

R A N D O M E F F E C T S
I N
S U R V I V A L A N A L Y S I S

Kelvin Kai Wing Yau

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

July 1995

ACKNOWLEDGEMENTS

During the course of preparation of this thesis, I have had assistance from a number of people and I would like to express my deep thanks to them.

I am indebted to my academic supervisor Professor C.A. McGilchrist, who has given me invaluable assistance and guided me throughout this thesis. His kindness, patience encouragement and enlightened suggestions has been greatly appreciated.

Special thanks go to my co-supervisors Dr. J. McCallum and Dr. A. H. Welsh for their valuable help and suggestions.

Thanks also go to my friend Mr. A. Saei for his assistance in computing and valuable discussion during my PhD study at NCEPH.

I would like to acknowledge gratefully the financial assistance from the National Health and Medical Research Council and the ANU PhD scholarship given by the National Centre for Epidemiology and Population Health.

Finally, I would like to apologize to my wife Sau Fong and my son Alex for the hardships I put them through over the past few years. I would also like to thank my parents for giving me the opportunity to achieve. To them, I dedicate this thesis.

ABSTRACT

A general method for fixed effect regression estimation in proportional hazards models was introduced by Cox (1972). In the past twenty years, this method has been widely used and applied to survival regression models throughout the clinical trial literature.

Problems arise with the Cox model when there is more than one observation of failure time for each patient or when we are interested in failure observations within a cluster. In each case, the independent event times assumption in the Cox model is no longer valid since dependence between failure time observations that come from the same patient or within each cluster is anticipated.

The aim of this thesis is to develop random effect survival models which extend Cox's partial likelihood method to handle multivariate failure time data based on the Generalised Linear Mixed Model (GLMM) approach.

The GLMM approach, on the one hand, preserves the cancellation property of the baseline hazard function in the partial likelihood expression. On the other hand, it provides predictions of random effects. Such predictions are useful to identify a high risk family or high risk individual when considering a genetic disease.

Further theoretical developments of the GLMM, where the random components are correlated, are expounded. Computable expressions are obtained for the likelihood function and its derivatives and also for the information matrix. Such developments serve as the basis of the estimation and inference procedure in the random effect survival models.

Various random effect survival models are developed for the analysis of different types of multivariate failure time data. The Chronic Granulomatous Disease (CGD) data, the litter matched tumorigenesis experiment data and the Dubbo study data are used to illustrate the application of these various models. The GLMM method is found to be successful in analysing multivariate failure time data. Consistent results are obtained when comparing with other methods. Simulation results further confirm the implementability of current method.

The development of the GLMM with correlated random components and the various random effect survival models established to analyse different types of multivariate failure time data are new and original techniques.

CONTENTS

	page
1 Introduction	1
1.1 Univariate survival analysis	1
1.2 Cox's proportional hazards model	2
1.3 The multivariate failure time problem	3
1.4 Different approaches	4
1.5 Research approach	6
1.6 Aim and outline of chapters	7
2 Review: Four types of multivariate failure time data	10
2.1 Type I: The failure occurrence in matched components of a patient	11
2.2 Type II: The occurrence of the same type of failure in the member of a family/cluster	13
2.3 Type III: The recurrent events in the same patient	18
2.4 Type IV: The occurrence of different failure events in the same patient	24
3 Review: Different approaches to the multivariate failure time problem	26
3.1 Univariate Cox's regression model in survival analysis	27
3.2 The frailty models approach	28
3.3 The marginal modelling approach	32
3.4 The counting processes approach	35
3.5 Generalised linear mixed model approach	38

4 Generalised linear mixed models	43
4.1 Introduction	44
4.2 Normally distributed random components	46
4.3 BLUP estimation	47
4.4 Some matrix results	49
4.5 Maximum likelihood estimation	54
4.6 Residual maximum likelihood estimation	57
4.7 Generalised linear mixed models	59
4.8 Application to the multivariate failure time problem	62
5 Time dependent frailty model	65
5.1 Introduction	65
5.2 Models	67
5.3 Estimation	71
5.4 Application	76
6 AR(1) frailty model	81
6.1 Model and estimation	81
6.2 Asymptotic variances	86
6.3 AR(1) frailty model	87
6.4 Application	90
7 Random block frailty model	93
7.1 Introduction	93
7.2 Baseline frailty model	95
7.3 Justification of the exponential relative risk function	100

7.4 Random block frailty model	102
7.5 Discussion	110
7.6 Simulations	112
8 Competing risk frailty model	127
8.1 Introduction	127
8.2 Models	129
8.3 Estimation	132
8.4 Application	138
8.5 Discussion	142
9 Further discussion	144
9.1 Preliminaries	144
9.2 Some remarks	145
9.3 Main difficulty	145
9.4 Unsolved problems	147
9.5 Research suggestions	149
Bibliography	151
Appendix I	167
Appendix II	180

CHAPTER ONE

INTRODUCTION

1.1 UNIVARIATE SURVIVAL ANALYSIS

Much research has been done on survival analysis, with various models being developed to analyse survival data. In survival analysis, the response variable is usually the death time of an individual in the study population. The main concern is to estimate the underlying death density function or equivalently the survivor function of the study population based on these death time observations.

There are two major types of models: non-regression and regression. In the former, there are both parametric and non-parametric methods for the estimation of death density or survivor function. For parametric methods, different distributions are used to model the survivor function. Some commonly used distributions are exponential, Weibull, Gamma and log-logistic. On the other hand, the Kaplan-Meier estimate is a non-parametric estimator of the survivor function. When regression models in survival analysis are considered, it is important to set up the relationship between the hazard of death and the explanatory variables as well as estimating the survivor distribution.

In the 1970's, a wide range of regression models in survival analysis were proposed. These may be classified broadly as the proportional hazards model and the accelerated failure time model. Techniques of

regression models used in survival analysis are well-documented in textbooks such as Kalbfleisch and Prentice (1980), Lawless (1982) and Cox and Oakes (1984). These may be classified as "univariate". Here, the term "univariate" refers to those survival data that have only one failure observation from each individual; and we can assume that the failure time observation of individuals are independent in the univariate situation.

1.2 COX'S PROPORTIONAL HAZARDS MODEL

In Cox (1972), a general method for fixed effect regression estimation in proportional hazards models was introduced and applied to survival analysis. There have been some hundreds of publications in the statistical literature applying and extending this basic technology and equally many applications in medical research. In fact, this method is now a standard medical research tool applied to survival regression models.

In Cox's proportional hazards model, a partial likelihood function is constructed. By maximizing this partial likelihood with respect to the regression parameters, we obtain the estimates of regression parameters and their asymptotic variances.

A convenient feature of the Cox model is that the baseline hazard function is cancelled out in constructing the partial likelihood and is then not involved in the estimation procedure of the regression parameters. Because of this feature, the Cox model is sometimes called a

semiparametric method since the distribution of the baseline hazard function is unspecified.

1.3 THE MULTIVARIATE FAILURE TIME PROBLEM

Problems arise with the Cox model when there is more than one observation of failure time for each patient or when we are interested in the failure observations within a cluster. Examples are the survival data that treat the failure to be the serious infection of certain disease in patients or the occurrence of a genetic disease in the relatives of a single proband.

These kinds of survival data are termed multivariate failure time data, for which there will be a detailed explanation in Chapter 2. What needs to be mentioned here, however, is the problem in analysing multivariate failure time data. Many regression variables, which are unrecorded and unrecognised as risk variables, may each contribute marginally to the total variability. The accumulation of such effects is often termed the "frailty" of the patient/cluster and this "frailty" accounts for correlations between observations on the same patient/cluster. This being so, when we consider the regression analysis in multivariate failure time data, the independence assumption in the univariate survival analysis is no longer valid.

The main difficulty in analysing multivariate failure time data is that the usual cancellation property of the baseline hazard function in

the partial likelihood procedure is lost if the marginal failure time distribution is obtained by integrating out the frailty components.

1.4 DIFFERENT APPROACHES

There are quite different approaches to solving the difficulty outlined above. Detailed explanation of these will be given in Chapter 3, though a brief description of them here is useful.

The derivation of certain failure time distributions has been done by Hougaard (1984, 1986a, 1986b, 1991) and Aalen (1992). Clayton and Cuzick (1985a) have provided a multivariate generalisation of survival analysis for some types of data. Developments of frailty models were given by Clayton (1978, 1991), Oakes (1982, 1989), Wild (1983) and Klein (1992).

A method for stratified data was investigated by Prentice, Williams and Peterson (1981) and has been extended by Wei, Lin and Weissfeld (1989) by modelling the marginal distribution. Such a marginal modelling approach has been further developed by Lee, Wei and Amato (1992), Lin and Wei (1992) and Lee, Wei and Ying (1993). Related work on multiple times to tumour can be found in Gail, Santner and Brown (1981).

The counting processes approach or the martingale approach aims to generalize the traditional Cox's model to the multivariate failure time problem based on the martingale and stochastic integral theory. The statistical theory of counting processes has been developed by Aalen

(1976, 1978). Andersen and Gill (1982) developed the multiplicative intensity model. This model may be seen as the multivariate version of Cox's regression model based on the martingale approach.

Based on this approach, Gamma distributed frailty models in survival analysis were developed by Nielsen, Gill, Andersen and Sørensen (1992). Self (1993) considered a time independent and a time dependent gamma distributed frailty model to analyse multiple failure time data. These two models are applied to the Chronic Granulomatous Disease (CGD) Data which appears in Fleming and Harrington (1991).

The Generalised Linear Mixed Model (GLMM) approach tries to combine the generalised linear model (GLM) technique of McCullagh and Nelder (1989) and the linear mixed model (LMM) to obtain a broader unified class of models which allows the non-identity link function with random effects in the linear predictor.

Schall (1991) developed an approximate ML and REML estimation algorithm. Breslow and Clayton (1993) considered both the Penalised quasi-likelihood (PQL) and the Marginal quasi-likelihood (MQL). Other comparable developments can be found in Solomon and Cox (1992) and Wolfinger and O'Connell (1993) using the Laplace approximation and the Psuedo-likelihood (PL) method respectively. Based on the quadratic approximation in the region of maximum, McGilchrist (1994) presented a detailed derivation of the estimators and its possible application area.

1.5 RESEARCH APPROACH

The approach adopted for this research follows the method described in McGilchrist (1994), and referred to as the GLMM mentioned above. Basically, the method starts with the Best Linear Unbiased Predictor (BLUP) of Henderson (1963, 1973, 1975), and then connects to the Residual Maximum Likelihood (REML) of Patterson and Thompson (1971). The link between BLUP and REML is outlined in Harville (1977) and is fully detailed in Thompson (1980), Fellner (1986, 1987) and Speed (1991).

In the framework of the GLMM, it turns out that random effect models in survival analysis can be seen as one application of the GLMM type of modelling as outlined in McGilchrist (1994). This thesis adopts the GLMM approach to develop the estimation and inference techniques in different random effect models in survival analysis.

When applying the GLMM method to random effect survival models, the cancellation property of the baseline hazard function in the partial likelihood expression is preserved. Another important feature of the GLMM approach is that other than the estimation of regression parameters, the prediction of random effects is also obtained. In some cases, the prediction of random effects are important and often of separate interest. The following example can be used to illustrate this importance: when considering a genetic disease it is as important to identify a high risk family or high risk individual as to estimate the risk variable parameters.

1.6 AIM AND OUTLINE OF CHAPTERS

The aim of this thesis is to develop random effect survival models that extend Cox's partial likelihood method to handle multivariate failure time data based on the Generalised Linear Mixed Model (GLMM) approach. Various random effect survival models are developed according to the different types of multivariate failure time data being classified in Chapter 2.

This Chapter is followed by a review of the multivariate failure time problem in survival analysis. The multivariate failure time data are classified into four types : (a) the failure occurrence in matched components of a patient, (b) the occurrence of the same type of failure in the member of a family/cluster, (c) the recurrent events in the same patient and (d) the occurrence of different failure events in the same patient.

In Chapter 3, initially the well known univariate Cox regression model in survival analysis is discussed briefly. This is followed by a review of different approaches to the multivariate failure time problem. These approaches are classified into : (a) the frailty models approach, (b) the marginal modelling approach, (c) the counting processes approach and (d) the GLMM approach.

Chapter 4 provides a further generalisation of the GLMM which allows possible correlation between the random effects. Estimation of

correlation parameters in the variance matrix of the random components as well as the fixed effect parameters , their asymptotic variances and the variance components are obtained. The development in this chapter serves as the basis of the estimation and inference procedure in the random effect survival models that established in the following chapters. According to the four types of multivariate failure time data, different random effect survival models are developed in Chapters 5 to 8.

Chapters 5 and 6 deal with Type III data (the recurrent events in the same patient). Chapter 5 provides three longitudinal models which are specialized for the time dependent frailty survival data and are developed to analyse the Chronic Granulomatous Disease (CGD) data set (Appendix I). In Chapter 6, the random effect at each failure time of a patient is treated as a time series process. In particular, the random effects are considered to follow an AR(1) process. The application of the AR(1) frailty model to the CGD data is also given.

Chapter 7 establishes the modelling of Type II data (the occurrence of the same type of failure in the member of a family/cluster). Two random effect survival models : (a) the baseline frailty model and (b) the random block frailty model are described. These can be used to analyse failure observations occurring in a cluster/family. A method to justify the usual exponential relative risk function is also proposed. The litter matched tumorigenesis experiment data (Appendix I) is used to illustrate these developments. Some simulation results of both the baseline frailty model and the random block frailty model are also made available. Moreover, the techniques developed in this chapter can also be used to

analyse Type I data (the failure occurrence in matched components of a patient). This will be explained in Section 7.5.

The competing risk frailty models are investigated in Chapter 8. They are developed for the analysis of Type IV data (the occurrence of different failure events in the same patient). The Dubbo study data (Appendix I) are used to illustrate these models. The data come from research on elderly people in Dubbo by the University of New South Wales Lipid Research Department, St. Vincent's Hospital, Sydney. In the Dubbo study data, both time to hospitalization and death are important outcome measures. It is noteworthy that, for each individual, the hospitalization could be a recurrent event and the occurrence of death should be later than hospitalization. Therefore, we have to model the dependence structure that comes from the multiple failure time observations in each individual and also the possible relationship between an individual's frailty in hospitalization and his frailty in death.

Chapter 9 details the difficulties faced during the research process, some unsolved problems and suggestions for further research. The three data sets, used to illustrate the random effect survival models developed in these Chapters, are provided in Appendix I. The corresponding APL programs for these models can be found in Appendix II.

CHAPTER TWO

REVIEW: FOUR TYPES OF MULTIVARIATE FAILURE TIME DATA

As stated in Chapter 1, this thesis focuses on the extension of Cox's partial likelihood method to multivariate failure time data based on the Generalised Linear Mixed Model approach. Essentially, Cox's method assumes the failure time observations to be independent, so that the estimation and inference technique follows the description in Section 3.1. But, in practice, there are some failure time data appearing which do not satisfy the independent event times assumption.

In some clinical trials data, when the failure time is not restricted to the death time, failure may refer to a recurrent serious infection, repeated hospitalization, the experience of visual loss in the left and right eyes in a patient, the occurrence of a genetic disease within the same family or the failure episode in the same cluster. These kinds of failure time data are called multivariate failure time data since each individual may experience more than one failure or there may be more than one failure observation in each cluster/family.

These multivariate failure time data have a common feature, which is the dependence or the correlation between the failure times within the same patient/cluster. Therefore, the main concern in analysing multivariate failure time data is to study the dependence between failure times and the effect of possible explanatory variables on the failure times in the presence of dependence.

The remainder of this chapter describes twelve multivariate failure time data sets taken from different studies. In each case, the objective is to explain the presence of dependence between the failure time observations that appear in different studies and the methods that have been used to analyse these data. For this reason, mention is made of main points only; further detail can be obtained by referring to the papers referenced.

The multivariate failure time data are classified into four types:

- (a) the failure occurrence in matched components of a patient,
- (b) the occurrence of the same type of failure in the member of a family/cluster,
- (c) the recurrent events in the same patient and
- (d) the occurrence of different failure events in the same patient.

2.1 TYPE I : THE FAILURE OCCURRENCE IN MATCHED COMPONENTS OF A PATIENT

In medical research, Type I data may be the experiences of visual loss in left and right eyes, or similarly, the time to failure of a right and left kidney.

Data set (1) -- The Diabetic Retinopathy study

Diabetic retinopathy is a complication associated with diabetes mellitus consisting of abnormalities in the microvasculature within the

retina of the eye. It is the major cause of visual loss in patients under 60 years of age in the United States and many industrialized countries.

The Diabetic retinopathy study (DRS) conducted by the National Eye Institute in the United States began in 1971. The study was concerned with the effectiveness of laser photocoagulation in delaying the onset of blindness in patients with diabetic retinopathy. A total of 1742 patients with diabetic retinopathy in both eyes and visual acuity of 20/100 or better in both eyes entered the study between 1972 and 1975. One eye of each patient was randomly selected for treatment (photocoagulation) and the other eye was observed without treatment. The patients were followed over several years for the occurrence of blindness in the left and right eyes. The study end point (blindness) was defined as being at the first occurrence of visual acuity less than 5/200.

The main purpose of the study was to assess the effectiveness of the laser photocoagulation treatment in the presence of possible dependence between the left and right eyes. Moreover, it was considered important to set up models to investigate the dependence between the two eyes within the same patient and the relationship between the treatment effect and the type of diabetes.

A subset of the original data ($N=197$), which was the 50% sample of the high-risk patients as defined by DRS criteria, has been analysed by Huster, Brookmeyer and Self (1989) using a fully parametric model and an independence working model in the analysis of paired censored survival data.

Further analysis of this subset of data has been done by Liang, Self and Chang (1993) using marginal modelling approach. Lin (1994) has provided a further discussion on this data set. Analyses of parametric and semiparametric frailty models in bivariate survival data have further been undertaken by Oakes (1994). In addition, two new methods, one building on the marginal models of Wei, Lin and Weissfeld (1989) and the other modifying the EM algorithm of Dempster, Laird and Rubin (1977) have been suggested in his paper.

2.2 TYPE II : THE OCCURRENCE OF THE SAME TYPE OF FAILURE IN THE MEMBERS OF A FAMILY/CLUSTER

Type II data may refer to the occurrence of a genetic disease among family members, the appearance of tumors in littermates exposed to a carcinogen, or more generally, the failure occurrence within a group with members who have experienced some common risk factors. The dependence structure is due to the fact that individuals within the same family/cluster share some risk factors due to genetic or environmental effects. Therefore, the individuals within the same family/cluster may be considered sharing a common, unobservable, random frailty.

Data set (2) -- The litter matched tumorigenesis experiment

The litter matched tumorigenesis experiment data were described by Mantel, Bohidar and Ciminera (1977) and Mantel and Ciminera (1979). In

the design, there were 50 male litters and 50 female litters. In each litter, one rat was treated with putative carcinogen while the other two rats served as controls.

The experiment was followed for 104 weeks and the failure time was the time to tumor occurrence or censoring as recorded to the nearest week. It was conceivable that the environmental conditions shared within litters would affect the risk of tumor formation. Therefore, the environmental conditions in each litter may contribute to the hazard function as a random litter effect.

Analysis of this data set based on the martingale approach using EM algorithm has been done by Nielsen, Gill, Andersen and Sørensen (1992). A subset of the data which considered only the female litters has been analysed by Hougaard (1986b) using the parametric Weibull margins model and a Cox type model. Clayton (1991) also analysed this subset of data using Gibb's sampling. In addition, a comparison of the estimated treatment effect and the likelihood ratio statistics in these various approaches has been provided by Nielsen, Gill, Andersen and Sørensen (1992).

Data set (3) -- The Schizophrenia study

The Schizophrenia study was first reported by Pulver and Liang (1991). This was a genetic epidemiology study of schizophrenia conducted by Dr. A.E. Pulver of John Hopkins University. 487 first-degree relatives (273 males, 214 females) of 93 female schizophrenic probands enrolled in

the study. For a single proband, the number of relatives ranged from 1 to 12.

Two main covariates considered were the gender of the relative and the proband's age, which was dichotomized at 16 years. The failure time was the age at diagnosis of effective illness for the relative. Here, the effective illness was defined as either depression, mania or both. 31 failure events were recorded out of the 487 relatives.

The main research question was whether the risk of effective illness of the relatives was associated with the age at onset of schizophrenia of probands while adjusting for the gender of the relatives. In this data set, it was anticipated that the relatives of the same proband share genetic effects and possibly environmental effects, so the times to effective illness were correlated among them.

The data set has been analysed by Liang, Self and Chang (1993) and Lin (1994). They both used the marginal approach to analyse the multivariate failure time data with a slightly different expression in the score function.

Data set (4) -- The Framingham study

The Framingham study was a cohort of 2336 men and 2873 women, aged between 30 and 62, that began in 1948. Details of the study can be found in Dawber (1980). Individuals had their first examination at study entry and subsequently recalled and examined every two years. The failure time

was the time to death. The various causes for death and the times to cardiovascular disease and cancer were recorded.

During the follow up period, possible risk variables such as gender, blood pressure, cholesterol levels, height and smoking behaviour were recorded in detail as well. The research interest centred on the assessment of the effect of various risk factors on the overall mortality, the cause-specific mortality and the time to the different types of diseases.

Analysis and discussion of this data set has been provided by Schatzkin, Cupples, Heeren, Morelock and Kannel (1984), Wolf, Abbott and Kannel (1987), Wolf, D'Agostino, Kannel and Belanger (1988) and Anderson, Castelli and Levy (1987) based on the independent event times assumption.

Klein (1992), however, did not employ the independent event times assumption when considering the random effects of Cox's regression model. The estimation process was based on the EM algorithm by treating the unobservable random frailty as missing value. Two possible cluster dependence structures were considered. The first cluster dependence structure was obtained by grouping individuals according to siblings and the second cluster dependence structure considered each married couple to be a cluster. The frailty in the first grouping could be considered as the combined effect of shared genes between siblings and the early environmental effects on siblings while the frailty in the second grouping was considered to be the late environmental effects that were shared by a married couple.

Data set (5) -- The Skin Allograft study

The Skin allograft study data was first published by Batchelor and Hackett (1970). This study was concerned with the differences in survival times of skin grafts from 16 severely burned patients as these related to closely or poorly matched HL-A transplantation antigen system between patients and donors. The patients received skin allografts from 2, 3 or 4 donors. Basically, the donors and the patients were matched for ABO blood groups. But they may have been either closely or poorly matched for the HL-A transplantation antigen system.

By the end of the study, all the skin allografts would have been destroyed due to the immune response of the patient. It was therefore important that the allografts survived as long as possible so as to reduce the chance of infection in the patient and to provide the time for the skin to grow in the same donor.

The main interest was to see whether there was any difference in the survival times for closely and poorly matched allografts. Moreover, due to the heterogeneity in patients, the immune response might well differ between patients. Therefore, other than the possible risk factors (e.g. matching type, amount of burn), the random patient effect also needed to be considered.

This data set has been analysed by Holt and Prentice (1974) and Kalbleisch and Prentice (1980) have analysed a subset of the data. Both

of these analyses used a model for survival in matched pairs. The counting processes frailty model approach to this data set was provided by Nielsen, Gill, Andersen and Sørensen (1991).

2.3 TYPE III : THE RECURRENT EVENTS IN THE SAME PATIENT

This type of data set is very common in medical research. Examples of multiple recurrent events are the sequence of asthmatic attacks, epileptic seizures, infection episodes, tumor recurrences or bleeding incidents in individual patients. The CGD data in Appendix I are of this type. For this type of data set, the analysis of regression effects in the presence of dependence is of concern. Such dependence is due to the multiple failure time observations that come from the same patient.

Data set (6) -- The CGD study

Chronic Granulomatous Disease (CGD), as described in Fleming and Harrington (1991), is a group of inherited rare disorders of the immune function characterized by recurrent pyogenic infections which usually present early in life and may lead to death in childhood.

Between October 1988 and March 1989, 128 eligible patients (most being children) with CGD were accrued by the International CGD Cooperative Study Group. The patients were followed for about one year. There were 63 patients in the treatment group and 65 patients in the control group,

with 203 serious infection/censoring observations recorded. The number of serious infections/censoring for the patients ranged from 1 to 8.

The aim of the trial was to investigate the effectiveness of γ -IFN in reducing the number of serious infections in CGD patients allowing for the variation in other risk variables. Other risk variables were age, sex, inheritance pattern, height, weight, hospital category, using corticosteroids and using prophylactic antibiotics at times of study entry.

This data set has been analysed by Fleming and Harrington (1991) using the Cox model and Andersen-Gill multiplicative intensity model. Both analyses are based on the information up to the interim cutoff. For the Cox model, only the time to first infection in each patient is used. The multiplicative intensity model is the multivariate version of the traditional Cox model developed through the martingale approach.

The full data set was given in the Appendix of Fleming and Harrington (1991). A comparison of the marginal model, the Andersen-Gill Markov and semi-Markov model and the model proposed by Prentice, Williams and Peterson (1981) was provided by Lin (1994). Moreover, the application of the time dependent frailty model to this data set based on the counting processes approach was provided by Self (1993). An analysis of this data set using the Poisson regression model in the counting processes framework can be found in Lindsey (1995).

Data set (7) -- Kidney patient data

The kidney patient data detail recurrent infections in kidney patients using portable dialysis. At entry of study, a catheter was inserted and remained in place until infection occurred at the point of insertion. When infection occurred, the catheter was removed and the infection would clear. Then, some weeks later, the catheter was reinserted.

The failure time is taken to be the time from the point of reinsertion until the next infection occurs. Therefore, during the study, several infections may be observed in each patient. Such times are correlated when they are repeated observations of the same patient. Censoring occurs when the study ends or when the catheter is removed for some reason not connected with infection.

82 patients were enrolled in the study with the number of serious infections/censoring in a patient ranging from 1 to 8. The aim of the study was to relate the incidence rate of infection to risk variables, such as age, sex and type of kidney disorder, and to explore the extent of variability among patients.

The recovery interval is assumed to be sufficiently long to make negligible carry over effects from one failure recurrence interval to the next. It seems reasonable to assume that patient frailty is a constant common factor in all observations on the same patient. Part of the data, the first 2 failure observations of 38 patients, has been analysed by

McGilchrist and Aisbett (1991b) and McGilchrist (1993) to give BLUP and REML estimations respectively assuming the random frailty terms are independent and constant over time.

Data set (8) -- Bladder cancer data

This recurrent bladder cancer data set appeared in Byar, Blackard and the VACURG (1977) and Byar (1980). The study was conducted by the Veterans Administration Cooperative Urological Research Group.

At the beginning of the trial, patients who had superficial bladder tumors were enrolled. These tumors were removed transurethrally. Patients were randomly allocated to one of the three treatments: placebo, thiotepa and pyridoxine. During the study, many patients experienced multiple recurrences of tumors and these tumors were removed at each visit. One of the research objectives was to evaluate the effectiveness of thiotepa in reducing the rate of tumor recurrence in bladder cancer patients as indicated in Byar (1980).

Analysis of this recurrent bladder cancer data by modelling the marginal distribution can be found in Wei, Lin and Weissfeld (1989). In this paper, however, only part of the data was used, which included the placebo and thiotepa group. Also, only up to the fourth failure time observations were considered in this paper. Moreover, a comparison of their method with that of Prentice, Williams and Peterson (1981) and Andersen and Gill (1982) was also presented. Recently, Lin and Wei (1992)

analysed the same set of data by extending the idea of Wei, Lin and Weissfeld (1989) to accelerated failure time models.

Data set (9) -- "Late infections" in Bone marrow transplant data

Between July 1970 and December 1976, 89 patients with aplastic anemia or acute leukemia treated by syngeneic (13 patients) or allogeneic (76 patients) marrow transplantation were studied to determine the incidence of late infections. Details of this study were described by Atkinson *et al.* (1979).

All infections, except the upper respiratory infections and varicella-zoster infections, occurring after 6 months from marrow transplantation were included. Possible risk variables were recorded for analysis. Of particular interest was the relationship between recurrent infections in long-term survivors and the chronic graft versus host disease (C-GVHD).

