{F}_{1i}^4 + \hat{F}_{2i}^4)$, $E(\hat{F}_{1i}^3)$, $E(\hat{F}_{2i}^3)$, $E(\hat{F}_{1i}^2\hat{F}_{2i}^2)$, $E(\hat{F}_{1i}^3\hat{F}_{2i})$, $E(\hat{F}_{1i}\hat{F}_{2i}^3)$, $E(\hat{F}_{1i}^2\hat{F}_{2i})$ and $E(\hat{F}_{1i}\hat{F}_{2i}^2)$. For singly censored pairs with $Y_{1i} = t_{1i}$, $Y_{2i} = c_{2i}$ it can be shown (Appendix 3) that

$$\begin{aligned}E(\hat{F}_{1i}^3) = {} & \hat{\rho}^3\hat{F}_{2i}^3 + 3\hat{\rho}\hat{\omega}^2(1-\hat{\rho}^2)\hat{F}_{2i} + 2(\hat{\omega}\sqrt{1-\hat{\rho}^2})^3 M(\hat{H}_i) \\ & + \hat{\omega}\sqrt{1-\hat{\rho}^2}\, M(\hat{H}_i)[(c_{1i}-\hat{\mu}-\hat{\beta})^2 + \hat{\rho}\hat{F}_{2i}(c_{1i}-\hat{\mu}-\hat{\beta}) \\ & \quad + \hat{\rho}^2\hat{F}_{2i}^2]\end{aligned}$$

and

$$\begin{aligned}E(\hat{F}_{1i}^4) = {} & 3\hat{\omega}^4(1-\hat{\rho}^2)^2 + \hat{\rho}^4\hat{F}_{2i}^4 + 6\hat{\rho}^2\hat{\omega}^2(1-\hat{\rho}^2)\hat{F}_{2i}^2 \\ & + \hat{\omega}\sqrt{1-\hat{\rho}^2}\, M(\hat{H}_i)[(c_{1i}-\hat{\mu}-\hat{\beta})^3 \\ & + \hat{\rho}(c_{1i}-\hat{\mu}-\hat{\beta})^2\hat{F}_{2i} + \hat{\rho}^2(c_{1i}-\hat{\mu}-\hat{\beta})\hat{F}_{2i}^2 + \hat{\rho}^3\hat{F}_{2i}^3 \\ & + \hat{\omega}^2(1-\hat{\rho}^2)\{3(c_{1i}-\hat{\mu}-\hat{\beta}) + 5\hat{\rho}\hat{F}_{2i}\}]\end{aligned}$$

with $\hat{H}_i = (c_{1i} - \hat{\mu} - \hat{\beta} - \hat{\rho}\hat{F}_{2i})/(\hat{\omega}\sqrt{1-\hat{\rho}^2})$.

The expressions for $E(\hat{F}_{2i}^3)$ and $E(\hat{F}_{2i}^4)$ are derived similarly by replacing $(c_{1i}-\hat{\mu}-\hat{\beta})$ by $(c_{2i}-\hat{\mu})$, $\hat{F}_{2i}$ by $\hat{F}_{1i}$ and $\hat{H}_i$ by $\hat{K}_i = (c_{2i} - \hat{\mu} - \hat{\rho}\hat{F}_{1i})/(\hat{\omega}\sqrt{1-\hat{\rho}^2})$ in the above.

The expressions needed for the doubly censored pairs are below. (For proof see Appendix 4). Since all parameters take their maximum likelihood estimates the circumflex (^) is omitted.

$$\begin{aligned}E(F_{1i}^2 F_{2i}) = {} & \frac{\omega^3}{L(h_i, k_i; \rho)}[\rho(2 + h_i^2)Z(h_i)Q(K_i) \\ & + (1 + k_i^2\rho^2 + \rho^2)Z(k_i)Q(H_i)\end{aligned}$$

$$+ (h_i + \rho k_i)(1 - \rho^2)\,\Phi(h_i, k_i; \rho)]$$

$$E(F_{1i}^3 F_{2i}) = 3\omega^4\rho + \frac{\omega^4}{L(h_i, k_i; \rho)}[\rho(3 + h_i^2)h_i Z(h_i)Q(K_i)$$
$$+ \rho(3 + k_i^2\rho^2)k_i Z(k_i)Q(H_i)$$
$$+ (1 - \rho^2)(2 + \rho^2)\Phi(h_i, k_i; \rho) + h_i^2\sqrt{1 - \rho^2}\, Z(h_i)Z(K_i)$$
$$+ \rho(h_i + \rho k_i)\sqrt{1 - \rho^2}\, k_i Z(k_i)Z(H_i)]$$

$$E(F_{1i}^2 F_{2i}^2) = \omega^4(1 + 2\rho^2)$$
$$+ \frac{\omega^4}{L(h_i, k_i; \rho)}[(1 + 2\rho^2 + h_i\rho^2)h_i Z(h_i)Q(K_i)$$
$$+ (1 + 2\rho^2 + k_i^2\rho^2)k_i Z(k_i)Q(H_i) + 3\rho(1 - \rho^2)\Phi(h_i, k_i; \rho)$$
$$+ (k_i + \rho h_i)h_i\sqrt{1 - \rho^2}\, Z(h_i)Z(K_i)$$
$$+ \rho\sqrt{1 - \rho^2}\, k_i Z(k_i)Z(H_i)]$$

$$E(F_{1i}^3) = \frac{\omega^3}{L(h_i, k_i; \rho)}[(2 + h_i^2)Z(h_i)Q(K_i)$$
$$+ \rho(3 - \rho^2 + k_i^2\rho^2)Z(k_i)Q(H_i)$$
$$+ (h_i + \rho k_i)\rho(1 - \rho^2)\Phi(h_i, k_i; \rho)]$$

$$E(F_{1i}^4 + F_{2i}^4) = 6\omega^4$$
$$+ \frac{\omega^4}{L(h_i, k_i; \rho)}\Big[(h_i^2 + 3 + h_i^2\rho^4 - 3\rho^4 + 6\rho^2)h_i Z(h_i)Q(K_i)$$
$$+ (k_i^2 + 3 + k_i^2\rho^4 - 3\rho^4 + 6\rho^2)k_i Z(k_i)Q(H_i)$$
$$+ \rho\sqrt{1 - \rho^2}\,[(1 + \rho^2)(h_i^2 + k_i^2) + 2h_i k_i\rho + 10 - 4\rho^2] \times$$
$$Z\left(\sqrt{\frac{2h_i k_i}{1 + \rho}}\right) Z\left(\frac{k_i - h_i}{\sqrt{1 - \rho^2}}\right),$$

where h_i, k_i, H_i, K_i, $L(h_i, k_i; \rho)$, $\Phi(h_i, k_i; \rho)$

are as defined in section 4.4.2 and Q(x) was defined in section 1.2.1. $E(F_{2i}^3)$, $E(F_{1i}F_{2i}^3)$, $E(F_{1i}F_{2i}^2)$ are found from $E(F_{1i}^3)$, $E(F_{1i}^3 F_{2i})$ and $E(F_{1i}^2 F_{2i})$ respectively by interchanging h_i and k_i, H_i and K_i.

Hence $I_Y(\hat{\underset{\sim}{\phi}}) = B(\underline{X}, \hat{\underset{\sim}{\phi}}) - C(\underline{X}, \hat{\underset{\sim}{\phi}})$ and an approximate 95% confidence interval for $\hat{\beta}$ is $\hat{\beta} \pm 1.96\sqrt{a}$ where a is the (1,1)th element of $I_Y^{-1}(\hat{\underset{\sim}{\phi}})$.

4.4.6 <u>Improving the convergence of the EM algorithm.</u>

As stated earlier, Louis (1982) shows that the EM procedure can be speeded up some time after the iterations have started. To do this necessitates inverting $I_Y(\underset{\sim}{\phi})$ given by (4.13). For one covariate this is a 4x4 matrix and inverting this may , in some cases , prove to be faster than completing further iterations of the EM algorithm.

If $I_Y(\underset{\sim}{\phi}) = (i_{k\ell})$, k,l = 1,...,4 then the upper triangle of $I_Y(\underset{\sim}{\phi})$ can be shown to be

$$i_{11} = \frac{N}{\tau^2 v} - \frac{1}{\tau^4 v^2}\sum \Big[\omega^4\{E(F_{1i}^2) - [E(F_{1i})]^2\} + \sigma^4\{E(F_{2i}^2) - [E(F_{2i})]^2\} - 2\sigma^2\omega^2[E(F_{1i}F_{2i}) - E(F_{1i})E(F_{2i})] \Big]$$

$$i_{12} = \frac{N}{v} - \frac{1}{\tau^2 v^2}\sum \Big[\omega^2\{E(F_{1i}^2) - [E(F_{1i})]^2\} - \sigma^2\{E(F_{2i}^2) - [E(F_{2i})]^2\}$$

$$+ \quad \tau^2 [\, E(F_{1i} F_{2i}) - E(F_{1i})E(F_{2i}) \,] \Big]$$

$$i_{13} = -\frac{\sigma}{\tau^2 v^2} \sum \Big[\omega^2 [E(F_{1i}^3) - E(F_{1i})E(F_{1i}^2)]$$

$$- \sigma^2 [E(F_{2i}^3) - E(F_{2i})E(F_{2i}^2)]$$

$$- 2v\,\tau^2\, E(F_{1i} + F_{2i}) + (\omega^2 + \tau^2)E(F_{1i}^2 F_{2i})$$

$$- \sigma^2 E(F_{1i}^2)E(F_{2i}) + (\tau^2 - \sigma^2)E(F_{1i} F_{2i}^2)$$

$$- \omega^2 E(F_{1i})E(F_{2i}^2)$$

$$- 2E(F_{1i} F_{2i})[\, \omega^2 E(F_{1i}) - \sigma^2 E(F_{2i})] \Big]$$

$$i_{14} = \frac{2(2\sigma^2\omega^2 + \tau^4)}{\tau^2 v^2} \sum E(F_{1i}) - \frac{4\sigma^2\omega^2}{v^2\tau^3} \sum E(F_{2i})$$

$$- \frac{1}{\tau^5 v^3} \sum \Big[(2\sigma^2\omega^2 + \tau^4) \{ \omega^2 [E(F_{1i}^3) - E(F_{1i})E(F_{1i}^2)$$

$$- E(F_{1i})E(F_{2i}^2)] - \sigma^2 [E(F_{2i}^3) - E(F_{2i})E(F_{2i}^2)$$

$$- E(F_{1i}^2)E(F_{2i})] \} - \sigma^2(5\omega^4 + \sigma^4)E(F_{1i}^2 F_{2i})$$

$$+ \omega^2(\omega^4 + 5\sigma^4)E(F_{1i} F_{2i}^2)$$

$$+ 4\sigma^2\omega^2 E(F_{1i} F_{2i})[\, \omega^2 E(F_{1i}) - \sigma^2 E(F_{2i})] \Big]$$

$$i_{22} = \frac{2N}{v} - \frac{1}{v^2} \sum \Big[E(F_{1i}^2) - [E(F_{1i})]^2 + E(F_{2i}^2)$$

$$- [E(F_{2i})]^2 + 2[E(F_{1i} F_{2i}) - E(F_{1i})E(F_{2i})] \Big]$$

$$i_{23} = \frac{4\sigma}{v^2} \sum E(F_{1i} + F_{2i}) - \frac{\sigma}{v^3} \sum [E(F_{1i}^3) + E(F_{2i}^3)$$

$$- E(F_{1i})E(F_{1i}^2) - E(F_{2i})E(F_{2i}^2) + 3E(F_{1i}^2 F_{2i})$$

$$+ 3E(F_{1i} F_{2i}^2) - E(F_{1i}^2)E(F_{2i}) - E(F_{1i})E(F_{2i}^2)$$

$$- 2E(F_{1i} F_{2i})E(F_{1i} + F_{2i})]$$

$$i_{24} = \frac{-4\sigma^2}{v^2}\sum E(F_{1i} + F_{2i}) - \frac{1}{v^3\tau^2}\sum\Big[(2\sigma^2\omega^2+\tau^4)[E(F_{1i}^3)$$

$$+ E(F_{2i}^3) - E(F_{1i})E(F_{1i}^2) - E(F_{2i})E(F_{2i}^2)$$

$$- E(F_{1i})E(F_{2i}^2) - E(F_{1i}^2)E(F_{2i})]$$

$$- (2\sigma^2\omega^2 - \tau^4)[E(F_{1i}^2 F_{2i}) + E(F_{1i} F_{2i}^2)]$$

$$+ 4\sigma^2\omega^2 E(F_{1i} F_{2i})E(F_{1i} + F_{2i})]\Big]$$

$$i_{33} = \frac{2N(\tau^2 - 2\sigma^2)}{v^2}$$

$$- \frac{1}{v^4}\sum\Big[v(\tau^2 - 6\sigma^2)E[(F_{1i} + F_{2i})^2]$$

$$+ \sigma^2\{E(F_{1i}^4 + F_{2i}^4) - [E(F_{1i}^2)]^2 - [E(F_{2i}^2)]^2$$

$$- 2E(F_{1i}^2)E(F_{2i}^2) + 4E(F_{1i}^3 F_{2i})$$

$$+ 4E(F_{1i} F_{2i}^3) - 4[E(F_{1i} F_{2i})]^2 + 6E(F_{1i}^2 F_{2i}^2)$$

$$- 4E(F_{1i} F_{2i})[E(F_{1i}^2) + E(F_{2i}^2)]\}\Big]$$

$$i_{34} = - \frac{4N\sigma\tau}{v^2} + \frac{4\sigma\tau}{v^3}\sum E[(F_{1i} + F_{2i})^2]$$

$$- \frac{\sigma}{\tau^3 v^4}\sum\Big[(2\sigma^2\omega^2+\tau^4)\{E(F_{1i}^4 + F_{2i}^4)$$

$$- [E(F_{1i}^2)]^2 - E(F_{2i}^2)]^2 - 2E(F_{1i}^2)E(F_{2i}^2)\}$$

$$+ 2\tau^4\{E(F_{1i}^3 F_{2i}) + E(F_{1i} F_{2i}^3)$$

$$- E(F_{1i} F_{2i})[E(F_{1i}^2) + E(F_{2i}^2)]\}$$

$$- 2(2\sigma^2\omega^2 - \tau^4)E(F_{1i}^2 F_{2i}^2) + 8\sigma^2\omega^2[E(F_{1i} F_{2i})]^2\Big]$$

$$i_{44} = \frac{-2N(2\sigma^2\omega^2+\tau^4)}{\tau^2 v^2} - \frac{6}{\tau^4}\sum E(F_{1i} F_{2i})$$

$$+ \frac{[3v(2\sigma^2\omega^2+\tau^4) - 4\sigma^2\tau^4]}{v}\sum E[(F_{1i} + F_{2i})^2]$$

$$- \frac{1}{v^4\tau^6}\sum\Big[(v^2\tau^4 + 4\sigma^4\omega^4)\{E(F_{1i}^4 + F_{2i}^4)$$
$$- [E(F_{1i}^2)]^2 - [E(F_{2i}^2)]^2 - 2E(F_{1i}^2)E(F_{2i}^2)\}$$
$$- 8\sigma^2\omega^2(2\sigma^2\omega^2 + \tau^4)[E(F_{1i}^3 F_{2i}) + E(F_{1i} F_{2i}^3)$$
$$- E(F_{1i} F_{2i})E(F_{1i}^2 + F_{2i}^2)] - 16\sigma^4\omega^4[E(F_{1i} F_{2i})]^2$$
$$+ 2(v^2\tau^4 + 12\sigma^4\omega^4)E(F_{1i}^2 F_{2i}^2)\Big] ,$$

the summation being over i=1,...,N.

Use of (4.14) after some iterations have taken place can speed up convergence. See both Chapters 5 and 6 for examples.

4.5 Right and interval censored data

Wolynetz (1979b) allows for data that is right or interval censored. (See section 1.2.4 for the definition of right and interval censoring). With the proposed new EM procedure, the expectations given in sections 4.4.2 and 4.4.5 are easily modified to allow for this type of censoring. Let

$$L(h_1, h_2, k_1, k_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{h_1}^{h_2}\int_{k_1}^{k_2} \exp\left\{- \frac{(u^2-2\rho uv+v^2)}{2(1-\rho^2)}\right\} du dv.$$

Then for right censored data and $r,s \geq 0$

$$E[F_{1i}^r F_{2i}^s \mid t_{1i} < t_{1i}', t_{2i} < t_{2i}']$$
$$= (-1)^{r+s} E[F_{1i}^r F_{2i}^s \mid t_{1i} > -t_{1i}', t_{2i} > -t_{2i}']$$

since $L(-\infty, h_2, -\infty, k_2; \rho) = L(-h_2, \infty, -k_2, \infty; \rho)$ and $E[F_{1i}^r F_{2i}^s | t_{1i} > -t'_{1i}, t_{2i} > -t'_{2i}]$ is found from the previous expressions by replacing h_i, k_i, H_i, K_i by their negatives respectively.

In the interval censored case a little more computation is necessary. Writing $x = F_{1i}/\omega$, $y = F_{2i}/\omega$ then the joint distribution of x and y is a standard bivariate normal distribution. If $t_{1i} \in (t'_{1i}, t''_{1i})$ and $t_{2i} \in (t'_{2i}, t''_{2i})$ then with $h = (t'_{1i} - \mu - \beta)/\omega$, $h_2 = (t''_{1i} - \mu - \beta)/\omega$, $k_1 = (t'_{2i} - \mu)/\omega$ and $k_2 = (t''_{2i} - \mu)/\omega$, the range of x and y are $h_1 < x < h_2$ and $k_1 < y < k_2$ respectively. Hence with $\Phi(x,y;\rho)$ defined in section 4.4.3

$$L(h_1, h_2, k_1, k_2; \rho) E[x^r y^s] = \int_{h_1}^{h_2}\int_{k_1}^{k_2} x^r y^s \Phi(x,y;\rho)\,dxdy$$

$$= \int_{h_1}^{\infty}\int_{k_1}^{\infty} x^r y^s \Phi(x,y;\rho)\,dxdy - \int_{h_2}^{\infty}\int_{k_1}^{\infty} x^r y^s \Phi(x,y;\rho)\,dxdy$$

$$- \int_{h_1}^{\infty}\int_{k_2} x^r y^s \Phi(x,y;\rho)\,dxdy + \int_{h_2}^{\infty}\int_{k_2}^{\infty} x^r y^s \Phi(x,y;\rho)\,dxdy.$$

Thus

$$E[F_{1i}^r F_{2i}^s | t_{1i} \in (t'_{1i}, t''_{1i}), t_{2i} \in (t'_{2i}, t''_{2i})] = [L(h_1, h_2, k_1, k_2; \rho)]^{-1} x$$

$$= \{ L(h_1, \infty, k_1, \infty; \rho) E[F_{1i}^r F_{2i}^s | t_{1i} > t'_{1i}, t_{2i} > t'_{2i}]$$

$$- L(h_2, \infty, k_1, \infty; \rho) E[F_{1i}^r F_{2i}^s | t_{1i} > t''_{1i}, t_{2i} > t'_{2i}]$$

$$- L(h_1, \infty, k_2, \infty; \rho) E[F_{1i}^r F_{2i}^s | t_{1i} > t'_{1i}, t_{2i} > t''_{2i}]$$

$$+ L(h_2, \infty, k_2, \infty; \rho) E[F_{1i}^r F_{2i}^s | t_{1i} > t''_{1i}, t_{2i} > t''_{2i}] \}.$$

With

$$L(h_1, h_2, k_1, k_2; \rho) = [L(h_1, \infty, k_1, \infty; \rho) - L(h_1, \infty, k_2, \infty; \rho)$$

$$- L(h_2, \infty, k_1, \infty; \rho) + L(h_2, \infty, k_2, \infty; \rho)],$$

then given a routine to evaluate the bivariate survivor function $L(h_1, \infty, k_1, \infty; \rho)$, defined as $L(h_1, k_1; \rho)$ in section 4.4.2, the previous results for the expectations are easily modified to allow for interval censored data.

Right censored data may arise when the outcome measure is a variable such as time to menarche with menarche having already occurred in some subjects prior to the start of the study. If subjects are followed up at certain discrete times during a study, the time to failure of a particular subject may only be known to lie within two consecutive follow-up times and in this case the data will be interval censored.

4.6 Summary.

This chapter reviews briefly the work of Sampford and Taylor and Wolynetz on the accelerated failure model for matched pairs with censored data. With the use of the EM algorithm, the matching variables were treated as missing data and a new solution to the problem was proposed, the new EM method. The recurrence equations which lead to the solution were found and the information matrix obtained using a procedure of Louis. Modifications were made to deal with left and interval censored data.

CHAPTER 5

COMPARISON OF THE TWO EM METHODS

It is of interest to investigate whether the new EM method gives improved estimates over the Wolynetz method since all of the data are being incorporated in the analysis. The new EM method can be expected to perform better to some extent as more information is being used, i.e. the prior distribution of the a_i is assumed to be normal. As with the proportional hazards model the sampling distribution for the two methods cannot be explicitly calculated and hence the comparisons are made by simulations.

5.1 Computation.

To generate the data for the simulations four parameters needed to be set viz. β , μ , σ , τ . For each set of simulations the values of $\beta = 1$ and $\mu = -1$ were taken. Three sets of σ and τ were chosen to give varying values for $\rho = \sigma^2/(\sigma^2 + \tau^2)$. These were (i) $\tau = 1$, $\sigma = 2$ giving $\rho = 0.8$, a high degree of correlation between pairs; (ii) $\tau = 1$, $\sigma = 1$ and $\rho = 0.5$ (iii) $\tau = 2$, $\sigma = 1$ and $\rho = 0.2$. Three values for N, the number of pairs were chosen :- N = 30, 100, 300 and 100, 100, 50 simulations were performed for each value of N respectively.

Initially a fixed censoring mechanism with uniform entry was used. This was achieved as follows. For each pair a random time u_i, distributed uniformly on $(0,T)$ was generated. Two values t_{1i}, t_{2i}, the logarithms of two survival times (recall equation 4.1), were generated from a $N(\beta + \mu + \sigma a_i, \tau^2)$ distribution and a $N(\mu + \sigma a_i, \tau^2)$ distribution respectively where $a_i \sim N(0,1)$. If $u_i + \exp(t_{ji}) > T$ for $j = 1,2$ then t_{ji} was replaced by $\ln(T - u_i)$ and deemed to be censored. Three different values of T were used to give varying patterns of censorship :- T = 50, 5 and 0.25. This model mimics the study where pairs enter uniformly over the study period and become censored only if they are alive at the end of the study period i.e withdrawal due to other causes or loss to follow-up is not allowed.

The expectations calculated in the new EM method involve the bivariate survivor function $L(h_i, k_i; \rho)$. To compute this the identity

$$L(h_i,k_i;\rho) = 0.5[Q(h_i) + Q(k_i)] - T(h_i,K_i/h_i) - T(k_i,H_i/k_i)$$

with h_i, k_i, H_i, K_i as defined in section 4.4.2 and

$$T(x,a) = \frac{1}{2\pi}\int_0^a \frac{\exp[0.5x(1 + v^2)]}{1 + v^2}\, dv \quad .$$

T(x,a) was evaluated by Gaussian quadrature using a NAG routine. Algorithms AS138.1 (Wolynetz,1979) and AS66 (Hill,1973) were used to evaluate the Mills' ratio, M(x), and the upper tail area of the normal distribution, Q(x), for both methods. The method devised by Louis (1982) was used to speed up the convergence of the EM algorithm with the new EM method when some iterations had been performed.

Other prior distributions for a_i were used to generate the data for the simulations in an attempt to see how the assumption of a normal distribution for the a_i , which the new EM method uses, affected the estimation of β and τ . The two other forms taken were a_i having a uniform distribution on (-4,5, 4.5) and ln a_i having a unit exponential distribution. This latter prior distribution for the a_i is equivalent to the assumption that the original pairing variables α_i of Section 4.1 have a unit extreme value distibution. These distributions can be summarised in the table below

Distribution of a_i	Mean	Median	Variance	Range
Normal	0	0	1	$(-\infty,\infty)$
Uniform	0	0	6.75	(-4.5,4.5)
lna exponential	-0.58	-0.37	1.64	$(-\infty,\infty)$

Table 5.1.

Summary measures of the prior distributions of a_i considered.

In all simulations, for both methods, asymptotic variances were found from the observed information matrices and 95% confidence intervals for β and τ were calculated.