In this data set, each patient might experience more than one infection. A subset of this data, the 76 patients receiving allogenic bone marrow transplantation data, were used to study the predictive factors for "late infection" in Prentice, Williams and Peterson (1981). In this paper, the method of stratification was employed to analyse this subset of data. Moreover, models considering "Gap time" and "Total time" as the failure time measure to construct the stratified hazard function were proposed.

Data set (10) -- Others

In fact, there are quite a large number of recurrent failure time data sets. Examples are the Multiple tumors Carcinogenesis experiment data given by Gail, Santner and Brown (1980), the Myocardial infarction data considered in Hougaard (1986a), the Bone marrow transplant data examined by Storb *et al.* (1990) and Pepe and Cai (1993).

Moreover, 3 multiple failure time data sets were reported in the epidemiology workshop held in the Monash University in July 1994. The first one was a project coordinated by the World Health Organisation (WHO). Beginning in 1984, the study recorded the subsequent cardiac events of over 5000 patients who were admitted to hospital for the first time with suspected Acute Myocardial Infarction (AMI). Patients were followed up for up to 8 years.

The second data set considered the recurrent hospital admissions for patients suffering from Lower Respiratory tract Illness (LRI). LRI is an important cause of mortality, morbidity and hospital admission in early childhood. The study was conducted by the Institute for Child Health Research in Western Australia. One of the research objectives was to investigate the effect of possible risk variables on the rate of hospital admission in LRI patients.

The third recurrent data set came from 150 Childhood Acute Lymphoblastic Leukaemia (ALL) patients. Data from these ALL patients were

derived from medical records in the Prince of Wales Children's Hospital, Sydney between 1983 and 1989.

2.4 TYPE IV : THE OCCURRENCE OF DIFFERENT FAILURE EVENTS IN THE SAME PATIENT

In some medical research, we may obtain the failure time data of different events in the same patient. For example, in studying elderly people, we may consider two types of failures, the hospitalisation and death. Since these two types of failures are related, it is necessary to develop survival models that allow more than one type of failure and establish the relationship between different types of failure.

Data set (11) -- The Dubbo study

The Dubbo study is a prospective study of the health of persons born before 1930 living in the New South Wales country town of Dubbo. An introduction and some preliminary analyses of this data set have been given by Simons *et. al.* (1990) and Simons *et. al.* (1991).

There were 1237 men and 1568 women enrolled in the study. The mean age of men and women was 69 and 70 respectively. Extensive examinations were performed in 1988-89 covering medical, social, physiological and lifestyle variables.

The study was followed for a period of 5 years, the two competing failure events were hospitalization and death. Note that the hospitalizations could be recurrent and the occurrence of death should not be earlier than hospitalizations. Research interest is to estimate the risk variable parameters for both hospitalization and death as well as to find out the relationship between the random effects of the two types of failures.

Data set (12) -- Colon cancer study

929 stage C Colon cancer patients were enrolled between March 1984 and October 1987. They were randomly assigned to observation, levamisole alone or levamisole combined with fluorouracil. Details of the study can be found in Moertel *et al.* (1990) and Fleming (1992).

The main research interest was to evaluate the role of levamisole combined with fluorouracil as adjuvant therapy for resected colon cancer. In this study, both the cancer recurrence and the survival time were considered to be important outcome measures.

Since the two types of failure, the cancer recurrence and death, may occur in the same patient, analysing techniques that will cope with the relationship between the two types of failure within the same patient need to be developed. An analysis of this data set based on the marginal approach has been given by Lin (1994).

CHAPTER THREE

REVIEW: DIFFERENT APPROACHES TO THE MULTIVARIATE FAILURE TIME PROBLEM

For the four types of multivariate failure time data mentioned in Chapter 2, the common feature is the dependence between the failure time observations that occurred within the same patient/cluster. With this dependence, extension of the traditional Cox method which assumed independence between events is needed. Methods used to analyse these multivariate failure time data should be able to handle such dependence. In the past twenty years, researchers have attempted to set up survival models which extend the univariate Cox regression model to solve the multivariate failure time problem.

In this Chapter, initially Cox's well-developed regression model in survival analysis in the univariate case is described. In developing this model, Cox proposed the partial likelihood method in the univariate case (1972, 1975). Here, univariate means that there is only one failure time corresponding to each individual and these observations can be assumed to be independent. Usually, in the medical context, the failure time refers to the survival time of patient. This is followed by a brief description of different approaches that have been used to tackle the multivariate failure time problem. They have been classified as (i) the frailty models approach, (ii) the marginal modelling approach, (iii) the counting processes approach and (iv) the generalised linear mixed model (GLMM) approach.

3.1 UNIVARIATE COX'S REGRESSION MODEL IN SURVIVAL ANALYSIS

In Cox's proportional hazards model (Cox 1972), the hazard function for the failure time T of n individuals associated with a $p \times 1$ vector of possibly time-varying covariates $\mathbf{X}_{n \times p} = [x_1, \dots, x_n]'$ is

$$3.1.1 \quad h(t_i; x_i) = \lambda(t_i) g(x_i; \beta)$$

where β is a p -dimensional vector of regression parameters

$\lambda(t)$ is the unspecified baseline hazard function

$g(x; \beta)$ is a positive-valued function and in most situation it is chosen to be $\exp(x' \beta)$

The log-partial likelihood function (Cox 1975) is given by

$$3.1.2 \quad l(\beta) = \sum_{i=1}^n D_i [\ln g(x_i; \beta) - \ln \sum_{j \in R_i} g(x_j; \beta)]$$

where $D_i = \begin{cases} 0 & \text{if patient } i \text{ is censored} \\ 1 & \text{if patient } i \text{ dies} \end{cases}$

is the censoring indicator

R_i is the risk set which is defined to be the set of individuals known to be alive just prior to T_i

If $g(x_i; \beta)$ is chosen to be $\exp(x_i' \beta)$, then the log-partial likelihood is given by

$$3.1.3 \quad l(\beta) = \sum_{i=1}^n D_i[x'_i\beta - \ln \sum_{j \in R_i} \exp(x'_j\beta)]$$

The estimation of the regression parameters β is given by maximizing the log-partial likelihood function in 3.1.2. Such a method is usually called a semi-parametric approach since the baseline hazard function is completely unspecified. A great convenience of this method is that the baseline hazard function is cancelled out in the partial likelihood expression and then not involved in the estimation process of the regression parameters.

If we denote the maximum partial likelihood estimator of β by $\hat{\beta}$, then, under some mild conditions, $\hat{\beta}$ is asymptotically normally distributed with mean β and variance estimated by $I(\hat{\beta})^{-1}$, where $I(\hat{\beta})^{-1} = -\partial^2 l(\beta)/\partial\beta\partial\beta'$ is the observed information matrix from the log-partial likelihood function. Consequently, the usual asymptotic results for likelihood ratio tests based on this log-partial likelihood can be applied. For the derivation of the asymptotic results, the maximum likelihood estimation of parametric regression models and the accelerated failure time regression models, refer to Kablefleish and Prentice (1980), Lawless (1982) and Cox and Oakes (1984).

3.2 THE FRAILTY MODELS APPROACH

Frailty models, following their introduction by Vaupel, Manton and Stallard (1979), have become common. Essentially, in the frailty models,

an extra random component is introduced into the proportional hazards model to explain individual variability.

For example, when we consider infection episodes, the failure time observations within the same patient are correlated because these observations come from the same patient. Therefore, in the proportional hazards model, a random patient effect u_i is introduced multiplicatively into the hazard function. The hazard function is given by

$$3.2.1 \quad h(t;i,j) = u_i \lambda(t) g(\beta)$$

where $h(\cdot)$, $\lambda(\cdot)$ and $g(\cdot)$ have the usual interpretation as in 3.1.1

u_i represents the random frailty of the i^{th} individual

The idea is that individuals have different frailties, and that after eliminating the effects of the corresponding risk variables, those who are most "frail" will experience the highest rates of infection. Since the frailty term is a random component, the frailty distribution can be modelled parametrically. Reviews on the heterogeneity in survival analysis and the analysis of event history data have been introduced by Aalen (1988) and Clayton (1988) respectively. A classification of multivariate survival data and different modelling suggestions can be found in Hougaard (1987). The relevance of these models in different situations have also been discussed in this paper.

The frailty model development starts from the matched survival data analysis of Holt and Prentice (1974). In this semiparametric model, the

association within a pair was treated as a separate nuisance hazard function and the covariate acted proportionally on that hazard for each pair. On the other hand, Clayton (1978) and Oakes (1982) provided a fully parametric model for matched survival data with Gamma distributed frailty. Besides the semiparametric nature, the main difference between the Holt-Prentice (HP) model and the Clayton-Oakes (CO) model is that the HP model applies the proportional hazards assumption to the conditional hazard rather than the marginal hazard.

Wild (1983) considered the Weibull specification in the baseline hazard function with the random component modelled as constant, independent Gamma distributed and unknown parameters. In this paper, the loss of efficiency problem in the HP model when there is appreciable censoring in the data, was also investigated.

In the bivariate situation, Clayton and Cuzick (1985a) considered estimation of the frailty parameter and covariate effects using a modified EM algorithm with the frailty assumed to be Gamma distributed. The illustration of the EM algorithm technique for Cox's regression model can be found in Clayton and Cuzick (1985b).

The extension of the fully parametric model of Clayton (1978) to incorporate covariate information has been suggested by Huster, Brookmeyer and Self (1989) in the analysis of the Diabetic retinopathy study. The marginal modelling approach based on the independent working model to the same set of data has also been provided in this paper.

Oakes (1989) considered the class of bivariate survival distributions that can be used to model the frailty within two observed survival times. Moreover, it has been shown in this paper that the observable bivariate distribution determines the unobservable frailty distribution up to a scale parameter.

Using the EM algorithm, Klein (1992) considered the estimation of fixed and random effects regression model in survival analysis. Based on a profile likelihood construction, the full likelihood consists of the observed failure time and the unobservable frailty by assuming the random frailties follow a Gamma distribution. In the E-step, the expectation of the likelihood with respect to the observable data is considered. In the M-step, a partial likelihood is constructed to estimate the covariate effects using the profile likelihood technique suggested by Johansen (1983). With the Gamma distributed frailty assumption, parameter estimates are obtained by iterating between the 2 steps.

The gamma distributed frailty model has also been considered by Vaupel, Manton and Stallard (1979), Lancaster (1979), Lancaster and Nickell (1980) and Vaupel and Yashin (1983). As indicated in Hougaard (1984), the Gamma distribution constitutes a very convenient family in modelling frailty distributions. Moreover, Vaupel and Yashin (1983) have proposed other distributions such as the Uniform, Weibull and Lognormal distribution to model the frailty.

In Hougaard (1984), the heterogeneity between individuals in the population has been modelled by the Inverse Gaussian frailty distribution.

Hougaard (1986a) considered a three parameter family of distributions on positive numbers to model the frailty distribution. This three parameter family $P(\alpha, \delta, \theta)$ includes the positive stable distribution ($\theta=0$), Gamma distribution ($\alpha=0$), degenerate distribution ($\alpha=1$) and the Inverse Gaussian distribution ($\alpha=1/2$). Basically, the family is derived from the positive stable distribution by introducing one more parameter.

A class of continuous multivariate lifetime distributions has been proposed by Hougaard (1986b). The dependence between individuals in a group is modelled by a group specific quantity, this quantity is assumed to follow a positive stable distribution. Moreover, the possibility of including covariates into the model is also discussed. Further work in modelling heterogeneity in survival data has been given by Hougaard (1991). Extension of the models suggested by Hougaard (1986a, 1986b) using the compound Poisson distribution can be found in Aalen (1992).

3.3 THE MARGINAL MODELLING APPROACH

The marginal modelling approach, sometimes called the "independence working model analysis", to multivariate survival analysis originated with Wei, Lin and Weissfeld (1989). Essentially, the method is analogous to that of Liang and Zeger (1986) for longitudinal data analysis. In Liang and Zeger (1986), a class of estimating equations which is an extension of generalised linear models to the analysis of longitudinal data is introduced. It has been shown that simple estimating equations can be

constructed to yield consistent and asymptotically normal estimates for regression parameters provided the marginal model is correctly specified.

The marginal modelling approach formulates the marginal distribution of multivariate failure times with the familiar Cox proportional hazards models while leaving the dependence structure between related failure times completely unspecified. So, in the marginal modelling approach, type-specific hazard function $h_j(t;i)$ is considered; this may be modelled as in 3.3.1 and 3.3.2

$$3.3.1 \quad h_j(t;i) = \lambda_o(t) \exp(x_{ij}'\beta_j)$$

$$3.3.2 \quad h_j(t;i) = \lambda_{oj}(t) \exp(x_{ij}'\beta_j)$$

where $h_j(t;i)$ is the hazard function corresponding to the j^{th} type failure

$\lambda_o(t)$ is the usual baseline hazard function

$\lambda_{oj}(t)$ is the type-specific baseline hazard function

β_j and x_{ij} correspond to the vector of j^{th} type failure regression parameter and risk variable of the j^{th} type failure in the i^{th} individual respectively

Note that in 3.3.1 and 3.3.2, there is no explicit modelling of the dependence structure between the failure events in the same patient/cluster. Moreover, 3.3.1 and 3.3.2 differ in that the baseline hazard functions for the different type of failures may or may not be identical. On the other hand, we may also consider the regression parameters β_j to be the same among the marginal submodels 3.3.1 and 3.3.2 (i.e. $\beta=\beta_j, \forall j$).

Prentice, Williams and Peterson (1981) proposed a stratified proportional hazards function to model the multivariate failure data from which it absorbed the model discussed by Gail, Santner and Brown (1980) as its 2-sample special case. This method of stratification was illustrated by applying the method to analyse the "late infection" in bone marrow transplant data. Moreover, in this paper, there is a discussion on using the "gap time" or the "total time" as the failure time measure.

Following this stratification technique, Wei, Lin and Weissfeld (1989) extended the method by modelling the marginal distributions. The method's further development was undertaken by Lee, Wei and Amato (1992). Related works by extending the marginal approach to accelerated failure time model can be found in Lin and Wei (1992) and Lee, Wei and Ying (1993).

Application of the independence working model to the Diabetic retinopathy data has been given by Huster, Brookmeyer and Self (1989). Moreover, a comparison of the marginal approach and the frailty model of Clayton (1978) (incorporating covariate information) in terms of efficiency was also considered in this paper. The analysis of the Schizophrenia and the Diabetic retinopathy data using the marginal approach has been provided by Liang, Self and Chang (1993). The method mentioned can be seen as an unstratified version of the marginal approach.

Pepe and Cai (1993) also considered the analysis of recurrent failure time problem. An intermediate approach between the intensity method of

Prentice, Williams and Peterson (1981) and the marginal method of Wei, Lin and Weissfeld (1989) was chosen by modelling the rate function. This method was applied to analyse a data set from a randomized clinical trial of bone marrow transplant patients treated at the Fred Hutchinson Cancer Research Centre for leukemia and aplastic anemia.

3.4 THE COUNTING PROCESSES APPROACH

Following the work of Aalen (1976, 1978), the theory of multivariate counting processes provided a general framework to analyse censored survival data. In fact, it unifies and extends many branches of non-parametric survival analysis. The development stems from modern martingale and stochastic integral theory.

It has been shown in Andersen and Gill (1982) that, based on the statistical theory of counting processes, the concept of partial likelihood (Cox 1972, 1975) in Cox's model can be treated as a special case of the multiplicative intensity model. Moreover, asymptotic properties can be derived in a natural way in this general framework.

Essentially, the counting processes approach consists of the modelling of an intensity function (or equivalently the hazard function in Cox's model) $\lambda_i(t)$ for the i^{th} individual. So, in the univariate case, the intensity function

$$3.4.1 \quad \lambda_i(t) = Y_i(t) \lambda_o(t) \exp\{x_i(t)' \beta\}$$

where $i = 1, \dots, n$ (n refers to the total number of individuals)

$$Y_i(t) = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ individual is under observation at time } t - \\ 0 & \text{otherwise} \end{cases}$$

$\lambda_o(t)$ is the usual baseline hazard function

β and $x_i(t)$ are as usual, correspond to the vector of regression parameter and risk variable of the i^{th} individual respectively

Basically, this model coincides with the Cox type model in the univariate situation and generalises to take into account the information from recurrent data in the multivariate case. As to the multiplicative intensity model considered in Andersen and Gill (1982), this permits the regression analysis of the intensity of a recurrent event allowing for complicated censoring patterns and time dependent covariates. Furthermore, in this paper, the asymptotic properties of the estimators were established using martingale techniques.

There has been a non-technical explanation of this approach in Gill (1984). A detailed review on the counting processes models for life history data has been provided by Andersen and Borgan (1985). The extension of the two-state survival models to multistate models in survival analysis by modelling the transition intensity has been explained in Andersen (1988). A systematic development of the counting processes approach and its application to multivariate failure time data can be found in two textbooks written by Fleming and Harrington (1991) and Andersen, Borgan, Gill and Keiding (1993). One of the criticisms of the counting processes approach is on the independence assumption of

transition events conditional on covariates as discussed in Clayton (1991).

The Maximum Likelihood (ML) estimation in frailty models using the martingale approach has been developed by Nielsen, Gill, Andersen and Sørensen (1992). The simplest and most commonly used model takes the intensity function of the k^{th} observation of the j^{th} type failure in the i^{th} cluster/individual as

$$3.4.2 \quad \lambda_{ijk}(t) = Z_i Y_{ijk}(t) \lambda_o(t) \exp\{x_{ijk}(t)' \beta_j\}$$

$$3.4.3 \quad \lambda_{ijk}(t) = Z_i Y_{ijk}(t) \lambda_{oj}(t) \exp\{x_{ijk}(t)' \beta_j\}$$

where Z_i is called the random frailty variable for the i^{th} cluster/individual

$$Y_{ijk}(t) = \begin{cases} 1 & \text{if the } (i,j,k) \text{ failure is under observation at time } t \\ 0 & \text{otherwise} \end{cases}$$

$\lambda_o(t)$ is the overall baseline hazard function

$\lambda_{oj}(t)$ is the j^{th} type failure baseline hazard function

β_j and $x_{ijk}(t)$ correspond to the vector of j^{th} type failure regression parameter and risk variable of the k^{th} observation of the j^{th} type failure in the i^{th} cluster/individual respectively

Moreover, according to the type of multivariate failure time data being considered, modification of these models is made possible by adjusting the indices of $\lambda(t)$, Z , $Y(t)$, $\lambda_o(t)$, $x(t)$ and β in 3.4.2 and 3.4.3.

In Nielsen, Gill, Andersen and Sørensen (1992), Z_i was assumed to be iid Gamma distributed. The risk variable parameters and the parameter in the frailty distribution are estimated by maximising the likelihood using the EM algorithm. This method of analysis has been illustrated through three data sets, the litter matched tumorigenesis experiment data, the skin allograft study data and the premature death in adult adoptees study data.

3.5 GENERALISED LINEAR MIXED MODEL APPROACH

Estimation and inference in the Linear Models (LM) with normal distributed error, are well developed. Let y be an observation vector with n components distributed according to the linear model

$$y = \mu + e \quad e \sim N(\theta, D)$$

$$E(y) = \mu = X\beta$$

where X is a $n \times p$ matrix of regression variables and β is a p -component vector of regression coefficient. The variance matrix of the error vector e is written as D , while D is considered as a function of some underlying parameter set θ .

Estimation and inference techniques on β and θ can be adapted from likelihood based methods such as Maximum Likelihood (ML) or Residual Maximum Likelihood (REML).

The appearance of the Generalised Linear Model (GLM) technique by Nelder and Wedderburn (1972) extends the traditional normal error linear model to non-linear responses. Setting the error vector to belong to the exponential family and connecting the mean to the linear predictor by a link function g allows the response to follow a family of distribution with non-identity link. Therefore

$$y = \mu + e \quad \text{distribution of } e \text{ belongs to the exponential family}$$

$$E(y) = \mu, \text{ with } g(\mu) = X\beta$$

Moreover, the introduction of quasi-likelihood in Wedderburn (1974) provided further flexibility to the GLM. This so-called quasi-likelihood can be constructed by specifying the dependence structure of the variance on μ without assuming the whole error distribution. Furthermore, such quasi-likelihood also has most of the appealing properties of the likelihood function.

The Linear mixed model (LMM) with Gaussian outcomes were considered by Laird and Ware (1982) and Ware (1985) in analysing longitudinal data. In LMM,

$$y = \mu + e \quad e \sim N(0, D)$$

$$E(y) = \mu = X\beta + Z\mu$$

where X, Z are the corresponding design matrices of β and u . β is the usual fixed effect parameter and u is a vector of random effects distributed as $N(0, A)$ with A containing possibly a parameter set ϕ .

In the LMM, the Best linear unbiased prediction (BLUP) method developed by Henderson (1963, 1973, 1975) and recently reviewed by Robinson (1991) can be used to estimate fixed effect coefficients, variance components and to predict random effects. Following the introduction of REML estimation by Patterson and Thompson (1971) and Thompson (1980), the extension of BLUP to ML and REML has been outlined in Harville (1977) and further developed in Fellner (1986, 1987) and Speed (1991). LMM with AR(1) or other structured covariance matrices have been considered by Mansour, Nordheim and Rutledge (1985), Jennrich and Schluchter (1986) and Chi and Reinsel (1989).

For the non-Gaussian outcome mixed model, Williams (1982), Stiratelli, Laird and Ware (1984) and Anderson and Aitkin (1985) considered the logit model, Ochi and Prentice (1984) proposed the probit model, Koch *et. al.* (1977) and Breslow (1984) examined the loglinear model. The first order Markov Chain model was developed by Zeger, Liang and Self (1985). Methods for maximising the joint likelihood for both fixed and random effects have been considered by Leonard (1972) in binomial data and Harville and Mee (1984) in ordered categorical data using the Baynesian argument. The Gilmour, Anderson and Rae (1985) method is somewhat analogous to the BLUP method in a mixed model for binomial data and has elements in common with the EM algorithm.

Analysis of longitudinal data for discrete and continuous outcomes using GLM by modelling the marginal distribution has been done by Liang and Zeger (1986) and Zeger and Liang (1986). Further work by Zeger, Liang

and Albert (1988) considered subject-specific (SS) and population-averaged (PA) models. In the SS model, the heterogeneity in regression parameters is explicitly modelled while the aggregate response for the population is the main concern in the PA model.

The Generalised linear mixed model (GLMM) approach attempts to combine the GLM and LMM technique into a unified framework. Therefore, in the GLMM, we have

$$y = \mu + e \quad \text{non-Gaussian outcome } y$$

$$E(y) = \mu, \text{ with } g(\mu) = X\beta + Zu$$

Different approximations in the development of the GLMM have been used. The linearisation argument has been given by Schall (1991). The quadratic approximation in the likelihood in the region of the maximum has been proposed by McGilchrist (1994). Breslow and Clayton (1993) provide the Penalised quasi-likelihood (PQL) approximation following Green (1987) and Marginal quasi-likelihood (MQL) following Goldstein (1991) and Zeger, Liang and Albert (1988). The Solomon and Cox (1992) method was based on the Laplace approximation. The Psuedo-likelihood method given by Wolfinger and O'Connell (1993) make use of the Gaussian approximation and Taylor's theorem. Further development of the GLMM approach, following the line of McGilchrist (1994), by considering possibly correlated random effects can be found in McGilchrist and Yau (1995a).

As outlined in McGilchrist (1994), by reprogramming the first and second derivative of the log-partial likelihood function on the

conditionally fixed random components, the GLMM technique can be applied to solve the multivariate failure time problem. Application of the GLMM approach to multivariate survival data can be found in McGilchrist and Aisbett (1991b), McGilchrist (1993), McGilchrist and Yau (1995b) and Yau and McGilchrist (1995a, 1995b).

CHAPTER FOUR

GENERALISED LINEAR MIXED MODELS

In the preceding Chapters, the four types of multivariate failure time data are described and the four different approaches to the multivariate failure time problem are mentioned. However, with its advantage in preserving the cancellation property and in predicting random components, the Generalised Linear Mixed Model (GLMM) approach is adopted. The cancellation property in the partial likelihood expression of the Cox model is a very convenient feature because it allows the baseline hazard function to be unspecified. The prediction of random components, as mentioned in Chapter 1, may be of interest when the identification of a high risk subpopulation is considered important.

The GLMM begins with the Linear Mixed Model with normally distributed random components. The Best Linear Unbiased Prediction (BLUP) method is used in the initial step of estimation and extends to obtain Maximum Likelihood (ML) and Residual Maximum Likelihood (REML) estimators. Then, based on the quadratic approximation of the likelihood in the region of the maximum, the techniques similar to those developed for normal theory models with random components are reflected to the GLMM.

This Chapter is devoted to the derivation of the GLMM, with a brief explanation of the application of the GLMM to the multivariate failure time problem in the last Section. The estimation and inference techniques developed in this chapter provides the foundation of development of the random effect survival models given in the subsequent chapters.

Moreover, the models considered here extend those given in McGilchrist (1994) by allowing possible correlation parameters in the variance matrix of the random components. The implementation of such extension for the analysis of multivariate failure time data is given in Chapter 6 in which the random effects of each individual is assumed to follow an AR(1) process.

4.1 INTRODUCTION

For Linear Mixed Models with normally distributed random components, Best Linear Unbiased Predictors (BLUP) were developed by Henderson (1963, 1973, 1975) for simultaneously estimating the fixed components of a mixed model as well as the realised values of random components of the model. Such estimators or predictors (for the random components) were shown to be best, linear, unbiased. The estimation procedure has been reviewed by Robinson (1991) who gives an extensive bibliography.

When the BLUP estimators of the random components are used to estimate the variances of those random components, it is found that the estimated variance components are severely biased towards zero. Harville (1977) has noted that the BLUP computational procedure may be used to find both Maximum Likelihood (ML) and Residual Maximum Likelihood (REML) estimators as described in Patterson and Thompson (1971). The relationship is fully developed in Thompson (1980) and further in Fellner (1986, 1987) as well as Speed (1991).

On the other hand, the Generalised Linear Model (GLM) technique of Nelder and Wedderburn (1972) introduced a much more flexible instrument for regression analysis in statistical modelling. It extends the classical Linear Model by specifying an appropriate link function as well as a particular choice of the error distribution in the exponential family. Hence, the GLM can be employed to analyse categorical, discrete and non-negative response data with non-normal error distribution.

It is then a natural extension to include random components that combine with the regression variables and act linearly in the GLM framework. This type of model is often called the Generalised Linear Mixed Model (GLMM). The GLMM starts with the BLUP estimators as the initial step and extends to finding ML and REML estimators. Parallel developments of the GLMM are given by Schall (1991), McGilchrist and Aisbett (1991a), Solomon and Cox (1992), Breslow and Clayton (1993), Wolfinger and O'Connell (1993), Wolfinger (1993), McGilchrist (1994) and McGilchrist and Yau (1995a).

The subsequent Sections extend the model considered in McGilchrist (1994) by putting it in a more general setting for use with correlated random components which often appear in a repeated measures setting. Applications of the approach have been made to multicentre clinical trials in McGilchrist and Zhaorong (1990); discordance data in Zhaorong, Matawie and McGilchrist (1992); threshold models in Zhaorong, McGilchrist and Jorgensen (1992); survival analysis in McGilchrist and Aisbett (1991b),

McGilchrist (1993), McGilchrist and Yau (1995b), Yau and McGilchrist (1995a, 1995b).

4.2 NORMALLY DISTRIBUTED RANDOM COMPONENTS

Let y be a response vector which is generated by the linear mixed model

$$y = \eta + e , \quad \eta = X\beta + Zu ,$$

where e is distributed as $N(0, \sigma^2 D)$, D is a known symmetric matrix of dimension equal to the number of observations n in the response vector y and X, Z are matrices of values of regression variables. The unknown parameter β has dimension v while the random component u may be partitioned

$$u' = [u'_1, u'_2, \dots, u'_k] , \quad Z = [Z_1, Z_2, \dots, Z_k]$$

where Z, u are partitioned conformally and u_j are independent random components distributed as $N[0, \sigma_j^2 A_j(\phi)]$. For convenience we write $\sigma_j^2 = \sigma^2 \theta_j$ and the σ^2, θ_j, ϕ are unknown parameters. The parameter ϕ may be a vector parameter of dimension p , which describes the covariance structure of the vectors u_j . For example, when u_j follows an ARMA(1,0) process with correlation ϕ ,

$$A_j(\phi) = \frac{1}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \dots & \phi^{n_j-1} \\ \cdot & 1 & \phi & \dots & \phi^{n_j-2} \\ \cdot & \cdot & 1 & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \phi^{n_j-1} & \phi^{n_j-2} & \phi^{n_j-3} & \dots & 1 \end{pmatrix}$$

Estimation of the parameters of this model is accomplished by firstly developing Best Linear Unbiased Prediction (BLUP) estimators and then using those estimators as an initial computation in finding Maximum Likelihood (ML) and Residual Maximum Likelihood (REML) estimators.

4.3 BLUP ESTIMATION

The loglikelihood of y taking u as conditionally fixed, is denoted by l_1 and the logarithm of the probability density function of u is denoted by l_2 . Expressions for these quantities are

$$l_1 = -(1/2)[n \ln 2\pi\sigma^2 + \ln |D| + \sigma^2(y-X\beta-Zu)'D^{-1}(y-X\beta-Zu)]$$

$$l_2 = -(1/2) \sum_{j=1}^k [v_j \ln 2\pi\sigma_j^2 + \ln |A_j(\phi)| + \sigma_j^2 u_j' A_j^{-1}(\phi) u_j]$$

where v_j is the dimension of u_j . BLUP estimates of the parameters and the realisations of the random components u , maximise $l=l_1+l_2$. For convenience let A be the block diagonal matrix

$$\mathbf{A} = \begin{bmatrix} \theta_1 \mathbf{A}_1 \\ \theta_2 \mathbf{A}_2 \\ \vdots \\ \theta_k \mathbf{A}_k \end{bmatrix}$$

The derivatives with respect to the parameters are

$$\begin{aligned}\frac{\partial l}{\partial \beta} &= \frac{\partial l}{\partial \beta} = \sigma^{-2} \mathbf{X}' \mathbf{D}^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u) \\ \frac{\partial l}{\partial u} &= \sigma^{-2} [\mathbf{Z}' \mathbf{D}^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u) - \mathbf{A}^{-1}u] \\ \frac{\partial l}{\partial \sigma^2} &= -(1/2)[n\sigma^{-2} - \sigma^{-4}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u)' \mathbf{D}^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}u)] \\ \frac{\partial l}{\partial \sigma_j^2} &= -(1/2)[v_j^{(s)} - \sigma_j^{-4} u_j' \mathbf{A}_j^{-1} u_j] \\ \frac{\partial l}{\partial \phi_s} &= -(1/2) \sum_{j=1}^k [v_j^{(s)} - \sigma_j^{-2} u_j' \mathbf{A}_j^{-1} (\partial \mathbf{A}_j / \partial \phi_s) \mathbf{A}_j^{-1} u_j]\end{aligned}$$

where $v_j^{(s)} = \text{tr}(\mathbf{A}_j^{-1} \partial \mathbf{A}_j / \partial \phi_s)$.

Equating the above derivatives to zero and solving gives the BLUP estimators, viz.

$$4.3.1 \quad \begin{bmatrix} \mathbf{X}' \mathbf{D}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{D}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{D}^{-1} \mathbf{X} & \mathbf{Z}' \mathbf{D}^{-1} \mathbf{Z} + \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{D}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{D}^{-1} \mathbf{y} \end{bmatrix}$$

$$4.3.2 \quad \tilde{\sigma}^2 = n^{-1} (\mathbf{y} - \mathbf{X}\tilde{\beta} - \mathbf{Z}\tilde{u})' \mathbf{D}^{-1} (\mathbf{y} - \mathbf{X}\tilde{\beta} - \mathbf{Z}\tilde{u})$$

$$4.3.3 \quad \tilde{\sigma}_j^2 = v_j^{-1} \tilde{u}_j' \mathbf{A}_j^{-1} \tilde{u}_j, \quad j = 1, 2, \dots, k$$

$$4.3.4 \quad \sum_{j=1}^k [v_j^{(s)} - \tilde{\sigma}_j^{-2} \tilde{u}_j' \mathbf{A}_j^{-1} (\partial \mathbf{A}_j / \partial \phi_s) \mathbf{A}_j^{-1} \tilde{u}_j] \Big|_{\phi=\tilde{\phi}} = 0, \quad s = 1, 2, \dots, p$$

The BLUP equation for ϕ_s may not be solvable explicitly. Using the first matrix equation and letting

$$\Sigma = D + ZAZ' , \quad K = D^{-1} - D^{-1}X(X'D^{-1}X)^{-1}X'D^{-1}$$

the matrix equation can be solved to give

$$4.3.5 \quad \hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$$

$$4.3.6 \quad \tilde{u} = (Z'KZ+A^{-1})^{-1}Z'Ky$$

Note that K is symmetric and $X'KX=0$ implies $KX=0$.

4.4 SOME MATRIX RESULTS

Definition Given any matrix A , the Moore-Penrose inverse A^+ of A is defined to be a matrix which satisfies

$$(a) \quad AA^+A = A$$

$$(c) \quad (AA^+)' = AA^+$$

$$(b) \quad A^+AA^+ = A^+$$

$$(d) \quad (A^+A)' = A^+A$$

Matrix results in this section rely much on the definition and properties of Moore-Penrose inverse. The existence and uniqueness properties of such an inverse is well-known and has been developed by Penrose (1955) based on the foundations laid by Moore (1920).

Let Σ and K be defined as in the previous section. A matrix Q is introduced in the following theorem.

Theorem 4.4.1. $Q = \Sigma^{-1}[I-X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}] = K(K\Sigma K)^{-1}K$

Proof: Let $\mathbf{W} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and since \mathbf{W} is a symmetric idempotent matrix, we can choose a rectangular matrix \mathbf{P} with columns the normalised eigenvectors corresponding to the unit eigenvalues of \mathbf{W} . Thus

$$\mathbf{W} = \mathbf{PP}' \text{ such that } \mathbf{P}'\mathbf{P} = \mathbf{I}, \mathbf{WP} = \mathbf{P}, \mathbf{P}'\mathbf{W} = \mathbf{P}'$$

and $\mathbf{P}(\mathbf{P}'\Sigma\mathbf{P})\mathbf{P}'$ is the Moore-Penrose inverse of $\mathbf{W}\Sigma\mathbf{W}$. This follows from

$$\mathbf{P}(\mathbf{P}'\Sigma\mathbf{P})\mathbf{P}'\mathbf{W}\Sigma\mathbf{W} = \mathbf{W}\Sigma\mathbf{W}\mathbf{P}(\mathbf{P}'\Sigma\mathbf{P})\mathbf{P}' = \mathbf{W}$$

Now \mathbf{Q} is also the Moore-Penrose inverse of $\mathbf{W}\Sigma\mathbf{W}$. This follows from

$$\mathbf{Q}\mathbf{W}\Sigma\mathbf{W} = \mathbf{W}\Sigma\mathbf{W}\mathbf{Q} = \mathbf{W} \text{ and } \mathbf{W}\mathbf{Q} = \mathbf{Q}\mathbf{W} = \mathbf{Q}.$$

Since the Moore-Penrose inverse is unique then $\mathbf{Q} = \mathbf{P}(\mathbf{P}'\Sigma\mathbf{P})\mathbf{P}'$. Since $\mathbf{KX=0}$, $\mathbf{P}'\mathbf{X}=0$ and both \mathbf{K}, \mathbf{W} have rank $n-v$, it follows that

$$\mathbf{Q} = \mathbf{P}(\mathbf{P}'\Sigma\mathbf{P})\mathbf{P}' = \mathbf{K}(\mathbf{K}\Sigma\mathbf{K})^{-1}\mathbf{K}.$$

Assuming that matrix \mathbf{A} is of full rank, the following results may be derived. Let

$$\begin{bmatrix} \mathbf{X}'\mathbf{D}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{D}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{D}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{D}^{-1}\mathbf{Z} + \mathbf{A}^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \cdot & \cdot \\ \cdot & \mathbf{T} \end{bmatrix}$$

where \mathbf{T} is that part of the inverse corresponding to $\mathbf{Z}'\mathbf{D}^{-1}\mathbf{Z} + \mathbf{A}^{-1}$ in the original matrix. Thus $\mathbf{T}^{-1} = \mathbf{Z}'\mathbf{K}\mathbf{Z} + \mathbf{A}^{-1}$. Also let $\mathbf{T}^* = (\mathbf{Z}'\mathbf{D}^{-1}\mathbf{Z} + \mathbf{A}^{-1})^{-1}$.

Theorem 4.4.2

$$(i) \quad \Sigma^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{Z}\mathbf{T}^*\mathbf{Z}'\mathbf{D}^{-1}$$

$$(ii) \quad \mathbf{Q} = \mathbf{K} - \mathbf{K}\mathbf{Z}\mathbf{T}\mathbf{Z}'\mathbf{K}$$

$$(iii) \quad \mathbf{Z}'\Sigma^{-1}\mathbf{Z} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{T}^*\mathbf{A}^{-1}$$

$$(iv) \quad \mathbf{Z}'\mathbf{Q}\mathbf{Z} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}$$

$$(v) \quad \mathbf{A}\mathbf{Z}'\Sigma^{-1}\mathbf{Z} = \mathbf{T}^*\mathbf{Z}'\mathbf{D}^{-1}\mathbf{Z}$$

$$(vi) \quad AZ'QZ = TZ'KZ$$

Proof.

$$\begin{aligned}
(i) \quad & \text{Since } \Sigma = D + ZAZ' \text{ and } T^* = (Z'D^{-1}Z + A^{-1})^{-1}, \\
& \Sigma (D^{-1} - D^{-1}ZT^*Z'D^{-1}) \\
&= (D + ZAZ') (D^{-1} - D^{-1}ZT^*Z'D^{-1}) \\
&= I - ZT^*Z'D^{-1} + ZAZ'D^{-1} - ZAZ'D^{-1}ZT^*Z'D^{-1} \\
&= I - ZT^*Z'D^{-1} + ZAZ'D^{-1} - ZA(Z'D^{-1}Z + A^{-1})T^*Z'D^{-1} + ZAA^{-1}T^*Z'D^{-1} \\
&= I - ZT^*Z'D^{-1} + ZAZ'D^{-1} - ZAZ'D^{-1} + ZT^*Z'D^{-1} \\
&= I
\end{aligned}$$

$$\text{Therefore, } \Sigma^{-1} = D^{-1} - D^{-1}ZT^*Z'D^{-1}.$$

$$\begin{aligned}
(ii) \quad & \text{Let } D = PP' \text{ where } P \text{ is of full rank and let } X = PX_1, Z = PZ_1 \text{ so} \\
& \text{that } \Sigma = P\Sigma_1P', \text{ where } \Sigma_1 = I + Z_1AZ'_1 \text{ and } K = P^{-1}K_1P^{-1}, \text{ where} \\
& K_1 = I - X_1(X'_1X_1)^{-1}X_1. \text{ Now } T^{-1} = Z'KZ + A^{-1} = Z'_1K_1Z_1 + A^{-1}, \text{ let } V_2 = K_1\Sigma_1K_1, \\
& V_1 = K_1 - K_1Z_1TZ'_1K_1, \text{ since } K_1 \text{ is idempotent,}
\end{aligned}$$

$$\begin{aligned}
V_1V_2 &= K_1\Sigma_1K_1 - K_1Z_1TZ'_1K_1\Sigma_1K_1 \\
&= K_1\Sigma_1K_1 - K_1Z_1TZ'_1K_1(I + Z_1AZ'_1)K_1 \\
&= K_1\Sigma_1K_1 - K_1Z_1TZ'_1K_1 - K_1Z_1TZ'_1K_1Z_1AZ'_1K_1 \\
&= K_1\Sigma_1K_1 - K_1Z_1TZ'_1K_1 - K_1Z_1T(A^{-1} + Z'_1K_1Z_1)AZ'_1K_1 + K_1Z_1TA^{-1}AZ'_1K_1 \\
&= K_1\Sigma_1K_1 - K_1Z_1TZ'_1K_1 - K_1Z_1AZ'_1K_1 + K_1Z_1TZ'_1K_1 \\
&= K_1\Sigma_1K_1 - K_1Z_1AZ'_1K_1 \\
&= K_1(\Sigma_1 - Z_1AZ'_1)K_1 \\
&= K_1IK_1 \\
&= K_1
\end{aligned}$$

$$\text{Similarly, } V_2V_1 = K_1.$$

$$\text{Now, } V_1V_2 = V_2V_1 = K_1, \text{ } K_1 \text{ is symmetric and idempotent, we then have}$$

$$(a) \quad V_1V_2V_1 = K_1V_1 = V_1 \quad (c) \quad (V_1V_2)' = K'_1 = K_1 = V_1V_2$$

$$(b) \quad V_2 V_1 V_2 = K_1 V_2 = V_2$$

$$(d) \quad (V_2 V_1)' = K_1' = K_1 = V_2 V_1$$

which satisfy the four Penrose conditions. Hence V_1 is a Moore-Penrose inverse of V_2 .

Since $V_2 = K_1 \Sigma_1 K_1$, $V_2 = (K_1 \Sigma_1 K_1)' = P^{-1}(K \Sigma K)^{-1}P^{-1}$, from Theorem 4.4.1

$$Q = K(K \Sigma K)^{-1}K$$

$$\begin{aligned} &= P^{-1}K_1 V_2 K_1 P^{-1} \\ &= P^{-1}V_1 V_2 V_2 V_1 P^{-1} \\ &= P^{-1}V_1 V_2 V_1 P^{-1} \\ &= P^{-1}V_1 P^{-1} \\ &= P^{-1}(K_1 - K_1 Z_1 T Z_1' K_1) P^{-1} \\ &= K - K Z T Z' K \end{aligned}$$

$$\text{Therefore, } Q = K(K \Sigma K)^{-1}K = K - K Z T Z' K$$

$$(iii) \text{ From (i), } \Sigma^{-1} = D^{-1} - D^{-1}ZT^*Z'D^{-1},$$

$$\begin{aligned} Z'\Sigma^{-1}Z &= Z'D^{-1}Z - Z'D^{-1}ZT^*Z'D^{-1}Z \\ &= Z'D^{-1}Z - (Z'D^{-1}Z + A^{-1})T^*Z'D^{-1}Z + A^{-1}T^*Z'D^{-1}Z \\ &= A^{-1}T^*Z'D^{-1}Z \\ &= A^{-1}T^*(Z'D^{-1}Z + A^{-1}) - A^{-1}T^*A^{-1} \\ &= A^{-1} - A^{-1}T^*A^{-1} \end{aligned}$$

$$(iv) \text{ From (ii), } Q = K - K Z T Z' K, \quad Z'QZ = Z'KZ - Z'K Z T Z' K Z, \text{ then follows similar simplification in (iii), we have } Z'QZ = A^{-1} - A^{-1}TA^{-1}.$$

$$(v) \text{ From (i), } \Sigma^{-1} = D^{-1} - D^{-1}ZT^*Z'D^{-1},$$

$$\begin{aligned} Z'\Sigma^{-1} &= Z'D^{-1} - Z'D^{-1}ZT^*Z'D^{-1} \\ &= Z'D^{-1} - Z'D^{-1} + A^{-1}T^*Z'D^{-1} \\ &= A^{-1}T^*Z'D^{-1} \end{aligned}$$

$$\text{Therefore, } AZ'\Sigma^{-1}Z = T^*Z'D^{-1}Z$$

$$(vi) \text{ From (ii), } Q = K - K Z T Z' K,$$

$$Z'Q = Z'K - Z'K Z T Z' K$$

$$= \mathbf{Z}'\mathbf{K} - \mathbf{Z}'\mathbf{K} + \mathbf{A}^{-1}\mathbf{T}\mathbf{Z}'\mathbf{K}$$

$$= \mathbf{A}^{-1}\mathbf{T}\mathbf{Z}'\mathbf{K}$$

Therefore, $\mathbf{A}\mathbf{Z}'\mathbf{Q}\mathbf{Z} = \mathbf{T}\mathbf{Z}'\mathbf{K}\mathbf{Z}$

Theorem 4.4.3

(i) $\mathbf{Z}'\Sigma^{-1}\mathbf{Z}_j = \theta_i^{-1}\mathbf{A}_i^{-1}\delta_{ij} - \theta_i^{-1}\theta_j^{-1}\mathbf{A}_i^{-1}\mathbf{T}_{ij}^*\mathbf{A}_j^{-1}$, where δ_{ij} = Kronecker delta.

(ii) $\mathbf{Z}'\mathbf{Q}\mathbf{Z}_j = \theta_i^{-1}\mathbf{A}_i^{-1}\delta_{ij} - \theta_i^{-1}\theta_j^{-1}\mathbf{A}_i^{-1}\mathbf{T}_{ij}\mathbf{A}_j^{-1}$.

Proof. Since $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_k]$ and $\mathbf{A} = \text{diag}[\theta_1\mathbf{A}_1 \ \theta_2\mathbf{A}_2 \ \dots \ \theta_k\mathbf{A}_k]$

(i) from Theorem 4.4.2 (iii), $\mathbf{Z}'\Sigma^{-1}\mathbf{Z} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{T}^*\mathbf{A}^{-1}$

$$\begin{aligned} \mathbf{Z}'\Sigma^{-1}\mathbf{Z}_j &= (\mathbf{i}, \mathbf{j}) \text{ block of } \mathbf{Z}'\Sigma^{-1}\mathbf{Z} \\ &= \{(\mathbf{i}, \mathbf{j}) \text{ block of } \mathbf{A}^{-1}\} - \{(\mathbf{i}, \mathbf{j}) \text{ block of } \mathbf{A}^{-1}\mathbf{T}^*\mathbf{A}^{-1}\} \\ &= \theta_i^{-1}\mathbf{A}_i^{-1}\delta_{ij} - \theta_i^{-1}\theta_j^{-1}\mathbf{A}_i^{-1}\mathbf{T}_{ij}^*\mathbf{A}_j^{-1} \end{aligned}$$

(ii) from Theorem 4.4.2 (iv), $\mathbf{Z}'\mathbf{Q}\mathbf{Z} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}$, by similar argument in (i), we obtain $\mathbf{Z}'\mathbf{Q}\mathbf{Z}_j = \theta_i^{-1}\mathbf{A}_i^{-1}\delta_{ij} - \theta_i^{-1}\theta_j^{-1}\mathbf{A}_i^{-1}\mathbf{T}_{ij}\mathbf{A}_j^{-1}$.

The following notation is useful in simplifying expressions. Let $\mathbf{T}^* = [\mathbf{T}_{ij}^*]$ be a partition of \mathbf{T}^* into blocks conformally to the partition of \mathbf{u} . We denote

$$\begin{aligned} v_j &= \text{tr } \mathbf{A}_j^{-1}\mathbf{A}_j, \quad v_j^{(s)} = \text{tr } \mathbf{A}_j^{-1}\partial\mathbf{A}_j/\partial\phi_s, \quad v_j^{(st)} = \text{tr } \partial\mathbf{A}_j^{-1}/\partial\phi_s \partial\mathbf{A}_j/\partial\phi_t, \\ r_j^* &= \theta_j^{-1} \text{tr } \mathbf{A}_j^{-1}\mathbf{T}_{jj}^*, \quad r_j^{*(s)} = \theta_j^{-1} \text{tr } \partial\mathbf{A}_j^{-1}/\partial\phi_s \mathbf{T}_{jj}^*, \\ r_j^{*(st)} &= \theta_j^{-1} \text{tr } \partial\mathbf{A}_j^{-1}/\partial\phi_s \mathbf{T}_{jj}^* \mathbf{A}_j^{-1} \partial\mathbf{A}_j/\partial\phi_t, \\ r_{ij}^* &= \text{tr } \mathbf{T}_{ij}^* \mathbf{A}_j^{-1} \mathbf{T}_{ji}^* \mathbf{A}_i^{-1}, \quad r_{ij}^{*(s)} = \text{tr } \mathbf{T}_{ij}^* \partial\mathbf{A}_j^{-1}/\partial\phi_s \mathbf{T}_{ji}^* \mathbf{A}_i^{-1}, \\ r_{ij}^{*(st)} &= \text{tr } \mathbf{T}_{ij}^* \partial\mathbf{A}_j^{-1}/\partial\phi_s \mathbf{T}_{ji}^* \partial\mathbf{A}_i^{-1}/\partial\phi_t \end{aligned}$$

If T^* is replaced by T in the above definitions, then r^* is replaced by r in all but the first line of the definitions. This is a definition of the equivalent $r, r^{(s)}, r^{(st)}$ terms with appropriate one or two subscripts.

Theorem 4.4.4

- (i) $\partial \ln |\Sigma| / \partial \theta_j = \text{tr } \Sigma^{-1} \partial \Sigma / \partial \theta_j = \text{tr } \Sigma^{-1} Z_j A_j Z'_j = \theta_j^{-1} (v_j - r_j^*)$
- (ii) $\partial \ln |\Sigma| / \partial \phi_s = \sum_{j=1}^k \theta_j \text{tr } \Sigma^{-1} Z_j \partial A_j / \partial \phi_s Z'_j = \sum_{j=1}^k (v_j^{(s)} + r_j^{*(s)})$
- (iii) $\partial \ln |K\Sigma K| / \partial \theta_j = \text{tr } (K\Sigma K)^{-1} K (\partial \Sigma / \partial \theta_j) K = \text{tr } Q Z_j A_j Z'_j = \theta_j^{-1} (v_j - r_j)$
- (iv) $\partial \ln |K\Sigma K| / \partial \phi_s = \text{tr } (K\Sigma K)^{-1} K (\partial \Sigma / \partial \phi_s) K = \sum_{j=1}^k \theta_j \text{tr } Q Z_j (\partial A_j / \partial \phi_s) Z'_j$
 $= \sum_{j=1}^k (v_j^{(s)} + r_j^{(s)})$

4.5 MAXIMUM LIKELIHOOD ESTIMATION

The loglikelihood function for y formed by integration over the distribution of u is

$$l_{ML} = -(1/2)[n \ln 2\pi\sigma^2 + \ln |\Sigma| + \sigma^{-2}(y-X\beta)' \Sigma^{-1} (y-X\beta)]$$

The derivatives with respect to the parameters β, σ^2, θ and ϕ are

$$\partial l_{ML} / \partial \beta = \sigma^{-2} X' \Sigma^{-1} (y-X\beta)$$

$$\partial l_{ML} / \partial \sigma^2 = -(1/2)[n\sigma^{-2} - \sigma^{-4} (y-X\beta)' \Sigma^{-1} (y-X\beta)]$$

$$\partial l_{ML} / \partial \theta_j = -(1/2)[\text{tr } \Sigma^{-1} \partial \Sigma / \partial \theta_j + \sigma^{-2} (y-X\beta)' (\partial \Sigma^{-1} / \partial \theta_j) (y-X\beta)]$$

$$\partial l_{ML} / \partial \phi_s = -(1/2)[\text{tr } \Sigma^{-1} \partial \Sigma / \partial \phi_s + \sigma^2 (y - X\beta)' (\partial \Sigma^{-1} / \partial \phi_s) (y - X\beta)]$$

giving

$$4.5.1 \quad \hat{\beta}_{ML} = \tilde{\beta} = H^{-1} X' \Sigma^{-1} y, \quad \text{where } H = X' \Sigma^{-1} X$$

$$4.5.2 \quad \hat{\sigma}_{ML}^2 = n^{-1} (y - X\hat{\beta}_{ML})' \Sigma^{-1} (y - X\hat{\beta}_{ML})$$

In general the equations for ϕ are not explicitly solvable although they may be for particular Σ . The information matrix I_{ML} is

$$\begin{bmatrix} \sigma^2 H & 0 & 0 \\ 0 & n/2\sigma^4 (1/2\sigma^2)[\text{tr } \Sigma^{-1} \partial \Sigma / \partial \theta_j] & (1/2\sigma^2)[\text{tr } \Sigma^{-1} \partial \Sigma / \partial \phi_t] \\ 0 & (1/2)\text{tr}[\Sigma^{-1} \partial \Sigma / \partial \theta_i \Sigma^{-1} \partial \Sigma / \partial \theta_j] & (1/2)\text{tr}[\Sigma^{-1} \partial \Sigma / \partial \theta_i \Sigma^{-1} \partial \Sigma / \partial \phi_t] \\ 0 & . & (1/2)\text{tr}[\Sigma^{-1} \partial \Sigma / \partial \phi_s \Sigma^{-1} \partial \Sigma / \partial \phi_t] \end{bmatrix}$$

In the current model the variance matrix has the form

$$\Sigma = D + ZAZ' = D + \sum_{j=1}^k \theta_j Z_j A_j(\phi) Z'_j$$

so that

$$\partial \Sigma / \partial \theta_j = Z_j A_j(\phi) Z'_j, \quad \partial \Sigma / \partial \phi_s = \sum_{j=1}^k \theta_j Z_j (\partial A_j / \partial \phi_s) Z'_j.$$

Note that if the ϕ parameter is not present in A_j then that derivative is zero.

Matrix Results

- (i) $\mathbf{Qy} = \Sigma^{-1}(y - \mathbf{X}\hat{\beta}_{ML}) = \mathbf{D}^{-1}[\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{D}^{-1}\mathbf{Z} + \mathbf{A}^{-1})^{-1}\mathbf{Z}'\mathbf{D}^{-1}](y - \mathbf{X}\hat{\beta}_{ML}) = \mathbf{D}^{-1}(y - \mathbf{X}\hat{\beta} - \mathbf{Z}\tilde{u})$
- (ii) $\mathbf{Z}'\mathbf{Qy} = \mathbf{Z}'\Sigma^{-1}(y - \mathbf{X}\hat{\beta}_{ML}) = \mathbf{Z}'\mathbf{D}^{-1}(y - \mathbf{X}\hat{\beta} - \mathbf{Z}\tilde{u}) = \mathbf{A}^{-1}\tilde{u}$
- (iii) $\mathbf{Z}'_j\mathbf{Qy} = \mathbf{Z}'_j\Sigma^{-1}(y - \mathbf{X}\hat{\beta}_{ML}) = \mathbf{Z}'_j\mathbf{D}^{-1}(y - \mathbf{X}\hat{\beta} - \mathbf{Z}\tilde{u}) = \theta_j^{-1}\mathbf{A}_j^{-1}\tilde{u}_j$
- (iv) $(y - \mathbf{X}\hat{\beta})' \partial \Sigma^{-1} / \partial \theta_j (y - \mathbf{X}\hat{\beta}) = -(y - \mathbf{X}\hat{\beta})' \Sigma^{-1} (\partial \Sigma / \partial \theta_j) \Sigma^{-1} (y - \mathbf{X}\hat{\beta}) = -\theta_j^{-2} \tilde{u}'_j \mathbf{A}_j^{-1} \tilde{u}_j$
- (v) $y'\mathbf{Q}(\partial \Sigma / \partial \theta_j)\mathbf{Qy} = \theta_j^{-2} \tilde{u}'_j \mathbf{A}_j^{-1} \tilde{u}_j$
- (vi) $(y - \mathbf{X}\hat{\beta})' \partial \Sigma^{-1} / \partial \phi_s (y - \mathbf{X}\hat{\beta}) = \sum_{j=1}^k \theta_j^{-1} \tilde{u}'_j (\partial \mathbf{A}_j^{-1} / \partial \phi_s) \tilde{u}_j$
- (vii) $y'\mathbf{Q}(\partial \Sigma / \partial \phi_s)\mathbf{Qy} = -\sum_{j=1}^k \theta_j^{-1} \tilde{u}'_j (\partial \mathbf{A}_j^{-1} / \partial \phi_s) \tilde{u}_j$

The proof is direct for each of the results. Using the above, and the matrix results of the previous section, we have

$$4.5.3 \quad \hat{\sigma}_{ML}^2 = n^{-1} y' \Sigma^{-1} (y - \mathbf{X}\hat{\beta}_{ML}) = n^{-1} y' \mathbf{D}^{-1} (y - \mathbf{X}\hat{\beta} - \mathbf{Z}\tilde{u}) = n^{-1} y' \mathbf{Qy}$$

$$\partial l_{ML} / \partial \theta_j \Big|_{\beta=\hat{\beta}} = -(1/2\theta_j)(v_j - r_j^* - \sigma_j^{-2} \tilde{u}'_j \mathbf{A}_j^{-1} \tilde{u}_j) = 0, \text{ giving}$$

$$4.5.4 \quad \hat{\sigma}_{j(ML)}^2 = \tilde{u}'_j \mathbf{A}_j^{-1} \tilde{u}_j / (v_j - r_j^*)$$

$$4.5.5 \quad \partial l_{ML} / \partial \phi_s \Big|_{\beta=\hat{\beta}} = -(1/2) \sum_{j=1}^k [v_j^{(s)} + r_j^{*(s)} + \hat{\sigma}_j^{-2} \tilde{u}'_j (\partial \mathbf{A}_j^{-1} / \partial \phi_s) \tilde{u}_j] \Big|_{\phi=\hat{\phi}} = 0,$$

giving equations which may be solved iteratively for ϕ . The information matrix, multiplied by 2, becomes

$$\begin{bmatrix} 2\sigma^{-2} \mathbf{H} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdot & n/\sigma^4 & \sigma_j^{-2} (v_j - r_j^*) & \sigma^{-2} \sum_{j=1}^k (v_j^{(t)} + r_j^{*(t)}) \\ \cdot & \cdot & [\theta_i^{-2} (v_i - 2r_i^*) \delta_{ij} + \theta_i^{-2} \theta_j^{-2} r_{ij}^*] & \theta_i^{-1} (v_i^{(t)} + 2r_i^{*(t)} - \sum_{j=1}^k \theta_i^{-1} \theta_j^{-1} r_{ij}^{(t)}) \\ \cdot & \vdots & \cdot & \sum_{j=1}^k [-v_j^{(st)} + 2r_j^{*(st)} + \sum_{m=1}^k \theta_j^{-1} \theta_m^{-1} r_{jm}^{(st)}] \end{bmatrix}$$

4.6 RESIDUAL MAXIMUM LIKELIHOOD ESTIMATION

Let l_{REML} be the loglikelihood function for residual maximum likelihood techniques. Since the matrix K defined in section 4.3 satisfies $KX=0$, an expression for l_{REML} given by Patterson and Thompson (1971) is

$$l_{\text{REML}} = -(1/2)[(n-v)\ln 2\pi\sigma^2 + \ln |K\Sigma K| + \sigma^2 y'K(K\Sigma K)^{-1}Ky]$$

where $|K\Sigma K|$ must be interpreted as the determinant of linearly independent rows and columns of $K\Sigma K$. The following development parallels Thompson (1980).