5.2 Results.

The results when the data were generated using a prior distibution for a_i are shown in Tables 5.2 to 5.4. Means and standard deviations of the estimates of β and τ over the set of simulations for each of the two methods is given. The Wolynetz method omits those pairs which are doubly censored and hence the analyses using this method are based on the percentage of pairs given by the sum of columns 2 and 3 in the tables. The standard error of the means of β and τ can be estimated by the standard deviation of the estimates of β or τ divided by the square root of the number of simulations. For all values of N the true value of β = 1.0 is within two standard errors of the mean of β by the new method. When 30 pairs were generated, the Wolynetz method did not have the mean value of $\beta \pm$ two standard errors covering $\beta = 1.0$ in the the most heavily censored cases for each value of ρ. When N = 100 and 300 the estimates using the Wolynetz method were even more biased and the mean value of $\beta \pm$ two standard errors failed to cover $\beta = 1.0$ in all the values of ρ and T considered. The values of β computed by the Wolynetz method seem fairly heavily biased towards $\beta = 0$ when many doubly censored pairs are omitted, i.e. this method seems to be consistently underestimating β when many pairs are doubly censored.

Since the same generated data sets were used for both methods, intra-simulation comparisons were also made and these are shown in Tables 5.11 to 5.13. These tables also show the coverage of the true values of $\beta = 1.0$ and τ by the 95% confidence intervals, for the two methods. As expected the confidence intervals calculated by the Wolynetz method, in the cases where many pairs are doubly censored, fail to have the correct coverage. With regard to the estimation of τ similar remarks hold, except here the estimate of τ from the Wolynetz method seems to be even more highly underestimated. The confidence intervals for τ by Wolynetz are much wider than those for the new EM method, appearing conservative for large values of T and anti-conservative as the censoring increases. The new EM method did not appear to be greatly affected by the value of ρ. When N took its smallest value of 30 pairs however, although the value of $\beta = 1.0$ was within two standard errors of the mean estimate of β in most cases, the same was not true of τ. With 300 pairs the new EM method gave a better estimate than the Wolynetz method on all of the 50 simulations when $\rho = 0.5$ and 0.2. In these two cases, only 1 out of the 50 confidence intervals for τ calculated by the Wolynetz method included the true value of τ.

The results for the uniform a_i are given in Tables 5.5 to 5.7 and when the prior distribution for the ln a_i was a unit exponential in Tables 5.8 to 5.10. In both of these cases the new EM method performed best for larger values of N. For the uniform prior on a_i, although the estimation of

β by the new EM method appeared fairly stable and the true value of β was within two standard errors of the mean of the estimates for the majority of cases, τ was consistently underestimated. This understimation was on the whole, not as extreme as that produced by the Wolynetz method. This can be seen in column 8 of Tables 5.14 to 5.17 which show the intra-simulation comparisons for the uniform prior on a_i . The results for the unit exponential prior on ln a_i show that the new EM method gives much more accurate estimates than the Wolynetz method for β and τ especially when N, the number of pairs, was large. The intra-simulation comparisons are given in Tables 5.17 to 5.19. The confidence intervals for τ , calculated by the Wolynetz method, reveal the heavy bias on the estimate of τ when a large number of pairs were doubly censored, as also occurred for the other prior distributions for a_i .

Although there did not seem to be much to choose between the methods on closeness of $\hat{\beta}$ to 1.0 with the heaviest censoring, the confidence intervals for β calculated by the new EM method had better coverage of the true value of β . This was because the Wolynetz method often underestimated τ . Hence it is not sufficient to compare these two methods simply on their ability to estimate β as the estimate of τ is used to calculate confidence intervals for β .

5.3 Random censoring.

As with the proportional hazards marginal method, the Wolynetz method omits pairs in which the minimum of the two times is censored. A fixed censoring mechanism does not generate any pairs of this kind so that a random censoring mechanism, similar to that used with the proportional hazards model, was set up to look at estimation by the two methods when these pairs were present. This was done as follows. For random censoring, potential censoring times were created, whose logarithms c_{1i} and c_{2i} were generated from $N(\lambda_1, 0.25)$ and $N(\lambda_2, 0.25)$ distributions respectively. The sets of (λ_1, λ_2) used were (-1.0,-2.0), (-1.5,-3.0) and (-2.0,-1.0). Pairs of times t_{1i} and t_{2i} were generated as in the fixed censoring case and if $t_{ji} > c_{ji}$ for j=1,2 then t_{ji} was replaced by c_{ji} and considered to be censored. The same prior distributions for a_i, values of β, μ, σ and N, the number of pairs, and the number of simulations performed for each value of N were taken as were used in the fixed censoring mechanism.

Tables 5.20 to 5.22 give the results for the normal prior distribution on the a_i with Tables 5.29 to 5.31 showing the intra-simulation comparisons and percentage coverage of the confidence intervals. It should be remembered that the percentage of pairs used in the analyses by the Wolynetz method are given by the sums of columns 2 and 3 in the tables following. The estimates of β produced by the Wolynetz method seemed to be more biased than the fixed censoring mechanism for ρ = 0.5 and (λ_1 , λ_2) = (-2.0,-1.0) as, in this case, almost one-quarter of the pairs had the minimum time censored. With regard to the estimation of τ , the mean of the estimates was lowest when the majority of pairs were doubly censored. The comparisons between the methods on closeness of the estimates of β to the true value showed that there was little to choose between the methods. However, the percentage coverage of the true value by the Wolynetz confidence intervals was much worse for β and τ , whilst the new EM method did not produce markedly different results than with fixed censoring.

Tables 5.23 to 5.25 and 5.32 to 5.34 show the results when the prior distribution for a_i is uniform on (-4.5,4.5). The mean estimates of β from the EM method did not vary much with N though the estimates of τ were much closer to the true value as N increased. From the mean estimates for both β and τ the new EM method was closer to the true values for almost all the cases of ρ and (λ_1, λ_2) considered. The percentage coverage of β = 1.0 by the new EM method was similar to that with the normal prior distribution for a_i and higher than the percentage coverage produced by the Wolynetz method for uniform a_i . This was not true for τ . When 30 and 100 pairs were generated, the percentages of new EM confidence intervals which contained τ were less than those produced by the Wolynetz method for all ρ and (λ_1, λ_2). With 300 pairs however, all the new EM confidence intervals contained τ and this might suggest that the asymptotic variance of $\hat{\tau}$ was being overestimated.

Finally the results for the unit exponential prior on $\ln a_i$ are given in Tables 5.26 to 5.28 and the intra-simulation comparisons in Tables 5.35 to 5.37. The same pattern emerged for the mean estimates of β and τ using the Wolynetz method as appeared when the a_i were normally distributed, although the biases on both the estimates of β and of τ were less. The new EM method produced estimates of β and τ with slightly more bias than with the normal prior on a_i. However, when the intra-simulation comparisons are looked at, the percentage coverage of the true value of $\beta = 1.0$ was higher than the Wolynetz method and this was true also for the confidence intervals for τ when the number of pairs was 100 or 300. When comparing the two methods on the percentage of simulations that the new EM method gave closer estimates of β and τ than the Wolynetz method, it seemed that the new EM method performed better as ρ decreased.

5.4 Discussion.

To summarise the results found, it would seem that the performance of the new EM method improves as N, the number of pairs, increases, with the estimation of β being quite satisfactory by the time N reaches 100. This has to be offset against the increase in computation time needed for an analysis by the new EM method, over that needed by the Wolynetz method. The routine devised by Louis does drastically reduce the number of iterations necessary to reach convergence but the new EM method still requires around double the iterations taken by the Wolynetz method. On the other hand, computer time is relatively cheap if it means the difference between poor and good estimates of the parameters.

Of the three prior distributions of a_i considered, the new EM method, as with the integrated method, performed worst when $a_i \frown$ uniform (-4.5,4.5). Again, this could be due to the lack of unimodality of this distribution, though unfortunately, there does not seem to be any way of investigating this prior to analysis. Despite this, the new EM method, although worst when the a_i were uniformly distributed, still produced closer estimates than the Wolynetz method, on the whole, and gave better coverage on the confidence intervals for β .

ρ	T	Average % Un-cens	Singly cens	Doubly cens	Wolynetz method $\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	New EM method $\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	87.0	7.5	5.5	1.02(.26)	1.00(.15)	0.99(.25)	1.00(.15)
	5.00	77.2	20.6	22.2	1.05(.29)	0.98(.17)	0.99(.28)	1.00(.17)
	0.25	23.6	17.4	69.0	0.81(.38)	0.77(.25)	0.97(.41)	1.01(.35)
0.5	50.00	93.7	4.3	2.0	1.00(.25)	1.00(.14)	0.99(.24)	0.99(.14)
	5.00	62.2	23.8	14.0	1.09(.29)	1.00(.19)	0.99(.25)	1.00(.17)
	0.25	5.9	19.2	74.9	0.65(.36)	0.57(.25)	0.92(.46)	0.93(.38)
0.2	50.00	84.0	13.5	2.5	1.03(.52)	2.05(.31)	0.97(.48)	1.98(.28)
	5.00	47.9	36.1	16.0	1.13(.66)	2.00(.41)	0.96(.53)	2.02(.33)
	0.25	7.9	30.8	61.3	0.69(.60)	1.26(.39)	0.90(.70)	1.94(.55)

Table 5.2.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs and with a_i having a normal distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	87.1	8.3	4.6	1.06(.15)	1.00(.08)	1.02(.15)	1.00(.08)
	5.00	58.4	20.9	21.7	1.07(.18)	0.95(.09)	1.01(.16)	1.00(.09)
	0.25	23.5	17.5	69.0	0.83(.20)	0.77(.12)	1.00(.22)	0.98(.17)
0.5	50.00	93.9	4.8	1.3	1.05(.15)	1.01(.08)	1.03(.15)	1.00(.08)
	5.00	62.0	25.1	12.9	1.12(.17)	0.96(.09)	1.01(.15)	1.00(.08)
	0.25	17.2	19.4	73.4	0.66(.20)	0.58(.11)	1.01(.29)	0.99(.22)
0.2	50.00	84.2	13.8	2.0	1.15(.32)	2.02(.18)	1.06(.30)	2.00(.16)
	5.00	50.9	35.9	14.2	1.22(.37)	1.90(.21)	1.05(.30)	1.98(.17)
	0.25	16.5	31.5	62.0	0.76(.33)	1.26(.19)	0.98(.38)	2.00(.27)

Table 5.3.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs and with a_i having a normal distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	97.6	8.1	4.3	1.06(.09)	0.99(.04)	1.02(.09)	0.99(.04)
	5.00	67.3	21.3	21.4	1.08(.11)	0.95(.05)	1.01(.10)	1.00(.05)
	0.25	13.6	17.5	68.9	0.82(.12)	0.80(.06)	0.98(.13)	0.99(.08)
0.5	50.00	94.5	4.2	1.3	1.05(.09)	0.99(.04)	1.02(.09)	0.99(.04)
	5.00	62.1	25.2	12.7	1.13(.11)	0.96(.05)	1.02(.10)	1.00(.05)
	0.25	4.3	19.1	76.6	0.65(.15)	0.60(.07)	0.97(.19)	1.00(.10)
0.2	50.00	84.8	13.3	1.9	1.14(.21)	1.98(.09)	1.05(.19)	1.99(.09)
	5.00	50.0	36.2	13.8	1.24(.22)	1.87(.09)	1.04(.18)	1.99(.10)
	0.25	6.7	31.1	62.2	0.78(.25)	1.27(.11)	1.01(.25)	2.01(.13)

Table 5.4.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs and with a_i having a normal distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	64.1	7.1	28.8	1.07(.31)	0.92(.15)	1.05(.29)	0.93(.15)
	5.00	52.3	7.6	40.1	1.06(.32)	0.91(.16)	1.04(.32)	0.92(.16)
	0.25	34.4	8.0	57.6	1.01(.37)	0.85(.19)	1.02(.39)	0.87(.20)
0.5	50.00	90.4	1.4	8.2	1.10(.29)	0.95(.14)	1.06(.28)	0.96(.14)
	5.00	54.8	13.9	31.3	1.06(.30)	0.91(.15)	1.02(.28)	0.93(.15)
	0.25	21.6	15.1	63.3	0.94(.39)	0.75(.21)	1.01(.44)	0.85(.27)
0.2	50.00	72.3	19.4	8.3	1.22(.61)	1.89(.29)	1.12(.55)	1.89(.26)
	5.00	48.6	24.7	26.7	1.17(.61)	1.77(.33)	1.07(.54)	1.86(.31)
	0.25	18.9	23.7	57.4	0.98(.71)	1.39(.38)	1.04(.70)	1.77(.50)

Table 5.5.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs and with a_i having a uniforn (-4.5,4.5) distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	65.2	7.5	27.3	1.04(.16)	0.97(.08)	1.02(.16)	0.98(.08)
	5.00	52.2	8.1	39.7	1.03(.18)	0.96(.09)	1.01(.17)	0.98(.09)
	0.25	34.9	8.2	56.9	1.00(.21)	0.92(.11)	1.00(.21)	0.96(.11)
0.5	50.00	80.5	11.8	7.7	1.08(.16)	0.97(.08)	1.04(.15)	0.98(.07)
	5.00	54.7	15.5	29.8	1.06(.17)	0.94(.10)	1.01(.16)	0.97(.09)
	0.25	21.9	15.0	63.1	0.94(.21)	0.83(.12)	0.97(.23)	0.95(.13)
0.2	50.00	74.8	18.5	7.7	1.16(.32)	1.93(.16)	1.06(.28)	1.97(.14)
	5.00	49.3	25.0	25.7	1.13(.36)	1.83(.17)	1.02(.30)	1.94(.17)
	0.25	18.9	23.8	57.3	0.99(.39)	1.52(.23)	1.02(.38)	1.92(.26)

Table 5.6.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs and with a_i having a uniform (-4.5,4.5) distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	65.4	7.6	27.0	1.03(.09)	0.97(.05)	1.00(.08)	0.99(.05)
	5.00	52.6	7.8	39.6	1.02(.11)	0.96(.05)	1.00(.10)	0.98(.05)
	0.25	35.6	7.8	56.6	0.98(.11)	0.94(.07)	0.97(.11)	0.98(.06)
0.5	50.00	80.0	12.1	7.9	1.04(.09)	0.98(.04)	1.01(.08)	1.00(.04)
	5.00	54.9	15.5	29.6	1.04(.10)	0.95(.05)	0.99(.09)	0.98(.05)
	0.25	22.1	15.0	62.9	0.92(.13)	0.86(.08)	0.95(.14)	0.97(.07)
0.2	50.00	74.1	18.0	7.9	1.09(.17)	1.94(.08)	1.00(.16)	1.98(.08)
	5.00	49.2	25.4	25.4	1.11(.20)	1.84(.11)	1.00(.18)	1.97(.10)
	0.25	19.3	23.5	57.2	0.95(.22)	1.59(.14)	0.96(.22)	1.94(.15)

Table 5.7.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs and with a_i having a uniform (-4.5,4.5) distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	94.1	4.3	1.6	1.06(.27)	0.97(.13)	1.05(.27)	0.96(.13)
	5.00	70.2	18.1	11.7	1.12(.32)	0.95(.14)	1.07(.32)	0.97(.14)
	0.25	27.1	17.9	55.0	0.95(.40)	0.83(.21)	1.04(.43)	0.92(.22)
0.5	50.00	96.1	2.9	1.0	1.06(.27)	0.96(.13)	1.05(.27)	0.96(.13)
	5.00	72.3	19.7	8.0	1.12(.33)	0.95(.15)	1.06(.30)	0.96(.14)
	0.25	12.6	23.1	64.3	0.81(.40)	0.71(.23)	1.01(.50)	0.92(.31)
0.2	50.00	88.9	9.8	1.3	1.15(.57)	1.96(.27)	1.10(.54)	1.91(.25)
	5.00	59.6	30.9	9.5	1.27(.68)	1.90(.32)	1.10(.58)	1.91(.27)
	0.25	13.8	33.8	52.4	0.97(.73)	1.40(.41)	1.12(.71)	1.86(.46)

Table 5.8.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs and with ln a_i having a unit exponential distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	93.7	4.7	1.6	1.04(.14)	0.96(.07)	1.03(.14)	0.99(.07)
	5.00	70.5	17.1	12.4	1.09(.17)	0.95(.08)	1.04(.16)	0.99(.07)
	0.25	27.7	18.5	53.8	0.97(.18)	0.87(.10)	1.04(.21)	0.98(.11)
0.5	50.00	96.6	2.7	0.7	1.03(.14)	0.99(.07)	1.02(.14)	0.98(.07)
	5.00	82.3	9.5	8.2	1.11(.17)	0.97(.08)	1.03(.15)	0.99(.07)
	0.25	12.1	23.7	64.2	0.85(.19)	0.73(.12)	1.07(.24)	0.98(.14)
0.2	50.00	89.4	9.5	1.1	1.09(.28)	1.98(.15)	1.05(.27)	1.96(.13)
	5.00	60.6	30.5	9.9	1.21(.39)	1.89(.18)	1.05(.31)	1.98(.15)
	0.25	13.4	34.2	52.4	1.00(.35)	1.44(.19)	1.09(.38)	1.98(.22)

Table 5.9.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs and with ln a_i having a unit exponential distribution.

		Average %			Wolynetz method		New EM method	
ρ	T	Un-cens	Singly cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	50.00	93.8	4.5	1.7	1.02(.08)	0.99(.04)	1.00(.08)	1.00(.04)
	5.00	70.1	17.3	12.6	1.07(.09)	0.97(.04)	1.02(.08)	1.00(.05)
	0.25	26.8	18.2	55.0	0.94(.12)	0.88(.07)	1.02(.13)	1.00(.06)
0.5	50.00	96.4	2.8	0.8	1.01(.08)	1.00(.04)	1.00(.08)	1.00(.04)
	5.00	72.3	19.4	8.3	1.08(.08)	0.97(.05)	1.00(.08)	1.00(.05)
	0.25	12.1	22.6	65.3	0.82(.12)	0.75(.07)	1.00(.16)	1.00(.08)
0.2	50.00	89.0	9.7	1.3	1.06(.17)	1.99(.08)	1.00(.15)	2.00(.08)
	5.00	58.0	31.0	10.0	1.16(.19)	1.90(.10)	1.00(.16)	2.01(.09)
	0.25	13.0	34.3	52.7	0.93(.22)	1.46(.11)	1.02(.23)	2.02(.14)

Table 5.10.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs and with ln a_{ι} having a unit exponential distribution.

ρ	T	Average % Un-cens	Average % Singly cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	87.0	7.5	5.5	45	97	95	47	89	98
	5.00	77.2	20.6	22.2	57	98	96	63	94	97
	0.25	23.6	17.4	69.0	55	96	87	65	84	84
0.5	50.00	93.7	4.3	2.0	64	97	96	57	90	99
	5.00	62.2	23.8	14.0	50	98	94	56	90	97
	0.25	5.9	19.2	74.9	60	93	69	78	83	68
0.2	50.00	84.0	13.5	2.5	62	98	96	49	92	99
	5.00	47.9	36.1	16.0	64	95	95	64	92	97
	0.25	7.9	30.8	61.3	58	96	88	80	87	72

Table 5.11.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution and the number of pairs equal to 30 .

ρ	T	Un-cens	Average % Singly cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ .
0.8	50.00	87.1	8.3	4.6	61	95	92	51	94	100
	5.00	58.4	20.9	21.7	57	95	91	65	95	100
	0.25	23.5	69.0	17.5	67	98	86	82	90	88
0.5	50.00	93.9	4.8	1.3	59	95	94	56	95	100
	5.00	62.0	25.1	12.9	64	96	87	56	96	100
	0.25	17.2	19.4	73.4	76	96	47	94	86	39
0.2	50.00	84.2	13.8	2.0	64	95	92	58	94	100
	5.00	50.9	35.9	14.2	69	93	88	66	97	100
	0.25	16.5	31.5	62.0	57	99	84	96	96	41

Table 5.12.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution and the number of pairs equal to 100 .

ρ	T	Un-cens	Average % Singly cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leqslant \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leqslant \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	97.6	8.1	4.3	66	96	84	52	92	100
	5.00	67.3	21.3	21.4	64	92	82	68	98	100
	0.25	13.6	17.5	68.9	84	98	66	98	98	70
0.5	50.00	94.5	4.2	1.3	64	98	90	58	94	100
	5.00	62.1	25.2	12.7	74	96	62	62	96	98
	0.25	4.3	19.1	76.6	88	90	18	100	98	2
0.2	50.00	84.8	13.3	1.9	66	92	82	60	94	100
	5.00	50.0	36.2	13.8	74	96	66	70	98	100
	0.25	6.7	31.1	62.2	80	96	72	100	98	2

Table 5.13.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leqslant \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leqslant \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution and the number of pairs equal to 300 .

ρ	T	Average % Un-cens	Average % Singly cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	64.1	7.1	28.8	52	93	92	48	89	98
	5.00	52.3	7.6	40.1	51	92	94	64	88	100
	0.25	34.4	8.0	57.6	49	94	91	60	79	93
0.5	50.00	90.4	1.4	8.2	65	94	90	47	90	99
	5.00	54.8	13.9	31.3	56	94	92	56	87	100
	0.25	21.6	15.1	63.3	47	89	91	67	78	86
0.2	50.00	72.3	19.4	8.3	71	92	89	47	90	99
	5.00	48.6	24.7	26.7	63	96	90	66	90	97
	0.25	18.9	23.7	57.4	50	94	89	83	80	79

Table 5.14.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution and the number of pairs equal to 30 .

ρ	T	Un-cens	Average % Singly cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	65.2	7.5	27.3	52	96	97	52	90	100
	5.00	52.2	8.1	39.7	53	96	97	60	92	100
	0.25	34.9	8.2	56.9	48	94	95	67	90	97
0.5	50.00	80.5	11.8	7.7	65	94	92	62	94	100
	5.00	54.7	15.5	29.8	63	96	94	70	93	97
	0.25	21.9	15.0	63.1	47	94	91	80	89	89
0.2	50.00	74.8	18.5	7.7	70	96	90	64	95	100
	5.00	49.3	25.0	25.7	62	96	88	74	90	99
	0.25	18.9	23.8	57.3	52	96	92	89	89	81

Table 5.15.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution and the number of pairs equal to 100 .

ρ	T	Average % Un-cens	Average % Singly cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	65.4	7.6	27.0	60	98	96	68	92	100
	5.00	52.6	7.8	39.6	62	94	92	56	90	98
	0.25	35.6	7.8	56.6	64	96	96	80	94	100
0.5	50.00	80.0	12.1	7.9	66	96	94	58	96	100
	5.00	54.9	15.5	29.6	58	96	96	64	94	100
	0.25	22.1	15.0	62.9	60	92	90	90	90	88
0.2	50.00	74.1	18.0	7.9	68	96	90	68	100	100
	5.00	49.2	25.4	25.4	62	98	88	88	94	96
	0.25	19.3	23.5	57.2	58	94	96	94	92	60

Table 5.16.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution and the number of pairs equal to 300 .

ρ	T	Un-cens	Average % Singly cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	94.1	4.3	1.6	67	94	92	56	91	100
	5.00	70.2	18.1	11.7	53	90	90	59	93	99
	0.25	27.1	17.9	55.0	42	89	85	58	85	90
0.5	50.00	96.1	2.9	1.0	73	93	92	66	90	100
	5.00	72.3	19.7	8.0	64	91	87	55	92	99
	0.25	12.6	23.1	64.3	46	89	84	73	86	81
0.2	50.00	88.9	9.8	1.3	62	93	92	39	89	100
	5.00	59.6	30.9	9.5	68	91	87	60	94	99
	0.25	13.8	33.8	52.4	47	93	86	80	84	80

Table 5.17.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with $\ln a_i$ having a unit exponential distribution and the number of pairs equal to 30 .

ρ	T	Un-cens	Average % Singly cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_w - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_w - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	93.7	4.7	1.6	58	96	92	53	93	100
	5.00	70.5	17.1	12.4	66	95	88	68	98	100
	0.25	27.7	18.5	53.8	42	97	97	79	95	96
0.5	50.00	96.6	2.7	0.7	59	98	97	49	93	100
	5.00	82.3	9.5	8.2	65	96	81	72	94	100
	0.25	12.1	23.7	64.2	57	99	85	87	95	75
0.2	50.00	89.4	9.5	1.1	61	97	95	50	95	99
	5.00	60.6	30.5	9.9	68	95	82	64	95	100
	0.25	13.4	34.2	52.4	37	95	95	96	96	70

Table 5.18.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with $\ln a_i$ having a unit exponential distribution and the number of pairs equal to 100 .

ρ	T	Average % Un-cens	Average % Singly cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leqslant \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leqslant \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	50.00	93.8	4.5	1.7	58	98	96	48	98	100
	5.00	70.1	17.3	12.6	64	98	90	50	94	94
	0.25	26.8	18.2	55.0	50	96	88	86	96	92
0.5	50.00	96.4	2.8	0.8	54	98	98	54	94	100
	5.00	72.3	19.4	8.3	66	98	86	60	96	100
	0.25	12.1	22.6	65.3	70	94	64	96	94	46
0.2	50.00	89.0	9.7	1.3	56	98	94	48	94	100
	5.00	58.0	31.0	10.0	70	98	84	70	96	100
	0.25	13.0	34.3	52.7	52	94	94	100	98	24

Table 5.19.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leqslant \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leqslant \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with $\ln a_i$ having a unit exponential distribution and the number of pairs equal to 300 .

ρ	Average % Un-cens	Singly cens	Min cens	Doubly cens	Wolynetz method $\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	New EM method $\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	23.2	12.5	5.8	58.5	0.97(.37)	0.82(.23)	0.97(.37)	0.99(.24)
	13.4	6.4	9.0	71.2	1.00(.42)	0.75(.26)	0.92(.50)	1.02(.36)
	17.7	18.0	14.8	49.5	0.68(.37)	0.88(.24)	0.97(.45)	0.99(.32)
0.5	12.1	15.1	8.8	64.0	0.93(.34)	0.70(.23)	0.94(.39)	1.01(.32)
	4.9	5.5	10.0	79.6	1.12(.43)	0.57(.26)	0.94(.61)	0.95(.38)
	8.0	19.3	23.3	49.4	0.46(.38)	0.73(.23)	1.04(.73)	0.95(.44)
0.2	13.7	29.2	8.5	48.6	0.91(.66)	1.44(.43)	0.95(.69)	1.99(.51)
	7.6	18.8	11.3	62.3	0.99(.63)	1.20(.42)	0.97(.76)	1.91(.56)
	11.8	31.1	14.0	43.1	0.79(.64)	1.41(.43)	1.04(.80)	1.94(.48)

Table 5.20.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs , with a_i having a normal distribution and with a random censoring mechanism.

	Average %				Wolynetz method		New EM method	
ρ	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	23.9	11.9	6.1	58.1	0.99(.23)	0.84(.12)	0.99(.23)	0.99(.13)
	13.8	6.5	9.5	70.2	1.07(.25)	0.77(.14)	0.99(.29)	1.00(.17)
	17.7	18.1	15.0	49.2	0.71(.21)	0.88(.10)	1.01(.27)	1.00(.15)
0.5	11.7	15.5	9.2	63.6	1.00(.21)	0.71(.11)	0.99(.23)	1.00(.17)
	4.0	5.3	9.9	80.8	1.16(.22)	0.58(.14)	0.99(.33)	1.02(.45)
	7.2	20.0	24.0	48.8	0.47(.19)	0.74(.11)	1.02(.32)	1.00(.19)
0.2	14.0	29.1	8.7	48.2	1.00(.39)	1.42(.20)	1.00(.39)	1.97(.28)
	6.8	18.3	12.1	62.8	1.10(.38)	1.24(.20)	1.00(.44)	1.98(.33)
	11.2	31.9	14.0	42.9	0.85(.35)	1.43(.21)	1.05(.45)	2.00(.26)

Table 5.21.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs , with a_i having a normal distribution and with a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	Wolynetz method $\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	New EM method $\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	23.8	12.4	6.3	57.5	0.99(.14)	0.84(.06)	1.00(.14)	1.00(.06)
	14.0	6.6	9.5	69.9	1.07(.14)	0.78(.07)	1.01(.15)	1.00(.09)
	17.9	18.3	14.8	49.0	0.72(.13)	0.89(.06)	1.00(.16)	1.01(.07)
0.5	11.5	16.3	9.2	63.0	1.01(.13)	0.71(.06)	1.01(.15)	1.01(.09)
	3.9	5.6	10.1	80.4	1.17(.12)	0.59(.07)	1.02(.17)	1.00(.14)
	7.2	20.6	23.8	48.4	0.49(.12)	0.74(.06)	1.00(.17)	1.00(.08)
0.2	13.8	30.1	8.6	47.5	1.02(.24)	1.47(.12)	1.03(.24)	2.00(.13)
	6.8	18.7	12.0	62.5	1.09(.24)	1.27(.11)	1.01(.25)	2.01(.18)
	11.6	32.3	13.7	42.4	0.86(.25)	1.45(.12)	1.04(.27)	2.01(.13)

Table 5.22.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs , with a_i having a normal distribution and with a random censoring mechanism.

		Average %			Wolynetz method		New EM method	
ρ	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	42.8	5.1	1.8	50.3	1.01(.38)	0.91(.19)	1.01(.37)	0.94(.19)
	38.2	3.0	3.9	54.9	1.01(.40)	0.89(.19)	1.01(.38)	0.92(.18)
	40.2	7.7	4.4	47.7	0.92(.35)	0.94(.18)	0.95(.38)	0.94(.19)
0.5	33.7	8.9	4.2	53.2	1.02(.40)	0.87(.19)	1.02(.37)	0.93(.19)
	24.7	6.4	7.4	61.4	1.03(.45)	0.81(.20)	1.04(.42)	0.92(.22)
	27.6	15.0	9.2	48.2	0.85(.35)	0.90(.17)	0.97(.43)	0.96(.21)
0.2	27.7	19.5	5.0	47.7	1.00(.78)	1.66(.43)	0.99(.69)	1.89(.41)
	19.8	16.4	7.7	56.1	1.02(.76)	1.53(.42)	1.01(.71)	1.90(.50)
	25.0	22.2	7.4	45.4	0.96(.75)	1.61(.39)	0.95(.79)	1.92(.40)

Table 5.23.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs , with a_i having a uniform (-4.5,4.5) distribution and with a random censoring mechanism.

	Average %				Wolynetz method		New EM method	
ρ	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	41.6	4.4	2.2	51.8	1.00(.21)	0.95(.10)	1.00(.20)	0.98(.09)
	37.1	3.2	3.6	56.1	1.00(.21)	0.94(.10)	1.01(.20)	0.97(.11)
	38.7	7.3	4.8	49.3	0.91(.19)	0.98(.09)	0.96(.22)	0.99(.10)
0.5	32.6	8.7	4.1	54.6	1.00(.22)	0.90(.10)	1.00(.20)	0.97(.11)
	24.1	6.3	7.2	62.4	1.03(.22)	0.85(.12)	1.04(.21)	0.95(.14)
	27.5	13.8	9.7	49.0	0.82(.20)	0.93(.11)	0.93(.23)	0.99(.11)
0.2	27.3	19.7	4.8	48.2	0.99(.42)	1.67(.21)	1.00(.37)	1.95(.23)
	19.7	16.3	7.7	56.3	1.04(.40)	1.59(.21)	1.03(.38)	1.95(.25)
	25.1	21.9	7.7	45.3	0.93(.41)	1.67(.20)	0.92(.41)	1.96(.22)

Table 5.24.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs , with a_i having a uniform (-4.5,4.5) distribution and with a random censoring mechanism.

	Average %				Wolynetz method		New EM method	
ρ	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	41.3	4.1	2.3	52.3	0.99(.13)	0.97(.05)	0.99(.12)	0.99(.05)
	36.7	3.0	3.7	56.6	0.99(.13)	0.97(.05)	1.01(.12)	1.00(.05)
	38.5	6.9	5.0	49.6	0.92(.12)	1.00(.05)	0.95(.13)	1.01(.05)
0.5	32.2	8.8	4.0	55.0	1.00(.13)	0.92(.05)	1.00(.12)	0.99(.05)
	23.5	6.4	7.4	62.7	1.03(.15)	0.87(.05)	1.03(.14)	0.99(.06)
	27.4	13.7	9.8	49.1	0.82(.13)	0.95(.06)	0.92(.14)	1.01(.06)
0.2	27.0	19.9	4.8	48.3	1.00(.26)	1.71(.09)	1.00(.23)	1.98(.09)
	19.5	16.4	7.7	56.4	1.02(.27)	1.60(.10)	1.02(.23)	1.98(.10)
	24.7	22.1	7.6	45.6	0.95(.24)	1.69(.11)	0.93(.25)	1.99(.11)

Table 5.25.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs , with a_i having a uniform distribution and with a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	Wolynetz method $\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	New EM method $\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	37.4	11.7	5.2	45.7	0.98(.41)	0.87(.20)	1.00(.40)	0.99(.22)
	28.1	6.5	9.0	56.4	1.01(.43)	0.84(.22)	0.97(.43)	0.96(.24)
	32.6	16.5	14.0	36.9	0.87(.34)	0.92(.18)	1.03(.45)	0.92(.21)
0.5	24.2	16.6	8.6	50.6	0.97(.39)	0.79(.19)	1.01(.40)	0.96(.22)
	12.2	7.4	11.7	68.7	1.03(.45)	0.69(.24)	0.89(.67)	0.99(.49)
	17.6	23.2	21.9	37.3	0.75(.34)	0.84(.20)	1.04(.64)	0.90(.24)
0.2	21.8	31.6	7.4	39.2	1.01(.74)	1.57(.40)	1.07(.74)	1.96(.38)
	13.4	21.2	12.3	53.1	0.98(.77)	1.39(.39)	0.99(.77)	1.92(.44)
	20.8	32.6	12.4	34.2	0.99(.69)	1.56(.41)	1.03(.74)	1.88(.37)

Table 5.26.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 30 pairs , with ln a_i having a unit exponential distribution and with a random censoring mechanism.

ρ	Average % Un-cens	Singly cens	Min cens	Doubly cens	Wolynetz method $\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	New EM method $\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	38.0	11.8	5.6	44.6	1.01(.20)	0.91(.10)	1.01(.19)	1.01(.10)
	27.6	7.3	9.3	55.8	1.01(.22)	0.87(.11)	0.99(.21)	1.01(.11)
	31.7	18.1	14.0	36.2	0.91(.20)	0.95(.10)	1.06(.20)	0.98(.11)
0.5	23.5	17.3	8.9	50.3	1.01(.20)	0.82(.10)	1.03(.18)	1.00(.11)
	12.3	7.8	11.8	68.1	1.08(.22)	0.72(.10)	0.98(.25)	0.99(.15)
	16.7	24.0	22.0	37.3	0.79(.22)	0.86(.11)	1.04(.22)	0.96(.12)
0.2	22.0	31.3	8.0	38.7	1.05(.39)	1.60(.19)	1.06(.34)	2.01(.21)
	13.0	21.9	12.0	53.1	1.04(.39)	1.43(.19)	1.02(.37)	2.02(.27)
	19.5	33.8	12.6	34.1	1.08(.39)	1.58(.19)	1.07(.34)	1.98(.22)

Table 5.27.

Means and standard deviations over 100 simulations of estimates of β and τ by the Wolynetz and new EM method for 100 pairs , with ln a_i having a unit exponential distribution and with a random censoring mechanism.

	Average %				Wolynetz method		New EM method	
ρ	Un-cens	Singly cens	Min cens	Doubly cens	$\hat{\beta}_W$ (S.D)	$\hat{\tau}_W$ (S.D)	$\hat{\beta}_E$ (S.D)	$\hat{\tau}_E$ (S.D)
0.8	38.1	11.8	6.0	44.1	1.02(.11)	0.92(.06)	1.03(.10)	1.03(.06)
	27.5	7.3	9.7	55.5	1.04(.11)	0.89(.05)	1.00(.10)	1.03(.06)
	31.7	18.2	14.1	36.0	0.94(.11)	0.97(.06)	1.09(.11)	1.00(.06)
0.5	23.1	17.7	9.5	49.7	1.03(.11)	0.83(.07)	1.03(.10)	1.03(.07)
	12.1	8.3	11.7	67.9	1.11(.11)	0.75(.07)	1.03(.13)	1.02(.09)
	16.5	24.3	22.3	36.9	0.83(.13)	0.89(.07)	1.06(.12)	0.99(.08)
0.2	22.1	31.7	8.2	37.9	1.08(.21)	1.66(.14)	1.07(.19)	2.04(.12)
	13.0	22.1	12.5	52.4	1.08(.20)	1.48(.11)	1.03(.18)	2.05(.14)
	19.1	34.8	12.5	33.6	1.14(.22)	1.63(.14)	1.10(.20)	2.03(.14)

Table 5.28.

Means and standard deviations over 50 simulations of estimates of β and τ by the Wolynetz and new EM method for 300 pairs , with ln a_i having a unit exponential distribution and with a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	23.3	12.5	5.8	58.5	51	98	95	67	89	89
	13.4	6.4	9.0	71.2	37	96	92	66	90	94
	17.7	18.0	14.8	49.5	59	96	75	59	83	93
0.5	12.1	15.1	8.8	64.0	49	98	94	71	88	81
	4.9	5.5	10.0	79.6	44	88	86	76	81	69
	8.0	19.3	23.3	49.4	76	85	49	58	78	86
0.2	13.7	29.2	8.5	48.6	50	95	92	75	89	82
	7.6	18.8	11.3	62.3	48	92	90	85	83	65
	11.8	31.1	14.0	43.1	54	94	87	85	85	80

Table 5.29.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution, the number of pairs equal to 30 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat\beta_E - 1\rvert \leqslant \lvert\hat\beta_W - 1\rvert$	% of new EM C.I. for $\hat\beta_E$ containing β	% of Wolynetz C.I. for $\hat\beta_W$ containing β	% with $\lvert\hat\tau_E - \tau\rvert \leqslant \lvert\hat\tau_W - \tau\rvert$	% of new EM C.I. for $\hat\tau_E$ containing τ	% of Wolynetz C.I. for $\hat\tau_W$ containing τ
0.8	23.9	11.9	6.1	58.1	51	90	89	77	93	92
	13.8	6.5	9.5	70.2	48	90	89	82	94	82
	17.7	18.1	15.0	49.2	73	91	68	63	89	98
0.5	11.7	15.5	9.2	63.6	43	90	90	86	92	66
	4.0	5.3	9.9	80.8	57	93	86	87	87	46
	7.2	20.0	24.0	48.8	88	91	19	79	91	79
0.2	14.0	29.1	8.7	48.2	44	91	81	91	92	63
	6.8	18.3	12.1	62.8	44	91	89	94	93	42
	11.2	31.9	14.0	42.9	41	92	86	89	97	66

Table 5.30.

Percentage of simulations with $\lvert\hat\beta_E - 1\rvert \leqslant \lvert\hat\beta_W - 1\rvert$, $\lvert\hat\tau_E - \tau\rvert \leqslant \lvert\hat\tau_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution, the number of pairs equal to 100 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	23.8	12.4	6.3	57.5	50	94	84	94	96	82
	14.0	6.6	9.5	69.9	54	98	80	92	94	58
	17.9	18.3	14.8	49.0	78	94	34	82	98	94
0.5	11.5	16.3	9.2	63.0	44	92	88	98	94	18
	3.9	5.6	10.1	80.4	74	96	68	96	92	4
	7.2	20.6	23.8	48.4	100	94	0	96	96	28
0.2	13.8	30.1	8.6	47.5	60	90	88	96	94	20
	6.8	18.7	12.0	62.5	52	96	80	98	92	2
	11.6	32.3	13.7	42.4	60	88	80	98	86	20

Table 5.31.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a normal distribution, the number of pairs equal to 300 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	42.8	5.1	1.8	50.3	65	92	90	59	87	96
	38.2	3.0	3.9	54.9	62	93	89	59	84	96
	40.2	7.7	4.4	47.7	49	89	95	45	84	99
0.5	33.7	8.9	4.2	53.2	60	92	93	63	83	95
	24.7	6.4	7.4	61.4	58	92	87	71	81	93
	27.6	15.0	9.2	48.2	45	92	90	52	92	99
0.2	27.7	19.5	5.0	47.7	68	94	85	71	84	91
	19.8	16.4	7.7	56.1	60	95	91	72	87	87
	25.0	22.2	7.4	45.4	46	91	88	71	88	89

Table 5.32.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution, the number of pairs equal to 30 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Average % Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	41.6	4.4	2.2	51.8	52	91	89	60	94	100
	37.1	3.2	3.6	56.1	50	96	95	58	94	99
	38.7	7.3	4.8	49.3	43	90	90	43	96	100
0.5	32.6	8.7	4.1	54.6	60	91	90	76	89	97
	24.1	6.3	7.2	62.4	59	95	92	76	85	93
	27.5	13.8	9.7	49.0	65	92	85	65	94	98
0.2	27.3	19.7	4.8	48.2	59	93	88	81	91	92
	19.7	16.3	7.7	56.3	54	94	89	85	89	86
	25.1	21.9	7.7	45.3	51	92	90	82	93	93

Table 5.33.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution, the number of pairs equal to 100 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Average % Min cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	41.3	4.1	2.3	52.3	62	92	92	60	100	100
	36.7	3.0	3.7	56.6	56	92	94	62	100	100
	38.5	6.9	5.0	49.6	58	88	88	54	100	100
0.5	32.2	8.8	4.0	55.0	62	94	94	82	100	100
	23.5	6.4	7.4	62.7	64	90	88	90	100	98
	27.4	13.7	9.8	49.1	76	88	64	62	100	100
0.2	27.0	19.9	4.8	48.3	60	94	92	98	100	88
	19.5	16.4	7.7	56.4	68	94	82	92	100	58
	24.7	22.1	7.6	45.6	48	92	86	88	100	84

Table 5.34.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with a_i having a uniform (-4.5,4.5) distribution, the number of pairs equal to 300 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	37.4	11.7	5.2	45.7	62	89	84	59	89	96
	28.1	6.5	9.0	56.4	48	91	91	66	88	94
	32.6	16.5	14.0	36.9	44	93	90	52	85	97
0.5	24.2	16.6	8.6	50.6	55	92	87	69	89	90
	12.2	7.4	11.7	68.7	47	89	86	74	84	80
	17.6	23.2	21.9	37.3	66	89	80	53	88	93
0.2	21.8	31.6	7.4	39.2	60	92	83	81	95	90
	13.4	21.2	12.3	53.1	56	91	82	81	91	86
	20.8	32.6	12.4	34.2	52	93	86	71	88	85

Table 5.35.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with ln a_i having a unit exponential distribution, the number of pairs equal to 30 and a random censoring mechanism.

ρ	Un-cens	Average % Singly cens	Min cens	Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	38.0	11.8	5.6	44.6	56	97	93	67	96	99
	27.6	7.3	9.3	55.8	58	98	92	68	98	97
	31.7	18.1	14.0	36.2	51	97	92	56	95	100
0.5	23.5	17.3	8.9	50.3	56	97	91	80	99	92
	12.3	7.8	11.8	68.1	48	97	92	89	97	82
	16.7	24.0	22.0	37.3	67	95	74	72	94	98
0.2	2.20	31.3	8.0	38.7	62	96	88	87	98	86
	13.0	21.9	12.0	53.1	53	97	88	92	94	74
	19.5	33.8	12.6	34.1	59	95	91	88	97	84

Table 5.36.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with ln a_i having a unit exponential distribution, the number of pairs equal to 100 and a random censoring mechanism.

ρ	Average % Un-cens	Average % Singly cens	Average % Min cens	Average % Doubly cens	% with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$	% of new EM C.I. for $\hat{\beta}_E$ containing β	% of Wolynetz C.I. for $\hat{\beta}_W$ containing β	% with $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$	% of new EM C.I. for $\hat{\tau}_E$ containing τ	% of Wolynetz C.I. for $\hat{\tau}_W$ containing τ
0.8	38.1	11.8	6.0	44.1	56	98	94	72	96	98
	27.5	7.3	9.7	55.5	70	100	96	76	96	96
	31.7	18.2	14.1	36.0	44	92	90	56	92	100
0.5	23.1	17.7	9.5	49.7	54	94	94	90	96	76
	12.1	8.3	11.7	67.9	66	94	80	98	94	48
	16.5	24.3	22.3	36.9	70	98	64	82	88	92
0.2	22.1	31.7	8.2	37.9	62	96	90	94	98	70
	13.0	22.1	12.5	52.4	64	100	94	98	98	28
	19.1	34.8	12.5	33.6	62	96	88	90	92	70

Table 5.37.

Percentage of simulations with $\lvert\hat{\beta}_E - 1\rvert \leq \lvert\hat{\beta}_W - 1\rvert$, $\lvert\hat{\tau}_E - \tau\rvert \leq \lvert\hat{\tau}_W - \tau\rvert$ and the percentage of simulations in which the confidence intervals of the new EM and Wolynetz methods contained the true value of β and τ, with $\ln a_i$ having a unit exponential distribution, the number of pairs equal to 300 and a random censoring mechanism.

CHAPTER 6

ANALYSIS OF SOME DATA SETS

Three data sets will be analysed. Firstly the skin graft data of Batchelor and Hackett (1970); secondly the data of Mantel, Bohidar and Ciminera (1977) and lastly the data of Sampford and Taylor (1959).

6.1 Skin graft data.

The data are given in Table 6.1 and consists of the survival times in days of closely matched and poorly matched skin grafts applied to the same individual. In two cases the second graft was a technical failure and in three cases more than two grafts were performed on the same person. The amount of burn is the full thickness burn as a percentage of the body surface. An asterisk denotes a censored observation. Ignoring initially the amount of burn, one covariate was considered viz. the closeness of the match and this was coded 0 for a close match and 1 for a poor match. Hence in the proportional hazards model exp(β) measures the 'relative risk' of receiving a poor match as opposed to a close match.

Case Number	Amount of burn	Match	Survival Time
1	22	Poor	19
2	20	Close	24
3	23	Poor Poor	18 18
4	30	Poor Close	29 37
5	20	Close Poor	19 13
6	25	Poor Poor	19 19
7	25	Poor Close Close	15 57* 57*
8	45	Close Poor	93 26
9	20	Close Poor	16 11
10	18	Close Poor	21-23 15-18
11	35	Close Poor	20 26
12	25	Poor Close	19-23 18
13	50	Poor Close Close Close	43 77 63 29
14	35	Poor Poor	28* 28*
15	30	Close Poor Poor	29 15 18
16	30	Poor Close	38-42 60*

Table 6.1.
Skin graft data Batchelor and Hackett (1970).

Where a range of times existed as in case 12, the middle time was taken for the marginal and integrated methods. In the non-parametric method, since the second graft time was not contained in the interval it was possible to determine the orderings of the times of the grafts for these cases without taking the middle time.

The non-parametric method demands pairs discordant with respect to the covariate. Hence cases 1, 2, 3, 6 and 14 are discarded in the analysis by this method. The log likelihood for the non-parametric method, given by 2.24, becomes

$$\ln L_N(\beta) = 10\beta - 9\ln(1 + e^{\beta}) - 2\ln(2 + e^{\beta}) - \ln(1 + 2e^{\beta}) - \ln(3 + e^{\beta})$$

This was maximised using a NAG routine, giving a maximum likelihood estimate of β, $\hat{\beta}_N = 1.39$, with standard deviation from the observed information of 0.707.

As at least pairs of observations are necessary for analysis by the marginal likelihood method, cases 1 and 2 are omitted. Using 2.27, the marginal likelihood gave $\hat{\beta}_M = 1.63$ with standard deviation 0.598 and $\hat{\eta}_M = 3.66$ with standard deviation 0.712.

When the whole data set is used the estimates from the integrated likelihood become $\hat{\beta}_I = 1.81$ with standard deviation 0.568 and $\hat{\eta}_I = 3.95$ with standard deviation 0.832. Both the marginal and integrated methods reject the null hypothesis of no difference in the survival times for the poor and close match although the non-parametric method does not. Approximate 95% confidence intervals for β , for the non-parametric, marginal and integrated methods are (0.00, 2.78), (0.46, 2.80) and (0.70, 2.92) respectively.

Since the likelihood function for β may not be normal in shape, a more useful measure than the confidence intervals already given may be plausibility intervals for β . For the covariate match, already considered, let

$$R(\beta) = \frac{\max_{\underline{\theta}} L(\beta, \underline{\theta})}{L(\hat{\beta}, \hat{\underline{\theta}})}$$

where $\underline{\theta}$ is the parameter space of all other parameters in the model i.e. for the marginal likelihood $\underline{\theta} = (\eta)$ and for the integrated likelihood $\underline{\theta} = (\eta, a, b)$. $R(\beta)$ is the relative likelihood function and a $100(1 - \gamma)$% plausibility interval for β can be found such that $R(\beta) \geq \gamma$.

Graphs of the relative likelihoods of β for each of the three methods, non-parametric, marginal and integrated, are shown in Figures 6.1 to 6.3 and 95% plausibility intervals for β are (-0.10, 3.37), (0.42, 2.88) and (0.84, 2.78) respectively. These demonstrate the skewness of the likelihood for the non-parametric method. The plausibility intervals for the marginal and integrated methods do not differ dramatically from the 95% confidence intervals. Contours (0.75, 0.25, 0.01) of relative likelihood of β and η for the marginal and integrated methods are shown in Figures 6.4 and 6.5 respectively. The contours for the integrated likelihood are slices through the relative likelihood function parallel to the plane $a = \hat{a}$, $b = \hat{b}$.