The first order derivatives of l_{REML} are

$$\begin{aligned}\partial l_{\text{REML}} / \partial \sigma^2 &= -(1/2)[(n-v)\sigma^{-2} - \sigma^{-4}y'K(K\Sigma K)^{-1}Ky] = -(1/2)[(n-v)\sigma^{-2} - \sigma^{-4}y'Qy], \\ \partial l_{\text{REML}} / \partial \theta_j &= -(1/2)\{\text{tr}[(K\Sigma K)^{-1}K\partial\Sigma/\partial\theta_j K] - \sigma^{-2}y'K(K\Sigma K)^{-1}K\partial\Sigma/\partial\theta_j K(K\Sigma K)^{-1}Ky\} \\ &= -(1/2)[\text{tr } Q\partial\Sigma/\partial\theta_j - \sigma^{-2}y'Q\partial\Sigma/\partial\theta_j Qy], \\ \partial l_{\text{REML}} / \partial \phi_s &= -(1/2)\{\text{tr}[(K\Sigma K)^{-1}K\partial\Sigma/\partial\phi_s K] - \sigma^{-2}y'K(K\Sigma K)^{-1}K\partial\Sigma/\partial\phi_s K(K\Sigma K)^{-1}Ky\} \\ &= -(1/2)[\text{tr } Q\partial\Sigma/\partial\phi_s - \sigma^{-2}y'Q\partial\Sigma/\partial\phi_s Qy]\end{aligned}$$

Thus

$$4.6.1 \quad \hat{\sigma}_{\text{REML}}^2 = (n-v)^{-1}y'Qy$$

and the REML information matrix I_{REML} is

$$\begin{bmatrix} (n-v)/2\sigma^4 & (1/2\sigma^2)\text{tr } Q \partial \Sigma / \partial \theta_j & (1/2\sigma^2)\text{tr } Q \partial \Sigma / \partial \phi_t \\ \cdot & (1/2) \text{tr } Q \partial \Sigma / \partial \theta_i Q \partial \Sigma / \partial \theta_j & (1/2) \text{tr } Q \partial \Sigma / \partial \theta_i Q \partial \Sigma / \partial \phi_t \\ \cdot & \cdot & (1/2) \text{tr } Q \partial \Sigma / \partial \phi_s Q \partial \Sigma / \partial \phi_t \end{bmatrix}$$

Using $\partial \Sigma / \partial \theta_j = Z_j A_j(\phi) Z'_j$, $\partial \Sigma / \partial \phi_s = \sum_{j=1}^k \theta_j Z_j (\partial A_j / \partial \phi_s) Z'_j$ and the matrix results in the previous sections, we have

$$\partial l_{\text{REML}} / \partial \theta_j = -(1/2\theta_j)[v_j - r_j - \sigma_j^{-2} \tilde{u}'_j A_j^{-1} \tilde{u}_j] = 0, \text{ giving}$$

$$4.6.2 \quad \hat{\sigma}_{j(\text{REML})}^2 = \tilde{u}'_j A_j^{-1} \tilde{u}_j / (v_j - r_j)$$

$$4.6.3 \quad \partial l_{\text{REML}} / \partial \phi_s = -(1/2) \sum_{j=1}^k [v_j^{(s)} + r_j^{(s)} + \hat{\sigma}_j^{-2} \tilde{u}'_j (\partial A_j^{-1} / \partial \phi_s) \tilde{u}_j] \Big|_{\phi=\hat{\phi}} = 0$$

Again this last equation may have to be solved iteratively for ϕ . The REML information matrix, multiplied by 2, is

$$\begin{bmatrix} (n-v)/\sigma^4 \sigma_j^{-2} (v_j - r_j) & \sigma^{-2} \sum_{j=1}^k (v_j^{(t)} + r_j^{(t)}) \\ \cdot & [\theta_i^{-2} (v_i - 2r_i) \delta_{ij} + \theta_i^{-2} \theta_j^{-2} r_{ij}] & \theta_i^{-1} (v_i^{(t)} + 2r_i^{(t)}) - \sum_{j=1}^k \theta_j^{-1} \theta_i^{-1} r_{ij}^{(t)} \\ \cdot & \cdot & \sum_{j=1}^k [-v_j^{(st)} + 2r_j^{(st)} + \sum_{m=1}^k \theta_j^{-1} \theta_m^{-1} r_{jm}^{(st)}] \end{bmatrix}$$

4.7 GENERALISED LINEAR MIXED MODELS

The extension of the theory to generalised linear mixed models has now been accomplished for a response vector y , which is not necessarily normally distributed, but has a distribution dependent on a vector quantity η which is related to vector regression variables through the equation

$$\eta = X\beta + Zu$$

If $f(y; \beta | u)$ is the probability (density) function of y conditional on fixed u , then $l_1 = \ln f(y; \beta | u)$ is the log-likelihood of y conditional on fixed u . Taking u to be normally distributed as in the previous sections, the logarithm of its probability density is l_2 as given in section 4.3. The sum $l = l_1 + l_2$ is then termed a penalised likelihood function and carries over the spirit of BLUP into a non-normal framework. In this sense l_2 is a penalty function for the conditional loglikelihood l_1 . The penalised likelihood estimators of β , u , equivalent to BLUP, are obtained by finding the likelihood derivatives

$$\partial l / \partial \beta = X' d l_1 / d \eta , \quad \partial l / \partial u_j = Z'_j d l_1 / d \eta - \sigma_j^{-2} A_j^{-1} u_j , \quad j=1,2,\dots,k.$$

and the second order derivatives of l are the same as the second order derivatives of l_1 except for

$$\partial^2 l / \partial u_j \partial u'_j = \partial^2 l_1 / \partial u_j \partial u'_j - \sigma_j^{-2} A_j^{-1} = -Z'_j B Z_j - \sigma_j^{-2} A_j^{-1}$$

where $B = -d^2 l_1 / d\eta d\eta'$. The Newton-Raphson iterative procedure for estimating β , u is

$$4.7.1 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_0 \end{bmatrix} + V \begin{bmatrix} X' \\ Z' \end{bmatrix} d l_1 / d\eta - V \begin{bmatrix} 0 \\ \sigma^{-2} A^{-1} u_0 \end{bmatrix}$$

$$\text{where } V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B[X, Z] + \begin{bmatrix} 0 & 0 \\ 0 & \sigma^{-2} A^{-1} \end{bmatrix}$$

If V is replaced by $E(V)$ then the iterative procedure becomes the method of scoring.

A heuristic approach in McGilchrist (1994) approximates l by a quadratic expression in the region of its maximum, viz.

$$l \approx \text{constant} + (1/2) \begin{bmatrix} \beta - \tilde{\beta} \\ u - \tilde{u} \end{bmatrix}' V \begin{bmatrix} \beta - \tilde{\beta} \\ u - \tilde{u} \end{bmatrix}.$$

In that case $\tilde{\beta}, \tilde{u}$ have approximately a joint normal distribution with mean β, u and variance matrix V .

An alternative formulation of the problem is given in McGilchrist and Aisbett (1991b) in which the component l_1 of the penalised likelihood procedure is replaced by the log-likelihood of $\hat{\beta}, \hat{u}$ as given by its approximate asymptotic distribution, viz. normal with mean β, u and variance matrix inverse given by the information matrix for $\hat{\beta}, \hat{u}$. In that case we may consider the BLUP estimation as having been derived from

the very approximate asymptotic distribution of $\hat{\beta}, \hat{u}$. If $I(\beta, u)$ is the sample information matrix for β, u derived from l_1 then

$$I(\beta, u) = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X, Z]$$

Replacing l_1 by l_1^* , the log-likelihood based on the approximate asymptotic distribution of $\hat{\beta}, \hat{u}$ and letting $y^* = X\hat{\beta} + Z\hat{u}$ gives

$$\begin{aligned} l_1^* &= \text{constant} - (1/2) \begin{bmatrix} \hat{\beta} - \beta \\ \hat{u} - u \end{bmatrix}' \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X, Z] \begin{bmatrix} \hat{\beta} - \beta \\ \hat{u} - u \end{bmatrix} \\ &= \text{constant} - (1/2)(y^* - X\beta - Zu)' B (y^* - X\beta - Zu) \end{aligned}$$

The formulation of the problem is now exactly as described for normal theory models with y^* replacing y , B in place of D^{-1} and $\sigma^2 = 1$ implying $\theta_j = \sigma_j^2$. It follows that estimators $\hat{\beta}, \hat{u}$ may be used to find ML and REML estimators.

The estimation procedure is as follows. For any given application, write down the loglikelihood l_1 as a function of $\eta = X\beta + Zu$, taking u to be conditionally fixed. Using initial estimates of θ, ϕ and letting $\beta = \beta_0, u = u_0$ be initial estimates of β, u solve equation 4.7.1 for $\tilde{\beta}, \tilde{u}$. Initial values are replaced by estimates in a new iteration and so on until convergence. Estimates of σ_j^2, ϕ are obtained from

$$4.7.2 \quad \hat{\sigma}_{j(ML)}^2 = \tilde{u}' A_j^{-1} \tilde{u}_j / (v_j - r_j^*)$$

$$4.7.3 \quad \hat{\sigma}_{j(REML)}^2 = \tilde{u}' A_j^{-1} \tilde{u}_j / (v_j - r_j)$$

and

$$4.7.4 \quad \sum_{j=1}^k [v_j^{(s)} + r_j^{*(s)} + \hat{\sigma}_j^{-2} \tilde{u}'_j \partial A_j^{-1} / \partial \phi_s \tilde{u}_j] = 0 \quad \text{for } \hat{\phi}_{s(ML)}, s=1,2,\dots,p$$

$$4.7.5 \quad \sum_{j=1}^k [v_j^{(s)} + r_j^{(s)} + \hat{\sigma}_j^{-2} \tilde{u}'_j \partial A_j^{-1} / \partial \phi_s \tilde{u}_j] = 0 \quad \text{for } \hat{\phi}_{s(REML)}, s=1,2,\dots,p$$

Information matrices are given for ML and REML in previous sections.

4.8 APPLICATION TO THE MULTIVARIATE FAILURE TIME PROBLEM

As indicated in Chapter 1, the main difficulty in analysing multivariate failure time data is when a proportional hazards function including a frailty component is constructed, the marginal failure time distribution formed by integrating out the random frailty component loses the simple properties of the original hazard function formulation, particularly the cancellation of the baseline hazard function in the partial likelihood procedure. The GLMM approach, as outlined in McGilchrist (1994), preserves the simple cancellation property of the baseline hazard function in the partial likelihood as in Cox proportional hazards model.

The recurrent failure time data (Type III) is considered here as an example to illustrate how the cancellation property is preserved in the GLMM approach. Suppose each individual has several failure times, the j^{th}

failure or censoring of the i^{th} individual is denoted by T_{ij} . The proportional hazards model is

$$h(t_{ij}; z_{ij}, U_{ij}) = \lambda(t_{ij})g(\eta_{ij}) \quad , \quad \eta_{ij} = x'_{ij}\beta + U_{ij}$$

where $h(\cdot)$, $\lambda(\cdot)$ and $g(\cdot)$ have the usual interpretation as in 3.1.1

x_{ij} and U_{ij} are the associated risk variable vector and random effect respectively

$$\text{Let } u' = (u_1' \ u_2' \ \dots \ u_M') \quad u_i' = (U_{i1} \ U_{i2} \ \dots \ U_{in_i})$$

$$\text{var } u = A$$

Let the failure/censoring times be arranged in ascending order and t_n be the n^{th} such time with

$$D_{ijn} = \begin{cases} 1 & \text{if patient } i \text{ with } j^{\text{th}} \text{ recurrence time fails at } t_n \\ 0 & \text{otherwise} \end{cases}$$

Then the probability that the j^{th} failure observation of patient i is the event that occurs at t_n is

$$p_{ijn} = h(t_{ij}; x_{ij}, U_{ij}) / \sum_{t_{kl} \geq t_{ij}} h(t_{kl}; x_{kl}, U_{kl})$$

$$= \lambda(t_{ij})g(\eta_{ij}) / \sum_{t_{kl} \geq t_{ij}} \lambda(t_{kl})g(\eta_{kl})$$

$$= g(\eta_{ij}) / \sum_{t_{kl} \geq t_{ij}} g(\eta_{kl})$$

Note that the baseline hazard function $\lambda(t_{ij})$ is cancelled out in this expression.

Hence, the log-partial likelihood on conditionally fixed U values gives

$$l_1 = \sum_n \sum_i \sum_j D_{ijn} \ln p_{ijn}$$

Followed by differentiating l_1 with respect to the parameters accordingly, the ML and REML estimation of the parameters are obtained from 4.7.1 to 4.7.5. Inference on parameters are then based on the information matrices given in Section 4.5 and 4.6. Therefore, the GLMM technique developed in this Chapter serves as a general tool to analyse different types of multivariate failure time data. Applications of this technique are provided in the subsequent Chapters.

CHAPTER FIVE

TIME DEPENDENT FRAILTY MODEL

5.1 INTRODUCTION

In Chapter 4, the estimation and inference techniques in the Generalised Linear Mixed Model (GLMM) provide a general method to analyse multivariate failure time data. In this and the following Chapter, the CGD data set described in Chapter 2 is considered. Note that the CGD data belongs to the Type III data, which is the multiple recurrent failure events in the same patient. Early work on the Maximum Likelihood (ML) and Residual Maximum Likelihood (REML) estimation of this type of data based on the GLMM can be found in McGilchrist (1993).

Recall that the CGD study was a placebo controlled randomized trial conducted by the International CGD Cooperative Study Group in the late 1980's. The main research interest was to study the ability of gamma interferon (γ -IFN) to reduce the rate of infections in CGD patients. During the study, each patient may experience several failures (infections). Censoring occurs usually due to the termination of the study.

The original CGD data set can be found in Fleming and Harrington (1993). The first data set contained in Appendix I is the modified CGD data set, it is arranged in such a format that it fits the APL programs

given in Appendix II. The format of the CGD data given in Appendix I is described as follows:

Column 1: patient number

Column 2: failure/censoring time, in days

Column 3: censoring indicator 0=censored 1=failure

Column 4: treatment 0=placebo 1= γ -IFN

Column 5: pattern of inheritance 0=autosomal recessive 1=X-linked

Column 6: age, in years

Column 7: height, in cm

Column 8: weight, in kg

Column 9: using corticosteroids at time of study 0=no 1=yes

Column 10: using prophylactic antibiotics at time of study 0=no 1=yes

Column 11: sex 1=male 2=female

Column 12: US-NIH 0=no 1=yes

Column 13: US-other 0=no 1=yes

Column 14: Europe-Amsterdam 0=no 1=yes

Column 15: time since first failure, in years

Columns 4 to 14 are the risk variables mentioned in Fleming and Harrington (1993). The hospital category is modified into Column 12 to 14 by choosing Europe-other as baseline. Column 15 is "time since first failure" variable, which is added here when considering time dependent frailty models (Model 1 to 3 in Section 5.2). The importance of including this variable in the analysis has been noted by Self (1993) and Lindsey (1995).

A time independent and three time dependent frailty models are described in the following Section. The time independent frailty model, which assumes the frailty of each patient is constant over time, is equivalent to the model described in McGilchrist (1993). For the three time dependent frailty models, a parameter representing change over time is introduced and is modelled into a fixed effect, a normally distributed random effect and a longitudinal effect in which the random component relates to the patient characteristics. The ML and REML estimators for these models are derived in Section 5.3. Its application to the CGD data is provided in Section 5.4. A comparable stochastic modelling of frailty models applied to this data set using the counting processes approach can be found in Self (1993). In this paper, the frailty process is described as a time independent and time dependent Gamma distributed frailty process.

5.2 MODELS

For the Type III data considered in this and the following Chapter, each patient is followed until an event termed a "failure" occurs with the possibility that there may be a sequence of several such failures for any given patient. Usually only the last of these failure times is censored due to the termination of study. Let

5.2.1 $h(t;i,j)$ = hazard function for j^{th} failure episode of patient i

at time t measured from the beginning of the current episode. We consider M patients with n_i failures for patient i so that $i=1,2,\dots,M$; $j=1,2,\dots,n_i$.

The hazard function at t for any given i,j may also depend on a quantity η_{ij} which is taken to be a linear combination of risk variables and possibly also random frailty components which reflect variation of the risk for a given patient which is not accounted for by the measured set of risk variables. The proportional hazards model is

$$5.2.2 \quad h(t;i,j) = \lambda(t)g(\eta_{ij}), \quad i = 1,2,\dots,M; \quad j = 1,2,\dots,n_i$$

where $\lambda(t)$ is a baseline hazard function and $g(\eta_{ij})$ is the combined effect of all risk components on that baseline hazard. For the Cox model, $g(\eta_{ij}) = \exp \eta_{ij}$.

The combined risk variable η_{ij} is derived from a v dimensional vector of measured risk variables x_{ij} and a random (frailty) component U_{ij} such that

$$5.2.3 \quad \eta_{ij} = x'_{ij}\beta + U_{ij}$$

where β is a vector of regression coefficients. The models considered here differ in the way the frailty components U_{ij} are considered. Note that here we take the only possible variation of the hazard function, in going from one failure episode to the next, to be a variation in the frailty of the patient.

Time independent frailty model: This model assumes the frailty for each patient to be the same for all episodes as considered by McGilchrist (1993) in analysing a kidney patient data set. So, the random frailty term is considered to be constant over time for each of the individuals.

Let

$$5.2.4 \quad U_{ij} = A_i, \quad A_i \text{ independent } N(0, \theta_1)$$

The term A_i can be interpreted as the frailty of the i^{th} patient. It then gives

$$5.2.5 \quad \eta_{ij} = x'_{ij}\beta + A_i$$

where β are fixed parameters and A_i are random components.

Model 1: In this model the frailty U_{ij} consists of a random patient effect plus a deterministic change over time which is the same for all patients. Specifically

$$5.2.6 \quad U_{ij} = A_i + \tau_{ij}\gamma, \quad A_i \text{ independent } N(0, \theta_1)$$

The frailty of patient i is denoted by A_i initially and changes over time deterministically where τ_{ij} is the time from the first serious failure until the time of current failure episode. For this model,

$$5.2.7 \quad \eta_{ij} = x'_{ij}\beta + \tau_{ij}\gamma + A_i$$

where β , γ are fixed parameters and A_i are random components.

Model 2: A sequence of failure observations on the same person is a repeated measure and, in this and the subsequent model, extensions to model 1 are similar to those considered in repeated measures analysis. For model 2, the coefficient of τ_{ij} is allowed to vary over patients giving

$$5.2.8 \quad U_{ij} = A_i + \tau_{ij}(\gamma + B_i), \quad A_i \text{ independent } N(0, \theta_1), \quad B_i \text{ independent } N(0, \theta_2).$$

Here, the frailty of patient i is denoted by A_i and changes over time linearly but with coefficient different for each patient. Thus

$$5.2.9 \quad \eta_{ij} = x'_{ij}\beta + \tau_{ij}\gamma + A_i + \tau_{ij}B_i$$

where β , γ are fixed parameters and A_i , B_i are random components.

Model 3: A further extension is to relate the coefficients of τ_{ij} for the different patients to the measured patient characteristics, giving

$$5.2.10 \quad U_{ij} = A_i + \tau_{ij}(\gamma + x'_{ij}\psi + B_i), \quad A_i, B_i \text{ distributed as in model 2.}$$

This frailty model would apply if the coefficient of the time trend in frailty could be predicted from patient characteristics, an obviously important aspect of the variation in frailty problem. It leads to

$$5.2.11 \quad \eta_{ij} = x'_{ij}\beta + \tau_{ij}\gamma + \tau_{ij}x'_{ij}\psi + A_i + \tau_{ij}B_i$$

where β, γ, ψ are fixed parameters and A_i, B_i are random components.

5.3 ESTIMATION

The basic structure of the estimation process is described in Chapter 4. In applying the method of estimation, we reorder the η_{ij} according to the failure time. After reordering, let

$$5.3.1 \quad \eta = Xb + Zu$$

where b, u represent vectors of fixed parameters and random components respectively. The entries in the matrices X and Z correspond to b and u and are reordered according to the reordering pattern of η .

The estimation process starts with a Newton-Raphson iterative procedure 5.3.2 which is analogous to 4.7.1. For fixed value of θ 's and given a set of initial values b_0, u_0 , the BLUP estimators maximize the partial likelihood of the observed survival times for conditionally fixed u . This partial likelihood is denoted by l_1 and is the usual expression given by Cox, viz.

$$l_1 = \sum_{i=1}^N D_i [\eta_i - \ln \sum_{j=i}^N \exp(\eta_j)]$$

where D_i is the failure/censoring indicator which is zero if censoring occurs and one if failure occurs, η_i is the i^{th} entry of the reordered vector η and N is the total number of observations $= \sum_{i=1}^M n_i$.

Equation 5.3.2 below gives the first step of the Newton-Raphson procedure which becomes iterative by replacing the initial values with the result of the previous iteration. Letting $A = \text{var } u$ we have

$$5.3.2 \quad \begin{bmatrix} \tilde{b} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} b_0 \\ u_0 \end{bmatrix} - V^{-1} \begin{bmatrix} 0 \\ A^{-1}u_0 \end{bmatrix} + V^{-1} [X \ Z]' \frac{dl_1}{d\eta_0}$$

$$\text{where } V = \begin{bmatrix} -\frac{\partial^2 l_1}{\partial b \partial b}, & -\frac{\partial^2 l_1}{\partial b \partial u} \\ -\frac{\partial^2 l_1}{\partial u \partial b}, & -\frac{\partial^2 l_1}{\partial u \partial u} + A^{-1} \end{bmatrix} = [X' \ Z'] \left(-\frac{d^2 l_1}{d\eta d\eta} \right) [X \ Z] + \begin{bmatrix} 0 & 0 \\ 0 & A^{-1} \end{bmatrix}$$

The expressions for $\frac{dl_1}{d\eta}$ and $-\frac{d^2 l_1}{d\eta d\eta}$ are simplified below:

$$\text{Letting } w_k = \exp \eta_k, \quad a_k = D_k / \sum_{j=k}^N w_j, \quad b_k = \sum_{j=1}^k a_j,$$

$$W = \text{diag}(w_1, w_2, \dots, w_N), \quad A = \text{diag}(a_1, a_2, \dots, a_N),$$

M = lower triangular matrix with ones on/below the principal diagonal,

$$B = \text{diag}(b_1, b_2, \dots, b_N) = \text{diag}(MAI), \quad I = \text{vector of ones},$$

$$d' = [D_1, D_2, \dots, D_N], \quad \eta' = [\eta_1, \eta_2, \dots, \eta_N]$$

then

$$5.3.3 \quad \frac{dl_1}{d\eta} = d - WMAI \quad \text{and} \quad -\frac{d^2 l_1}{d\eta d\eta} = WB - WMA^2 M' W$$

In general, for given θ 's, 5.3.2 gives exactly the same estimates of b in BLUP, ML and REML; but since we have different estimating equations for θ in different methods, different values of b estimates are obtained from the three different methods in each of the models.

Time independent frailty model and Model 1: The estimation procedure of these two models are essentially the same except the different components contained in the fixed effect parameter b . In time independent frailty model, b corresponds to the parameters β while it corresponds to the parameters β and γ in model 1. In both models, the random effect u contains the random effects A_i described in the previous section. The matrix V is then partitioned conformally to $b|u$ as

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \quad V^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Letting $W_{22}^1 = V_{22}$, results in Section 4.5 give

$$5.3.4 \quad \text{var } \tilde{b} = A_{11}$$

$$5.3.5 \quad \hat{\theta}_{1(\text{ML})} = M^{-1}(\text{tr } W_{22} + \tilde{u}'\tilde{u})$$

$$5.3.6 \quad \text{var } \hat{\theta}_{1(\text{ML})} = 2\theta_1^2[M - 2\theta_1^{-1}\text{tr } W_{22} + \theta_1^{-2}\text{tr}(W_{22}^2)]^{-1}$$

For REML estimation, results in Section 4.6 give

$$5.3.7 \quad \hat{\theta}_{1(\text{REML})} = M^{-1}(\text{tr } A_{22} + \tilde{u}'\tilde{u})$$

$$5.3.8 \quad \text{var } \hat{\theta}_{1(\text{REML})} = 2\theta_1^2[M - 2\theta_1^{-1}\text{tr } A_{22} + \theta_1^{-2}\text{tr}(A_{22}^2)]^{-1}$$

where M is the number of patients.

Models 2 and 3: In model 2 the fixed parameter b corresponds to β, γ as in model 1 but the random component vector u consists of two components u_1, u_2 corresponding to the random components A_i and B_i respectively. The V matrix is partitioned conformally to $b | u_1 | u_2$ as

$$V = \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{bmatrix} \quad V^{-1} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

Letting $W = \begin{bmatrix} V_{22} & V_{23} \\ V_{32} & V_{33} \end{bmatrix}$, $W^{-1} = \begin{bmatrix} W_{22} & W_{23} \\ W_{32} & W_{33} \end{bmatrix}$ we have

5.3.9 $\text{var } \tilde{b} = A_{11}$.

Model 3 differs from model 2 only in the composition of the fixed parameter vector b which then contains β, γ, ψ so that the dimension of b in Model 2 is $v+1$ while it is $2v+1$ in model 3, where v is number of risk variables ($\dim \beta$). The partition of the V matrix is therefore the same for model 3 as it is for model 2 so that, subject only to change in the X matrix between the two models, the estimation process is essentially the same.