Both the marginal and integrated methods reject the exponential model for the hazard rate. The estimates of the ratio of expected survival times of poor match over close match is $\exp(-\hat{\beta} / \hat{\eta})$ and these are 0.63 for the marginal method based on the 14 cases and 0.63 for the integrated method using the whole data. The corresponding approximate 95% confidence intervals using the delta method are (0.45, 0.83) and (0.47, 0.79).

Holt and Prentice, in their analysis of this data considered as a second covariate the interaction of match and amount of burn. Since amount of burn is the same within all n-tuplets this is implicitly allowed for in the pairing variables α_i. Using the two covariates match (coded as previously) and match x burn, the non-parametric method gives $\hat{\beta}_1$ (match) = 1.66 with standard deviation 1.946 and $\hat{\beta}_2$ (match x burn) = -0.02 with standard deviation 0.057. The results for the marginal method are $\hat{\beta}_1$ (match) = 1.04 with standard deviation 1.607, $\hat{\beta}_2$ (match x burn) = 0.02 with standard deviation 0.046 and $\hat{\eta}$ = 3.72 with standard deviation 0.739. The integrated method resulted in estimates $\hat{\beta}_1$ (match) = 2.77, S.D. = 1.736; $\hat{\beta}_2$ (match x burn) = -0.03, S.D. = 0.053 and $\hat{\eta}$ = 3.75, S.D. = 0.894. The Wilks likelihood ratio statistics ($-2\ln(L(\beta_1)/L(\beta_1, \beta_2))$

for the three methods were .00746, .00434 and 0.00073 (non-parametric, marginal and integrated methods respectively). Hence all three methods show no strong evidence of an interaction term in the model.

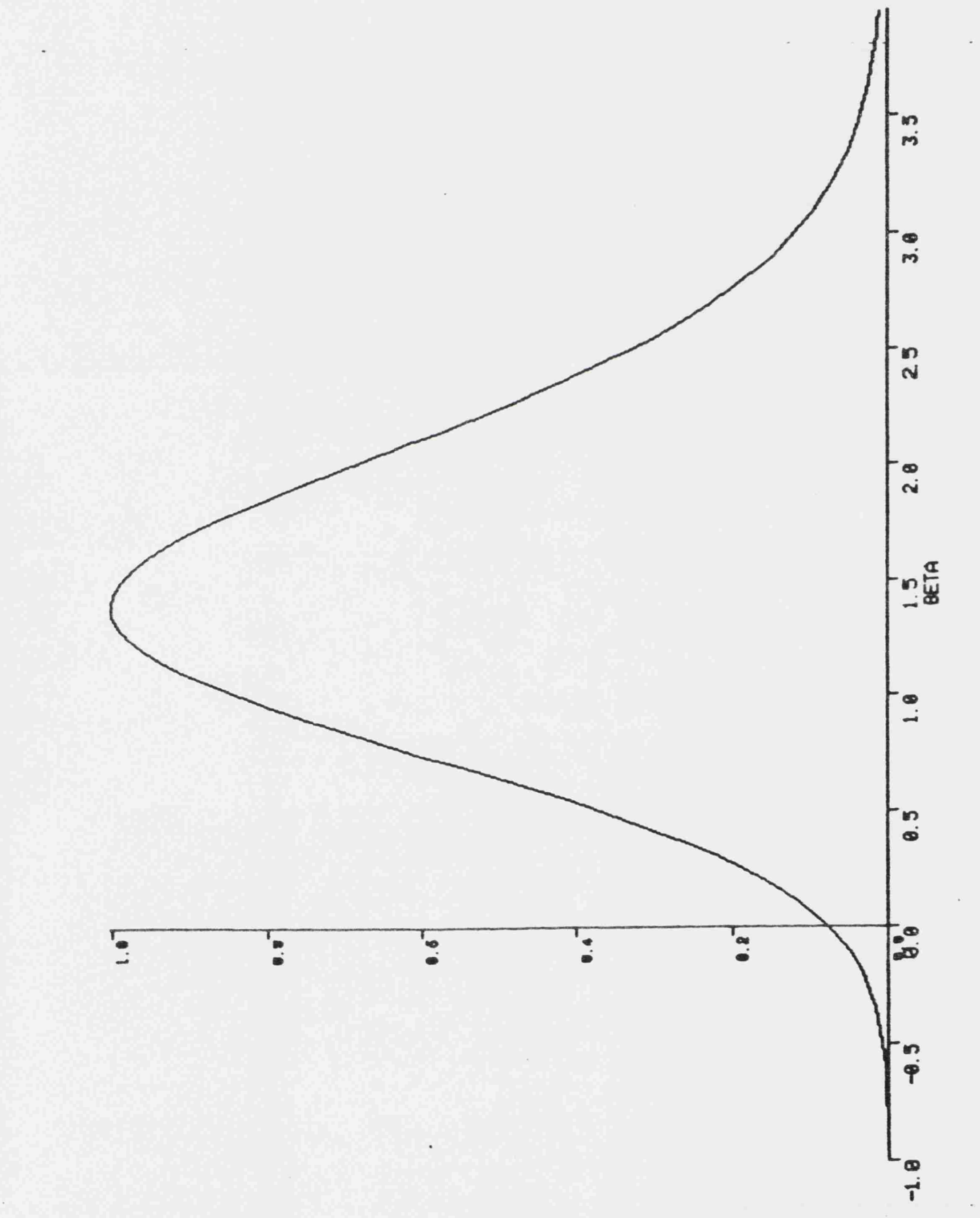

Figure 6.1.

Relative likelihood of β for the non-parametric method using the data of Batchelor and Hackett.

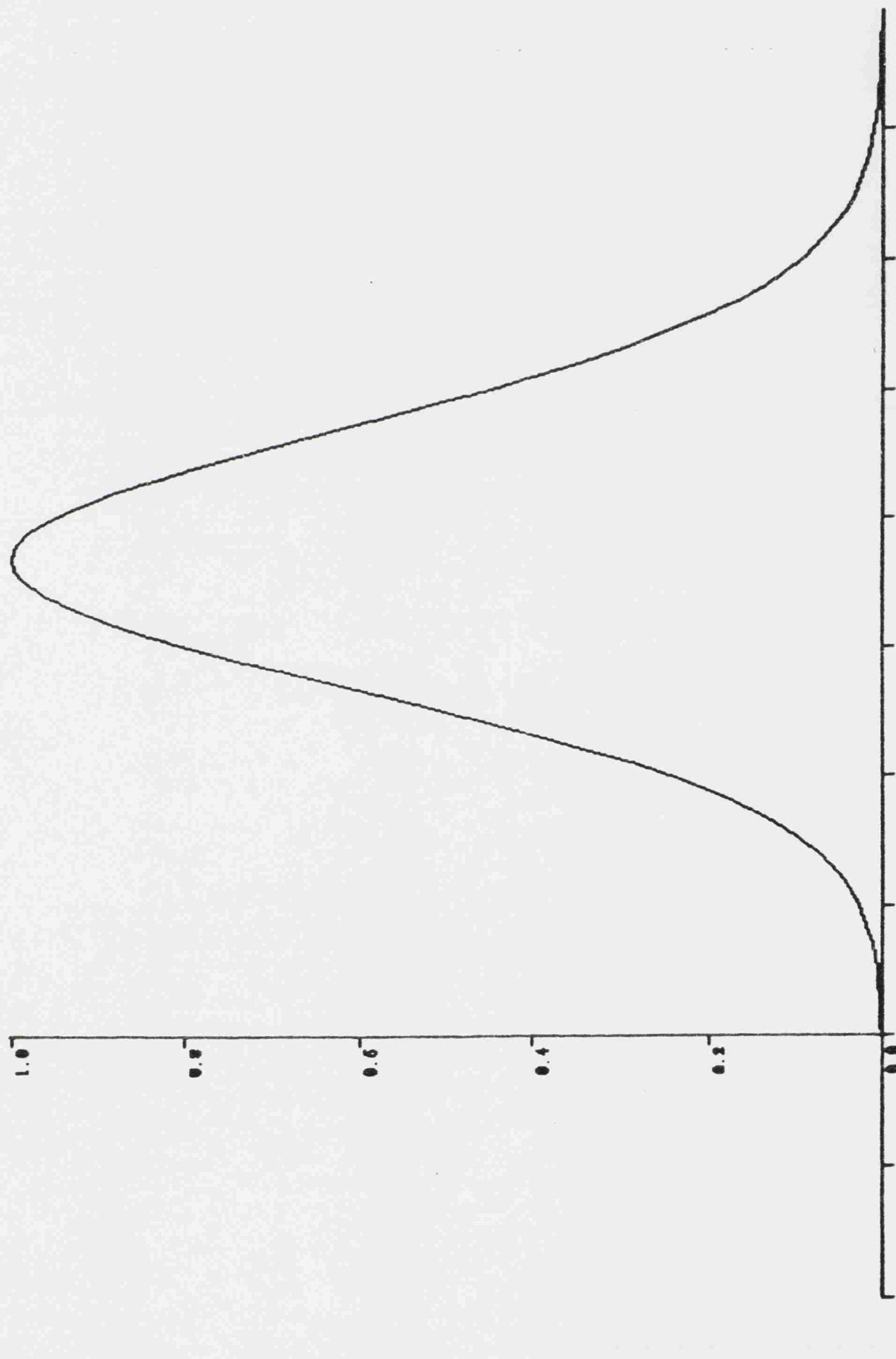

Figure 6.2.

Relative likelihood of β for the marginal method using the data of Batchelor and Hackett.

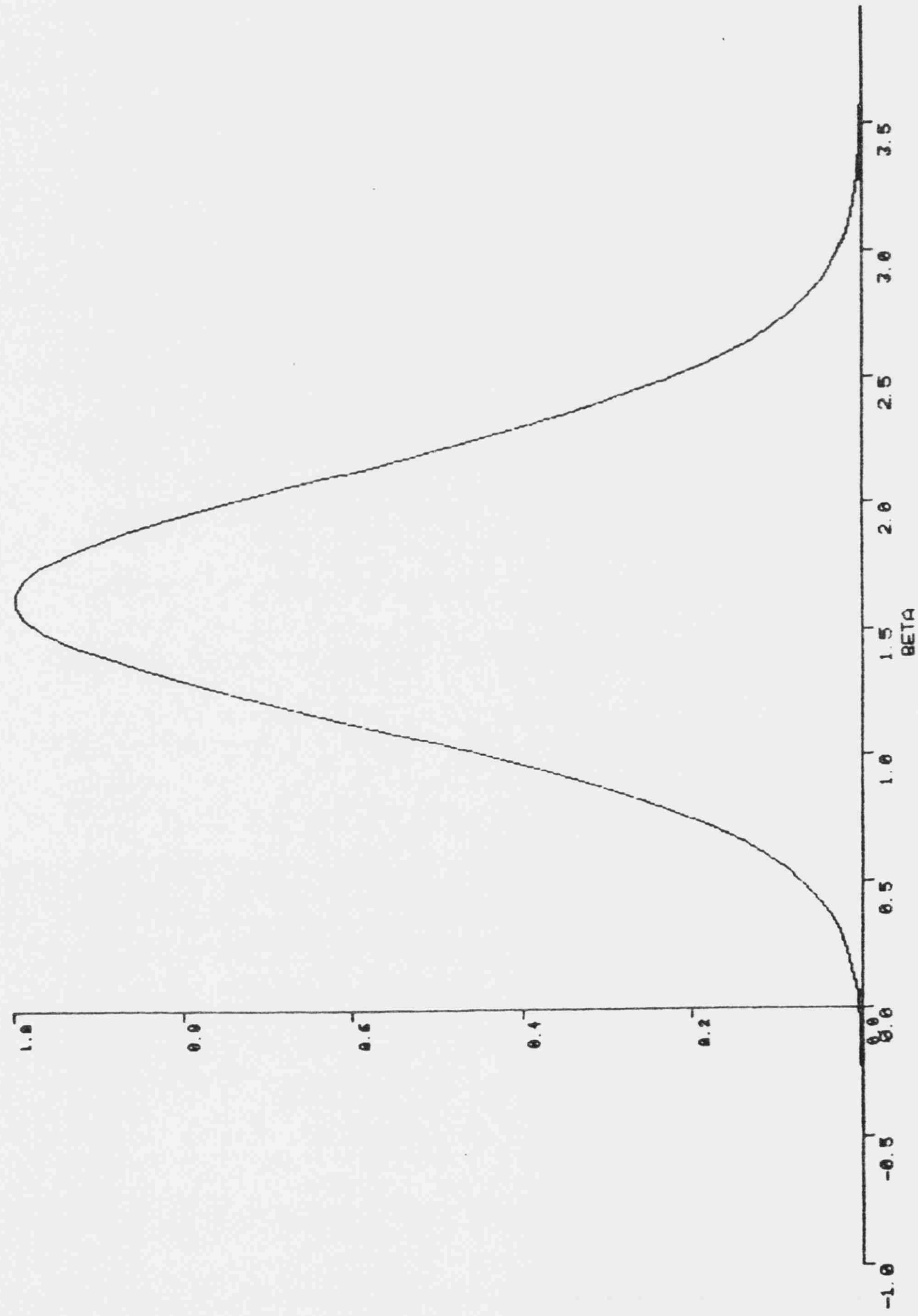

Figure 6.3.

Relative likelihood of β for the integrated method using the data of Batchelor and Hackett.

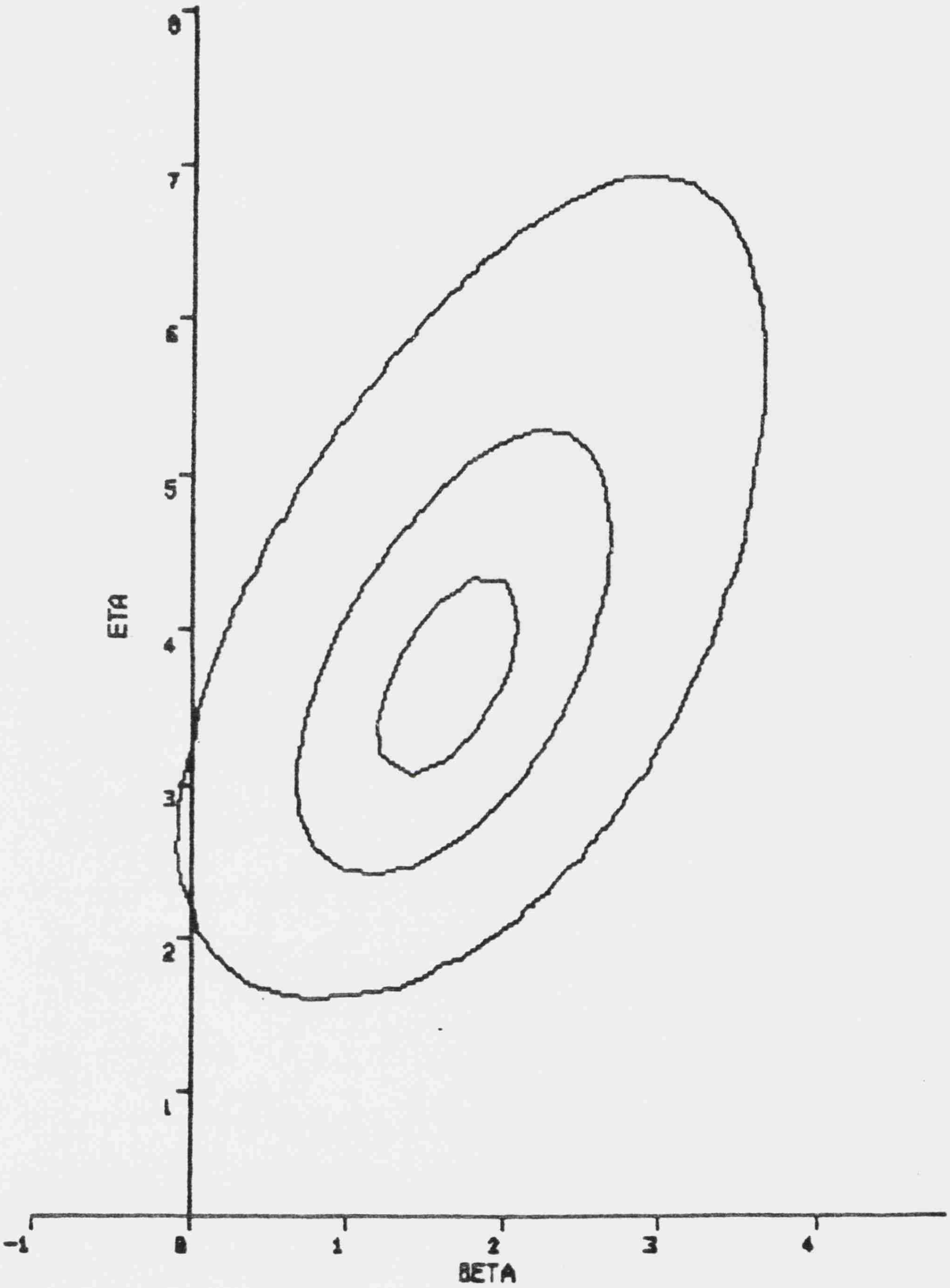

Figure 6.4.

Contours (0.75,0.25,0.01) of relative likelihood for the marginal method using the data of Batchelor and Hackett.

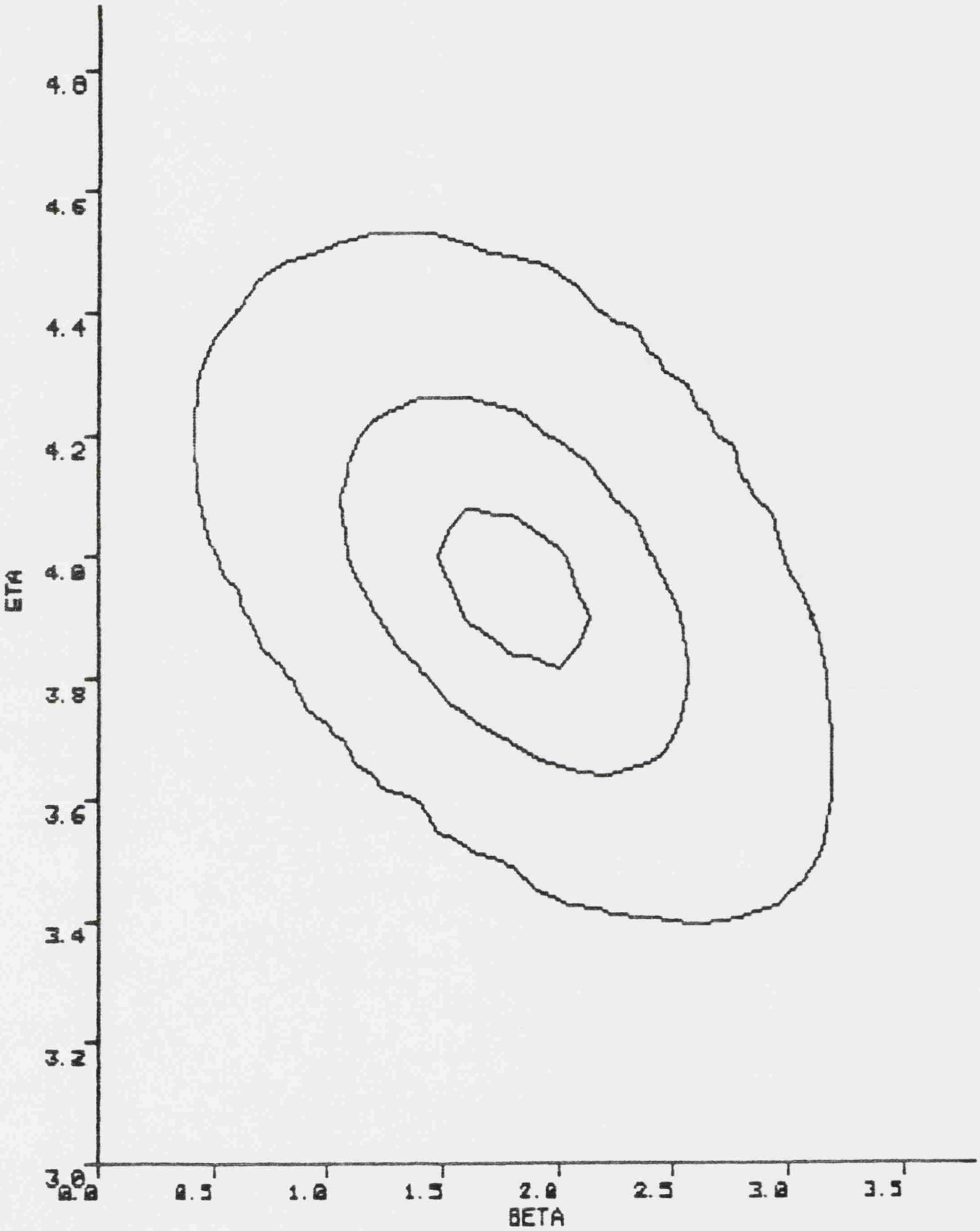

Figure 6.5.

Contours (0.75,0.25,0.01) of β and η of relative likelihood with $a = \hat{a}$, $b = \hat{b}$, for the integrated method using the data of Batchelor and Hackett.

6.2 Data of Mantel et al.

Table 6.2 shows the number of weeks to death without tumour or to tumour appearance of a group of treated rats and their two litter-matched controls. The whole data-set comprised three hundred rats, fifty male litters and fifty female litters of size three and these were divided at random so that one of the rats from each litter received a drug believed to induce tumours. Table 6.2 gives the data for the female litters. Of the males treated, none displayed tumours whilst only two of the control males did. Hence if the treatment is tumorigenic it would seem to be confined to females. All rats were sacrificed at the end of 104 weeks.

For all three methods of analysis the single covariate treated/control was taken with $z_{1i} = 1$ (treated rat) and $z_{ji} = 0$ j=2,3 (control) for i=1,...,50. The outcome measure was time to tumour, with death without tumour counting as a censored observation. No differentiation was made between censoring due to sacrifice and censoring due to death before the end of the study period.

The non-parametric method analyses only 20 litters since 30 of the litters have either all three members censored or the minimum of the three times censored.

Litter No.	Treated rat	Control 1	Control 2
01	101 D	49 T	104 D
03	104 D	102 D	104 D
05	104 D	104 D	104 D
07	77 D	97 D	79 D
09	89 D	104 D	104 D
11	88 T	96 T	104 D
13	104 T	94 D	77 T
15	96 T	104 D	104 D
17	82 D	77 D	104 D
19	70 T	104 D	77 D
21	89 T	91 D	90 D
23	91 D	70 D	92 D
25	39 T	45 D	50 T
27	103 T	69 D	91 D
29	93 D	104 D	103 D
31	85 D	72 D	104 D
33	104 D	63 D	104 D
35	104 D	104 D	74 D
37	81 D	104 D	69 D
39	67 T	104 D	68 T
41	104 D	104 D	104 D
43	104 D	104 D	104 D
45	104 D	83 D	40 T
47	87 D	104 D	104 D
49	104 D	104 D	104 D
51	89 D	104 D	104 D
53	78 D	104 D	104 D
55	104 D	81 T	64 T
57	86 T	55 T	94 D
59	34 T	104 D	54 T
61	76 D	87 D	74 D
63	103 T	73 T	84 T
65	102 T	104 D	80 D
67	80 T	104 D	73 D
69	45 T	79 D	104 D
71	94 T	104 D	104 D
73	104 D	104 D	104 D
75	104 D	101 T	94 D
77	76 D	84 T	78 T
79	80 T	81 T	76 D
81	72 T	95 D	104 D
83	73 T	104 D	66 T
85	92 T	104 D	102 T
87	104 D	98 D	73 D
89	55 D	104 D	104 D
91	49 D	83 D	77 D
93	89 T	104 D	104 D
95	88 D	79 D	99 D
97	103 T	91 D	104 D
99	104 D	104 D	79 T

Table 6.2.
Time to response of female rats in 50 litters .D denotes week of death without tumour, T week of tumour appearance.

The log likelihood given by 2.24 becomes

$$\ln L_N(\beta) = 14\beta - 20\ln(e^{\beta} + 2) - 4\ln(e^{\beta} + 1)$$

which is maximised at $\beta = \hat{\beta}_N$, the solution of

$$5\exp(2\beta) - 7\exp(\beta) - 14 = 0 ,$$

and the standard deviation of $\hat{\beta}_N$ is given by

$$\frac{[\exp(\beta) + 2][\exp(\beta) + 1]}{\sqrt{44\exp(3\beta) + 96\exp(2\beta) + 56\exp(\beta)}} .$$

Hence $\hat{\beta}_N = 0.92$ with standard deviation 0.417. The marginal method gave $\hat{\beta}_M = 0.76$ with standard deviation 0.320, and $\hat{\eta}_M = 1.84$ with standard deviation 0.462. This analysis was based on 27 litters since 23 litters had all three members' observations censored. The integrated method utilised all the data and gave $\hat{\beta}_I = 0.91$ with standard deviation 0.322, $\hat{\eta}_I = 3.93$ with standard deviation 0.569.