The development of ML and REML estimates of θ is derived from 4.7.2 and 4.7.3 respectively. Its asymptotic variances are given in Section 4.5 and Section 4.6. Expressions for the estimates of θ 's which are similar to those in model 1 are obtained as

$$5.3.10 \quad \hat{\theta}_{1(ML)} = M^{-1}(\operatorname{tr} W_{22} + \tilde{u}_1' \tilde{u}_1)$$

$$5.3.11 \quad \hat{\theta}_{2(ML)} = M^{-1}(\operatorname{tr} W_{33} + \tilde{u}_2' \tilde{u}_2)$$

We have also the variance matrix for $\hat{\theta}_{1(ML)}, \hat{\theta}_{2(ML)}$ as

$$\begin{bmatrix} (1/2\theta_1^2)[M - 2\theta_1^{-1}\operatorname{tr} W_{22} + \theta_1^{-2}\operatorname{tr} W_{22}^2] & (1/2\theta_1^2\theta_2^2)\operatorname{tr} W_{32}W_{23} \\ (1/2\theta_1^2\theta_2^2)\operatorname{tr} W_{32}W_{23} & (1/2\theta_2^2)[M - 2\theta_2^{-1}\operatorname{tr} W_{33} + \theta_2^{-2}\operatorname{tr} W_{33}^2] \end{bmatrix}^{-1}$$

Similarly, for REML estimation,

$$5.3.12 \quad \hat{\theta}_{1(REML)} = M^{-1}(\operatorname{tr} A_{22} + \tilde{u}_1' \tilde{u}_1)$$

$$5.3.13 \quad \hat{\theta}_{2(REML)} = M^{-1}(\operatorname{tr} A_{33} + \tilde{u}_2' \tilde{u}_2)$$

We have also the variance matrix for $\hat{\theta}_{1(REML)}, \hat{\theta}_{2(REML)}$ as

$$\begin{bmatrix} (1/2\theta_1^2)[M - 2\theta_1^{-1}\operatorname{tr} A_{22} + \theta_1^{-2}\operatorname{tr} A_{22}^2] & (1/2\theta_1^2\theta_2^2)\operatorname{tr} A_{32}A_{23} \\ (1/2\theta_1^2\theta_2^2)\operatorname{tr} A_{32}A_{23} & (1/2\theta_2^2)[M - 2\theta_2^{-1}\operatorname{tr} A_{33} + \theta_2^{-2}\operatorname{tr} A_{33}^2] \end{bmatrix}^{-1}$$

In each case, we start with a set of initial values b_0, u_0 and θ_0 's in equation 5.3.2 and then iterate alternately using equation 5.3.2 and the equation(s) of θ 's. Estimated asymptotic variance of \tilde{b} and $\hat{\theta}$'s in ML and REML are given by putting the estimates into the appropriate expression.

5.4 APPLICATION

All the models described in Section 5.2 are fitted and estimates with standard errors are reported in Table 5.4.1. In our analysis, we use the whole set of variables, but only those variables which are statistically significant are listed.

Table 5.4.1. ML and REML estimates of parameters (with standard errors) for the time independent and the three time dependent frailty models described in section 5.2.

Time independent frailty model

Estimate (S.E.)

Variables	ML	REML
γ -IFN	-1.1495 (0.3228)**	-1.1435 (0.3463)**
Inheritance	-0.7122 (0.3589)*	-0.7120 (0.3988)
Age	-0.0885 (0.0423)*	-0.0897 (0.0457)*
Corticosteroid	2.1182 (0.8149)**	2.1668 (0.9184)*
$\hat{\theta}_1$	0.4850 (0.2632)	0.7993 (0.3556)

Model 1

Estimate (S.E.)

Variables	ML	REML
γ -IFN	-1.1228 (0.3052)**	-1.1218 (0.3325)**
Inheritance	-0.6614 (0.3276)*	-0.6742 (0.3756)
Age	-0.0849 (0.0401)*	-0.0870 (0.0441)*
Corticosteroid	1.9998 (0.7274)**	2.0909 (0.8589)*
$\tilde{\gamma}$	1.1462 (0.4913)*	0.8770 (0.5082)
$\hat{\theta}_1$	0.2373 (0.2114)	0.5933 (0.3162)

Model 2

Estimate (S.E.)

Variables	ML	REML
γ -IFN	-1.1232 (0.3053)**	-1.1226 (0.3330)**
Inheritance	-0.6674 (0.3279)*	-0.6821 (0.3766)
Age	-0.0848 (0.0400)*	-0.0871 (0.0442)*
Corticosteroid	2.0040 (0.7261)**	2.1002 (0.8603)*
$\tilde{\gamma}$	1.1456 (0.4970)*	0.8678 (0.5163)
$\hat{\theta}_1$	0.2323 (0.2256)	0.5951 (0.3240)
$\hat{\theta}_2$	0.0730 (0.8750)	0.1023 (1.1946)

Model 3

Estimate (S.E.)

Variables	ML	REML
γ -IFN	-1.1879 (0.3450)**	-1.1722 (0.3751)**
Inheritance	-0.9258 (0.3779)*	-0.9222 (0.4281)*
Age	-0.0822 (0.0411)*	-0.0872 (0.0462)
Corticosteroid	2.1325 (0.8061)**	2.2897 (0.9581)*
$\tilde{\gamma}$	-2.4659 (7.4474)	-2.2842 (8.1247)
$\hat{\theta}_1$	0.2650 (0.2316)	0.7576 (0.3718)
$\hat{\theta}_2$	0.0610 (0.9452)	0.0991 (1.9960)

* significant at 5% level

** significant at 1% level.

In the time independent frailty model, the frailty of a patient is assumed to be constant. In model 1, the frailty of a patient is represented by an individual random effect which may change linearly over time but with the coefficient of the linear term fixed for all patients. Model 2 allows that coefficient of the linear time effect to vary randomly from patient to patient while model 3 extends model 2 such that variation of the linear time effect coefficient may be related also to patient characteristics. As can be seen from the results of the analysis, in model 1, the longitudinal parameter (γ) is significant at 5% level in ML and at 10% level in REML. Such significance of γ is also reflected by the considerable reduction in the variability of patient's frailty (θ_1) from time independent frailty model to model 1. Moreover, for models 2 and 3, the estimates of θ_2 are small compared to their standard errors. Also all the regression parameters in ψ of model 3 are not statistically significant. Hence, for this data, it is not necessary to extend the model from the deterministic time effect on frailty, as given by model 1, to the more complex models of random coefficient repeated measure effects given in models 2 and 3. Never-the-less, the extension given here is useful in that there is an indication of how such models may be fitted and the method is then available to encourage its consideration in future data collection.

When we examine the significant variables, the presence of γ -IFN decreases the rate of serious infection in CGD patients significantly. The X-linked pattern of inheritance is significant in ML but not in REML. Age is another significant variable which shows that the rate of serious

infection is higher in children than in adults. Another important variable is corticosteroid, the analysis shows that the using of corticosteroid at time of study entry will increase the rate of serious infection in CGD patients. The longitudinal parameter (γ) is significant in ML at 5% level and is significant in REML at 10% level indicating that as the time to failure increases, the rate of infection in CGD patient also increases.

The data have also been analysed in Fleming and Harrington (1991) using Cox's model and Andersen-Gill multiplicative intensity model. While not exactly comparable, an idea of the robustness of the estimation may be obtained by a comparison of the estimates and standard errors of $\hat{\beta}$ (the γ -IFN) parameter in Table 5.4.2. The estimation of β is consistent in the four models using ML or REML. It also agrees with previous analyses using the Cox model and Andersen-Gill multiplicative intensity model. The current analysis gives a slightly smaller value of standard error and which in turn provides narrower 95% confidence interval for the hazard ratio in $\exp \beta$.

Table 5.4.2. Estimates (with standard errors) of the effect of γ -IFN together with an estimate and confidence interval for the equivalent hazard ratio.

	$\hat{\beta}$	S.E.	Z Stat	$\exp \hat{\beta}$	95% C.I. hazard ratio
Cox model	-1.2063	0.4398	-2.7428	0.2993	(0.1264, 0.7088)
AG model	-1.2765	0.3774	-3.3824	0.2790	(0.1332, 0.5846)
TIF model (ML)	-1.1495	0.3228	-3.5424	0.3168	(0.1683, 0.5964)
TIF model (REML)	-1.1435	0.3463	-3.3021	0.3187	(0.1617, 0.6283)
Model 1 (ML)	-1.1228	0.3052	-3.6789	0.3254	(0.1789, 0.5918)
Model 1 (REML)	-1.1218	0.3325	-3.3738	0.3257	(0.1697, 0.6249)
Model 2 (ML)	-1.1232	0.3053	-3.6790	0.3252	(0.1788, 0.5917)
Model 2 (REML)	-1.1226	0.3330	-3.3712	0.3254	(0.1694, 0.6251)
Model 3 (ML)	-1.1879	0.3450	-3.4432	0.3049	(0.1550, 0.5995)
Model 3 (REML)	-1.1722	0.3751	-3.1250	0.3097	(0.1485, 0.6460)

CHAPTER SIX

AR(1) FRAILTY MODEL

In Chapter 5, a time independent frailty model and three time dependent models are introduced. These models are mainly developed for Type III data, the multiple recurrent failure time data. The three time dependent frailty models are used when the frailty of each patient is considered to be varying over time.

In this Chapter, the modelling of recurrent failure time data is developed from another direction. Since a sequence of failure observations on the same patient is a repeated measure, the frailty of each patient is considered to follow a time series process. Hence, a correlation parameter ϕ may be present in the variance matrix of the random components. A method of estimation for correlated frailty models in survival analysis is described in the following Section. The AR(1) frailty model, which considers the frailty of each patient follows an AR(1) process, will be provided in Section 6.3. It is then applied to analyse the effectiveness of gamma interferon (γ -IFN) in reducing the number of serious infections in Chronic Granulomatous Disease (CGD) patients.

6.1 MODEL AND ESTIMATION

For Type III data, patients are followed over time and failure times are recorded. In this context, failure refers to serious infection by a

certain disease. In such a case, each patient may have more than one failure observation.

A proportional hazards model for the j^{th} observation of patient i at time t is given by

$$6.1.1 \quad h(t;i,j) = \lambda(t) g(\eta_{ij}) ,$$

$$i = 1, 2, \dots, M; \quad j = 1, 2, \dots, n_i; \quad \sum_{i=1}^M n_i = N$$

where M is the number of patients and N is the total number of observations. The function $\lambda(t)$ is the baseline hazard function and for the Cox hazard function, $g(\eta_{ij}) = \exp(\eta_{ij})$. The combined risk variable η_{ij} is a linear combination of the known risk variables contained in a v dimensional vector x_{ij} having fixed regression coefficient β , together with a random component U_{ij} , viz.

$$6.1.2 \quad \eta_{ij} = x'_{ij}\beta + U_{ij}$$

The random component U_{ij} is the residual variation in risk for the (i,j) failure time not accounted for by the regression on known risk variables and is termed the frailty of the patient. It is modelled as a time series with its variance given by θ and covariance structure parametrized by a vector ϕ . The frailty vector u is taken to be distributed as $N[\theta, \theta A]$, where $u' = [u'_1, u'_2, \dots, u'_M]$; $u'_i = [U_{i1}, U_{i2}, \dots, U_{in_i}]$ and the frailties on different patients are taken to be independent, so that

$$A = \begin{pmatrix} A_1(\phi) & 0 & \dots & 0 \\ 0 & A_2(\phi) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_M(\phi) \end{pmatrix}$$

BLUP estimation maximizes the sum of two components, $l = l_1 + l_2$, where l_1 is the partial likelihood of failure times taking u fixed and l_2 is the logarithm of the probability density function of u , viz.

$$l_2 = \text{constant} - (1/2)(N \ln \theta + \sum_{i=1}^M \ln |A_i| + \theta^{-1} u' A^{-1} u)$$

By reordering the failure/censoring times (T) of the observations, η_i represents the η value corresponding to the i^{th} occurring failure/censoring time (T_i). The reorganised values η_i form a vector denoted by η and matrix X has rows of the correspondingly reorganised x_{ij} vectors, while Z is a matrix such that Zu gives the reorganised vector of U_{ij} values. Thus $\eta = X\beta + Zu$. For Cox proportional hazards model with u conditionally fixed,

$$l_1 = \sum_{i=1}^N D_i [\eta_i - \ln \sum_{j=i}^N \exp(\eta_j)]$$

where D_i is the failure/censoring indicator which is zero if censoring occurs and one if failure occurs. The first order derivatives are

$$\partial l / \partial \beta = \partial l_1 / \partial \beta$$

$$\partial l / \partial u_j = \partial l_1 / \partial u_j - \theta^{-1} A_j^{-1} u_j$$

For given values of θ and ϕ , the Newton-Rapson iterative procedure gives the solution of these equations for β , u as the iterative application of

$$6.1.3 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_0 \end{bmatrix} + V^{-1}[X, Z]'(dl_1/d\eta_0) - V^{-1} \begin{bmatrix} 0 \\ \theta^{-1}A^{-1}u_0 \end{bmatrix}$$

where β_0 , u_0 are initial values and $\tilde{\beta}$, \tilde{u} closer approximations to the BLUP estimates. Equation 6.1.3 above is analogous to 4.7.1 in Chapter 4. In equation 6.1.3, the vector $\eta_0 = X\beta_0 + Zu_0$ and if $B = -d^2l_1/d\eta d\eta'$

$$V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X, Z] + \begin{bmatrix} 0 & 0 \\ 0 & \theta^{-1}A^{-1} \end{bmatrix} \text{ evaluated at } \eta = \eta_0.$$

Expressions for $dl_1/d\eta$ and $-d^2l_1/d\eta d\eta'$ are provided in 5.5.3

Let $V^{-1} = \begin{bmatrix} V_1 & \cdot \\ \cdot & T \end{bmatrix}$ and $[Z'BZ + \theta^{-1}A^{-1}]^{-1} = T^*$

From equations 4.7.2 and 4.7.3, the maximum likelihood $\hat{\theta}_{ML}$ and residual maximum likelihood $\hat{\theta}_{REML}$ estimators of θ are respectively given by

$$6.1.4 \quad \hat{\theta}_{ML} = N^{-1}(\text{tr } A^{-1}T^* + \tilde{u}'A^{-1}\tilde{u})$$

$$6.1.5 \quad \hat{\theta}_{REML} = N^{-1}(\text{tr } A^{-1}T + \tilde{u}'A^{-1}\tilde{u})$$

and a similar derivation following equation 4.7.4 yields estimating equations for the ML estimator of ϕ as

$$6.1.6 \quad \text{tr } A^{-1} \frac{\partial A}{\partial \phi} = \hat{\theta}_{ML}^{-1} [\tilde{u}' A^{-1} \frac{\partial A}{\partial \phi} A^{-1} \tilde{u} + \text{tr } T^* A^{-1} \frac{\partial A}{\partial \phi} A^{-1}]$$

or equivalently,

$$6.1.7 \quad \text{tr } A \frac{\partial A^{-1}}{\partial \phi} = \hat{\theta}_{ML}^{-1} [\tilde{u}' \frac{\partial A^{-1}}{\partial \phi} \tilde{u} + \text{tr } T^* \frac{\partial A^{-1}}{\partial \phi}]$$

and an estimating equation follows from 4.7.5 for the REML estimator of ϕ as

$$6.1.8 \quad \text{tr } A^{-1} \frac{\partial A}{\partial \phi} = \hat{\theta}_{REML}^{-1} [\tilde{u}' A^{-1} \frac{\partial A}{\partial \phi} A^{-1} \tilde{u} + \text{tr } T A^{-1} \frac{\partial A}{\partial \phi} A^{-1}]$$

or equivalently,

$$6.1.9 \quad \text{tr } A \frac{\partial A^{-1}}{\partial \phi} = \hat{\theta}_{REML}^{-1} [\tilde{u}' \frac{\partial A^{-1}}{\partial \phi} \tilde{u} + \text{tr } T^* \frac{\partial A^{-1}}{\partial \phi}]$$

In section 6.3, these estimating equations are particularised to the case of the frailty having a first order autoregressive structure. In that case, equations 6.1.7 and 6.1.9 reduce to a cubic equation for ϕ which then forms the basis of a Newton-Raphson iterative method to estimate ϕ .

For given θ and ϕ , the ML and REML estimates of β are the same as the BLUP estimates and are obtained using 6.1.3 for a given set of initial θ , ϕ values. Once \tilde{u} is obtained, equations 6.1.4 and 6.1.5 may be used to estimate θ and equations 6.1.7 and 6.1.9 to estimate ϕ . These estimates of θ and ϕ may then serve as initial values for a further iteration of

the whole process. Asymptotic variances of these estimators are given in the next Section.

6.2 ASYMPTOTIC VARIANCES

The asymptotic variance matrix for $\hat{\beta}$ is V_1 which is part of the inverse of the matrix V as defined in section 6.1, i.e.

$$6.2.1 \quad \text{var } \hat{\beta} = V_1$$

The asymptotic variance matrices for ML and REML estimators of θ, ϕ are analogous to those given in Section 4.5 and 4.6. For ML,

$$\text{var} \begin{pmatrix} \hat{\theta} \\ \hat{\phi} \end{pmatrix} = \begin{pmatrix} (1/2)\text{tr}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta} \Sigma^{-1}\frac{\partial \Sigma}{\partial \theta}) & (1/2)\text{tr}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta} \Sigma^{-1}\frac{\partial \Sigma}{\partial \phi_i}) \\ (1/2)\text{tr}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \phi_i} \Sigma^{-1}\frac{\partial \Sigma}{\partial \theta}) & (1/2)\text{tr}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \phi_i} \Sigma^{-1}\frac{\partial \Sigma}{\partial \phi_j}) \end{pmatrix}^{-1}$$

where $\Sigma = B^{-1} + \theta ZAZ'$.

For REML, letting $Q = \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$,

$$\text{var} \begin{pmatrix} \hat{\theta} \\ \hat{\phi} \end{pmatrix} = \begin{pmatrix} (1/2)\text{tr}(Q \frac{\partial \Sigma}{\partial \theta} Q \frac{\partial \Sigma}{\partial \theta}) & (1/2)\text{tr}(Q \frac{\partial \Sigma}{\partial \theta} Q \frac{\partial \Sigma}{\partial \phi_i}) \\ (1/2)\text{tr}(Q \frac{\partial \Sigma}{\partial \phi_i} Q \frac{\partial \Sigma}{\partial \theta}) & (1/2)\text{tr}(Q \frac{\partial \Sigma}{\partial \phi_i} Q \frac{\partial \Sigma}{\partial \phi_j}) \end{pmatrix}^{-1}$$

These asymptotic variance matrices of θ and ϕ provide the expressions of the estimated standard error of $\hat{\theta}$ and $\hat{\phi}$.

6.3 AR(1) FRAILTY MODEL

In some situations, the frailty of each patient may not be constant over time and must be modelled as a time series with a correlation structure. In this section, the frailties of the repeated failure times in each patient are considered to follow an AR(1) process. We have then

$$A_i(\phi) = \frac{1}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \dots & \phi^{n_i-1} \\ . & 1 & \phi & \dots & \phi^{n_i-2} \\ : & . & 1 & \dots & . \\ . & . & . & \dots & . \\ \phi^{n_i-1} & \phi^{n_i-2} & \phi^{n_i-3} & \dots & 1 \end{pmatrix}$$

The simplification of equation 6.1.4 (θ_{ML}) and 6.1.7 (ϕ_{ML}) in the AR(1) frailty model is given below:

Three matrices I_i , J_i and K_i are defined as the following. Each of them is symmetric with n_i rows and columns.

I_i is the identity matrix

J_i has diagonals of ones above and below the principal diagonal but has all other elements zero

K_i has only two non-zero elements, one at each end of the principal diagonal

Using that notation for the AR(1) frailty model,

$$6.3.1 \quad \mathbf{A}_i^{-1} = (1+\phi^2)\mathbf{I}_i - \phi\mathbf{J}_i - \phi^2\mathbf{K}_i$$

$$6.3.2 \quad \partial \mathbf{A}_i^{-1} / \partial \phi = 2\phi\mathbf{I}_i - \mathbf{J}_i - 2\phi\mathbf{K}_i$$

If \mathbf{T}_i^* is the block diagonal component of \mathbf{T}^* partitioned conformally to the partition of u and

$$\Lambda = \tilde{u} \quad \tilde{u}' = \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & \ddots & \Lambda_M \end{pmatrix}$$

then using

$$(i) \quad \sum_i \text{tr}[\mathbf{I}_i(\mathbf{T}_i^* + \Lambda_i)] = \mathbf{U}_2$$

$$(ii) \quad \sum_i \text{tr}[\mathbf{J}_i(\mathbf{T}_i^* + \Lambda_i)] = 2\mathbf{U}_3$$

$$(iii) \quad \sum_i \text{tr}[\mathbf{K}_i(\mathbf{T}_i^* + \Lambda_i)] = \mathbf{U}_4$$

enables the ML estimating equation for θ to be written as

$$\begin{aligned} \hat{\theta}_{ML} &= N^{-1}(\text{tr } \mathbf{A}^{-1}\mathbf{T}^* + \tilde{u}'\mathbf{A}^{-1}\tilde{u}) \\ &= N^{-1} \sum_{i=1}^M \text{tr}[\mathbf{A}_i^{-1}(\mathbf{T}_i^* + \Lambda_i)] \\ &= N^{-1}[(1+\phi^2)\mathbf{U}_2 - 2\phi\mathbf{U}_3 - \phi^2\mathbf{U}_4] \end{aligned}$$

Hence, we have

$$6.3.3 \quad \hat{\theta}_{ML} = N^{-1}[(1+\phi^2)\mathbf{U}_2 - 2\phi\mathbf{U}_3 - \phi^2\mathbf{U}_4]$$

From equation 6.1.7, the ML estimating equation for ϕ may similarly be written in terms of the submatrices as

$$6.3.4 \quad \sum_{i=1}^M \text{tr} (\partial A_i^{-1}/\partial \phi) A_i = \hat{\theta}_{ML}^{-1} \left[\sum_{i=1}^M \text{tr}[(\partial A_i^{-1}/\partial \phi)(T_i^* + \Lambda_i)] \right]$$

Since $\text{tr} (\partial A_i^{-1}/\partial \phi) A_i = -2\phi/(1-\phi^2)$, equation 6.3.4 becomes

$$6.3.5 \quad -2M\phi/(1-\phi^2) = \hat{\theta}_{ML}^{-1}(2\phi U_2 - 2U_3 - 2\phi U_4)$$

and by substituting the expression for $\hat{\theta}_{ML}$ (6.3.3) in equation 6.3.5, the estimating equation for ϕ_{ML} is obtained as a cubic equation.

$$6.3.6 \quad f(\phi) = C_1\phi^3 + C_2\phi^2 + C_3\phi + C_4 = 0$$

$$\text{where } C_1 = (N - M)(U_2 - U_4), \quad C_2 = (2M - N)U_3$$

$$C_3 = NU_4 - (N + M)U_2, \quad C_4 = NU_3$$

In solving the cubic equation for ϕ_{ML} , we may use Cardan's formula for getting the exact solution. Or, in practice, the Newton-Raphson approximation is sufficiently good, viz.

$$6.3.7 \quad \hat{\phi} = \phi_0 - [f(\phi_0)/f'(\phi_0)]$$

The derivation for REML estimators $\hat{\theta}_{REML}$ and $\hat{\phi}_{REML}$ is along parallel lines and results in T^* being replaced by T and T_i^* by T_i in the above maximum likelihood equations 6.3.3 and 6.3.7. Asymptotic variances for $\hat{\theta}$ and $\hat{\phi}$ are given in section 6.2.

6.4 APPLICATION

The AR(1) frailty model developed in the previous section is applied to analyse the CGD data in Fleming and Harrington (1991). Recall that the CGD study is a placebo controlled randomized trial of Gamma interferon (γ -IFN) in Chronic granulomatous disease and the aim of the trial is to test the efficacy of Gamma interferon (γ -IFN) in preventing the disease.

Similar to model 1 of Chapter 5, a longitudinal parameter γ is introduced into the fixed effect regression parameter estimation in modelling the time dependent structure of the frailty. The longitudinal parameter γ , as before, corresponds to the risk variable τ_{ij} which is defined to be the time from the first failure until the j^{th} failure time for patient i . Moreover, we assume that the frailty in each patient follows an AR(1) process with correlation parameter ϕ . Then, equations 6.1.3, 6.3.3 and 6.3.7 are used iteratively to perform the analysis. We use the whole set of risk variables in our analysis, but only those significant variables are listed. Results for fixed effect regression parameter are similar to those obtained in Chapter 5. Both ML and REML estimates are given in Table 6.4.1.

Table 6.4.1. ML and REML estimates of parameters (with standard errors) for the AR(1) frailty model.

Variables	Estimate (S.E.)	
	ML	REML
γ -IFN	-1.1777 (0.3019)**	-1.2980 (0.3433)**
Inheritance	-0.7057 (0.3143)*	-0.8038 (0.3691)*
Age	-0.0893 (0.0400)*	-0.0986 (0.0474)*
Corticosteroid	2.1210 (0.7201)**	2.4833 (0.9200)**
$\tilde{\gamma}$	1.5082 (0.5024)**	1.7025 (0.5801)**
$\hat{\theta}$	0.2870 (0.2556)	1.1772 (0.4014)
$\hat{\phi}$	-0.0008 (5.0136)	-0.0005 (2.3024)

* significant at 5% level ** significant at 1% level

The treatment γ -IFN has the highest significance. It indicates that the application of γ -IFN significantly reduces the rate of infection in CGD patients. Consistent results appear in ML and REML estimation which shows that X-linked inheritance pattern, age and the presence of corticosteroid in the entry of study are also significant variables. The longitudinal parameter γ is highly significant. It indicates that as the time to failure increases, the infection rate in CGD patients also increases, a result which agrees with the analysis given in Chapter 5 which finds a highly significant time dependent frailty. In this analysis, the frailty correlation parameter ϕ is not significant in both ML and REML estimation. One possible reason for such insignificance of ϕ is discussed in Section 9.2. The estimate of the effect of γ -IFN is in

broad agreement with the results obtained in Fleming and Harrington (1991) who use the Cox model and the Andersen-Gill multiplicative model. A list of the treatment effect (γ -IFN) estimates obtained by these models, the models considered in Chapter 5 and the AR(1) frailty model is given in Table 6.4.2.

Table 6.4.2. Estimates (with standard errors) of the effect of γ -IFN together with an estimate and confidence interval for the equivalent hazard ratio.

	$\hat{\beta}$	S.E.	Z Stat	$\exp \hat{\beta}$	95% C.I. hazard ratio
Cox model	-1.2063	0.4398	-2.7428	0.2993	(0.1264, 0.7088)
AG model	-1.2765	0.3774	-3.3824	0.2790	(0.1332, 0.5846)
TIF model (ML)	-1.1495	0.3228	-3.5424	0.3168	(0.1683, 0.5964)
TIF model (REML)	-1.1435	0.3463	-3.3021	0.3187	(0.1617, 0.6283)
Model 1 (ML)	-1.1228	0.3052	-3.6789	0.3254	(0.1789, 0.5918)
Model 1 (REML)	-1.1218	0.3325	-3.3738	0.3257	(0.1697, 0.6249)
Model 2 (ML)	-1.1232	0.3053	-3.6790	0.3252	(0.1788, 0.5917)
Model 2 (REML)	-1.1226	0.3330	-3.3712	0.3254	(0.1694, 0.6251)
Model 3 (ML)	-1.1879	0.3450	-3.4432	0.3049	(0.1550, 0.5995)
Model 3 (REML)	-1.1722	0.3751	-3.1250	0.3097	(0.1485, 0.6460)
AR(1) model (ML)	-1.1777	0.3019	-3.9010	0.3080	(0.1704, 0.5566)
AR(1) model (REML)	-1.2980	0.3433	-3.7809	0.2730	(0.1393, 0.5352)

CHAPTER SEVEN

RANDOM BLOCK FRAILTY MODEL

7.1 INTRODUCTION

While Chapters 5 and 6 centre on the development of models for Type III data (multiple recurrent events in the same patient), the modelling of Type II data (the occurrence of the same type of failure in the member of a family/cluster) is the focus of this Chapter. Basically, the models developed in this Chapter are based on the results derived from the Generalised Linear Mixed Model (GLMM) as given in Chapter 4. The method is used to analyse data from a litter matched tumorigenesis experiment.

As mentioned in Chapter 2, this set of data was first presented and analysed by Mantel, Bohidar and Ciminera (1977) and by Mantel and Ciminera (1979). In the experiment, there are 50 male litters and 50 female litters. In each litter, one rat was treated with putative carcinogen while the other two rats served as control. The experiment was followed for 104 weeks and the measure of failure time was the time to tumor occurrence or censoring as recorded to the nearest week. One of the main research questions is to investigate whether the exposure to carcinogen significantly increases the failure rate of rats while allowing for possible correlation between the failure time observations within the same litter.