Table 6.3 summarises these results together with the 95% plausibility intervals calculated as for the previous data set. The relative likelihoods of β for the three methods are shown in Figures 6.6 to 6.8 with Figures 6.9 and 6.10 showing contours of relative likelihood of β and η for the marginal and integrated methods similar to those for the previous data set. Also shown in the table are the estimates of the ratio of survival times of treated rat to control, $\exp(-\hat{\beta} / \hat{\eta})$, for the marginal and integrated methods together with their approximate 95% confidence intervals.

The 95% confidence interval and 95% plausibility interval for β from the integrated method are very similar and firmly reject the hypothesis of no treatment effect. Comparisons between these intervals for the other two methods demonstrate the skewness of the likelihood and cast doubt upon the rejection of the null hypothesis $\beta = 0$ by the 95% confidence intervals. For example, the 95% confidence interval for β using the non-parametric method was (0.22, 1.88) whilst the 95% plausibility interval for β was (-0.09, 1.99) which covers $\beta = 0$. The marginal method seems to underestimate η and therefore the estimated ratio of survival times of the treated rat to its control, $\exp(-\hat{\beta} / \hat{\eta})$, 79% for the integrated method as opposed to 66% for the marginal method. This reflects the results shown in Chapter 3 from the simulations that omission of pairs (or in this case triplets) in which all members are censored, considerably biases the estimation of η.

With regard to the computation of the likelihoods for maximisation, the integrated method extends the most easily to triplets or, indeed, to data of the previous type which has n-tuplets with differing values of n. The non-parametric method requires the times within each n-tuplet to be ordered. The marginal likelihood is the most difficult to program for this data set since some of the triplets have the minimum time censored and hence the final term in the log likelihood given by 2.27 has to be evaluated.

	Non-Par. Method	Marginal Method	Integrated Method
$\hat{\beta}$ (S.D)	0.92(.417)	0.76(.320)	0.91(.322)
$\hat{\eta}$ (S.D)	-	1.84(.462)	3.93(.569)
95% C.I. for β	(0.22,1.88)	(0.13,1.38)	(0.28,1.54)
95% C.I. for η	-	(0.93,2.74)	(2.81,5.05)
95% plausibility interval for β	(-0.09,1.99)	(-0.07,1.59)	(0.27,1.47)
$\exp(-\hat{\beta}/\hat{\eta})$	-	0.66	0.79
95% C.I. for $\exp(-\beta/\eta)$	-	(0.39,0.93)	(0.66,0.93)
No.of litters used in analysis	20	27	50

Table 6.3.

Analyses of data of Mantel et al.

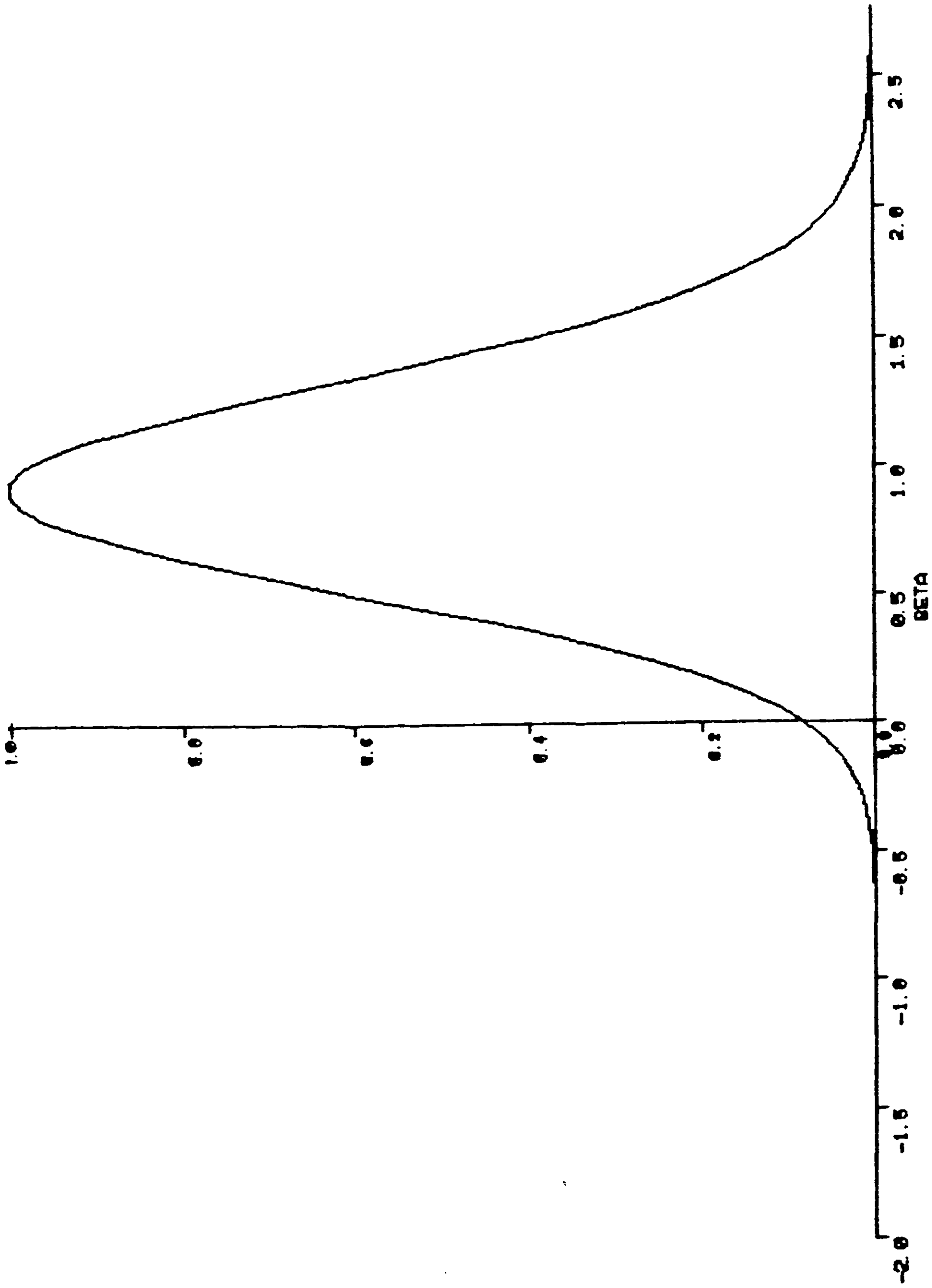

Figure 6.6.

Relative likelihood of β for the Non-parametric method using the data of Mantel et al.

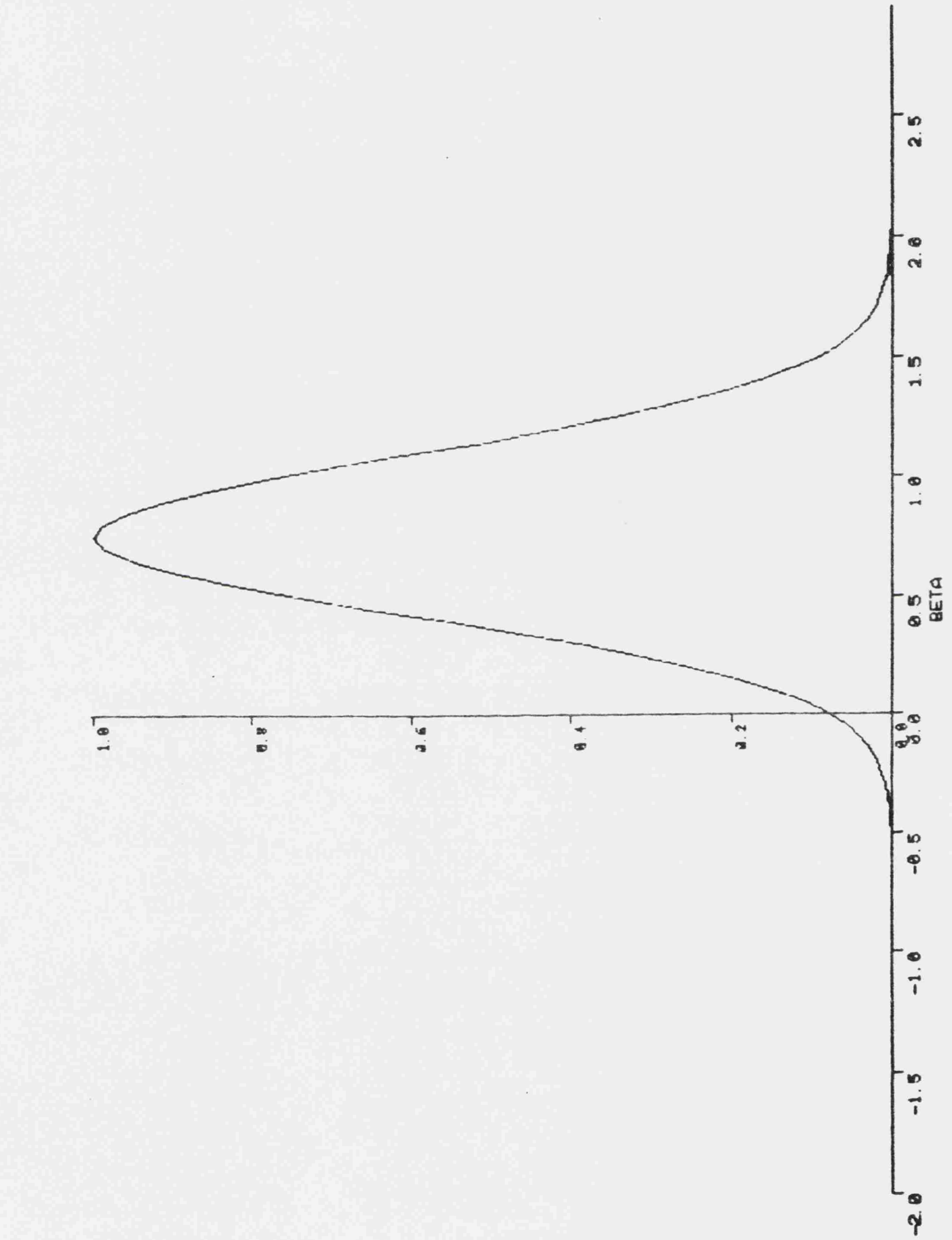

Figure 6.7.

Relative likelihood of β for the marginal method using the data of Mantel et al.

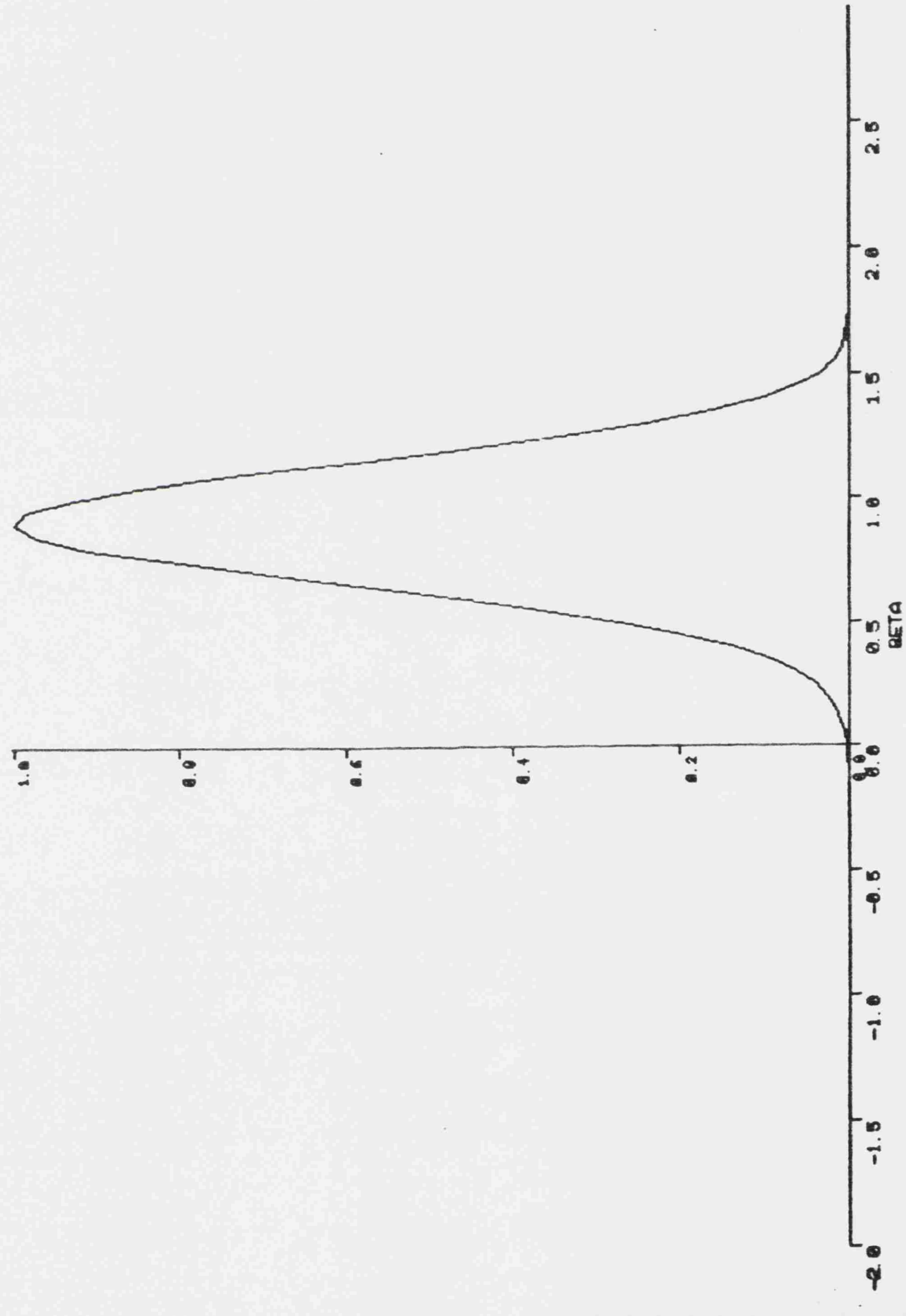

Figure 6.8.

Relative likelihood of β for the integrated method using the data of Mantel et al.

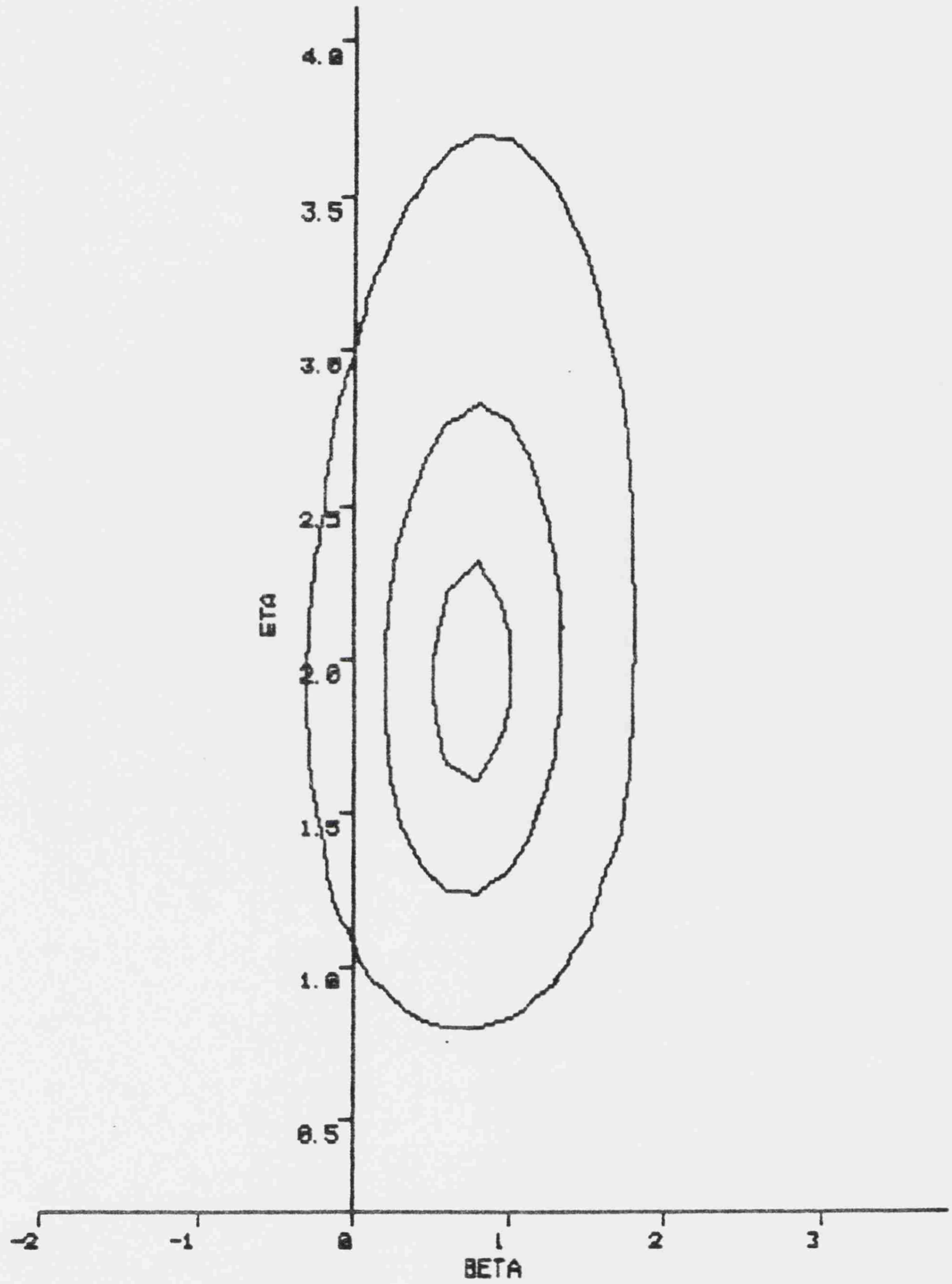

Figure 6.9.

Contours (0.75,0.25,0.01) of relative likelihood for the marginal method using the data of Mantel et al.

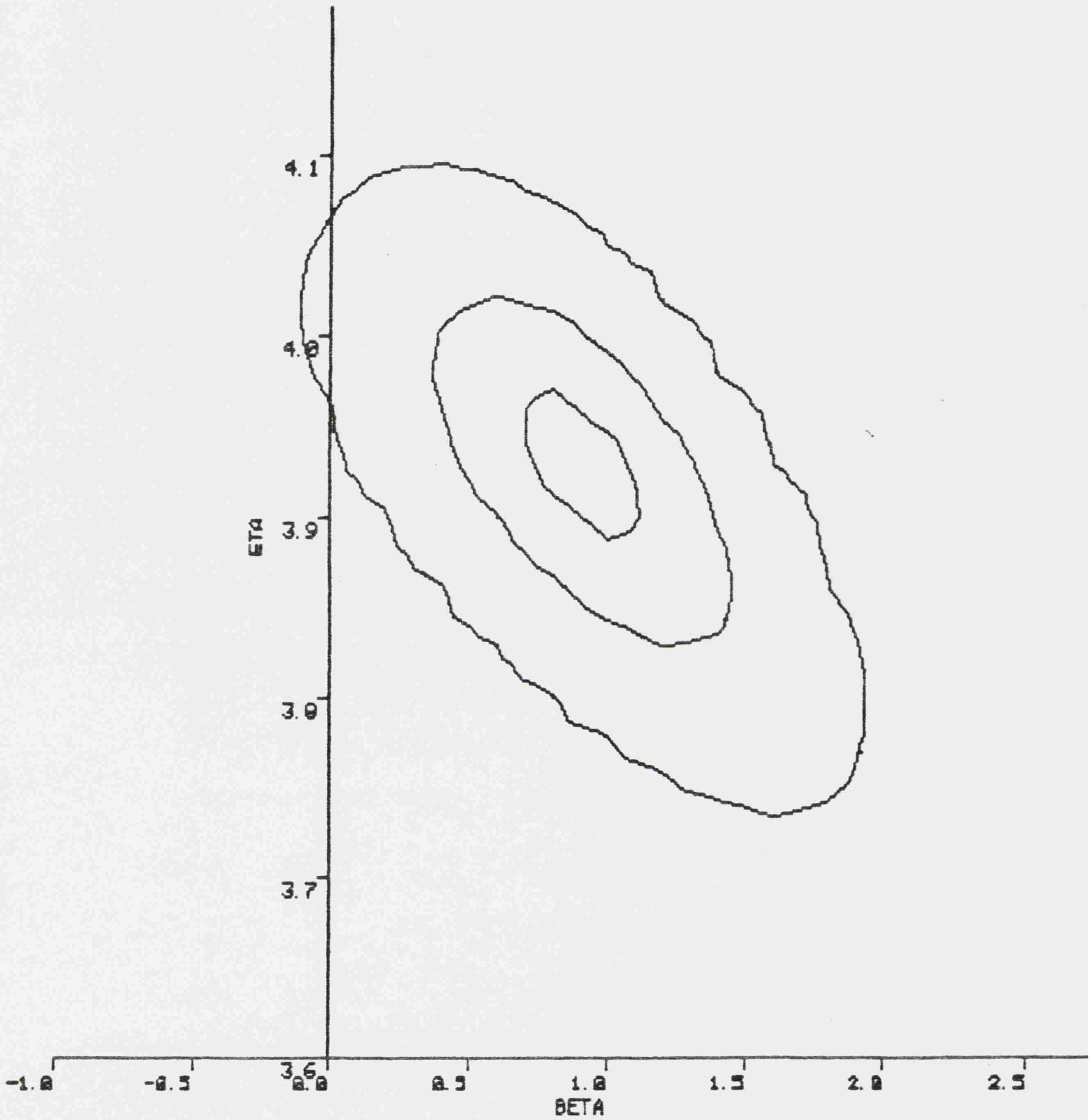

Figure 6.10.

Contours (0.75,0.25,0.01) of β and η of relative likelihood with $a = \hat{a}$, $b = \hat{b}$, for the integrated method using the data of Mantel et al.

6.3 Data of Sampford and Taylor.

The final data set is that given in the 1959 paper by Sampford and Taylor. Table 6.4 shows the logarithms of the survival times of rats treated with carbon tetrachloride and vitamin B12 and their litter mates treated only with carbon tetrachloride.The experiment was performed to investigate whether vitamin B12 affected the rate of action of carbon tetrachloride. Observation ceased after 16 hours (ln16 = 2.98). The new EM method is used to see if the results differ greatly from the Wolynetz method (which is a reparametrized version of the original method of Sampford and Taylor).

B12	Control
2.73	>2.98
2.80	>2.98
2.01	2.84
2.19	2.76
2.34	2.83
2.61	2.73
2.51	2.62
2.65	2.70
2.72	2.76
2.79	2.82
2.90	2.79
2.78	2.64
2.78	2.48
2.97	2.64
2.74	2.31
2.96	2.51
>2.98	2.68
>2.98	>2.98
>2.98	>2.98
>2.98	>2.98

Table 6.4.
Logs of Survival times in hours of rats treated with carbon tetrachloride and vitamin B12 and their controls treated with carbon tetrachloride only.

The covariate, z_{jl} , vitamin treated/control was taken to be 1 for treated and zero for control. Using the Wolynetz method the estimates and 95% confidence intervals for β and τ are -0.054 (-0.25, 0.14) and 0.283 (0.06, 0.50). This analysis is based on 17 litters, ignoring the three doubly censored pairs. The new EM method produced estimates of β and τ of -0.042 and 0.135 with 95% confidence intervals of (-0.20, 0.12) and (0.17, 0.34) respectively. Both show insufficient evidence to reject the null hypothesis of $\beta = 0$. Even with only three pairs of observations doubly censored, there is some difference between the estimates of τ from the two methods. This difference has an effect on the expected ratio of survival times of the treated group to the control and the approximate 95% confidence interval for this measure. For the Wolynetz method these were 1.03 with approximate 95% confidence interval of (0.78, 1.48) whilst the new EM method gave results of 0.98 and (0.84, 1.27). Obviously with only fifteen percent of the pairs doubly censored, this data set does not show off the merits of the integrated method to the full.

CHAPTER 7

CONCLUSIONS AND SCOPE FOR FURTHER WORK

7.1 Conclusions.

Previous work by Holt and Prentice with the proportional hazards model revealed the need for a method of analysis for matched censored data which produced less biased estimates of the parameters measuring the effects of the covariates in the model, especially when the amount of censoring was large. With this in mind, the integrated method was put forward and in Chapter 2 equations producing the maximum likelihood estimates of the parameters and the calculations necessary to produce the observed information matrix were derived. This method used the added assumption that the matching variables came from a gamma distribution. One advantage of this method was that all the data were included in the analysis whereas previous methods omitted pairs in which the rank order of the times within pairs or the ratio of times within pairs was indeterminate. The results of Holt and Prentice and Wild showed that the marginal method produced very biased results for the estimation of β with heavy censoring when the exponential model was assumed. Wild showed that the integrated method performed better.