A subset of the data, for the female rats litter, is provided in the second data set of Appendix I. It is arranged in such a format that it fits the APL programs given in Appendix II. The format of this data set given in Appendix I is described as follows:

Column 1: litter number

Column 2: failure/censoring time, in weeks

Column 3: censoring indicator 0=censored 1=failure

Column 4: treatment 0=control 1=carcinogen

This subset of data was analysed by Hougaard (1986b) who used a Cox type model and a parametric model with Weibull margins. The Bayesian approach using Gibb's sampling was given by Clayton (1991). The consideration of Gamma distributed frailty models in the counting processes framework was given by Nielsen, Gill, Andersen and Sørensen (1992).

Section 7.2 considers the baseline frailty model in which each litter has a random litter effect. The justification of the exponential relative risk function is given in Section 7.3. The random block frailty model further extends the baseline frailty model by considering both individual and litter effects to be random in Section 7.4. The comparison of the results from different methods and further discussion of frailty models are given in Section 7.5. Simulation results of the baseline frailty model and the random block frailty model are presented in Section 7.6.

As we will see in Section 7.5, estimation of treatment effect parameter (carcinogen effect) has good agreement with previous analyses obtained in the literature though the dependence structure within a litter is modelled in different ways. The variance component estimation provides the estimated dispersion of the random effects. Moreover, an important feature of the GLMM method is the prediction of random effects. Such prediction is useful, for instance, in identifying high risk families and individuals when we are interested in the risk of occurrence of a family disease in the study population.

7.2 BASELINE FRAILTY MODEL

For the data from the litter matched tumorigenesis experiment described in Section 7.1, one may conceive that the genetic and environmental conditions shared within litters affect the risk of tumor formation, so the most intuitive random effect modelling is to assume that an unobservable random frailty is present in each litter.

Therefore, for a total of M litters, T_{ij} is the observable failure/censoring time for the j^{th} individual in the i^{th} litter. In the proportional hazards model, the hazard function

$$h(t;i,j) = \lambda(t) \exp(\eta_{ij}) , \quad \eta_{ij} = x'_{ij}\beta + U_i , \\ i=1,2,\dots,M , \quad j=1,2,\dots,n$$

where x_{ij} is a vector of risk variables corresponding to the j^{th} individual in the i^{th} litter, β is a vector parameter and U_i is the unobservable random effect of the i^{th} litter taken to be iid $N(0, \theta)$ while $\lambda(t)$ is the usual unspecified baseline hazard function. The total number of observations is $N=nm$. Let $u' = [U_1, U_2, \dots, U_M]$.

The estimation and inference procedure is described in Chapter 4. Essentially the setup of the baseline frailty model is the same as the time independent frailty model given in Section 5.2. The difference is in the type of multivariate failure time data being considered. The recurrent time to infection for the CGD data (Type III data) is analysed by assuming independent random patient effects in the time independent frailty model while Type II data is considered here with independent litter effects in the baseline frailty model.

Estimators of β and u are found by maximizing the Best Linear Unbiased Prediction (BLUP) likelihood in the initial step and then extended to obtain Maximum Likelihood (ML) and Residual Maximum Likelihood (REML) estimators of β , u and θ . Estimated standard errors of β and θ are also given. BLUP estimates, for a given initial value of θ , maximise $l_1 + l_2$, where

l_1 = partial loglikelihood of failure times with u conditionally fixed

$$l_2 = -(1/2)[M \ln 2\pi\theta + (1/\theta) \sum_{i=1}^M U_i^2]$$

Taking β_0 , u_0 as initial values for β , u , the following equations give an iterative solution for the BLUP estimates, in which the result of each iteration is taken as the initial value of the next.

$$7.2.1 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_0 \end{bmatrix} - V^{-1} \begin{bmatrix} 0 \\ \theta^{-1} u_0 \end{bmatrix} + V^{-1} [X \ Z] ' \frac{dl_1}{d\eta_0}$$

$$\text{where } V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X \ Z] + \begin{bmatrix} 0 & 0 \\ 0 & \theta^{-1} I \end{bmatrix}, \quad B = -(d^2 l_1 / d\eta d\eta'), \quad \eta = X\beta + Zu.$$

Here I is the identity matrix and X , Z are the design matrices of β , u respectively after reordering according to the failure/censoring time and η_0 is the value of η corresponding to β_0 , u_0 . The matrix V is partitioned conformally to $\beta|u$ as

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \quad V^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Let $W_{22}^{-1} = V_{22}$; we have

$$7.2.2 \quad \text{var } \tilde{\beta} = A_{11},$$

$$7.2.3 \quad \hat{\theta}_{(ML)} = M^{-1} (\text{tr } W_{22} + \tilde{u}' \tilde{u}),$$

$$7.2.4 \quad \text{var } \hat{\theta}_{(ML)} = 2\theta^2 (M - 2\theta^{-1} \text{tr } W_{22} + \theta^{-2} \text{tr } W_{22}^2)^{-1}$$

For REML estimation, we need only to replace W_{22} by A_{22} to obtain the estimate $\hat{\theta}_{(REML)}$ and its asymptotic variance, viz.

$$7.2.5 \quad \hat{\theta}_{(REML)} = M^{-1} (\text{tr } A_{22} + \tilde{u}' \tilde{u}),$$

$$7.2.6 \quad \text{var } \hat{\theta}_{(\text{REML})} = 2\theta^2(M - 2\theta^{-1}\text{tr } A_{22} + \theta^{-2}\text{tr } A_{22}^2)^{-1}$$

The application of the baseline frailty model to the data from the litter matched tumorigenesis experiment is given as below:

Estimates of parameters (SE)

Parameter	β	θ
ML estimates	0.9169 (0.3229)	0.4253 (0.3361)
REML estimates	0.9170 (0.3229)	0.4299 (0.3378)

The ML and REML estimates agree with each other and show a significant carcinogen effect with an estimated hazard ratio of 2.502.

The number of failures observed in each litter and the corresponding litter effect prediction are shown in Table 7.2.1. The largest 12 predicted litter effects are highlighted with (**). Comparison with Column 2 of Table 7.2.1 shows that those 12 litter effects match with the litters that have 2 or more failures observed. This finding agrees with the intuitive idea that the more *frail* litters have more failure observations observed.

Table 7.2.1 ML and REML frailty prediction for the data from the Litter Matched Tumorigenesis Experiment in the baseline frailty model

Litter	Number of failure	Frailty Prediction		Litter	Number of failure	Frailty Prediction	
		ML	REML			ML	REML
1	1	0.06	0.06	26	0	-0.31	-0.31
2	0	-0.36	-0.36	27	0	-0.26	-0.26
3	0	-0.37	-0.37	28	2	0.32**	0.33**
4	0	-0.17	-0.17	29	2	0.47**	0.47**
5	0	-0.30	-0.30	30	2	0.59**	0.60**
6	2	0.34**	0.34**	31	0	-0.15	-0.15
7	2	0.29**	0.29**	32	3	0.62**	0.62**
8	1	-0.02	-0.02	33	1	0.01	0.01
9	0	-0.23	-0.23	34	1	0.13	0.13
10	1	0.16	0.16	35	1	0.20	0.21
11	1	0.09	0.09	36	1	-0.01	-0.02
12	0	-0.22	-0.22	37	0	-0.38	-0.39
13	2	0.79**	0.80**	38	1	-0.04	-0.04
14	1	0.04	0.04	39	2	0.55**	0.55**
15	0	-0.31	-0.31	40	2	0.52**	0.53**
16	0	-0.23	-0.23	41	1	0.12	0.12
17	0	-0.33	-0.33	42	2	0.49**	0.50**
18	0	-0.34	-0.34	43	2	0.30**	0.30**
19	0	-0.22	-0.22	44	0	-0.32	-0.32
20	2	0.51**	0.52**	45	0	-0.23	-0.23
21	0	-0.38	-0.39	46	0	-0.10	-0.10
22	0	-0.38	-0.39	47	1	-0.00	-0.00
23	1	0.04	0.04	48	0	-0.23	-0.23
24	0	-0.30	-0.30	49	1	-0.05	-0.05
25	0	-0.38	-0.39	50	1	-0.04	-0.04

** The highest 12 litter effect predictions in ML and REML

7.3 JUSTIFICATION OF THE EXPONENTIAL RELATIVE RISK FUNCTION

In our hazard function specification in Section 7.2, the relative risk function is considered to be exponential. In fact, such exponential restriction can be relaxed to consider a class of more general relative risk functions. Hence, the hazard function

$$h(t; i, j) = \lambda(t) g(\eta_{ij}) , \quad \eta_{ij} = x'_{ij}\beta + U_i ,$$

$$\text{where } g(\eta_{ij}) = \exp f(\eta_{ij})$$

The partial likelihood conditional on fixed u is given by

$$l_1 = \sum_{i=1}^N D_i [f(\eta_i) - \ln \sum_{j=i}^N \exp f(\eta_j)]$$

where η_i is the value of η obtained when η_{ij} are arranged in increasing order of the failure/censoring times T_i , $i = 1, 2, \dots, N$. D_i denotes the corresponding censoring indicator.

Let $f(\eta_i) = \ln(1+k\eta_i)^{1/k}$, note that when $k \rightarrow 0$, $f(\eta_i) = \eta_i$ which corresponds to the exponential relative risk function and when $k=1$, $f(\eta_i) = \ln(1+\eta_i)$ giving the linear relative risk function. We then have

$$dl_1/d\eta = F(d-WMAI) , \quad -d^2 l_1/d\eta d\eta' = [kD+(1-k)WB]F^2 - FWMA^2M'WF$$

where

$$w_k = \exp f(\eta_k) = (1+k\eta_k)^{1/k}, \quad a_k = D_k / \sum_{j=k}^N w_j, \quad b_k = \sum_{j=1}^k a_j,$$

$$d = [D_1, D_2, \dots, D_N]', \quad W = \text{Diag}(w_1, w_2, \dots, w_N),$$

$$A = \text{Diag}(a_1, a_2, \dots, a_N), \quad I = \text{vector of ones},$$

M = lower triangular matrix with ones on/below the principal diagonal,

$$B = \text{Diag}(b_1, b_2, \dots, b_N) = \text{Diag}(MAI), \quad D = \text{Diag}(D_1, D_2, \dots, D_N),$$

$$\eta = (\eta_1, \eta_2, \dots, \eta_N)', \quad F^1 = \text{Diag}(1+k\eta_1, 1+k\eta_2, \dots, 1+k\eta_N).$$

Substituting these two expressions in equation 7.2.1, the estimation of β , \tilde{u} can be obtained by providing initial values of β , u , k and θ . The value of k can be estimated by maximizing the BLUP likelihood ($l_1 + l_2$) for given β , u and θ . In fact, as the value of l_2 is not affected by the choice of k , the maximization procedure does not involve l_2 . The derivative expression of l_1 with respect to k is quite complicated; we decide to estimate k by maximizing l_1 through a linear search from 0 to 1. Hence our iterative scheme is

Step 1: Given initial values β_0 , u_0 , k_0 and θ_0 , use 7.2.1 to estimate β , \tilde{u} until convergent.

Step 2: Estimate \tilde{k} by maximizing l_1 through linear search to 2 decimal places.

Step 3: Replace k_0 by \tilde{k} , repeat Step 1 and 2 until $k_0 = \tilde{k}$.

Step 4: Estimate $\hat{\theta}$ by 7.2.3 (for ML) or 7.2.5 (for REML).

Step 5: Replace θ_0 by $\hat{\theta}$, repeat Step 1, 2 and 3 until convergent.

Such a scheme is applied to our data set. In ML, the value of l_1 decreases monotonically from -167.35 to -170.91, when k goes from 0 to 1 in the final step of estimation. Similarly, l_1 decreases from -167.25 to -170.83 for REML. That means the BLUP likelihood is maximized when $k=0$ and confirms that the exponential relative risk function is correctly specified at least within the family of functions we consider. In fact, we have chosen different starting values of k and they all coincide with a final estimate of $\tilde{k}=0$, $\tilde{\beta}(\text{S.E.})=0.917(0.323)$ and $\hat{\theta}(\text{S.E.})=0.43(0.34)$ in both ML and REML.

7.4 RANDOM BLOCK FRAILTY MODEL

One of the problems of the baseline frailty model is that it assumes the hazards of C1 and C2 individual in each litter to be the same. When we examine the original data set in Table 7.4.1, we see that some of the individuals are more *frail* than the others in terms of their short failure times as highlighted in Table 7.4.1 with (*).

Table 7.4.1 Survival time (weeks) of exposed (E) and control rats (C1,C2) in each of 50 litters

Litter	E	C1	C2	Litter	E	C1	C2
1	101.0@	49.0*	104.0@	26	89.0@	104.0@	104.0@
2	104.0@	102.0@	104.0@	27	78.0@	104.0@	104.0@
3	104.0@	104.0@	104.0@	28	104.0@	81.0	64.0*
4	77.0@	97.0@	79.0@	29	86.0	55.0*	94.0@
5	89.0@	104.0@	104.0@	30	34.0*	104.0@	54.0*
6	88.0	96.0	104.0@	31	76.0@	87.0@	74.0@
7	104.0	94.0@	77.0	32	102.8	73.0	83.9
8	95.9	104.0@	104.0@	33	101.9	104.0@	80.0@
9	82.0@	77.0@	104.0@	34	79.9	104.0@	73.0@
10	70.0	104.0@	77.0@	35	45.0*	79.0@	104.0@
11	88.9	91.0@	90.0@	36	94.0	104.0@	104.0@
12	91.0@	70.0@	92.0@	37	104.0@	104.0@	104.0@
13	39.0*	45.0@	50.0*	38	104.0@	101.0	94.0@
14	102.9	69.0@	91.0@	39	76.0@	84.0	78.0
15	93.0@	104.0@	103.0@	40	80.0	80.9	76.0@
16	85.0@	72.0@	104.0@	41	72.0	95.0@	104.0@
17	104.0@	63.0@	104.0@	42	72.9	104.0@	66.0*
18	104.0@	104.0@	74.0@	43	92.0	104.0@	102.0
19	81.0@	104.0@	69.0@	44	104.0@	98.0@	73.0@
20	67.0	104.0@	68.0	45	55.0@	104.0@	104.0@
21	104.0@	104.0@	104.0@	46	49.0@	83.0@	77.0@
22	104.0@	104.0@	104.0@	47	89.0	104.0@	104.0@
23	104.0@	83.0@	40.0*	48	88.0@	79.0@	99.0@
24	87.0@	104.0@	104.0@	49	103.0	91.0@	104.0@
25	104.0@	104.0@	104.0@	50	104.0@	104.0@	79.0

@ Right-censored times.

* Individuals corresponding to the 10 smallest failure times.

From Clayton (1991), with some modifications to the original data as described in his paper.

Although the baseline frailty model can correctly identify the more *frail* litters, it may not be good enough to model the individual variability. We therefore consider the random block frailty model that models the litter variability and the individual variability as well. The hazard function is given by

$$h(t; i, j) = \lambda(t) \exp(\eta_{ij}) \quad \eta_{ij} = x' \beta + E_{ij} + F_i$$

where E_{ij} is the j^{th} individual effect in litter i distributed iid $N(0, \theta_1)$, and F_i is the i^{th} litter effect distributed iid $N(0, \theta_2)$. Let

$$\begin{aligned} e'_i &= [E_{i1}, E_{i2}, \dots, E_{in}] , \quad e' = [e'_1, e'_2, \dots, e'_M] , \\ f' &= [F_1, F_2, \dots, F_M] , \quad u' = [e' \ f'] \end{aligned}$$

We have

$$\text{var } u = A = \begin{bmatrix} \theta_1 I_N & 0 \\ 0 & \theta_2 I_M \end{bmatrix}$$

and β, u can be estimated by

$$7.4.1 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_0 \end{bmatrix} - V^{-1} \begin{bmatrix} 0 \\ A^{-1} u_0 \end{bmatrix} + V^{-1} [X \ Z]' \frac{dl_1}{d\eta_0}$$

$$\text{where } V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X \ Z] + \begin{bmatrix} 0 & 0 \\ 0 & A^{-1} \end{bmatrix}, \quad B = -(d^2 l_1 / d\eta d\eta'), \quad \eta = X\beta + Zu$$

The matrices \mathbf{X} , \mathbf{Z} are the design matrices of β , u respectively after the reordering according to the failure/censoring times. The \mathbf{Z} matrix here is different from that in the baseline frailty model since we are considering a different random effect vector u . The expression of $dl_1/d\eta$ and $-(d^2l_1/d\eta d\eta')$ are the same as in the baseline frailty model. The \mathbf{V} matrix is partitioned conformally to $\beta | u | f$ as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{V}_{13} \\ \mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{V}_{23} \\ \mathbf{V}_{31} & \mathbf{V}_{32} & \mathbf{V}_{33} \end{bmatrix}, \quad \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}$$

and

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{22} & \mathbf{W}_{23} \\ \mathbf{W}_{32} & \mathbf{W}_{33} \end{bmatrix}, \quad \mathbf{W}^{-1} = \begin{bmatrix} \mathbf{W}_{22} & \mathbf{W}_{23} \\ \mathbf{W}_{32} & \mathbf{W}_{33} \end{bmatrix}$$

We then have

$$7.4.2 \quad \text{var } \hat{\beta} = \mathbf{A}_{11}$$

$$7.4.3 \quad \hat{\theta}_{1(\text{ML})} = N^{-1}(\text{tr } \mathbf{W}_{22} + \tilde{e}'\tilde{e}),$$

$$7.4.4 \quad \hat{\theta}_{2(\text{ML})} = M^{-1}(\text{tr } \mathbf{W}_{33} + \tilde{f}'\tilde{f})$$

and the asymptotic variance matrix of these two ML estimators is

$$\left[\begin{array}{cc} (1/2\theta_1^2)(N - 2\theta_1^{-1}\text{tr } \mathbf{W}_{22} + \theta_1^{-2}\text{tr } \mathbf{W}_{22}^2) & (1/2\theta_1^2\theta_2^2)\text{tr } \mathbf{W}_{32}\mathbf{W}_{23} \\ . & (1/2\theta_2^2)(M - 2\theta_2^{-1}\text{tr } \mathbf{W}_{33} + \theta_2^{-2}\text{tr } \mathbf{W}_{33}^2) \end{array} \right]^{-1}$$

Replacing \mathbf{W}_{kk} by \mathbf{A}_{kk} in the above equations correspondingly, gives similar results in REML, viz.

$$7.4.5 \quad \hat{\theta}_{1(\text{REML})} = N^{-1}(\text{tr } \mathbf{A}_{22} + \tilde{\mathbf{e}}'\tilde{\mathbf{e}}),$$

$$7.4.6 \quad \hat{\theta}_{2(\text{REML})} = M^{-1}(\text{tr } \mathbf{A}_{33} + \tilde{\mathbf{f}}'\tilde{\mathbf{f}})$$

and the asymptotic variance matrix of these two REML estimators is

$$\begin{bmatrix} (1/2\theta_1^2)(N - 2\theta_1^{-1}\text{tr } \mathbf{A}_{22} + \theta_1^{-2}\text{tr } \mathbf{A}_{22}^2) & (1/2\theta_1^2\theta_2^2)\text{tr } \mathbf{A}_{32}\mathbf{A}_{23} \\ . & (1/2\theta_2^2)(M - 2\theta_2^{-1}\text{tr } \mathbf{A}_{33} + \theta_2^{-2}\text{tr } \mathbf{A}_{33}^2) \end{bmatrix}^{-1}$$

When the random block frailty model is applied to the data from the litter matched tumorigenesis experiment, we obtain the results given as below:

Estimates of parameters (S.E.)

Parameter	β	θ_1	θ_2
ML estimate	0.9190 (0.3281)	0.0933 (0.5507)	0.4404 (0.4277)
REML estimate	0.9193 (0.3288)	0.1041 (0.5523)	0.4458 (0.4297)

The estimate β agrees with our previous results in the baseline frailty model. The variance component estimate of litter effect, $\hat{\theta}_2$, can be compared with $\hat{\theta}_1$ in the baseline frailty model. The agreement is good. Moreover, the variance estimate in the individual variability is given by $\hat{\theta}_1$ and, while this is not significantly different from zero, it is interesting to consider the estimates of individual effects.

Table 7.4.2 and 7.4.3 present the ML and REML prediction of litter and individual effects in the random block frailty model respectively. Both sets of prediction are similar. The random effect predictions correctly identify the most *frail* litters (***) as well as the most *frail* individuals (*), which are highlighted in Table 7.4.2 and 7.4.3. Those 12 litters with largest litter effect prediction match with those 12 litters given in Table 7.2.1 and in turn with those 12 litters that have 2 or more failures observed. Moreover, on comparing with Table 7.4.1, those 10 individuals with the largest frailty prediction match exactly with those 10 individuals with the smallest failure times.

Table 7.4.2 Prediction of litter and individual effects by the Random Block Frailty Model (ML)

Litter	Litter Effect	E	C1	C2	Litter	Litter Effect	E	C1	C2
1	0.07	-0.05	0.09*	-0.03	26	-0.31	-0.03	-0.02	-0.02
2	-0.36	-0.04	-0.01	-0.02	27	-0.27	-0.01	-0.02	-0.02
3	-0.37	-0.04	-0.02	-0.02	28	0.35**	-0.09	0.08	0.09*
4	-0.18	-0.01	-0.02	-0.01	29	0.47**	0.04	0.09*	-0.03
5	-0.30	-0.03	-0.02	-0.02	30	0.61**	0.09*	-0.05	0.09*
6	0.34**	0.04	0.06	-0.04	31	-0.15	-0.01	-0.01	-0.01
7	0.29**	0.00	-0.02	0.08	32	0.62**	-0.02	0.08	0.07
8	-0.03	0.05	-0.03	-0.03	33	0.00	0.04	-0.03	-0.01
9	-0.23	-0.02	-0.01	-0.02	34	0.13	0.07	-0.03	-0.01
10	0.16	0.08	-0.03	-0.01	35	0.21	0.09*	-0.01	-0.03
11	0.08	0.05	-0.02	-0.02	36	-0.02	0.05	-0.03	-0.03
12	-0.22	-0.03	-0.00	-0.01	37	-0.39	-0.05	-0.02	-0.02
13	0.82**	0.09*	-0.00	0.09*	38	-0.04	-0.06	0.07	-0.02
14	0.03	0.03	-0.01	-0.02	39	0.56**	-0.03	0.07	0.08
15	-0.32	-0.03	-0.02	-0.02	40	0.53**	0.05	0.07	-0.01
16	-0.24	-0.02	0.00	-0.02	41	0.12	0.07	-0.02	-0.03
17	-0.33	-0.05	0.00	-0.02	42	0.50**	0.06	-0.04	0.09*
18	-0.34	-0.05	-0.02	0.00	43	0.29**	0.04	-0.04	0.06
19	-0.22	-0.02	-0.02	0.00	44	-0.32	-0.05	-0.01	-0.01
20	0.52**	0.07	-0.05	0.08	45	-0.23	-0.01	-0.02	-0.02
21	-0.39	-0.05	-0.02	-0.02	46	-0.10	0.00	-0.01	-0.01
22	-0.39	-0.05	-0.02	-0.02	47	-0.01	0.05	-0.03	-0.03
23	0.05	-0.07	-0.01	0.09*	48	-0.23	-0.03	-0.01	-0.02
24	-0.30	-0.02	-0.02	-0.02	49	-0.05	0.03	-0.01	-0.03
25	-0.39	-0.05	-0.02	-0.02	50	-0.03	-0.06	-0.03	0.08

* The largest 10 individual effects (they match with the 10 smallest failure time individuals).

** The largest 12 litter effects (they match with those litters that have 2 or more failures observed).

Table 7.4.3 Prediction of litter and individual effects by the Random Block Frailty Model (REML)

Litter	Litter Effect	E	C1	C2	Litter	Litter Effect	E	C1	C2
1	0.07	-0.06	0.10*	-0.03	26	-0.32	-0.03	-0.02	-0.02
2	-0.37	-0.05	-0.02	-0.02	27	-0.27	-0.02	-0.02	-0.02
3	-0.38	-0.05	-0.02	-0.02	28	0.35**	-0.10	0.09	0.10*
4	-0.18	-0.02	-0.02	-0.01	29	0.48**	0.04	0.10*	-0.03
5	-0.31	-0.03	-0.02	-0.02	30	0.62**	0.10*	-0.06	0.10*
6	0.34**	0.05	0.07	-0.04	31	-0.15	-0.02	-0.01	-0.01
7	0.29**	0.00	-0.02	0.09	32	0.63**	-0.02	0.09	0.08
8	-0.03	0.05	-0.03	-0.03	33	0.00	0.04	-0.03	-0.01
9	-0.24	-0.02	-0.01	-0.02	34	0.13	0.07	-0.03	-0.01
10	0.17	0.08	-0.04	-0.01	35	0.21	0.10*	-0.01	-0.04
11	0.09	0.06	-0.02	-0.02	36	-0.02	0.05	-0.03	-0.03
12	-0.22	-0.03	-0.01	-0.01	37	-0.39	-0.05	-0.02	-0.02
13	0.83**	0.10*	-0.00	0.10*	38	-0.04	-0.07	0.08	-0.02
14	0.03	0.03	-0.01	-0.02	39	0.56**	-0.03	0.08	0.09
15	-0.32	-0.03	-0.02	-0.02	40	0.53**	0.05	0.08	-0.01
16	-0.24	-0.03	-0.01	-0.02	41	0.12	0.08	-0.02	-0.03
17	-0.33	-0.05	-0.00	-0.02	42	0.51**	0.07	-0.05	0.10*
18	-0.34	-0.05	-0.02	-0.01	43	0.29**	0.04	-0.04	0.07
19	-0.22	-0.02	-0.02	-0.01	44	-0.32	-0.05	-0.02	-0.01
20	0.53**	0.08	-0.05	0.09	45	-0.24	-0.01	-0.02	-0.02
21	-0.39	-0.05	-0.02	-0.02	46	-0.10	-0.01	-0.01	-0.01
22	-0.39	-0.05	-0.02	-0.02	47	-0.01	0.06	-0.03	-0.03
23	0.06	-0.08	-0.01	0.10*	48	-0.23	-0.03	-0.01	-0.02
24	-0.31	-0.03	-0.02	-0.02	49	-0.06	0.03	-0.02	-0.03
25	-0.39	-0.05	-0.02	-0.02	50	-0.03	-0.07	-0.03	0.09

* The largest 10 individual effects (they match with the 10 smallest failure time individuals).

** The largest 12 litter effects (they match with those litters that have 2 or more failures observed).

7.5 DISCUSSION

As mentioned in Section 7.1, the tumour data have been analysed in the literature using different methods. A comparison of results is given as below:

	$\hat{\beta}$	S.E.	L.R.	$\hat{\xi}$	L.R.	$\hat{\theta}_1$	S.E.	$\hat{\theta}$ or $\hat{\theta}_2$	S.E.
Hougaard Cox	0.898	0.317	7.87	---	---	---	---	---	---
Weibull	0.944	0.327	8.28	---	---	---	---	---	---
Clayton	0.919	---	---	0.50	---	---	---	---	---
Nielsen <i>et al.</i>	0.904	---	7.16	0.47	1.52	---	---	---	---
Model 1 ML	0.917	0.323	---	---	---	---	0.43	0.34	0.34
REML	0.917	0.323	---	---	---	---	0.43	0.34	0.34
Model 2 ML	0.919	0.328	---	---	---	0.09	0.55	0.44	0.43
REML	0.919	0.329	---	---	---	0.10	0.55	0.45	0.43

The Gamma frailty distribution was considered by Clayton (1991) and by Nielsen, Gill, Andersen and Sørensen (1992). So ξ is the Gamma distribution's parameter which models the intralitter correlation. L.R. is the likelihood ratio test statistic for testing the corresponding null hypothesis. On comparing the results from different methods, they all give consistent estimates of β which shows a significant increase in failure rate for the rats who are exposed to carcinogen.

An important feature of the current approach when compared with other methods is the prediction of random effects. Such prediction will be important for early detection of the high risk subgroups in the study population. For example, looking for the significance of risk variables is as important as identifying the high risk families and individuals for some kinds of family disease.

Generalisations of the method described in this Chapter are possible in different ways. Firstly, the models given in section 7.2 and 7.4 are not restricted to Type II data. With corresponding adjustment in the design matrix of the random component, we can apply the method to analyse Type I data. In fact, Types I and II data are similar when we treat a patient in Type I data as a cluster in Type II data. The basic structure of the modelling of these two types of data are essentially the same. However, before actually implementing the method, one should consider the possible change in the model to fit the particular structure of a data set.

Secondly, the method to justify the exponential relative risk function described in Section 7.3 is not restricted to the baseline frailty model only. Essentially, the technique can be carried to any kind of frailty model without altering the development in Section 7.3. The only modification that needs to be made is the corresponding adjustment in the matrix, $\mathbf{A} = \text{var } u$.

Lastly, Model 2 can be seen as a two-stage random block frailty model. In some practical situations, a k-stage random block frailty model may need to be considered. As we can see in Section 7.4, the expressions for the θ 's have a symmetric structure and by working through the information matrix, we obtain the corresponding variance matrix of the θ estimators. So, the generalisation to a k-stage model follows exactly the same route.

7.6 SIMULATIONS

In normal error models, it has been noticed that ML estimators of the variance components in a regression model are often negatively biased. Such bias becomes more serious as the number of regression variables increases. The REML procedure following Thompson (1980) has been proposed as a method of reducing such biases.

This section provides preliminary exploratory work to compare the performance of ML and REML estimators of variance components in the GLMM when applied to the analysis of multivariate failure time data. In particular, it includes a small simulation comparing ML and REML estimators allowing the variance of random effects to increase and varying the number of regression variables in the baseline frailty model and the random block frailty model.

For the baseline frailty model, the hazard function is given by

$$h(t;i,j) = \lambda(t) \exp(\eta_{ij}) , \quad \eta_{ij} = x'_{ij}\beta + U_i , \\ i=1,2,\dots,30 , j=1,2,3$$

That is, there are 30 litters and each litter has 3 individuals. Let $\lambda(t)=0.1$, component of x_{ij} is random selected as 0 or 1 , the litter effect U_i as independent $N(0,\theta)$. Assuming there is no censored observation, θ is chosen to be 1 or 4 and the dimension of x_{ij} is one or four.

Simulation results of ML and REML estimators in the baseline frailty model are provided in Tables 7.6.1 and 7.6.2 respectively. For both ML and REML, the estimation of regression parameters are rarely significantly biased in all cases. The ML estimator of variance component tends to be negatively biased as the number of regression variables increases from one to four. Such bias becomes more serious with increasing θ . The REML estimator of variance component are relatively stable and tends to be slightly positively biased. Moreover, very good agreement between the standard error of estimates over simulations and the average standard error of estimates is obtained in both ML and REML.

Table 7.6.1 Estimated biases and standard errors for 100 simulations of ML estimation in the baseline frailty model

Simulation 1: $\theta=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.008 (0.037)	0.438	0.371
θ	1	0.029 (0.046)	0.369	0.456
Mean square error in estimation of frailties				0.361

Simulation 2: $\theta=1$, x four component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.034 (0.053)	0.437	0.534
β_2	-0.5	-0.064 (0.047)	0.434	0.472
β_3	0.8	0.063 (0.045)	0.441	0.454
β_4	-0.8	0.019 (0.039)	0.440	0.392
θ	1	-0.144 (0.038)	0.325	0.385
Mean square error in estimation of frailties				0.427

Simulation 3: $\theta=4$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.096 (0.068)	0.765	0.678
θ	4	-0.050 (0.148)	1.145	1.478
Mean square error in estimation of frailties				0.693

Simulation 4: $\theta=4$, x four component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	-0.121 (0.073)	0.739	0.729
β_2	-0.5	0.034 (0.079)	0.733	0.788
β_3	0.8	-0.091 (0.083)	0.730	0.827
β_4	-0.8	0.105 (0.084)	0.740	0.844
θ	4	-0.839 (0.107)	0.937	1.074
Mean square error in estimation of frailties				1.027

* Each simulation gives a S.E. of estimate. Tabular value is the average S.E.

† Tabular value is the S.E. of the 100 simulated estimates.

Table 7.6.2 Estimated biases and standard errors for 100 simulations of REML estimation in the baseline frailty model

Simulation 1: $\theta=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.020 (0.043)	0.450	0.431
θ	1	0.106 (0.043)	0.396	0.432
Mean square error in estimation of frailties				0.369

Simulation 2: $\theta=1$, x four component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	-0.007 (0.053)	0.471	0.532
β_2	-0.5	-0.046 (0.043)	0.483	0.432
β_3	0.8	0.085 (0.059)	0.480	0.592
β_4	-0.8	-0.027 (0.048)	0.481	0.479
θ	1	0.101 (0.048)	0.419	0.478
Mean square error in estimation of frailties				0.458

Simulation 3: $\theta=4$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.106 (0.078)	0.770	0.782
θ	4	-0.071 (0.141)	1.160	1.411
Mean square error in estimation of frailties				0.726

Simulation 4: $\theta=4$, x four component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.073 (0.094)	0.843	0.943
β_2	-0.5	0.103 (0.079)	0.830	0.787
β_3	0.8	0.021 (0.093)	0.840	0.933
β_4	-0.8	0.089 (0.076)	0.844	0.757
θ	4	0.176 (0.160)	1.299	1.601
Mean square error in estimation of frailties				1.046

* Each simulation gives a S.E. of estimate. Tabular value is the average S.E.

† Tabular value is the S.E. of the 100 simulated estimates.

For the random block frailty model, the hazard function is given by

$$h(t;i,j) = \lambda(t) \exp(\eta_{ij}), \quad \eta_{ij} = x'_{ij}\beta + E_{ij} + F_i, \\ i=1,2,\dots,30, \quad j=1,2,3$$

That is, there are 30 litters and each litter has 3 individuals. Let $\lambda(t)=0.1$, component of x_{ij} is random selected as 0 or 1, the litter effect F_i as independent $N(0, \theta_2)$ and the individual effect E_{ij} as independent $N(0, \theta_1)$. Assuming there is no censored observation, θ_1 and θ_2 are chosen to be 1 or 4 and the dimension of x_{ij} is one or four.

Simulation results of ML and REML estimators in the random block frailty model are provided in Tables 7.6.3 and 7.6.4 respectively. As in the baseline frailty model, the estimation of regression parameters are rarely significantly biased in both ML and REML. Good agreement is also obtain between the standard error of estimates over simulations and the average standard error of estimates in both ML and REML.

The ML estimation of variance component (θ_1 and θ_2) tends to be negatively biased. The degree of biasedness seems to be not affected by the number of regression variables and increases as the value of the variance parameter (θ_1 and θ_2) increases. The bias in the REML estimation of variance component also tends to be negative when the variance parameter increases but such bias is reduced when the number of regression variables increases. Moreover, REML estimator of variance component also has a small tendency of positive biasedness when the variance parameter equals to 1.

Table 7.6.3 Estimated biases and standard errors for 100 simulations of ML estimation in the random block frailty model

Simulation 1: $\theta_1=1$, $\theta_2=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.041 (0.053)	0.466	0.527
θ_1	1	-0.062 (0.014)	0.401	0.142
θ_2	1	-0.147 (0.048)	0.458	0.478
Mean square error in estimation of frailties				0.850

Simulation 2: $\theta_1=1$, $\theta_2=1$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.032 (0.050)	0.477	0.496
β_2	-0.5	-0.030 (0.049)	0.475	0.490
β_3	0.8	-0.021 (0.048)	0.476	0.480
β_4	-0.8	-0.067 (0.053)	0.476	0.525
θ_1	1	-0.053 (0.019)	0.404	0.186
θ_2	1	-0.295 (0.038)	0.426	0.378
Mean square error in estimation of frailties				0.919

Simulation 3: $\theta_1=1$, $\theta_2=4$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.192 (0.081)	0.711	0.809
θ_1	1	-0.184 (0.021)	0.378	0.212
θ_2	4	-0.965 (0.128)	0.998	1.276
Mean square error in estimation of frailties				1.273

Simulation 4: $\theta_1=1$, $\theta_2=4$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.010 (0.070)	0.706	0.705
β_2	-0.5	0.069 (0.067)	0.707	0.671
β_3	0.8	-0.072 (0.079)	0.712	0.794
β_4	-0.8	0.226 (0.069)	0.711	0.689
θ_1	1	-0.182 (0.021)	0.378	0.210
θ_2	4	-1.415 (0.107)	0.882	1.074
Mean square error in estimation of frailties				1.525

Simulation 5: $\theta_1=4$, $\theta_2=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.160 (0.046)	0.424	0.460
θ_1	4	-2.751 (0.030)	0.468	0.295
θ_2	1	-0.522 (0.042)	0.410	0.418
Mean square error in estimation of frailties			2.289	

Simulation 6: $\theta_1=4$, $\theta_2=1$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	-0.276 (0.046)	0.437	0.456
β_2	-0.5	0.102 (0.046)	0.438	0.462
β_3	0.8	-0.228 (0.049)	0.443	0.492
β_4	-0.8	0.221 (0.045)	0.441	0.451
θ_1	4	-2.739 (0.037)	0.474	0.372
θ_2	1	-0.604 (0.035)	0.398	0.346
Mean square error in estimation of frailties			2.455	

Simulation 7: $\theta_1=4$, $\theta_2=4$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.339 (0.062)	0.552	0.620
θ_1	4	-2.841 (0.028)	0.440	0.283
θ_2	4	-2.568 (0.071)	0.618	0.710
Mean square error in estimation of frailties			3.112	

Simulation 8: $\theta_1=4$, $\theta_2=4$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	-0.183 (0.061)	0.576	0.618
β_2	-0.5	0.240 (0.059)	0.571	0.594
β_3	0.8	-0.289 (0.062)	0.575	0.619
β_4	-0.8	0.213 (0.057)	0.575	0.565
θ_1	4	-2.852 (0.033)	0.438	0.332
θ_2	4	-2.659 (0.067)	0.595	0.672
Mean square error in estimation of frailties			3.645	

* Each simulation gives a S.E. of estimate. Tabular value is the average S.E.

† Tabular value is the S.E. of the 100 simulated estimates.

Table 7.6.4 Estimated biases and standard errors for 100 simulations of REML estimation in the random block frailty model

Simulation 1: $\theta_1=1$, $\theta_2=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	0.048 (0.052)	0.495	0.519
θ_1	1	-0.047 (0.013)	0.400	0.131
θ_2	1	0.033 (0.045)	0.505	0.452
Mean square error in estimation of frailties				0.854

Simulation 2: $\theta_1=1$, $\theta_2=1$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.196 (0.063)	0.637	0.629
β_2	-0.5	-0.163 (0.056)	0.633	0.556
β_3	0.8	0.178 (0.061)	0.635	0.608
β_4	-0.8	-0.058 (0.059)	0.630	0.589
θ_1	1	1.263 (0.124)	0.642	1.239
θ_2	1	0.472 (0.089)	0.792	0.886
Mean square error in estimation of frailties				1.378

Simulation 3: $\theta_1=1$, $\theta_2=4$, x one component

Parameter value	True β	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
	0.5	-0.047 (0.074)	0.747	0.738
θ_1	1	-0.137 (0.021)	0.386	0.210
θ_2	4	-0.628 (0.113)	1.110	1.132
Mean square error in estimation of frailties				1.180

Simulation 4: $\theta_1=1$, $\theta_2=4$, x four components

Parameter value	True β_1	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
	0.5	-0.037 (0.076)	0.896	0.756
β_2	-0.5	-0.011 (0.085)	0.908	0.853
β_3	0.8	0.008 (0.098)	0.904	0.983
β_4	-0.8	-0.023 (0.088)	0.911	0.879
θ_1	1	0.555 (0.069)	0.515	0.692
θ_2	4	0.497 (0.205)	1.564	2.051
Mean square error in estimation of frailties				1.659

Simulation 5: $\theta_1=4$, $\theta_2=1$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.135 (0.050)	0.467	0.501
θ_1	4	-2.284 (0.041)	0.567	0.407
θ_2	1	-0.395 (0.051)	0.511	0.512
Mean square error in estimation of frailties				1.940

Simulation 6: $\theta_1=4$, $\theta_2=1$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.115 (0.066)	0.670	0.658
β_2	-0.5	-0.008 (0.063)	0.667	0.627
β_3	0.8	0.020 (0.064)	0.667	0.638
β_4	-0.8	0.015 (0.068)	0.667	0.683
θ_1	4	0.320 (0.087)	1.043	0.867
θ_2	1	0.003 (0.063)	0.935	0.634
Mean square error in estimation of frailties				1.581

Simulation 7: $\theta_1=4$, $\theta_2=4$, x one component

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β	0.5	-0.253 (0.069)	0.608	0.688
θ_1	4	-2.529 (0.042)	0.498	0.420
θ_2	4	-2.178 (0.097)	0.763	0.965
Mean square error in estimation of frailties			2.760	

Simulation 8: $\theta_1=4$, $\theta_2=4$, x four components

Parameter	True value	Average bias (S.E. in brackets)	Average of S.E.* of estimates	S.E. of estimates† over simulations
β_1	0.5	0.018 (0.087)	0.871	0.867
β_2	-0.5	0.095 (0.084)	0.871	0.836
β_3	0.8	-0.024 (0.086)	0.869	0.858
β_4	-0.8	0.129 (0.087)	0.871	0.867
θ_1	4	-0.416 (0.083)	0.887	0.827
θ_2	4	-0.720 (0.127)	1.432	1.268
Mean square error in estimation of frailties			2.191	

* Each simulation gives a S.E. of estimate. Tabular value is the average S.E.

† Tabular value is the S.E. of the 100 simulated estimates.

CHAPTER EIGHT

COMPETING RISK FRAILTY MODEL

8.1 INTRODUCTION

This Chapter considers the modelling of Type IV data (the occurrence of different failure events in the same individual) using the Generalised Linear Mixed Model (GLMM). As mentioned in Chapter 1, the modelling of this type of data is the most complicated among the four types of multivariate failure time data. The complexity is due to the fact that more than one type of failure event may occur in each individual.

The data set being considered in this Chapter is the Dubbo study data described in Chapter 2. In the Dubbo study, each individual is followed until an "failure" (hospital admission or death) occurs with the possibility that there may be a sequence of several such failures (hospital admission) for any given individuals. Possible risk variables for each individual are recorded in detail. The two types of failure times are the time to hospitalization and time to death. Hence, additional to the modelling of the dependence structure that comes from the multiple time to hospitalization in each individual, the possible relationship between an individual's frailty in hospitalization and his frailty in death may also need to be considered.

In order to illustrate the competing risk frailty model, only a subset of the Dubbo study data is used. It contains the failure time

observations and the risk variables of male individuals of age 70. Such subset of the data is provided in Appendix I. As before, it is arranged in a format that fits the APL programs given in Appendix II. The Dubbo study data given in Appendix I is described as follows:

Column 1: patient number

Column 2: failure/censoring time (hospitalization or death), in days

Column 3: censoring indicator H 0=censored 1=hospitalized

Column 4: censoring indicator D 0=censored 1=death

Column 5: on BP medication 0=no 1=yes

Column 6: prior CHD 0=no 1=yes

Column 7: cholesterol, in mmol/l

Column 8: prior diabetes 0=no 1=yes

Column 9: one disability 0=no 1=yes

Column 10: more than one disability 0=no 1=yes

Column 11: married 0=no 1=yes

Column 12: former smoker 0=no 1=yes

Column 13: current smoker 0=no 1=yes

The failure time given in column 2 is defined to be the time between subsequent hospitalization/death events. These times to subsequent failure events are used to construct the risk sets in the log-partial likelihood expression. The construction of the risk sets for the two types of failure events (hospitalization and death) is investigated in the next section. Columns 3 and 4 are the censoring indicators for hospitalization and death respectively. Columns 5 to 13 are the risk variables. In fact, the original Dubbo study data records many other

variables such as body mass index, BP reading, CESD depression scale, education level, friends, alone, regular sport and Self Rated Health (SRH) etc. The variables given in Appendix I are those which are possibly important predictors of hospitalization and death based on univariate Cox regression model in survival analysis of the full Dubbo study data.

The focus of this Chapter, however, is neither to provide a complete survival analysis with random effects of the full Dubbo study data nor to develop a general method to handle Type IV data. The attempt here is to establish competing risk frailty models with two types of failure events (hospitalization and death) for a subset of the Dubbo study data (male individuals of age 70) which may then be possible to generalise to develop models with more than two types of failure events.

Two competing risk frailty models are considered in the following section. Estimation procedures of these two models are developed in section 8.3. The application of these two models to the subset of the Dubbo study data given in Appendix I is provided in section 8.4. Some discussions of further research on the Dubbo study data are given in the last section.

8.2 MODELS

In the Dubbo study, individuals born before 1930 living in the New South Wales country town of Dubbo were enrolled and extensive examinations were performed in 1988-89 covering medical, social, physiological and

lifestyle variables. Individuals were followed for a period of 5 years with detailed record of hospitalization and death events. Let

$h_H(t_H; i, j)$ = hazard function for j^{th} hospitalization event of individual i

$h_D(t_D; i, j)$ = hazard function for death event of individual i

where t_H is the time measured from the current failure event (in days) and t_D is the time measured since individual's study entry (in days).

We consider M individuals with n_i failure events for individual i so that $i=1,2,\dots,M$; $j=1,2,\dots,n_i$ and the total number of observations $N = \sum_{i=1}^M n_i$. A proportional hazards model gives

$$h_H(t_H; i, j) = \lambda_H(t_H) g(\eta_{Hij})$$

$$h_D(t_D; i, j) = \lambda_D(t_D) g(\eta_{Dij})$$

where $\lambda_H(t_H)$ and $\lambda_D(t_D)$ are the baseline hazard functions for hospitalization and death respectively and for the Cox hazard function $g(\cdot) = \exp(\cdot)$. The combined risk variables η_{Hij} and η_{Dij} are given by

$$\eta_{Hij} = x'_{ij} \beta_H + U_{Hi}$$

$$\eta_{Dij} = x'_{ij} \beta_D + U_{Di}$$

where x_{ij} is a v dimensional vector, β_H and β_D are the fixed effect regression coefficients for hospitalization and death respectively. U_{Hi}

and U_{Di} correspond to the frailties for hospitalization and death of individual i respectively.

Two models are considered. Model 1 assumes that the frailties for hospitalization and death are equal. Model 2 extends model 1 by relating the frailties for hospitalization and death by a parameter γ . So, by choosing $\gamma=1$, model 1 becomes a special case of model 2. Let

$$\eta^* = (\eta_H^{**} \quad \eta_D^{**})'$$

$$\eta_H^* = (\eta_{H11} \dots \eta_{H1n_1} \dots \eta_{HM1} \dots \eta_{HMn_M})'$$

$$\eta_D^* = (\eta_{D11} \dots \eta_{D1n_1} \dots \eta_{DM1} \dots \eta_{DMn_M})'$$

Model 1: In this model, the frailties for hospitalization and death of individual i are assumed equal. That is, $U_i = U_{Hi} = U_{Di}$ for all i . We have

$$\eta^* = \begin{bmatrix} \eta_H^* \\ \eta_D^* \end{bmatrix} = \begin{bmatrix} X_1^* & \mathbf{0} \\ \mathbf{0} & X_1^* \end{bmatrix} \begin{bmatrix} \beta_H \\ \beta_D \end{bmatrix} + \begin{bmatrix} Z_1^* \\ Z_1^* \end{bmatrix} u_1$$

where $X_1^* = (x'_{ij})_{N \times v}$

$$Z_1^* = \text{diag}[I_1 \dots I_M]$$

I_i is a vector of 1's with dimension n_i

$$u_1 = (U_1 \ U_2 \ \dots \ U_M)' \text{ with } U_i \text{ independent } N(0, \theta)$$

Model 2: In this model, the equality restriction on the frailties for hospitalization and death is relaxed and these frailties are related by a parameter γ . That is, $U_{Hi} = \gamma U_{Di}$. Let $U_{Di} = U_i$ and $U_{Hi} = \gamma U_i$, we have

$$\eta^* = \begin{bmatrix} \eta_H^* \\ \eta_D^* \end{bmatrix} = \begin{bmatrix} X_1^* & 0 \\ 0 & X_1^* \end{bmatrix} \begin{bmatrix} \beta_H \\ \beta_D \end{bmatrix} + \begin{bmatrix} \gamma Z_1^* \\ Z_1^* \end{bmatrix} u_1$$

where X_1^* , Z_1^* and u_1 are defined as in model 1

8.3 ESTIMATION

Let $\eta = (\eta_H' \ \eta_D')'$

$$\eta_H = (\eta_{H1} \ \eta_{H2} \ \dots \ \eta_{HN})'$$

$$\eta_D = (\eta_{D1} \ \eta_{D2} \ \dots \ \eta_{DN})'$$

where η_H is the reordered vector of η_H^* according to the ascending order of T_H and similarly η_D is the reordered vector of η_D^* according to the ascending order of T_D . We have

$$\eta = \begin{bmatrix} \eta_H \\ \eta_D \end{bmatrix} = X \begin{bmatrix} \beta_H \\ \beta_D \end{bmatrix} + Zu_1 = X\beta + Zu_1$$

where $\beta = (\beta_H \ \beta_D)'$ and X , Z are the reorganised design matrices according to the reordered pattern of η_H and η_D . Remember that the only difference between models 1 and 2 before reordering is on the design matrix of u_1 . So, after reordering, the only difference between models 1 and 2 is still on the matrix Z .

The log-partial likelihood on conditionally fixed random components is given by

$$l_1 = \sum_{i=1}^N D_{Hi} [\eta_{Hi} - \ln \sum_{j=i}^N \exp(\eta_{Hj})] + \sum_{i=1}^N D_{Di} [\eta_{Di} - \ln \sum_{j=i}^N \exp(\eta_{Dj})]$$

where D_{Hi} and D_{Di} are respectively the censoring indicators for hospitalization and death of the corresponding i^{th} reordered failure event.

Model 1: The estimation procedure given below can be used in model 1 or when the γ value is known in model 2. For given value of θ , the Newton-Raphson procedure gives the solution of these equations for β , u_1 as the iterative application of

$$8.3.1 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u}_1 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_{10} \end{bmatrix} + V^{-1}[X, Z]'(dl_1/d\eta_0) \cdot V^{-1} \begin{bmatrix} 0 \\ \theta^{-1}u_{10} \end{bmatrix}$$

where β_0 , u_{10} are initial values and $\tilde{\beta}$, \tilde{u}_1 closer approximations to the BLUP estimates. In equation 8.3.1, the vector $\eta_0 = X\beta_0 + Z u_{10}$ and if $B = -d^2 l_1 / d\eta d\eta'$

$$V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X, Z] + \begin{bmatrix} 0 & 0 \\ 0 & \theta^{-1} I \end{bmatrix} \quad \text{evaluated at } \eta = \eta_0.$$

$$dl_1/d\eta = \begin{bmatrix} dl_1/d\eta_H \\ dl_1/d\eta_D \end{bmatrix} \quad \text{and} \quad -d^2 l_1 / d\eta d\eta' = \begin{bmatrix} -d^2 l_1 / d\eta_H d\eta'_H & 0 \\ 0 & -d^2 l_1 / d\eta_D d\eta'_D \end{bmatrix}$$

Letting $\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{V}_1 & \cdot \\ \cdot & \mathbf{T} \end{bmatrix}$ and $[\mathbf{Z}'\mathbf{B}\mathbf{Z} + \theta^{-1}\mathbf{I}]^{-1} = \mathbf{T}^*$

From equations 4.7.2 and 4.7.3, the maximum likelihood $\hat{\theta}_{ML}$ and residual maximum likelihood $\hat{\theta}_{REML}$ estimators of θ are respectively given by

$$8.3.2 \quad \hat{\theta}_{ML} = M^{-1}(\text{tr } \mathbf{T}^* + \tilde{u}'_1 \tilde{u}_1)$$

$$8.3.3 \quad \hat{\theta}_{REML} = M^{-1}(\text{tr } \mathbf{T} + \tilde{u}'_1 \tilde{u}_1)$$

Asymptotic variances of these estimators are derived from the information matrix given in sections 4.5 (for ML) and 4.6 (for REML). The matrix \mathbf{V} is then partitioned conformally to $\beta | u_1$ as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \quad \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

Letting $\mathbf{W}_{22}^{-1} = \mathbf{V}_{22}$, we have

$$8.3.4 \quad \text{var } \beta = \mathbf{A}_{11}$$

$$8.3.5 \quad \text{var } \hat{\theta}_{(ML)} = 2\theta^2[M - 2\theta^{-1}\text{tr } \mathbf{W}_{22} + \theta^{-2}\text{tr}(\mathbf{W}_{22}^2)]^{-1}$$

$$8.3.6 \quad \text{var } \hat{\theta}_{(REML)} = 2\theta^2[M - 2\theta^{-1}\text{tr } \mathbf{A}_{22} + \theta^{-2}\text{tr}(\mathbf{A}_{22}^2)]^{-1}$$

Model 2: The linear predictor before reordering is written as

$$\eta^* = \begin{bmatrix} \eta_H^* \\ \eta_D^* \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1^* & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_1^* \end{bmatrix} \begin{bmatrix} \beta_H \\ \beta_D \end{bmatrix} + \begin{bmatrix} \gamma \mathbf{Z}_1^* \\ \mathbf{Z}_1^* \end{bmatrix} u_1$$

or equivalently,

$$\eta^* = \begin{bmatrix} \eta_H^* \\ \eta_D^* \end{bmatrix} = \begin{bmatrix} X_1^* & 0 \\ 0 & X_1^* \end{bmatrix} \begin{bmatrix} \beta_H \\ \beta_D \end{bmatrix} + Z_1^* u$$

where $u' = (g'_1 \ g'_2) = (\gamma u'_1 \ u'_1)$ and the vector u satisfies the following restrictions

$$G'u = G'_1 g_1 + G'_2 g_2 = 0$$

where $G' = (G'_1 \ G'_2)$, $G'_1 = \gamma^{-1} I_1$, $G'_2 = -I_1$ and I_1 is an $M \times M$ identity matrix.