However neither Holt and Prentice nor Wild considered estimation with the Weibull model and Chapter 3 investigated this by means of simulations since closed forms for the estimates do not exist. The simulations showed that the marginal method produced less biased estimates of β with heavy censoring but very biased estimates of the scale parameter η. The non-parametric method performed well overall, as Holt and Prentice had previously shown, although the integrated method gave estimates as least as good as the non-parametric method in most cases.

Both Holt and Prentice and Wild only considered the fixed censoring scheme where observations are censored only if they reach the end of the study period without having failed. In practice this is not a very realistic scheme as observations are often liable to censoring from other influences such as withdrawal of the subject, loss to follow-up or death from other causes. Thus a random censoring mechanism was also investigated, which is capable of producing pairs in which the minimum of the times is censored and these pairs are omitted from analysis by the non-parametric and marginal methods. Of two of the pairs of values of the censoring parameters (λ_1 , λ_2) considered, one set generated a large number of pairs with the minimum time within pairs censored and few doubly censored pairs whilst the other generated a large number of doubly censored pairs and few with the minimum time censored. The total number of pairs omitted by the Wolynetz method, i.e. the sum of the doubly censored pairs and those with the minimum time censored, were approximately equal in both cases. The non-parametric method gave much more heavily biased results in the case of the majority of pairs omitted having the minimum time censored rather than the case with the majority of pairs omitted being doubly censored. The integrated method did not seem to suffer to the same extent.

The integrated method seemed to perform best when the number of pairs was large, although the estimates were at least as good as the estimates produced by the other two methods for all sample sizes considered. The assumption that the prior distribution of the matching variables was gamma, seemed fairly robust to the actual distribution, apart from the case when the matching variables were uniformly distributed and here the estimation was more biased. One disadvantage with the integrated method which should be mentioned is that it requires appreciably more computer time than the other two methods considered. This, however, can be offset against the improved estimation and, if the data consist of n-tuplets of size greater than two, this difference in computer time diminishes since the integrated method extends more easily to this type of data.

A similar problem existed with the normal theory accelerated failure model and a similar solution to the marginal method of Holt and Prentice was suggested by Sampford and Taylor, involving conditioning out the matching variables. Wolynetz tightened up the estimation by using the EM algorithm and treating the censored observations as missing data, although the basic method was the same. Both the analyses of Wolynetz and Sampford and Taylor omitted the same type of censored observations as did the marginal method for the proportional hazards model, namely pairs of the type (c), (d), (g), (h) and (i) in Figure 2.1. In Chapter 4 a similar solution was proposed to the integrated method but this time the normal form of the likelihood made use of the EM algorithm possible. The problem was formulated with the matching variables being treated as missing data along with the censored observations. An assumption was made that the matching variables came from a normal distribution with unknown mean and variance and these were then estimated. The estimates of the parameters in the model together with their asymptotic variances from the observed information matrix were derived in Chapter 4.

There had been no previous work stating that the estimation of the parameters by the Wolynetz method might be subject to bias with heavy censoring, but it is suggested by the results of Holt and Prentice with the marginal method. Hence in Chapter 5, comparisons were made between the two techniques, that due to Wolynetz and the proposed new EM method. Simulations were again necessary since closed forms for the estimates did not exist. Both fixed and random censoring mechanisms were looked at and various prior distributions for the matching variables were taken to produce the data for the simulations in order to assess the robustness of the assumption of a normal distribution for the pairing variables. Very similar results were found to those of the marginal and integrated methods. The Wolynetz method gave very biased estimates of the scale parameter τ with heavy censoring whilst the bias using the new EM method was slight. The new EM method seemed robust to changes in the distribution of the pairing variables although the uniform prior gave the greatest bias. However this bias was still less than that using the Wolynetz method. Random censoring produced even greater bias on the Wolynetz estimates when a large number of pairs had the minimum time censored. The new EM method was fairly easily extended to cope with right and interval censored data, as was the Wolynetz method. The same comments with regard to computer time held as were stated previously for the integrated method. The method of Louis was used to improve the convergence of the new EM method and this did reduce the number of iterations required. However the new EM method was still the slower of the two methods.

As more information is being used, it is to be expected that the proposed new solutions, the integrated and the new EM methods, should give better estimates than the previous methods. It has been shown, however, that the gain in accuracy of the estimation is considerable, even with quite modest amounts of censoring. Another by-product of the new methods is that the strength of association of times within pairs can be estimated also, with other factors in the form of covariates being allowed for. This was sketched out in section 3.4.

7.2 Further work.

Although some of the theory presented in Chapters 2 and 4 relate to one binary covariate only, there is no reason why the covariate cannot take any values and this is shown, to some extent, by the fitting of the interaction term in the model of the first data set of Chapter 6. Similarly multiple covariates can easily be handled.

The obvious extension to the work of Chapter 4 is to allow for data in the form of triplets or n-tuplets of greater size. The normal theory accelerated failure model was solved only for matched pairs. Although Wolynetz extended the work of Sampford and Taylor to deal with more than one covariate, he does not extend it to allow for n-tuplets with $n > 2$. The extension would be quite natural in the new EM method but this would involve the evaluation of expectations of truncated multivariate normal distributions. Approximations may perhaps be found for the integrals and these approximations used in the method, although this would almost certainly increase the number of iterations necessary for convergence. For the design with one case and multiple controls it might be possible to overcome the difficulty of multivariate expectations. For ease consider the problem of one case and two controls i.e. data of the form of Mantel et al in given Chapter 6. Let the survival times for the case and two controls be t_{1i} , t_{2i} and t_{3i} respectively. A new survival time $t_i' = (t_{2i} + t_{3i})/2$ can be formed from the means of the times of the two controls. If the case and two control times are assumed to be trivariate normal, as the normal theory model assumes, with means $\mu + \beta + \sigma a_i$, $\mu + \sigma a_i$, $\mu + \sigma a_i$ and variances τ^2 then t_{1i} and t_i' will be bivariate normal with means $\mu + \beta + \sigma a_i$, $\mu + \sigma a_i$ and variances τ^2 and $\tau^2/2$. If t_{2i} and/or t_{3i} are censored i.e. the three cases $t_{2i} = t_{2i}^*$, $t_{3i} > t_{3i}^*$; $t_{2i} > t_{2i}^*$, $t_{3i} = t_{3i}^*$ and $t_{2i} > t_{2i}^*$, $t_{3i} > t_{3i}^*$ then t_i' is known to be at least $(t_{2i}^* + t_{3i}^*)/2$. Because the variances of t_{1i} and t_i' are no longer the same, the theory of Chapter 4 needs to be reworked to allow for this, but the same principles hold.

The difficulty that arises with most multivariate problems is that of testing the goodness-of-fit of the assumed model. The proportional hazards assumption may be tested, as in the two sample case, by the inclusion of a time dependent factor which might take the form of the grouping factor multiplied by time. However residual plots to test the Weibull or normal assumptions are not easy in this type of problem as the residuals will be correlated within n-tuplets. If the design is balanced, i.e. all the n-tuplets are of, say, size two with the case as the first member of the pair and the control as the second member, it might be worthwhile to plot the residuals separately for the cases and then the controls. For larger n-tuplets with one case as the first member and multiple controls, the residuals can be plotted for the cases, then the controls who are second members of the n-tuplets followed by the third members, etc. although this does not seem wholly satifactory especially if the n-tuplets are of different sizes, as in the data of Batchelor and Hackett analysed in Chapter 6. The integrated method can be generalised for any distribution of survival times, not just exponential or Weibull. If the general form for the hazard rate, given by (2.1) is taken and the prior distribution of the matching variables is again assumed to be gamma with parameters a and b then the log likelihood (2.17) becomes

$$L\ (\underline{t}_1, \underline{t}_2, \beta, \underline{\phi}, a, b) = \sum [\ \ln \Gamma (b + \delta_{1i} + \delta_{2i}) - \ln \Gamma(b)] + \beta \sum (\delta_{1i} z_{1i} + \delta_{2i} z_{2i}) + Nb \ln a + \sum \delta_{1i} \ln [\lambda_0(t_{1i} ; \underline{\phi})] + \sum \delta_{2i} \ln [\lambda_0(t_{2i} ; \underline{\phi})] + \sum (b + \delta_{1i} + \delta_{2i}) \ln [\Lambda_0(t_{1i} ; \underline{\phi}) e^{\beta z_{1i}} + \Lambda_0(t_{2i} ; \underline{\phi}) e^{\beta z_{2i}} + a\ ]$$

where $\underline{\phi}$ is the vector of parameters of the failure time distribution, Λ_0 is the integrated hazard function and the summation is taken over i=1,...,N. Thus in the case of the Weibull distribution $\underline{\phi} = (\eta)$. Hence other parametric models of failure time may be fitted, the only constraint being the ability to express the integrated hazard in a closed form. The Weibull assumption may then be tested by fitting a more general parametric distribution of which the Weibull is a special case. Although, in principle this is a simple matter, in practice it is likely that the differentiation necessary to form the maximum likelihood equations will be difficult for more general models.

In Chapter 1 the different censoring mechanisms were discussed and it was assumed that the likelihood factored out so that the distributional form of the censoring times did not contribute to the kernel of the likelihood. In the case of unmatched data, this requires that the survival times and potential censoring times for individuals are independent. When this is extended to n-tuplets the relationships within the n-tuplets need also to be taken into account. Certainly if all the T_{ji} and C_{ji}, the survival times and censoring times of the j-th individual in the i-th n-tuplet, were all independent then the likelihood would factor out in the required way. However the T_{ji} within an n-tuplet are not likely to be independent, except conditional upon the 'matching' variable, and it is likely that the C_{ji} will not be either. This points to the belief that it might be necessary only for the joint distributions of the $(T_{1i}, \ldots, T_{ni})$ and $(C_{1i}, \ldots, C_{ni})$ to be independent. However weaker assumptions on the relationships between the T_{ji} and C_{ji} may suffice.

In some cases the assumption of non-informative censoring may not be valid. Little work has been done on this topic for the unmatched data case, although the theory derived for competing risks may be applied by considering the failure time and censoring time instead of the, more usual, failure times from two causes. Some joint distribution for the T_{ji} and C_{ji} may be taken, introducing extra parameters which would need to be estimated. The assumptions on the distributional form of the joint distribution of the T_{ji} and C_{ji} are, however, impossible to test with the usual kind of data set and it is not known therefore, how robust the methods, which already exist, are for the unmatched data case. Another problem that is likely to arise is the complexity of the likelihood with joint distributions between the T_{ji} , $j=1,...,n$ and between the T_{ji} and C_{ji} .

CHAPTER 8

APPENDICES

8.1 Appendix 1.

Let $E[B(\underline{X}, \hat{\Phi}) \mid \underline{X} \in \mathcal{R}] = (b_{k\ell})$ for k,l = 1,...,4. Then from page 4-16 with the summation being taken over i = 1,...,N

$$b_{11} = N/\hat{\tau}^2 \ , \quad b_{22} = N/\hat{\sigma}^2 \ ,$$

$$b_{12} = b_{13} = b_{24} = b_{34} = 0,$$

$$b_{13} = \frac{2}{\hat{\tau}^2} E[\Sigma \ (t_{1i} - \beta - a_i) \mid \underline{X} \in \mathcal{R} \ , \Phi = \hat{\Phi}]$$

$$b_{23} = \frac{2}{\hat{\sigma}^2} E[\Sigma \ (a_i - \mu) \mid \underline{X} \in \mathcal{R} \ , \Phi = \hat{\Phi}]$$

$$b_{33} = -N/\hat{\sigma}^2 + \frac{3}{\hat{\sigma}^4} E[\Sigma \ (a_i - \mu)^2 \mid \underline{X} \in \mathcal{R} \ , \Phi = \hat{\Phi}]$$

$$b_{44} = -2N/\hat{\tau}^2 + \frac{3}{\hat{\tau}^4} E[\Sigma \ [(t_{1i} - \beta - a_i)^2 + (t_{2i} - a_i)^2] \mid \underline{X} \in \mathcal{R} \ , \Phi = \hat{\Phi}].$$

The expectation is taken over a_i and the censored times.

Conditional upon t_{1i} , t_{2i} and Φ

$$a_i \sim N\left(\mu + \frac{\sigma^2(F_{1i} + F_{2i})}{v}, \frac{\sigma^2\tau^2}{v}\right)$$

where $F_{1i} = t_{1i} - \mu - \beta$, $F_{2i} = t_{2i} - \mu$ and $v = 2\sigma^2 + \tau^2$. Also, at $\Phi = \hat{\Phi}$, the following identities hold from equating (p+1) and (p) in equations (4.18)

$$\sum E(\hat{F}_{1i}) = \sum E(\hat{F}_{2i}) = 0 \tag{8.1}$$
$$\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^2] = 2N\hat{v} \tag{8.2}$$
$$\sum E[(\hat{F}_{1i} - \hat{F}_{2i})^2] = 2N\hat{\tau}^2 \tag{8.3}$$
$$\sum E[\hat{F}_{1i}^2 + \hat{F}_{2i}^2] = 2N\hat{\omega}^2 \tag{8.4}$$
$$\sum E[\hat{F}_{1i}\hat{F}_{2i}] = N\hat{\sigma}^2 \tag{8.5}$$

with $\hat{F}_{1i} = t_{1i} - \hat{\mu} - \hat{\beta}$, $\hat{F}_{2i} = t_{2i} - \hat{\mu}$ and the summations from i = 1,...,N.

Hence

$$b_{13} = \frac{2}{\hat{\tau}^2} \sum E[(t_{1i} - \beta - a_i) \mid \underline{X} \in R, \Phi = \hat{\Phi}]$$

$$= \frac{2}{\hat{\tau}^2} \sum E\left\{\left[F_{1i} - \frac{\sigma^2(F_{1i} + F_{2i})}{v}\right] \middle| \Phi = \hat{\Phi}\right\}$$

$$= 0.$$

Similarly

$$b_{22} = \frac{2}{\hat{\sigma}^2} \sum E[(a_i - \mu) \mid \underline{X} \in R, \Phi = \hat{\Phi}]$$

$$= \frac{2}{\hat{\sigma}^2} \sum E\left[\frac{\sigma^2(F_{1i} + F_{2i})}{v} \middle| \Phi = \hat{\Phi}\right]$$

$$= 0.$$

Now

$$b_{33} = -N/\hat{\sigma}^2$$

$$+ \frac{3}{\hat{\sigma}^4} \sum E[(a_i - \mu)^2 | \underline{X} \in R , \Phi = \hat{\Phi}]$$

$$= -N/\hat{\sigma}^2$$

$$+ \frac{3}{\hat{\sigma}^4} \sum E\left[\frac{\sigma^2\tau^2}{v} + \frac{\sigma^4(F_{1i} + F_{2i})^2}{v^2} \middle| \Phi = \hat{\Phi} \right]$$

$$= -N/\hat{\sigma}^2 + 3N\hat{\tau}^2/(\hat{\sigma}^2 \hat{v}) + 6N/\hat{v}$$

$$= 2N/\hat{\sigma}^2$$

and

$$b \quad = -2N/\hat{\tau}^2$$

$$+ \frac{3}{\hat{\tau}^4} \sum E\left\{ [(t_{1i} - \beta - a_i)^2 + (t_{2i} - a_i)^2] \middle| \underline{X} \in R, \Phi = \hat{\Phi} \right\}$$

$$= -2N/\hat{\tau}^2$$

$$+ \frac{3}{\hat{\tau}^4} \sum E[\sum_{j=1}^{2} (a_i - \mu - F_{ji})^2 \quad | \underline{X} \in R , \Phi = \hat{\Phi}]$$

$$= -2N/\hat{\tau}^2$$

$$+ \frac{3}{\hat{\tau}^4} \sum E[2(a_i - \mu)^2 - 2(a_i - \mu)(F_{1i} + F_{2i})$$

$$+ (F_{1i}^2 + F_{2i}^2) \quad | \underline{X} \in R , \Phi = \hat{\Phi}]$$

$$= -2N/\hat{\tau}^2$$

$$+ \frac{3}{\hat{\tau}^4} \sum E\left[\frac{2\sigma^2\tau^2}{v} + \frac{2\sigma^4(F_{1i} + F_{2i})^2}{v^2} - \frac{2\sigma^2(F_{1i} + F_{2i})^2}{v} \right.$$

$$\left. + (F_{1i}^2 + F_{2i}^2) \middle| \Phi = \hat{\Phi} \right]$$

$$= \frac{-2N}{\hat{\tau}^2} + \frac{3}{\hat{\tau}^4}\left[\frac{2N\hat{\sigma}^2\hat{\tau}^2}{\hat{v}} + \frac{4N\hat{\sigma}^4}{\hat{v}} - 4N\hat{\sigma}^2 + N(\hat{v} + \hat{\tau}^2) \right]$$

$$= 4N/\hat{\tau}^2 \ .$$

Thus $E[B(\underline{X}, \Phi) \mid \underline{X} \in R] = \mathrm{diag}(N/\hat{\tau}^2, N/\hat{\sigma}^2, 2N/\hat{\sigma}^2, 4N/\hat{\tau}^2)$ with the off diagonal elements being zero.

8.2 Appendix 2.

The symmetric matrix C is defined as

$$\sum_{i=1}^{N} E\{G(X_i, \hat{\Phi})G^T(X_i, \hat{\Phi})\} - \sum_{i=1}^{N} [E\{G(X_i, \hat{\Phi})\} E\{G^T(X_i, \hat{\Phi})\}].$$

The gradient vector $G(X_i, \Phi)$ is given by

$$\left(\frac{(t_{1i} - \beta - a_i)}{\tau^2}, \frac{(a_i - \mu)}{\sigma^2}, \frac{[(a_i - \mu)^2 - \sigma^2]}{\sigma^3}, \frac{[(t_{1i} - \beta - a_i)^2 + (t_{2i} - a_i)^2 - 2\tau^2]}{\tau^3} \right)^T$$

$$= \left(\frac{(F_{1i} - u_i)}{\tau^2}, \frac{u_i}{\sigma^2}, \frac{(u_i^2 - \sigma^2)}{\sigma^3}, \frac{[2u_i^2 - 2u_i(F_{1i} + F_{2i}) + F_{1i}^2 + F_{2i}^2]}{\tau^3} \right)^T$$

where $u_i = a_i - \mu$, $F_{1i} = t_{1i} - \mu - \beta$, $F_{2i} = t_{2i} - \mu$.

The following results will be required in the calculation of C.

$$E[u_i \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{\hat{\sigma}^2}{\hat{V}} E(\hat{F}_{1i} + \hat{F}_{2i})$$

$$E[u_i^2 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{\hat{\sigma}^2 \hat{\tau}^2}{\hat{V}} + \frac{\hat{\sigma}^4}{\hat{V}} E[(\hat{F}_{1i} + \hat{F}_{2i})^2]$$

$$E[u_i^3 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{3\hat{\sigma}^4 \hat{\tau}^2}{\hat{V}^2} E(\hat{F}_{1i} + \hat{F}_{2i}) + \frac{\hat{\sigma}^6}{\hat{V}^3} E[(\hat{F}_{1i} + \hat{F}_{2i})^3]$$

$$E[u_i^4 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{3\hat{\sigma}^4\hat{\tau}^4}{\hat{v}^2} + \frac{6\hat{\sigma}^6\hat{\tau}^2 E[(\hat{F}_{1i} + \hat{F}_{2i})^2]}{\hat{v}^3} + \frac{\hat{\sigma}^8 E[(\hat{F}_{1i} + \hat{F}_{2i})^4]}{\hat{v}^4}$$

using (8.1), and

$$\sum E[u_i \mid \underline{X} \in R, \Phi = \hat{\Phi}] = 0$$

$$\sum E[u_i^2 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = N\hat{\sigma}^2$$

$$\sum E[u_i^3 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{\hat{\sigma}^6 \sum E[(\hat{F}_{1i} + \hat{F}_{2i})^3]}{\hat{v}^3}$$

$$\sum E[u_i^4 \mid \underline{X} \in R, \Phi = \hat{\Phi}] = \frac{3N\hat{\sigma}^4\hat{\tau}^2(\hat{\tau}^2 + 4\hat{\sigma}^2)}{\hat{v}^2} + \frac{\hat{\sigma}^8}{\hat{v}^4}\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^4]$$

with $E(\hat{F}_{1i}) = E(t_{1i} - \mu - \beta \mid t_{1i} > c_{1i}, \Phi = \hat{\Phi})$ and the summation being over i=1,...,N.

Thus the terms of the upper triangle of the matrix C are

$$C(1,1) = \frac{1}{\hat{\tau}^4}\sum\Big[E[(F_{1i} - u_i)^2 \mid \underline{X} \in R, \Phi = \hat{\Phi}] - \{E[(F_{1i} - u_i) \mid \underline{X} \in R, \Phi = \hat{\Phi}]\}^2\Big]$$

$$= \frac{1}{\hat{\tau}^4}\left[\sum E(\hat{F}_{1i}^{\,2}) - 2\,\frac{\hat{\sigma}^2}{\hat{v}}\sum E[\hat{F}_{1i}(\hat{F}_{1i} + \hat{F}_{2i})] + N\hat{\sigma}^2 - \sum[\hat{\omega}^2 E(\hat{F}_{1i}) - \hat{\sigma}^2 E(\hat{F}_{2i})]^2\right]$$

with $\hat{\omega}^2 = \hat{\sigma}^2 + \hat{\tau}^2$.

$$C(1,1) = \frac{1}{\hat{\tau}^4}\left[N\hat{\sigma}^2 + \frac{\hat{\tau}^2}{\hat{v}}\sum E(\hat{F}_{1i}^{\,2}) - \frac{N\hat{\sigma}^2}{\hat{v}}(\hat{v} - \hat{\tau}^2) - \frac{\hat{\omega}^4}{\hat{v}^2}\sum[E(\hat{F}_{1i})]^2 + \frac{2\hat{\sigma}^2\hat{\omega}^2}{\hat{v}^2}\sum E(\hat{F}_{1i})E(\hat{F}_{2i}) - \frac{\hat{\sigma}^4}{\hat{v}^2}\sum[E(\hat{F}_{2i})]^2\right]$$

using the identity (8.5). Then

$$C(1,1) = \frac{1}{\hat{\tau}^4}\sum\left[\frac{\hat{\tau}^2\hat{\sigma}^2}{\hat{v}} + \frac{\hat{\tau}^2}{\hat{v}}E(\hat{F}_{1i}^{\,2}) - \frac{\hat{\omega}^4}{\hat{v}^2}[E(\hat{F}_{1i})]^2 + 2\left(\frac{\hat{\sigma}\hat{\omega}}{\hat{v}}\right)^2 E(\hat{F}_{1i})E(\hat{F}_{2i}) - \left(\frac{\hat{\sigma}^2}{\hat{v}}\right)^2[E(\hat{F}_{2i})]^2\right]$$

$$C(1,2) = \frac{1}{\hat{\tau}^2\hat{\sigma}^2}\sum\left[E[u_i(F_{1i} - u_i) \mid \underline{X} \in R, \underline{\Phi} = \hat{\underline{\Phi}}] - E[u_i \mid \underline{X} \in R, \underline{\Phi} = \hat{\underline{\Phi}}]E[(F_{1i} - u_i) \mid \underline{X} \in R, \underline{\Phi} = \hat{\underline{\Phi}}]\right]$$

$$= \frac{1}{\hat{\tau}^2\hat{\sigma}^2}\left[\frac{\hat{\sigma}^2}{\hat{v}}\sum E[\hat{F}_{1i}(\hat{F}_{1i} + \hat{F}_{2i})] - N\hat{\sigma}^2 - \frac{\hat{\sigma}^2}{\hat{v}}\sum E(\hat{F}_{1i} + \hat{F}_{2i})[\hat{\omega}^2 E(\hat{F}_{1i}) - \hat{\sigma}^2 E(\hat{F}_{2i})]\right]$$

$$C(1,2) = \frac{1}{\hat{\tau}^2\hat{v}}\sum\left[E(\hat{F}_{1i}^{\,2}) - \hat{\omega}^2 - \frac{\hat{\omega}^2}{\hat{v}}[E(\hat{F}_{1i})]^2 - \frac{\hat{\tau}^2}{\hat{v}}E(\hat{F}_{1i})E(\hat{F}_{2i}) + \frac{\hat{\sigma}^2}{\hat{v}}[E(\hat{F}_{2i})]^2\right]$$

$$C(1,3) = \frac{1}{\hat{\sigma}^3\hat{\tau}^2}\sum\Big[E[(F_{1i} - u_i)(u_i^2 - \sigma^2) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]$$

$$- E[(F_{1i} - u_i) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]E[(u_i^2 - \sigma^2) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]\Big]$$

$$= \frac{1}{\hat{\sigma}^3\hat{\tau}^2}\sum\Big[E[(F_{1i} u_i^2 - u_i^3) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]$$

$$- E[(F_{1i} - u_i) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]E[u_i^2 \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]$$

from (8.1)

$$= \frac{1}{\hat{\sigma}^3\hat{\tau}^2}\Big[\frac{\hat{\sigma}^4}{\hat{v}^2}\sum E[\hat{F}_{1i}(\hat{F}_{1i} + \hat{F}_{2i})^2] - \frac{\hat{\sigma}^6}{\hat{v}^3}\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^3]$$

$$- \frac{\hat{\sigma}^4}{\hat{v}^2}\sum E(\hat{F}_{1i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2] + \frac{\hat{\sigma}^6}{\hat{v}^3}\sum E(\hat{F}_{1i} + \hat{F}_{2i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2]$$

Thus

$$C(1,3) = \frac{\hat{\sigma}}{\hat{v}^3\hat{\tau}^2}\sum\Big[\hat{\omega}^2[E(\hat{F}_{1i}^3) - E(\hat{F}_{1i})E(\hat{F}_{1i}^2) - 2E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i})$$

$$- E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] + (\hat{\omega}^2 + \hat{\tau}^2)E(\hat{F}_{1i}^2\hat{F}_{2i})$$

$$- (\hat{\sigma}^2 - \hat{\tau}^2)E(\hat{F}_{1i}\hat{F}_{2i}^2) + \hat{\sigma}^2[E(\hat{F}_{1i}^2)E(\hat{F}_{2i})$$

$$+ 2E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i}) + E(\hat{F}_{2i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^3)]\Big]$$

$$C(1,4) = \frac{1}{\hat{\tau}^5}\sum\Big[E\big\{(F_{1i} - u_i)[2u_i^2 - 2u_i(F_{1i} + F_{2i})$$

$$+ F_{1i}^2 + F_{2i}^2 \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]\big\}$$

$$- E[(F_{1i} - u_i) \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]E[2u_i^2 - 2u_i(F_{1i} + F_{2i})$$

$$+ F_{1i}^2 + F_{2i}^2 \mid \underline{X}\in R, \underline{\Phi} = \hat{\underline{\Phi}}]\Big]$$

$$= \frac{1}{\hat{\tau}^5}\Big[\sum E(\hat{F}_{1i}^3) + 2E(\hat{F}_{1i}\hat{F}_{2i}^2) - 2\frac{\hat{\sigma}^6}{\hat{v}^3}\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^3]$$

$$+ \frac{2\hat{\sigma}^4}{\hat{v}^2}\sum E[(2\hat{F}_{1i} + \hat{F}_{2i})(\hat{F}_{1i} + \hat{F}_{2i})^2] - \sum E(\hat{F}_{1i})E(\hat{F}_{2i}^2)$$

$$- \frac{\hat{\sigma}^2}{\hat{v}} \sum E[(\hat{F}_{1i} + \hat{F}_{2i})(3\hat{F}_{1i}^2 + 2\hat{F}_{1i}\hat{F}_{2i} + \hat{F}_{2i}^2)]$$

$$- \frac{2\hat{\sigma}^2}{\hat{v}}\left(\frac{\hat{\sigma}^2}{\hat{v}} + 1\right) \sum E(\hat{F}_{1i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2] - \sum E(\hat{F}_{1i})E(\hat{F}_{1i}^2)$$

$$+ \frac{2\hat{\sigma}^4}{\hat{v}^2}\left(\frac{\hat{\sigma}^2}{\hat{v}} - 1\right) \sum E(\hat{F}_{1i} + \hat{F}_{2i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2]$$

$$- \frac{\hat{\sigma}^2}{\hat{v}} \sum E(\hat{F}_{1i}^2 + \hat{F}_{2i}^2)E(\hat{F}_{1i} + \hat{F}_{2i}) \qquad \text{using (8.1).}$$

Hence

$$\begin{aligned} C(1,4) = \frac{1}{\hat{v}^3\hat{\tau}^5} \sum \Big[& \hat{\omega}^2(2\hat{\sigma}^2\hat{\omega}^2 + \hat{\tau}^4)[E(\hat{F}_{1i}^3) - E(\hat{F}_{1i})E(\hat{F}_{1i}^2) \\ & - E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] - \hat{\sigma}^2(5\hat{\omega}^4 + \hat{\sigma}^4) E(\hat{F}_{1i}^2 \hat{F}_{2i}) \\ & - \hat{\omega}^2(\hat{\omega}^4 + 5\hat{\sigma}^4) E(\hat{F}_{1i}\hat{F}_{2i}^2) \\ & - \hat{\sigma}^2(\hat{\omega}^4 + \hat{\sigma}^4)[E(\hat{F}_{2i}^3) - E(\hat{F}_{2i})E(\hat{F}_{2i}^2) \\ & + E(\hat{F}_{1i}^2)E(\hat{F}_{2i})] + 4\hat{\sigma}^2\hat{\omega}^4 E(\hat{F}_{1i})E(\hat{F}_{1i}\hat{F}_{2i}) \\ & - 4\hat{\sigma}^4\hat{\omega}^2 E(\hat{F}_{2i})E(\hat{F}_{1i}\hat{F}_{2i}) \Big]. \end{aligned}$$

$$C(2,2) = \frac{1}{\hat{\sigma}^4} \sum \Big[E(u_i^2 \mid \underline{X} \in R, \Phi = \hat{\Phi}) - [E(u_i \mid \underline{X} \in R, \Phi = \hat{\Phi})]^2 \Big]$$

$$C(2,2) = \frac{N}{\hat{\sigma}^2} - \frac{1}{\hat{v}^2} \sum \Big[[E(\hat{F}_{1i})]^2 + 2E(\hat{F}_{1i})E(\hat{F}_{2i}) + [E(\hat{F}_{2i})]^2 \Big]$$

$$\begin{aligned} C(2,3) &= \frac{1}{\hat{\sigma}^5} \sum \Big[E[u_i(u_i^2 - \sigma^2) \mid \underline{X} \in R, \Phi = \hat{\Phi}] \\ &\qquad - E[u_i \mid \underline{X} \in R, \Phi = \hat{\Phi}]E[(u_i^2 - \sigma^2) \mid \underline{X} \in R, \Phi = \hat{\Phi}] \Big] \\ &= \frac{1}{\hat{\sigma}^5} \left[\frac{\hat{\sigma}^6}{\hat{v}^3} \sum E[(\hat{F}_{1i} + \hat{F}_{2i})^3] - \frac{\hat{\sigma}^6}{\hat{v}^3} \sum E(\hat{F}_{1i} + \hat{F}_{2i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2] \right] \end{aligned}$$

using (8.1), giving

$$C(2,3) = \frac{\hat{\sigma}}{\hat{v}^3}\sum\Big[E(\hat{F}_{1i}^3+\hat{F}_{2i}^3) + 3\,E(\hat{F}_{1i}^2\,\hat{F}_{2i}) + 3\,E(\hat{F}_{1i}\,\hat{F}_{2i}^2)$$

$$-\quad E(\hat{F}_{1i})E(\hat{F}_{1i}^2) - 2\,E(\hat{F}_{1i})E(\hat{F}_{1i}\,\hat{F}_{2i})$$

$$-\quad E(\hat{F}_{1i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^2)E(\hat{F}_{2i})$$

$$-\quad 2\,E(\hat{F}_{2i})E(\hat{F}_{1i}\,\hat{F}_{2i}) - E(\hat{F}_{1i}\,\hat{F}_{2i}^2)\Big]$$

$$C(2,4) = \frac{1}{\hat{\sigma}^2\hat{\tau}^3}\sum\Big[E[2u_i^3 - 2u_i^2(F_{1i} + F_{2i})$$

$$+ u_i\,(F_{1i}^2 + F_{2i}^2)\,|\,\underline{X}\in R,\ \Phi=\hat{\Phi}\,]$$

$$- E[u_i\,|\,\underline{X}\in R,\ \Phi=\hat{\Phi}\,]E[2u_i^2 - 2u_i\,(F_{1i} + F_{2i})$$

$$+ F_{1i}^2 + F_{2i}^2\,|\,\underline{X}\in R,\ \Phi=\hat{\Phi}\,]\Big]$$

$$= \frac{1}{\hat{\sigma}^2\hat{\tau}^3}\Big[\frac{2\hat{\sigma}^6}{\hat{v}^3}\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^3] - \frac{2\hat{\sigma}^4}{\hat{v}^2}\sum E[(\hat{F}_{1i} + \hat{F}_{2i})^2]$$

$$+ \frac{\hat{\sigma}^2}{\hat{v}}\sum E[(\hat{F}_{1i}^2 + \hat{F}_{2i}^2)(\hat{F}_{1i} + \hat{F}_{2i})]$$

$$- \frac{2\hat{\sigma}^4}{\hat{v}^2}\left(\frac{\hat{\sigma}^2}{\hat{v}} - 1\right)\sum E(\hat{F}_{1i} + \hat{F}_{2i})E[(\hat{F}_{1i} + \hat{F}_{2i})^2]$$

$$- \frac{\hat{\sigma}^2}{\hat{v}}\sum E(\hat{F}_{1i} + \hat{F}_{2i})E(\hat{F}_{1i}^2 + \hat{F}_{2i}^2)\Big] \qquad \text{giving}$$

$$C(2,4) = \frac{1}{\hat{v}^3\hat{\tau}^3}\sum\Big[(2\hat{\sigma}^2\hat{\omega}^2 + \hat{\tau}^4)[\,E(\hat{F}_{1i}^3 + \hat{F}_{2i}^3)$$

$$-\quad E(\hat{F}_{1i})E(\hat{F}_{2i}^2) - E(\hat{F}_{2i})E(\hat{F}_{2i}^2) - E(\hat{F}_{1i}^2)E(\hat{F}_{2i})$$

$$+\quad E(\hat{F}_{1i})E(\hat{F}_{2i}^2)] - (2\hat{\sigma}^2\hat{\omega}^2 - \hat{\tau}^4)[\,E(\hat{F}_{1i}^2\hat{F}_{2i})$$

$$+\quad E(\hat{F}_{1i}\,\hat{F}_{2i}^2)] + 4\hat{\sigma}^2\hat{\omega}^2[\,E(\hat{F}_{1i})E(\hat{F}_{1i}\,\hat{F}_{2i})$$

$$+\quad E(\hat{F}_{2i})E(\hat{F}_{1i}\,\hat{F}_{2i})\,]\Big]$$

$$C(3,3) = \frac{1}{\hat{\sigma}^6}\sum\Big[E[(u_i^2 - \sigma^2)^2\,|\,\underline{X}\in R,\ \Phi=\hat{\Phi}\,]$$

$$-\ \{E[(u_i^2 - \sigma^2)\,|\,\underline{X}\in R,\ \Phi=\hat{\Phi}\,]\}^2\Big]$$

$$= \frac{1}{\hat\sigma^6}\left[\frac{3N\hat\sigma^4\hat\tau^2(\hat\tau^2+4\hat\sigma^2)}{\hat v^2} + \frac{\hat\sigma^8}{\hat v^4}\sum E[(\hat F_{1i}+\hat F_{2i})^4]\right.$$

$$- N\hat\sigma^4 - \frac{N\hat\sigma^4\hat\tau^4}{\hat v^2} - \frac{4N\hat\sigma^6\hat\tau^2}{\hat v^2} + \frac{\hat\sigma^8}{\hat v^4}\sum\{E[(\hat F_{1i}+\hat F_{2i})^2]\}^2$$

$$\left. + N\hat\sigma^4\right]$$ using (8.1) and (8.2). Thus

$$C(3,3) = \frac{2N\hat\tau^2(\hat v+2\hat\sigma^2)}{\hat\sigma^2\hat v^2} + \frac{\hat\sigma^2}{\hat v^4}\sum\left[E(\hat F_{1i}^4+\hat F_{2i}^4) + 4E(\hat F_{1i}^3\hat F_{2i})\right.$$

$$+ 6E(\hat F_{1i}^2\hat F_{2i}^2) + 4E(\hat F_{1i}\hat F_{2i}^3) - [E(\hat F_{1i}^2)]^2$$

$$- [E(\hat F_{2i}^2)]^2 - 4[E(\hat F_{1i}\hat F_{2i})]^2 - 4E(\hat F_{1i}^2)E(\hat F_{1i}\hat F_{2i})$$

$$\left. - 4E(\hat F_{2i}^2)E(\hat F_{1i}\hat F_{2i}) - E(\hat F_{1i}^2)E(\hat F_{2i}^2)\right]$$

$$C(3,4) = \frac{1}{\hat\sigma^3\hat\tau^3}\sum\left[E[2u_i^4 - 2u_i^3(F_{1i}+F_{2i})\right.$$

$$+ u_i^2(F_{1i}^2+F_{2i}^2) - 2u_i^2\sigma^2 + 2u_i\sigma^2(F_{1i}+F_{2i})$$

$$- \sigma^2(F_{1i}^2+F_{2i}^2) \mid \underline{X}\in R, \phi=\hat\phi]$$

$$- E[(u_i^2-\sigma^2) \mid \underline{X}\in R, \phi=\hat\phi]E[2u_i^2$$

$$\left. - 2u_i(F_{1i}+F_{2i}) + F_{1i}^2 + F_{2i}^2 \mid \underline{X}\in R, \phi=\hat\phi]\right]$$

$$= \frac{1}{\hat\sigma^3\hat\tau^2}\left[\frac{6N\hat\sigma^4\hat\tau^2(\hat\tau^2+4\hat\sigma^2)}{\hat v^2} + \frac{2\hat\sigma^8}{\hat v^4}\sum E[(\hat F_{1i}+\hat F_{2i})^4]\right.$$

$$- \frac{12N\hat\sigma^4\hat\tau^2}{\hat v} - \frac{2\hat\sigma^4}{\hat v^2}\sum E[(\hat F_{1i}+\hat F_{2i})^3] + \frac{2N\hat\sigma^2\hat\tau^2\hat\omega^2}{\hat v} - \frac{2N\hat\sigma^4\hat\tau^4}{\hat v^2}$$

$$+ \frac{\hat\sigma^4}{\hat v^2}\sum E[(\hat F_{1i}^2+\hat F_{2i}^2)(\hat F_{1i}+\hat F_{2i})^2] - \frac{8N\hat\sigma^6\hat\tau^2}{\hat v^2} + \frac{4N\hat\sigma^4\hat\tau^2}{\hat v}$$

$$- \frac{2\hat\sigma^6}{\hat v^3}\left(\frac{\hat\sigma^2}{\hat v} - 1\right)\sum\{E[(\hat F_{1i}+\hat F_{2i})^2]\}^2 - \frac{2N\hat\sigma^2\hat\tau^2\hat\omega^2}{\hat v}$$

$$\left. - \frac{\hat\sigma^4}{\hat v^2}\sum E[(\hat F_{1i}+\hat F_{2i})^2]E[(\hat F_{1i}^2+\hat F_{2i}^2)]\right]$$

using (8.1) - (8.5), thus

$$C(3,4) = \frac{-4N\hat\sigma\hat\tau}{\hat v^2} + \frac{\hat\sigma}{\hat v^4\hat\tau^2}\sum\Big[(2\hat\sigma^2\hat\omega^2 + \hat\tau^4)\{E(\hat F_{1i}^4 + \hat F_{2i}^4)$$
$$- [E(\hat F_{1i}^2)]^2 - [E(\hat F_{2i}^2)]^2 - 2E(\hat F_{1i}^2)E(\hat F_{2i}^2)\}$$
$$+ 2\hat\tau^4[E(\hat F_{1i}^3\hat F_{2i}) + E(\hat F_{1i}\hat F_{2i}^3)$$
$$- E(\hat F_{1i}^2)E(\hat F_{1i}\hat F_{2i}) - E(\hat F_{2i}^2)E(\hat F_{1i}\hat F_{2i})]$$
$$+ 8\hat\sigma\hat\omega^2[E(\hat F_{1i}\hat F_{2i})]^2 - 2(2\hat\sigma^2\hat\omega^2 - \hat\tau^4)E(\hat F_{1i}^2\hat F_{2i}^2)\Big]$$

$$C(4,4) = \frac{1}{\hat\tau^6}\sum\Big[E[4u_i^4 - 8u_i^3(F_{1i} + F_{2i})$$
$$+ 8u_i^2(F_{1i}^2 + F_{2i}^2 + F_{1i}F_{2i}) - 4u_i(F_{1i} + F_{2i})(F_{1i}^2 + F_{2i}^2)$$
$$+ (F_{1i}^2 + F_{2i}^2)^2 \mid \underline{X}\in R, \phi = \hat\phi]$$
$$- \{E[2u_i^2 - 2u_i(F_{1i} + F_{2i}) + F_{1i}^2 + F_{2i}^2 \mid \underline{X}\in R, \phi = \hat\phi]\}^2\Big]$$
$$= \frac{1}{\hat\tau^6}\Big[\frac{12N\hat\sigma^4\hat\tau^2(\hat\tau^2 + 4\hat\sigma^2)}{\hat v^2} + \frac{4\hat\sigma^8}{\hat v^4}\sum E[(\hat F_{1i} + \hat F_{2i})^4]$$
$$- \frac{48N\hat\sigma^4\hat\tau^2}{\hat v} - \frac{8\hat\sigma^6}{\hat v^3}\sum E[(\hat F_{1i} + \hat F_{2i})^4] + \frac{8N\hat\sigma^2\hat\tau^2(2\hat\omega^2 + \hat\sigma^2)}{\hat v}$$
$$+ \frac{\hat\sigma^4}{\hat v^2}\sum E[(\hat F_{1i} + \hat F_{2i})^2(\hat F_{1i}^2 + \hat F_{2i}^2 + \hat F_{1i}\hat F_{2i})] - \frac{4N\hat\sigma^4\hat\tau^4}{\hat v^2} - \frac{16N\hat\sigma^6\hat\tau^2}{\hat v^2}$$
$$- \frac{4\hat\sigma^2}{\hat v}\sum E[(\hat F_{1i} + \hat F_{2i})^2(\hat F_{1i}^2 + \hat F_{2i}^2)] - \frac{4\hat\sigma^8}{\hat v^4}\sum\{E[(\hat F_{1i} + \hat F_{2i})^2]\}^2$$
$$- \frac{4\hat\sigma^4}{\hat v^2}\sum\{E[(\hat F_{1i} + \hat F_{2i})^2]\}^2 + \frac{8N\hat\sigma^4\hat\tau^2}{\hat v} + \frac{8\hat\sigma^6}{\hat v^3}\sum\{E[(\hat F_{1i} + \hat F_{2i})^2]\}^2$$
$$- \frac{8N\hat\sigma^2\hat\tau^2\hat\omega^2}{\hat v} - \frac{4\hat\sigma^4}{\hat v^2}\sum E[(\hat F_{1i}^2 + \hat F_{2i}^2)]E[(\hat F_{1i} + \hat F_{2i})^2]$$
$$+ \frac{4\hat\sigma^2}{\hat v}\sum E[(\hat F_{1i} + \hat F_{2i})^2]E[(\hat F_{1i}^2 + \hat F_{2i}^2)] \quad \text{using (8.1) - (8.5).}$$

Thus

$$C(4,4) = \frac{8N\hat\sigma^2\hat\omega^2}{\hat\tau^2\hat v^2} + \frac{1}{\hat v^4\hat\tau^2}\sum\Big[(\hat v^2\hat\tau^4 + 4\hat\sigma^4\hat\omega^2)\{E(\hat F_{1i}^4 + \hat F_{2i}^4)$$

$$
\begin{aligned}
&- [E(\hat{F}_{1i}^{2})]^2 - [E(\hat{F}_{2i}^{2})]^2 \\
&- 8\hat{\sigma}^2\hat{\omega}^2(\hat{\tau}^4 + 2\hat{\sigma}^2\hat{\omega}^2)\,[\,E(\hat{F}_{1i}^{3}\hat{F}_{2i}) + E(\hat{F}_{1i}\hat{F}_{2i}^{3}) \\
&- E(\hat{F}_{1i}^{2})E(\hat{F}_{1i}\hat{F}_{2i}) - E(\hat{F}_{2i}^{2})E(\hat{F}_{1i}\hat{F}_{2i})] \\
&+ 2(\hat{v}^2\hat{\tau}^2 + 12\hat{\sigma}^4\hat{\omega}^4)\,E(\hat{F}_{1i}^{2}\hat{F}_{2i}^{2}) - 16\hat{\sigma}^4\hat{\omega}^4[E(\hat{F}_{1i}\hat{F}_{2i})]^2 \\
&- 2(\hat{v}^2\hat{\tau}^4 + 4\hat{\sigma}^4\hat{\omega}^4)\,E(\hat{F}_{1i}^{2})E(\hat{F}_{2i}^{2})\Big]\,.
\end{aligned}
$$

8.3 Appendix 3.

The expressions for the matrix C involve calculation of the expectations $E(F_{1i}^3)$, $E(F_{1i}^4)$ in the case of singly censored pairs of the data. Consider the pairs which have t_{1i} censored at c_{1i} and t_{2i} observed. Since all parameters are maximum likelihood estimates the circumflex ($\wedge$) will be omitted.

The joint distribution of t_{1i} and t_{2i} is bivariate normal with means $\mu + \beta$, μ , variances $\sigma^2 + \tau^2$ and correlation $\rho = \sigma^2/(\sigma^2 + \tau^2)$. If $F_{1i} = t_{1i} - \mu - \beta$, $F_{2i} = t_{2i} - \mu$ and $\omega^2 = \sigma^2 + \tau^2$ then the joint distribution of $x_i = F_{1i}/\omega$ and $y_i = F_{2i}/\omega$ is standard bivariate normal with correlation ρ and $h_i < x_i < \infty$ and $-\infty < y_i < \infty$, where $h_i = (c_{1i} - \mu - \beta)/\omega$.

Define $G_r(b)$ as follows

$$G_r(b) = \int_b^\infty u^r \, Z(u)du$$

where $Z(u) = (2\pi)^{-1/2}\exp(-u^2/2)$ is the usual univariate standard normal ordinate. Then by integration, the following recurrence result is obtained

$$G_r(b) = b^{r-1}Z(b) + (r-1)G_{r-2}(b)$$

$$G_0(b) = Q(b) \qquad (8.6)$$

where $Q(b) = \int_b^{\infty} Z(u)du.$

Hence conditional upon t_{1i} being censored at c_{1i} and t_{2i} being observed to fail and with $H_i = (h_i - \rho y_i)/\sqrt{1-\rho^2}$

$$E(x_i^3) = \frac{1}{\sqrt{1-\rho^2}\,Q(H_i)} \int_{h_i}^{\infty} x_i^3 \; Z\left(\frac{x_i - \rho y_i}{\sqrt{1-\rho^2}}\right) dx_i$$

$$= \frac{1}{Q(H_i)} \int_{H_i}^{\infty} (\sqrt{1-\rho^2}\,u + y_i)^3 \, Z(u)du$$

$$= [Q(H_i)]^{-1} [\rho^3 y_i^3 \, G_0(H_i) + 3\rho^2\sqrt{1-\rho^2}\, y_i^2 G_1(H_i)$$
$$+ 3\rho y_i (1-\rho^2) G_2(H_i) + (1-\rho^2)^{3/2} G_3(H_i)] .$$

Now from (8.6)

$$G_0(H_i) = Q(H_i)$$
$$G_1(H_i) = Z(H_i)$$
$$G_2(H_i) = H_i Z(H_i) + Q(H_i)$$
$$G_3(H_i) = H_i^2 Z(H_i) + 2Z(H_i)$$
$$G_4(H_i) = [H_i^3 + H_i^2 + 3H_i + 2)Z(H_i) + 3Q(H_i) .$$

Thus, writing the Mills' ratio $Z(u)/Q(u) = M(u)$

$$E(x_i^3) = \rho^3 y_i^3 + 3\rho^2\sqrt{1-\rho^2}\, y_i M(H_i)$$
$$+ 3\rho(1-\rho^2) y_i [H_i M(H_i) + 1]$$
$$+ (1-\rho^2)^{3/2} [H_i^2 M(H_i) + 2M(H_i)]$$

and as $x_i = F_{1i}/\omega$, $y_i = F_{2i}/\omega$ and $H_i = (c_{1i} - \mu - \beta - \rho F_{2i})/(\omega\sqrt{1-\rho^2})$

$$E(F_{1i}^3) = \rho^3 F_{2i}^3 + 3\rho\omega^2(1-\rho^2)F_{2i} + 2(\omega\sqrt{1-\rho^2})^3 M(H_i) + \omega\sqrt{1-\rho^2}\, M(H_i)[(c_{1i}-\mu-\beta)^2 + \rho F_{2i}(c_{1i}-\mu-\beta) + \rho^2 F_{2i}^2].$$

Similarly $E(F_{1i}^4) = \omega^4 E(x_i^4)$ and

$$E(x_i^4) = \frac{1}{Q(H_i)} \int_{H_i}^{\infty} (\sqrt{1-\rho^2}\, u + \rho y_i)^4 Z(u)du$$

$$= [Q(H_i)]^{-1} [\rho^4 y_i^4 G_0(H_i) + 4\rho^3 y_i^3 \sqrt{1-\rho^2}\, G_1(H_i) + 6\rho^2 y_i^2 (1-\rho^2) G_2(H_i) + 4\rho y_i (1-\rho^2)^{3/2} G_3(H_i) + (1-\rho^2)^2 G_4(H_i)]$$

$$= \rho^4 y_i^4 + 4\rho^3 y_i^3 \sqrt{1-\rho^2}\, M(H_i) + 6\rho^2 y_i^2 (1-\rho^2)[H_i M(H_i) + 1] + 4\rho y_i (1-\rho^2)^{3/2}[H_i^2 M(H_i) + 2M(H_i)] + (1-\rho^2)^2 [(H_i^3 + H_i^2 + 3H_i + 2)M(H_i) + 3].$$

Thus

$$E(F_{1i}^4) = 3\omega^4(1-\rho^2)^2 + \rho^4 F_{2i}^4 + 6\rho^2\omega^2(1-\rho^2)F_{2i}^2 + \omega\sqrt{1-\rho^2}\, M(H_i)[(c_{1i}-\mu-\beta)^3 + \rho(c_{1i}-\mu-\beta)^2 F_{2i} + \rho^2(c_{1i}-\mu-\beta)F_{2i}^2 + \rho^3 F_{2i}^3 + \omega^2(1-\rho^2)\{3(c_{1i}-\mu-\beta) + 5\rho F_{2i}\}].$$

8.4 Appendix 4.

The expressions for the matrix C involve calculation of the expectations $E(F_{1i}^2 F_{2i})$, $E(F_{1i}^3 F_{2i})$, $E(F_{1i}^2 F_{2i}^2)$, $E(F_{1i}^3)$ and $E(F_{1i}^4 + F_{2i}^4)$ in the case of doubly censored pairs of data i.e. those pairs with t_{1i} censored at c_{1i} and t_{2i} censored at c_{2i}. Since all parameters are maximum likelihood estimates the circumflex ($\wedge$) will be omitted.

The joint distribution of t_{1i} and t_{2i} is bivariate normal with means $\mu + \beta$, μ, variances $\sigma^2 + \tau^2$ and correlation $\rho = \sigma^2/(\sigma^2 + \tau^2)$. If $F_{1i} = t_{1i} - \mu - \beta$, $F_{2i} = t_{2i} - \mu$ and $\omega^2 = \sigma^2 + \tau^2$ then the joint distribution of $x_i = F_{1i}/\omega$ and $y_i = F_{2i}/\omega$ is standard bivariate normal with correlation ρ and $h_i < x_i < \infty$ and $k_i < y_i < \infty$ where $h_i = (c_{1i} - \mu - \beta)/\omega$ and $k_i = (c_{2i} - \mu)/\omega$.

Let $\xi_{r,s} = E(x_i^r y_i^s)$. Then Shah and Parikh (1964) give the following recurrence formulae

$$\begin{aligned}\xi_{r,s} &= \frac{h_i^{r-1}}{L} Z(h_i) G_s(k_i, h_i\rho, \sqrt{1-\rho^2}) \\ &+ \frac{k_i^{s-1}}{L} Z(k_i) G_r(h_i, k_i\rho, \sqrt{1-\rho^2}) \\ &- \frac{h_i^{r-1} k_i^{s-1}\sqrt{1-\rho^2}}{\sqrt{2\pi}\, L} Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) \\ &+ (r-1)(s-1)(1-\rho^2)\xi_{r-2,s-2} + \rho(r+s-1)\xi_{r-1,s-1} \\ &\qquad r \geqslant 1,\ s \geqslant 1 \qquad (8.7)\end{aligned}$$

and

$$\xi_{r,0} = \frac{h_i^{r-1}}{L} Z(h_i)Q(K_i) + \frac{\rho}{L} Z(k_i)G_{r-1}(h_i, k_i\rho, \sqrt{1-\rho^2}) + (r-1)\xi_{r-2,0} \quad (8.8)$$

$$\xi_{0,s} = \frac{k_i^{s-1}}{L} Z(k_i)Q(H_i) + \frac{\rho}{L} Z(h_i)G_{s-1}(k_i, h_i\rho, \sqrt{1-\rho^2}) + (s-1)\xi_{0,s-2} \quad (8.9)$$

with $H_i = (h_i - \rho k_i)/\sqrt{1-\rho^2}$, $K_i = (k_i - \rho h_i)/\sqrt{1-\rho^2}$, $Z(u)$, $Q(u)$ as defined in section 4.4.2 and $L = L(h_i, k_i; \rho)$ of section 4.4.2. The function $G_r(a,b,c)$ is defined as

$$G_r(a,b,c) = \frac{1}{c}\int_a^\infty u^r Z\left(\frac{u-b}{c}\right)du$$

and equation (3.2) of Shah and Parikh gives the recurrence result for $G_r(a,b,c)$ as

$$G_r(a,b,c) = a^{r-1} cZ\left(\frac{a-b}{c}\right) + (r-1)c^2 G_{r-2}(a,b,c) + bG_{r-1}(a,b,c)$$

with

$$G_0(a,b,c) = Q\left(\frac{a-b}{c}\right). \quad (8.10)$$

Hence

$$E(F_{1i}^2 F_{2i}) = \omega^3 E(x_i^3 y_i) = \omega^3 \xi_{2,1} .$$

$$\xi_{2,1} = \frac{1}{L}\Bigg[h_i Z(h_i) G_1(k_i, h_i\rho, \sqrt{1-\rho^2})$$

$$+ Z(k_i) G_2(h_i, k_i\rho, \sqrt{1-\rho^2})$$

$$- \frac{h_i\sqrt{1-\rho^2}}{\sqrt{2\pi}} Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) + 2\rho L\xi_{1,0}\Bigg].$$

Now $$\xi_{1,0} = \frac{1}{L}\left[Z(h_i)Q(K_i) + \rho Z(k_i)Q(H_i)\right]$$

from (4.14) of Shah and Parikh. From (8.10) above is obtained

$$G_1(k_i, h_i\rho, \sqrt{1-\rho^2}) = \sqrt{1-\rho^2}\, Z(K_i) + h_i\rho\, Q(K_i)$$

and

$$G_2(k_i, h_i\rho, \sqrt{1-\rho^2}) = \sqrt{1-\rho^2}(k_i + h_i\rho) Z(K_i)$$
$$+ (1-\rho^2 + h_i^2\rho^2) Q(K_i) . \qquad (8.11)$$

Thus

$$E(F_{1i}^2 F_{2i}) = \frac{\omega^3}{L}\Bigg[\sqrt{1-\rho^2}\, h_i Z(h_i) Z(K_i)$$
$$+ \rho(h_i^2 + 2) Z(h_i) Q(K_i)$$
$$+ \sqrt{1-\rho^2}(h_i + k_i\rho) Z(k_i) Z(H_i)$$
$$+ (1-\rho^2 + k_i^2\rho^2 + 2\rho^2) Z(k_i) Q(H_i)$$
$$- \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}} h_i Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right)\Bigg] .$$

Now

$$Z(h_i) Z(K_i) = Z(k_i) Z(H_i) = \sqrt{1-\rho^2}\,\Phi(h_i, k_i; \rho)$$
$$= \sqrt{2\pi}\, Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) \qquad (8.12)$$

where $\Phi(h_i, k_i; \rho)$ is defined in section 4.4.2. Therefore

$$E(F_{1i}^2 F_{2i}) = \frac{\omega^3}{L} [\rho (2 + h_i^2) Z(h_i) Q(K_i)$$
$$+ (1 + k_i^2 \rho^2 + \rho^2) Z(k_i) Q(H_i)$$
$$+ (h_i + \rho k_i)(1 - \rho^2) \Phi(h_i, k_i; \rho)] \quad .$$

$$E(F_{1i}^3 F_{2i}) = \omega^4 E(x_i^3 y_i) = \omega^4 \xi_{3,1}$$

and

$$\xi_{3,1} = \frac{1}{L} \Bigg[h_i^2 Z(h_i) G_1(k_i, h_i \rho, \sqrt{1-\rho^2})$$
$$+ Z(k_i) G_3(h_i, k_i \rho, \sqrt{1-\rho^2})$$
$$- \frac{h_i^2 \sqrt{1-\rho^2}}{\sqrt{2\pi}} Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) + 3\rho L \xi_{2,0} \Bigg]$$

with

$$\xi_{2,0} = \frac{1}{L} \Bigg[h_i Z(h_i) Q(K_i)$$
$$+ \rho Z(k_i) G_1(h_i, k_i \rho, \sqrt{1-\rho^2}) + L \Bigg] \quad .$$

Using (8.10) to substitute for $G_3(h_i, k_i\rho, \sqrt{1-\rho^2})$ and (8.12) gives

$$\xi_{3,1} = 3\rho + \frac{1}{L} \Bigg[3\rho h_i Z(h_i) Q(K_i)$$
$$+ (2 + \rho^2) Z(k_i) G_1(h_i, k_i \rho, \sqrt{1-\rho^2})$$
$$+ h_i^2 Z(h_i) G_1(k_i, h_i \rho, \sqrt{1-\rho^2})$$
$$+ k_i \rho Z(k_i) G_2(h_i, k_i \rho, \sqrt{1-\rho^2}) \Bigg]$$

and using (8.11), $E(F_{1i}^3 F_{2i})$ is found to be

$$E(F_{1i}^3 F_{2i}) = 3\omega^4\rho + \frac{\omega^4}{L}[\rho(3 + h_i^2)h_i Z(h_i)Q(K_i)$$
$$+ \rho(3 + k_i^2\rho^2)k_i Z(k_i)Q(H_i)$$
$$+ (1 - \rho^2)(2 + \rho^2)\Phi(h_i, k_i; \rho) + k_i\sqrt{1 - \rho^2}\, Z(k_i)Z(H_i)$$
$$+ \rho(k_i + \rho h_i)\sqrt{1 - \rho^2}\, h_i Z(h_i)Z(K_i)] .$$

Now $E(F_{1i}^2 F_{2i}^2) = \omega^4 E(x_i^2 y_i^2) = \omega^4 \xi_{2,2}$ with

$$\xi_{2,2} = \frac{1}{L}\Bigg[h_i Z(h_i)G_2(k_i, h_i\rho, \sqrt{1-\rho^2})$$
$$+ k_i Z(k_i)G_2(h_i, k_i\rho, \sqrt{1-\rho^2}) + (1 - \rho^2)L$$
$$- \frac{h_i k_i\sqrt{1-\rho^2}}{\sqrt{2\pi}} Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) + 3\rho L \xi_{1,1}\Bigg]$$

and

$$\xi_{1,1} = \frac{1}{L}\Bigg[Z(h_i)G_1(k_i, h_i\rho, \sqrt{1-\rho^2})$$
$$+ Z(k_i)G_1(h_i, k_i\rho, \sqrt{1-\rho^2})$$
$$- \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}} Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right) + \rho L\Bigg] .$$

Thus

$$\xi_{2,2} = (1 + 2\rho^2) + \frac{1}{L}\Bigg[h_i Z(h_i)G_2(k_i, h_i\rho, \sqrt{1-\rho^2})$$
$$+ 3\rho Z(h_i)G_1(k_i, h_i\rho, \sqrt{1-\rho^2})$$
$$+ k_i Z(k_i)G_2(h_i, k_i\rho, \sqrt{1-\rho^2})$$
$$+ 3\rho Z(k_i)G_1(h_i, k_i\rho, \sqrt{1-\rho^2})$$

$$- \frac{(h_i k_i + 3\rho)}{\sqrt{2\pi}} \sqrt{1-\rho^2}\, Z\left(\sqrt{\frac{h_i^2 - 2\rho h_i k_i + k_i^2}{1-\rho^2}}\right)\Bigg]$$

and using (8.11) and (8.12) gives

$$\begin{aligned}
\xi_{2,2} &= (1 + 2\rho^2) \\
&+ \frac{1}{L}\Big[\sqrt{1-\rho^2}[h_i k_i + \rho (h_i^2 + k_i^2) + 3\rho] Z(h_i) Z(K_i) \\
&+ (h_i + h_i^3 \rho^2 + 2h_i \rho^2) Z(h_i) Q(K_i) \\
&+ (k_i + k_i^3 \rho^2 + 2k_i \rho^2) Z(k_i) Q(H_i)\Big] .
\end{aligned}$$

Thus

$$\begin{aligned}
E(F_{1i}^2 F_{2i}^2) &= \omega^4 (1 + 2\rho^2) \\
&+ \frac{\omega^4}{L} [(1 + 2\rho^2 + h_i^2\rho^2) h_i Z(h_i) Q(K_i) \\
&+ (1 + 2\rho^2 + k_i^2\rho^2) k_i Z(k_i) Q(H_i) + 3\rho (1-\rho^2) \Phi(h_i, k_i; \rho) \\
&+ (k_i + \rho h_i) h_i \sqrt{1-\rho^2}\; Z(h_i) Z(K_i) \\
&+ \rho\sqrt{1-\rho^2}\; k_i Z(k_i) Z(H_i)] .
\end{aligned}$$

$E(F_{1i}^3) = \omega^3 \xi_{3,0}$ and using (8.8)

$$\begin{aligned}
\xi_{3,0} &= \frac{h_i^2}{L} Z(h_i) Q(K_i) + \frac{\rho}{L} Z(k_i) G_2 (h_i, k_i \rho, \sqrt{1-\rho^2}) \\
&+ 2\xi_{1,0} \\
&= \frac{1}{L}\Big[(2 + h_i^2) Z(h_i) Q(K_i) \\
&+ \rho\sqrt{1-\rho^2}(h_i + k_i \rho) Z(k_i) Z(H_i) \\
&+ \rho (3 - \rho^2 + k_i^2\rho^2) Z(k_i) Q(H_i)\Big] .
\end{aligned}$$

Therefore

$$E(F_{1i}^{3}) = \frac{\omega^3}{L}\left[(2 + h_i^2)Z(h_i)Q(K_i)\right.$$

$$+ \rho(3 - \rho^2 + k_i^2\rho^2)Z(k_i)Q(H_i)$$

$$\left. + (h_i + \rho k_i)\rho(1 - \rho^2)\Phi(h_i, k_i; \rho)\right] \quad .$$

Finally

$$E(F_{1i}^{4} + F_{2i}^{4}) = \omega^4(\xi_{4,0} + \xi_{0,4}) \text{ with}$$

$$\xi_{4,0} + \xi_{0,4} = \frac{1}{L}\left[h_i^3 Z(h_i)Q(K_i)\right.$$

$$+ \rho Z(k_i)G_3(h_i, k_i\rho, \sqrt{1-\rho^2}) + k_i^3 Z(k_i)Q(H_i)$$

$$\left. + \rho Z(h_i)G_3(k_i, h_i\rho, \sqrt{1-\rho^2}) + 3(\xi_{2,0} + \xi_{0,2})\right].$$

Now

$$G_3(k_i, h_i\rho, 1-\rho^2) = k_i^2\sqrt{1-\rho^2}\, Z(K_i)$$

$$+ 2(1-\rho^2)G_1(k_i, h_i\rho, \sqrt{1-\rho^2})$$

$$+ h_i\rho\, G_2(k_i, h_i\rho, \sqrt{1-\rho^2})$$

$$= [k_i^2 + 2(1-\rho^2) + h_i\rho(k_i + h_i\rho)]\sqrt{1-\rho^2}\, Z(K_i)$$

$$+ h_i\rho\,[3(1-\rho^2) + h_i^2\rho^2]Q(K_i)$$

and

$$\xi_{2,0} + \xi_{0,2} = \frac{1}{L}\left[h_i Z(h_i)Q(K_i) + k_i Z(k_i)Q(H_i)\right.$$

$$+ \rho Z(k_i)G_1(h_i, k_i\rho, \sqrt{1-\rho^2})$$

$$\left. + \rho Z(h_i)G_1(k_i, h_i\rho, \sqrt{1-\rho^2}) + 2L\right]$$

$$= 2 + \frac{1}{L}\left[h_i(1+\rho^2)Z(h_i)Q(K_i)\right.$$

$$\left. + k_i(1+\rho^2)Z(k_i)Q(H_i) + 2\rho\sqrt{1-\rho^2}\, Z(h_i)Z(K_i)\right].$$

Thus

$$\int^{\rho}_{4,0} + \int^{\rho}_{0,4} = 6 + \frac{1}{L}\Big[\{h_i^3 + 3h_i(1+\rho^2)$$
$$+ \rho^2 h_i[3(1-\rho^2) + h_i^2\rho^2]\} Z(h_i)Q(K_i)$$
$$+ \{k_i^3 + 3k_i(1+\rho^2) + \rho^2 k_i[3(1-\rho^2)$$
$$+ k_i^2\rho^2]\} Z(k_i)Q(H_i) + \rho\sqrt{1-\rho^2}\,[h_i^2$$
$$+ 4(1-\rho^2) + k_i\rho(h_i + k_i\rho) + k_i^2$$
$$+ h_i\rho(k_i + h_i\rho) + 6]Z\left(\sqrt{\frac{2h_i k_i}{1+\rho}}\right) Z\left(\frac{k_i - h_i}{\sqrt{1-\rho^2}}\right)\Big]$$

as

$$Z(h_i)Z(K_i) = Z\left(\sqrt{\frac{2h_i k_i}{1+\rho}}\right) Z\left(\frac{k_i - h_i}{\sqrt{1-\rho^2}}\right).$$

Hence

$$E(F_{1i}^4 + F_{2i}^4) = 6\omega^4$$
$$+ \frac{\omega^4}{L}\Big[(h_i^2 + 3 + h_i^2\rho^4 - 3\rho^4 + 6\rho^2)h_i Z(h_i)Q(K_i)$$
$$+ (k_i^2 + 3 + k_i^2\rho^4 - 3\rho^4 + 6\rho^2)k_i Z(k_i)Q(H_i)$$
$$+ \rho\sqrt{1-\rho^2}\,[(1+\rho^2)(h_i^2 + k_i^2) + 2h_i k_i\rho + 10 - 4\rho^2] \times$$
$$Z\left(\sqrt{\frac{2h_i k_i}{1+\rho}}\right) Z\left(\frac{k_i - h_i}{\sqrt{1-\rho^2}}\right)\Big].$$

CHAPTER 9

REFERENCES

Baker,R. and Nelder J. (1978). The GLIM System, Release 3. Oxford: Numerical Algorithms Group.

Batchelor,J.R and Hackett,M. (1970). H.L.A. matching in the treatment of burned patients with skin allografts. Lancet, 2 ,581-83.

Bennett,S and Whitehead,J. (1981). Fitting logistic and log-logistic regression models to censored data using GLIM. GLIM newsletter, No. 4, 12-19.

Cederlof,R., Fiberg,L and Hrubec,Z (1969). Cardiovascular and Respiratory symptoms in Relation to Tobacco Smoking. A study on American Twins. Arch. Env. Health 18 ,934.

Clayton, D.G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65, 141-51.

Cox, D.R. (1972). Regression models and life tables (with discussion). J. R. Statist. Soc. B , 34, 187-220.

De Faire,U. (1974). Ischaemic Heart Disease in death discordant twins. A study on 205 male and female pairs. Acta Med. Scand. Suppl. 568.

Dempster,A.P., Laird,N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. J. R. Statist. Soc. B, 39, 1-38.

Fischer,M., Harrald, B., and Hauge,M (1969). A Danish twin study of Schizophrenia. Brit. J. Psychiat., 115, 981-90.

Fisher,R.A. (1958). Statistical Methods and Scientific Inference 2nd edition. Edinburgh,Oliver and Boyd.

Hill, I.D. (1973). Algorithm AS 66. The normal integral. Appl. Statist., 22, 424-7.

Holt, J.D. and Prentice, R.L. (1974). Survival analyses in twin studies and matched pair experiments. Biometrika , 61, 17-30.

Kalbfleisch, J.D. and Prentice, R.L. (1973). Marginal likelihoods based on Cox's regression and life model. Biometrika 60, 267-78.

Kuhback,B. and Tiilikainen,A. (1975). Results in necrokidney transplantation to pair recipients. Scand. J. Urol.

Nephrol. Suppl. 29,117-20.

Liljefors,I (1970). Coronary Heart Disease in Male Twins. Acta. Med. Scand. Suppl. 511.

Louis, T.A. (1982). Finding the observed information matrix when using the EM algorithm. J. R. Statist. Soc. B, 44, 226-233.

Mak, T.K. and Ng,K.W. (1981). Analysis of familial data: Linear-model approach. Biometrika, 68, 457-61.

Mantel,N., Bohidar, N.R. and Ciminera, J.L. (1977). Mantel-Haenszel analyses of litter-matched time to response data, with modifications for recovery of interlitter information. Cancer Res., 37, 3863-68.

Mantel, N. and Ciminera, J.L. (1979). Use of Logrank Scores in the analysis of litter matched data on time to tumour appearance. Cancer Res. 39, 4308-15.

NAG Library (Mark 10) (1983). Oxford Numerical Algorithms Group.

Neyman, J. and Scott, E.L. (1948). Consistent estimates based on partially consistent observations. Econometrika, 16, 1-32.

Oakes, D. (1982). A model for association in bivariate survival data. J. R. Statist. Soc. B, 44, 414-422.

Roger,J.H. and Peacock,S.D. (1983). Fitting the scale as a GLIM parameter for Weibull, extreme value, logistic and log-logistic regression models with censored data. GLIM newsletter, No. 6, 30-37.

Sampford, M.R. and Taylor,J. (1959). Censored observations in randomised block experiments. J. R. Statist. Soc. B, 21, 214-37

Shah, S.M. and Parikh, N.T. (1964). Moments of singly and doubly truncated standard bivariate normal distribution. Vidya (Gujarat University), 7, 82-91.

The Diabetic Retinopathy Study Research Group (1981). Diabetic Retinopathy Study, Investigative Ophthalmology and Visual Science, 21, part 2, 149-226.

Wei, L.J. (1980). A generalised Gehan and Gilbert test for the two-sample problem in the presence of right censoring. JASA, 75, 634-638.

WHO (1966). The use of Twins in Epidemiological Studies. Report of a WHO meeting of investigators. Acta Gen. Med. et Gemellol (Rome) 15, 109.

Wild, C.J. (1983). Failure time models with matched data. Biometrika, 70, 633-41.

Wolynetz, M.S. (1979a). AS 138 . Maximum likelihood estimation from confined and censored Normal data. Appl. Statist, 28, 185-95.

Wolynetz,M.S. (1979b). AS 139. Maximum Likelihood Estimation in a linear model from confined and censored Normal data. Appl. Statist., 28, 195-206.

Woolson,R.F and Lachenbruch, P.A. (1980). Rank tests for censored matched pairs. Biometrika, 67, 597-607.

MODELS FOR PROGNOSTIC VARIABLES IN MATCHED GROUPS WITH CENSORED DATA

C. Jagger

(Submission for the degree of Doctor of Philosophy, University of Leicester, 1984)

ABSTRACT

Previous methods of analysis of matched data have used some form of conditioning to eliminate the terms describing the matching, usually considered as nuisance parameters. If the outcome measured can be subject to censoring, these methods break down for those n-tuplets containing only censored observations. These cases have previously been omitted from analyses, by arguing that these n-tuplets do not contain any information about the parameters representing the effect of the prognostic variables although in some studies this can result in ignoring a large part of the data set and the estimation of the parameters can be subject to bias.

These problems have been overcome by introducing a prior distribution on the matching variables. Two classes of model were considered, the proportional hazards model with Weibull (and as a special case exponential) failure times and the normal theory accelerated failure model. Two different censoring mechanisms, fixed and random, were also investigated.

Although omitting n-tuplets with all observations censored was known to produce biased estimates of the location parameter with exponential failure times, it was shown that this bias was transferred to the scale parameter when the failure times were Weibull. Previous rank methods which gave little bias with a fixed censoring mechanism were found to give much more unreliable estimates when the censoring mechanism was random. The new method compared well with the previous methods, producing less biased results with both censoring mechanisms. The robustness of the assumptions of the prior distribution of the matching variables was investigated.

The new method of analysis for data from the accelerated failure model was based upon the EM algorithm. Similar results were found to those of the proportional hazards model.

Extensions to other failure time distributions for the proportional hazards model and further research areas are discussed.