If u is taken to have a normal distribution with zero mean and variance matrix $\theta_1[I_2 - G(G'G)^{-1}G']$ where I_2 is a $2M \times 2M$ identity matrix, then $G'u$ has zero variance consistent with the restrictions. In that case,

$$\begin{aligned} \text{var } g_2 &= \theta_1[I_1 - G_2(G'G)^{-1}G'_2] \\ &= \theta_1[I_1 - G(I_1 + G'G)^{-1}G'] \\ &= \theta_1[I_1 + GG']^{-1} \end{aligned}$$

where $G = -(G_2G_1^{-1}) = \gamma I_1$. Hence, $\text{var } g_2 = \text{var } u_1 = \theta_1(1+\gamma^2)^{-1}I_1$. Note that $\text{var } u_1 = \theta_1 I_1$ in model 1 while here $\text{var } u_1 = \theta_1(1+\gamma^2)^{-1}I_1$ in model 2. The BLUP likelihood is then constructed as

$$l = l_1 + l_2 = (l_{1H} + l_{1D}) + l_2$$

$$\text{where } l_{1H} = \sum_{i=1}^N D_{Hi} [\eta_{Hi} - \ln \sum_{j=i}^N \exp(\eta_{Hj})]$$

$$l_{1D} = \sum_{i=1}^N D_{Di} [\eta_{Di} - \ln \sum_{j=i}^N \exp(\eta_{Dj})]$$

$$l_2 = \text{constant} - (1/2)[M \ln 2\pi \theta_1 (1+\gamma^2)^{-1} + \ln |I_1| + \theta_1^{-1} (1+\gamma^2) u'_1 u_1]$$

Consequent derivatives are

$$\partial l / \partial \beta = \partial l_1 / \partial \beta$$

$$\partial l / \partial u_1 = \partial l_1 / \partial u_1 - \theta_1^{-1} (1+\gamma^2) u_1$$

$$\partial l / \partial \theta_1 = \partial l_2 / \partial \theta_1 = -(1/2)[M \theta_1^{-1} - \theta_1^{-2} u'_1 u_1]$$

$$\partial l / \partial \gamma = \partial l_{1H} / \partial \gamma - (1/2)[2\gamma \theta_1^{-1} u'_1 u_1 - 2M\gamma (1+\gamma^2)^{-1}]$$

So, the estimation structure is essentially the same as in model 1 with corresponding adjustment in var u_1 . Therefore, for given θ_1 and γ , β and u_1 are estimated by maximizing the BLUP likelihood,

$$8.3.7 \quad \begin{bmatrix} \tilde{\beta} \\ \tilde{u}_1 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ u_{10} \end{bmatrix} + V^{-1} [X, Z]' (dl_1 / d\eta_0) - V^{-1} \begin{bmatrix} 0 \\ \theta_1^{-1} (1+\gamma^2) u_{10} \end{bmatrix}$$

In equation 8.3.7, the vector $\eta_0 = X\beta_0 + Z u_{10}$ and if $B = -d^2 l_1 / d\eta d\eta'$

$$V = \begin{bmatrix} X' \\ Z' \end{bmatrix} B [X, Z] + \begin{bmatrix} 0 & 0 \\ 0 & \theta_1^{-1} (1+\gamma^2) I_1 \end{bmatrix} \quad \text{evaluated at } \eta = \eta_0.$$

As before,

$$dl_1/d\eta = \begin{bmatrix} dl_1/d\eta_H \\ dl_1/d\eta_D \end{bmatrix} \text{ and } -d^2l_1/d\eta d\eta' = \begin{bmatrix} -d^2l_1/d\eta_H d\eta'_H & 0 \\ 0 & -d^2l_1/d\eta_D d\eta'_D \end{bmatrix}$$

Letting $\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{V}_1 & \cdot \\ \cdot & \mathbf{T} \end{bmatrix}$ and $[\mathbf{Z}'\mathbf{B}\mathbf{Z} + \theta_1^{-1}\mathbf{I}_1]^{-1} = \mathbf{T}^*$

the maximum likelihood $\hat{\theta}_{1(ML)}$ and residual maximum likelihood $\hat{\theta}_{1(REML)}$ estimators of θ_1 are respectively given by

$$8.3.8 \quad \hat{\theta}_{1(ML)} = \mathbf{M}^{-1}(\text{tr } \mathbf{T}^* + \tilde{u}'_1 \tilde{u}_1)$$

$$8.3.9 \quad \hat{\theta}_{1(REML)} = \mathbf{M}^{-1}(\text{tr } \mathbf{T} + \tilde{u}'_1 \tilde{u}_1)$$

Note that the definition of \mathbf{V} here is different from that given in model 1, so the matrices \mathbf{T}^*, \mathbf{T} defined here and those matrices defined in the subsequent paragraph in deriving the asymptotic variances are also different from that given in model 1.

Asymptotic variances of these estimators are derived from the information matrix given in sections 4.5 (for ML) and 4.6 (for REML). As before, the matrix \mathbf{V} is then partitioned conformally to $\beta | u_1$ as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \quad \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

Letting $\mathbf{W}_{22}^{-1} = \mathbf{V}_{22}$, we have

$$8.3.10 \quad \text{var } \beta = \mathbf{A}_{11}$$

$$8.3.11 \quad \text{var } \hat{\theta}_{1(\text{ML})} = 2\theta_1^2[M - 2\theta_1^{-1}\text{tr } W_{22} + \theta_1^{-2}\text{tr}(W_{22}^2)]^{-1}$$

$$8.3.12 \quad \text{var } \hat{\theta}_{1(\text{REML})} = 2\theta_1^2[M - 2\theta_1^{-1}\text{tr } A_{22} + \theta_1^{-2}\text{tr}(A_{22}^2)]^{-1}$$

Estimation of γ is by choosing a value $\tilde{\gamma}$ which maximizes the BLUP likelihood l . As we can see from the expression of $\partial l / \partial \gamma$, the partial derivative $\partial l_{\text{IH}} / \partial \gamma$ is complicated and a closed form expression of γ estimate is not obtainable, so $\tilde{\gamma}$ is chosen to maximize l through a linear search to 2 decimal places. Moreover, since l_{ID} does not involve γ , to maximize l is equivalent to maximize

$$8.3.13 \quad l_{\text{IH}} + (1/2)[M \ln(1+\gamma^2) - \hat{\theta}_1^{-1}(1+\gamma^2)\tilde{u}'_1\tilde{u}_1] .$$

As a result, our iterative scheme is

Step 1: Given initial values β_0 , u_{10} , θ_{10} and γ_0 , use equation 8.3.7 to estimate $\tilde{\beta}$ and \tilde{u}_1 until convergent.

Step 2: Estimate $\hat{\theta}_1$ by equations 8.3.8 (for ML) or 8.3.9 (for REML).

Step 3: Replace θ_{10} by $\hat{\theta}_1$, repeat Steps 1 and 2 until convergent.

Step 4: Estimate $\tilde{\gamma}$ by maximizing the expression 8.3.13 through linear search to 2 decimal places.

Step 5: Replace γ_0 by $\tilde{\gamma}$, repeat Step 1,2 and 3 until $\gamma_0 = \tilde{\gamma}$.

8.4 APPLICATION

The two models described in section 8.3 are applied to analyse a subset of the Dubbo study data given in Appendix I. In this subset of

data, $M=62$ (number of individuals) and $N=239$ (total number of failure/censoring events). The ML and REML estimation for model 1 are given in Table 8.4.1.

Table 8.4.1. ML and REML estimates of parameters (with standard errors) for model 1.

ML estimation in Model 1

Variable	β_H	β_D
On BP medication	0.5941(0.2819)*	0.0738(0.9763)
Prior CHD	0.6358(0.2876)*	0.0286(0.9796)
Cholesterol	0.0007(0.0067)	-0.0095(0.0154)
Prior diabetes	-0.5408(0.4344)	-0.4155(1.2537)
One disability	0.6275(0.3396)@	2.7014(1.1285)*
More than one disability	0.5348(0.3608)	0.5175(1.1251)
Married	-0.2008(0.3505)	-3.0728(1.0431)**
Former smoker	0.8010(0.3019)**	2.4645(1.3319)@
Current smoker	0.5946(0.3913)	1.3639(1.5050)
θ	0.3965(0.1537)	

$$\text{var } u_1 = \theta = 0.3965$$

REML estimation in Model 1

Variable	β_H	β_D
On BP medication	0.5940(0.2969)*	0.0702(0.9854)
Prior CHD	0.6379(0.3057)*	0.0325(0.9884)
Cholesterol	0.0008(0.0071)	-0.0098(0.0157)
Prior diabetes	-0.5334(0.4577)	-0.4382(1.2702)
One disability	0.6285(0.3575)@	2.7316(1.1414)*
More than one disability	0.5317(0.3832)	0.5253(1.1345)
Married	-0.2151(0.3706)	-3.1089(1.0558)**
Former smoker	0.8009(0.3171)*	2.4860(1.3455)@
Current smoker	0.5926(0.4112)	1.3602(1.5195)
θ	0.5093(0.1945)	

$$\text{var } u_1 = \theta = 0.5093$$

@ significant at 10% * significant at 5% ** significant at 1%

Significant variables are consistent in both ML and REML estimations. Results show that at 5% significance level, individuals who are on BP medication, former smokers or with prior CHD have a higher risk in hospitalization while those individuals who are unmarried or with one disability have a higher risk in death. It should be noted that these results can only apply to our subset of data but should not be generalised to the full Dubbo study data. The estimated variance of individual's frailty is 0.3965 in ML and 0.5093 in REML. Table 8.4.2 below shows the results of the ML and REML estimation for model 2.

Table 8.4.2. ML and REML estimates of parameters (with standard errors) for model 2.

ML estimation in Model 2

Variable	β_H	β_D
On BP medication	0.5948(0.2687)*	0.0838(0.9601)
Prior CHD	0.6350(0.2713)*	0.0392(0.9649)
Cholesterol	0.0005(0.0063)	-0.0093(0.0151)
Prior diabetes	-0.5497(0.4137)	-0.3769(1.2282)
One disability	0.6273(0.3240)@	2.6385(1.1069)*
More than one disability	0.5396(0.3406)	0.4775(1.1112)
Married	-0.1861(0.3325)	-2.9948(1.0219)**
Former smoker	0.7986(0.2882)**	2.4237(1.3089)@
Current smoker	0.5966(0.3736)	1.3594(1.4795)
θ_1	0.6199(0.1189)	

$$\gamma = 1.07$$

$$\text{var } u_1 = \theta_1(1+\gamma^2)^{-1} = 0.2890$$

REML estimation in Model 2

Variable	β_H	β_D
On BP medication	0.5898(0.3093)@	0.0907(0.9538)
Prior CHD	0.6363(0.3203)*	0.0836(0.9618)
Cholesterol	0.0005(0.0074)	-0.0100(0.0150)
Prior diabetes	-0.5260(0.4766)	-0.3797(1.2307)
One disability	0.6276(0.3726)@	2.6076(1.1051)*
More than one disability	0.5356(0.4010)	0.4067(1.1110)
Married	-0.2274(0.3865)	-2.9549(1.0187)**
Former smoker	0.7863(0.3292)*	2.4186(1.3071)@
Current smoker	0.5916(0.4268)	1.3322(1.4745)
θ_1	0.8859(0.1170)	

$$\gamma = 1.35$$

$$\text{var } u_1 = \theta_1(1+\gamma^2)^{-1} = 0.3139$$

@ significant at 10% * significant at 5% ** significant at 1%

ML and REML estimated regression coefficients are similar to those obtained in model 1. Estimate of γ is 1.07 in ML and 1.35 in REML indicate that the frailties for hospitalization and death are of comparable magnitude in this group of individuals. Estimated variance of individual's frailty for death is 0.2890 in ML and 0.3139 in REML which are less than those obtained in model 1. Such decrease in $\text{var } u_1$ from model 1 to model 2 is expected since the γ estimates in both ML and REML are greater than one.

8.5 DISCUSSION

The derivation of competing risk frailty models are given in the previous sections. Such models are applied to analyse a subset of the Dubbo study data. There are some points worthy of notice. Firstly, the models considered in section 8.2 are for those data which involve two types of failure events only, generalisation to problems with more than two types of failure events is possible following the development in section 8.3.

Secondly, the estimation of γ in section 8.3 is by maximizing the BLUP likelihood through linear search to 2 decimal places. Such estimation method is primitive and requires much computer workspace especially when the data set becomes large. Approximate ML and REML estimation methods to estimate γ are still being investigated.

Thirdly, since the parameter γ represents the relationship between individual's frailty for hospitalization and death, it is of interest to find out the change of γ across sex and age groups. Intuitively, we may expect that γ will decrease from younger age groups to older age groups and will be close to one in older age groups. Moreover, we may also expect that for the same age group, the parameter γ will be higher in female than in male since the life expectancy of females is longer than male.

Lastly, the problem which initiates the analysis of the Dubbo study data using a mixed model is on the investigation of the relationship between individual's frailty prediction and a variable called Self-Rated Health (SRH). Such variable has been hypothesized to be an important indicator of survival in elderly people and predict survival independently of other indicators. Here, we aim to look for the relationship between the frailty prediction for hospitalization and SRH. We expect that individual's frailty prediction will be high for those individuals who anticipate a "poor" SRH at their study entry.

CHAPTER NINE

FURTHER DISCUSSION

9.1 PRELIMINARIES

In the previous chapters, the application of the GLMM is found to be useful to analyse multivariate failure time data. On one hand, the method is a generalisation of Cox regression model by including random components in the linear predictor. On the other hand, it makes use of the quadratic approximation in the likelihood in the region of the maximum.

Although the method is based on an approximation in the region of the maximum, it works quite well in analysing multivariate failure time data. As we can see in previous chapters, consistent results in the estimation of fixed effect regression coefficients are obtained. When compared to other approaches, the current method assumes the random individual effect to be independent identically normally distributed with variance θ . In most cases, such variance of random effects is significantly different from zero, so it is inappropriate to reduce the mixed model to a fixed effect model by ignoring random components.

Random effect prediction is another important feature of the current method. As indicated in chapters 1 and 7, these predictions are useful, for instance, in identifying a high risk family or high risk individual when considering a genetic disease.

9.2 SOME REMARKS

In developing the three time dependent frailty models in chapter 5, the parameter representing change over time is modelled into a fixed effect, a normally distributed random effect and a longitudinal effect in which the random component relates to the patient characteristics. Although models 2 and 3 considered in this chapter do not show significant advantage over model 1 in analysing the CGD data, these models are useful in indicating how such models may be fitted and the method is then available to encourage its consideration in future data collection.

In chapter 6, an AR(1) frailty model is fitted to analyse the CGD data. We see that the estimate of correlation parameter ϕ in the AR(1) process is not significant and is very close to zero. Preliminary simulations show that the method used in this chapter in the estimation of ϕ is biased and tends to give an estimate of ϕ close to zero. Extensive simulations are required to validate this effect. On the other hand, methods to reduce the bias in the estimation of ϕ are now being investigated.

9.3 MAIN DIFFICULTY

In the random effect regression model, the dimension of the matrix involved in each iteration is greatly inflated due to the fact that random components are considered to be conditionally fixed in the BLUP procedure.

Moreover, the reordering process in the Cox model further complicates the problem. In most cases, the design matrix before reordering has certain pattern and allows analytic simplification of the matrix involved. Unfortunately, the reordering mechanism destroys the original pattern of the design matrix and gives a completely structureless design matrix which prohibits further simplification. Such effect results in the impossibility of handling large data sets. In fact, for the analysis of moderate size data (~ 500 individuals), the capability limit of computer is usually exceeded.

Another problem is about the iteration speed. The programming language being used in this research is APL. With a mathematics co-processor, APL has a relatively high speed. But due to the multiplication and inversion of large dimensional matrices involved in each iteration, it takes quite a long time for the whole estimation procedure to be completed. For example, in the analysis of the CGD data using the AR(1) frailty model considered in chapter 6, it takes a few days for the estimation procedure to be completed. In the simulation study of the random block frailty model considered in chapter 7, the number of random effects is 120 in each simulation and it takes about three months to obtain those results. The computation is carried out on a 486 microcomputer.

While we are expecting the improvement in computer technology may help us to solve this problem, efforts are now being made to develop a procedure which preserves the original pattern of the design matrix so that it allows analytic simplification. We hope that such development may

help to save computer time and workspace, and hence make it possible to analyse large data sets.

As a whole, the main difficulty in this research is in the handling of large data sets, due to the limited workspace and the slow iteration speed which are available in our computers.

9.4 UNSOLVED PROBLEMS

During the research process, there are two important problems coming up which remain unsolved. The first one is in the development of unbiased estimating equations in the variance components estimation. In the linear model,

$$y = X\beta + e \quad e \sim N(0, \sigma^2 \Sigma(\phi))$$

where y is an observation vector with N components

X is an $N \times p$ matrix of regression variables having rank v

β is a p -component vector of regression coefficients

ϕ is a vector of correlation parameters.

For given ϕ , ML and REML estimates of β is given by

$$\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} y$$

the estimation of variance parameter σ^2 is given by

$$\hat{\sigma}_{(ML)}^2 = N^{-1} S^2$$

$$\hat{\sigma}_{(REML)}^2 = (N-v)^{-1} S^2$$

$$\text{where } S^2 = (y - \hat{X}\hat{\beta})' \Sigma^{-1} (y - \hat{X}\hat{\beta})$$

It is well-known that the estimating equation of σ^2 is asymptotically negatively biased in ML and asymptotically unbiased in REML. Note that in the above development, ϕ is assumed to be known and given. In practice, ϕ is unknown and needs to be estimated. Therefore, the iteration procedure becomes (i) β and σ^2 is estimated based on a starting value of ϕ , (ii) ϕ is estimated by maximizing the likelihood based on the estimated value of β and σ^2 obtained previously, (iii) such estimated value of ϕ is substituted back to the estimating equations of β and σ^2 and (iv) these iterations continue until the stopping criteria is met.

Although the REML technique removes the unbiasedness in the estimation of σ^2 for given values of ϕ , there is no guarantee of such unbiasedness when in fact ϕ is being estimated by maximizing the likelihood and such values are ultimately used to obtain estimates of β and σ^2 . Since the development of the GLMM relies on the link between BLUP, ML and REML estimation, the biasedness problem in the linear model will transfer to the GLMM. So, it is important to develop a set of simultaneous unbiased estimating equations of β , σ^2 and ϕ in linear model and carry those results to the development of the GLMM.

The second important problem is in the development of Goodness of fit tests and model diagnostic tools in the GLMM. In chapter 4, we see that estimates of parameters are obtained by repeated iterations of equations 4.7.1 to 4.7.5. Their corresponding estimated standard errors are provided by the estimated information matrices given in sections 4.5 and 4.6. Inference procedures are then performed by comparing the parameter estimates with their corresponding estimated standard errors. The development of the GLMM will be more complete if we can have a standard procedure such as the likelihood ratio test for choosing a parsimonious model and model diagnostic tools such as the residual analysis.

9.5 RESEARCH SUGGESTIONS

Other than the two unsolved problems mentioned above, further research is suggested in the following. Firstly, the random components considered in the AR(1) frailty model of chapter 6 are assumed to follow an AR(1) process. It is then a natural extension to consider the random components to follow in general an ARIMA(p,d,q) process. The ML and REML estimating equations of the parameters in the variance matrix are provided by equations 4.7.4 and 4.7.5 respectively.

Secondly, extensive simulations are required to justify the performance of the estimation procedure in the baseline frailty model and the random block frailty model described in chapter 7. Moreover, simulations are called for to investigate the biasedness in the estimation

of the correlation parameter ϕ of the AR(1) frailty model described in section 6.3.

Thirdly, the technique of the justification of the exponential relative risk function mentioned in section 7.3 can be extended to apply to other frailty models without altering the development of that section. The only adjustment to be made is in the variance matrix of the random components.

Lastly, in the competing risk frailty models developed in chapter 8, the estimation procedure of the parameter γ that relates the two types of failure makes use of the maximization of the BLUP likelihood by linear search to 2 decimal places. We see in section 8.3 that the estimation of γ is established by extending the random effect vector. Such extension of the random effect vector results in restrictions imposed on this extended vector. This problem exposes us to a broader class of models called "the GLMM with restricted random components". Efforts are now being made to develop the ML and REML estimating equations of γ in the GLMM with restricted random components.

BIBLIOGRAPHY

- Aalen, O.O. (1976). Statistical theory for a family of counting processes. Institute of Mathematical Statistics. University of Copenhagen.
- Aalen, O.O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics* 6, 701-726.
- Aalen, O.O. (1988). Heterogeneity in survival analysis. *Statistics in Medicine* 7, 1121-1137.
- Aalen, O.O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Ann. Appl. Prob.* 2, 951-972.
- Andersen, P.K. (1988). Multistate models in survival analysis: A study of nephropathy and mortality in diabetes. *Statistics in Medicine* 7, 661-670.
- Andersen, P.K. and Borgan, O. (1985). Counting process models for life history data : A review (with discussion). *Scand. J. of Statistics* 7, 161-171.
- Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993). Statistical models based on counting processes. Springer-Verlag, New York.

Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes : A large sample study. *Annals of Statistics* 10, 1100-1120.

Anderson, D.A. and Aitkin, M. (1985). Variance component models with binary response: Interviewer variability. *J. R. Statist. Soc. B* 47, 203-210.

Anderson, K.W., Castelli, W.P. and Levy, D. (1987). Cholesterol and mortality, 30 years of follow-up from the Framingham Study. *J. Amer. Medical Assn.* 257, 2176-2180.

Atkinson, K., Storb, R., Prentice, R.L., Weiden, P.L., Witherspoon, R.P., Sullivan, K., Noel, D. and Thomas, E.D. (1979). Analysis of late infections in 89 long-term survivors of bone marrow transplantation. *Blood* 53, 720-731.

Batchelor, J.R. and Hackett, M. (1970). HL-A matching in treatment of burned patients with skin allografts. *Lancet* 19, 581-583.

Breslow, N.E. (1984). Extra-Poisson variation in loglinear models. *Appl. Statist.* 33, 38-44.

Breslow, N.E. and Clayton, D.G. (1993). Approximate inference in generalised linear mixed models. *J. Amer. Statist. Assn.* 88, 9-25.

Byar, D.P. (1980). The Veterans Administration study of chemoprophylaxis for recurrent stage I bladder tumors: comparisons of placebo, pyridoxine

and topical thiotepea. In Bladder Tumors and Other Topics in Urological Oncology, Ed. M. Pavone-Macaluso, P.H. Smith and F. Edsmyr, 363-370. Plenum, New York.

Byar, D.P., Blackard, C. and the VACURG (1977). Comparisons of placebo, pyridoxine, and topical thiotepea in preventing recurrence of stage I bladder cancer. *Urology* 10, 556-561.

Cai, J. and Prentice, R.L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika* 82, 151-164.

Chi, E.M. and Reinsel, G.C. (1989). Models for longitudinal data with random effects and AR(1) errors. *J. Amer. Statist. Assn.* 84, 452-459.

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, 141-151.

Clayton, D.G. (1988). The analysis of event history data: A review of progress and outstanding problems. *Statistics in Medicine* 7, 819-841.

Clayton, D.G. (1991). A Monte Carlo method for Bayesian inference in frailty models. *Biometrics* 47, 467-485.

Clayton, D.G. and Cuzick, J. (1985a). Multivariate generalisations of the proportional hazards model. *J. R. Statist. Soc. A* 148, 82-117.

Clayton, D.G. and Cuzick, J. (1985b). The EM algorithm for Cox's regression model using GLIM. *Appl. Statist.* 34, 148-156.

Cox, D.R. (1972). Regression models and life tables (with discussion). *J. R. Statist. Soc. B* 34, 187-220.

Cox, D.R. (1975). Partial likelihood. *Biometrika* 62, 269-276.

Cox, D.R. and Oakes, D. (1984). Analysis of survival data. Chapman and Hall, London.

Dawber, T.R. (1980). The Framingham Study: The epidemiology of Atherosclerotic disease. Harvard University Press, Cambridge, Massachusetts.

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, 1-22.

Fellner, W.H. (1986). Robust estimation of variance components. *Technometrics* 28, 51-60.

Fellner, W.H. (1987). Sparse matrices, and the estimation of variance components by likelihood methods. *Commun. Statist. -Simula. B* 16, 439-463.

Fleming, T.R. (1992). Evaluating therapeutic interventions: some issues and experiences (with discussion). *Statistical Science* 7, 428-456.

Fleming, T.R. and Harrington, D.P. (1991). Counting processes and survival analysis. John Wiley & Sons, New York.

Gail, M.H., Santner, T.J., and Brown, C.C. (1981). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics* 36, 255-266.

Gill, R.D. (1984). Understanding Cox's regression model: A martingale approach. *J. Amer. Statist. Assn.* 79, 441-447.

Gilmour, A.R., Anderson, R.D. and Rae, A.L. (1985). The analysis of binomial data by a generalised linear mixed model. *Biometrika* 72, 593-599.

Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. *Biometrika* 78, 45-51.

Green, P.J. (1987). Penalized likelihood for general semi-parametric regression models. *Int. Statist. Rev.* 55, 245-259.

Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *J. Amer. Statist. Assn.* 72, 320-340.

Harville, D.A. and Mee, R.W. (1984). A mixed model procedure for analysing ordered categorical data. *Biometrics* 40, 393-408.

Henderson, C.R. (1963). Selection index and expected genetic advance. In Statistical Genetics and Plant Breeding. W.D. Hanson and H.F. Robinson eds. Publication 982, 141-163. Washington: Nat. Acad. Sc., Nat. Res. Council.

Henderson, C.R. (1973). Sire evaluation and genetic trends. In Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr Jay Lush, 10-41. champaign, Illinois: Amer. Soc. Animal Science.

Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423-447.

Holt, J.D. and Prentice, R.L. (1974). Survival analyses in twin studies and matched pair experiments. *Biometrika* 61, 17-30.

Hougaard, P. (1984). Life table methods for heterogeneous populations : Distributions describing the heterogeneity. *Biometrika* 71, 75-83.

Hougaard, P. (1986a). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387-396.

Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika* 73, 671-678.

Hougaard, P. (1987). Modelling multivariate survival. *Scand. J. of Statistics* 14, 291-304.

Hougaard, P. (1991). Modelling heterogeneity in survival data. *J. Appl. Prob.* 28, 695-701.

Huster, W.J., Brookmeyer, R. and Self, S.G. (1989). Modelling paired survival data with covariates. *Biometrics* 45, 145-156.

Jennrich, R.I. and Schluchter, M.D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 42, 805-820.

Johansen, S. (1983). An extension of Cox's regression model. *Int. Statist. Rev.* 51, 258-262.

Kalbfleisch, J.D. and Prentice, R.L. (1980). *The statistical analysis of failure time data*. John Wiley & Sons, New York.

Klein, J.P. (1992). Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* 48, 795-806.

Koch, G.C., Landis, J.R., Freeman, J.L., Freeman, D.H. and Lehman, R.G. (1977). A general methodology for the analysis of repeated measurements of categorical data. *Biometrics* 33, 133-158.

Laird, N.M. (1978). Empirical Bayes methods for two-way contingency tables. *Biometrika* 65, 581-590.

- Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics* 38, 963-974.
- Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica* 47, 939-956.
- Lancaster, A. and Nickell, S.J. (1980). The analysis of re-employment probabilities for the unemployed. *J. R. Statist. Soc. A* 143, 141-165.
- Lawless, J.F. (1982). Statistical models and methods for lifetime data. John Wiley & Sons, New York.
- Lee, E.W., Wei, L.J. and Amato, D.A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival analysis: State of the Art*, Ed. J.P. Klein and P.K. Goel, 237-247. Kluwer Academic Publishers, Dordrecht.
- Lee, E.W., Wei, L.J. and Ying, Z. (1993). Linear regression analysis for highly stratified failure time data. *J. Amer. Statist. Assn.* 88, 557-565.
- Leonard, T. (1972). Bayesian methods for binomial data. *Biometrika* 59, 581-589.
- Liang, K.Y., Self, S.G. and Chang, Y.C. (1993). Modelling marginal hazards in multivariate failure time data. *J. R. Statist. Soc. B* 55, 441-453.

Liang, K.Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalised linear models. *Biometrika* 73, 13-22.

Lin, J.S. (1994). Analysis of multivariate failure time data: Marginal modelling approaches. The XVIIth International Biometric Conference (invited paper), 255-273.

Lin, J.S. and Wei, L.J. (1992). Linear regression analysis for multivariate failure time observations. *J. Amer. Statist. Assn.* 87, 1071-1097.

Lindsey, J.K. (1995). Fitting parametric counting processes by using log-linear models. *Appl. Statist.* 44, 201-212.

Mansour, H., Nordheim, E.V. and Rutledge, J.J. (1985). Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. *Biometrics* 41, 287-294.

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 Research* 37, 3863-3868.

Mantel, N. and Ciminera, J.L. (1979). Use of logrank scores in the analysis of litter-matched data on time to tumor appearance. *Cancer Research* 39, 4308-4315.

McCullagh, P. and Nelder, J.A. (1989). Generalized linear models, Second Ed. London: Chapman and Hall.

McGilchrist, C.A. (1993). REML estimation for survival models with frailty. *Biometrics* 49, 221-225.

McGilchrist, C.A. (1994). Estimation in generalised mixed models. *J. R. Statist. Soc. B* 56, 61-69.

McGilchrist, C.A. and Aisbett, C.W. (1991a). Restricted BLUP for mixed linear models. *Biom. J.* 33, 131-141.

McGilchrist, C.A. and Aisbett, C.W. (1991b). Regression with frailty in survival analysis. *Biometrics* 47, 461-466.

McGilchrist, C.A. and Yau, K.K.W. (1995a). Generalised linear mixed models. *Commun. Statist. - Theory Meth.* (revising).

McGilchrist, C.A. and Yau, K.K.W. (1995b). Survival analysis with time dependent frailty using a longitudinal model. *Aust. J. of Statistics.* (revising).

McGilchrist, C.A. and Zhaorong, J. (1990). Multicentre clinical trials and variance components. *Biom. J.* 32, 545-550.

Moertel, C.G., Fleming, T.R., McDonald, J.S., Haller, D.G., Laurie, J.A., Goodman, P.J., Ungerleider, J.S., Emerson, W.A., Tormey, D.C., Glick,

J.H., Veeder, M.H. and Mailliard, J.A. (1990). Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. N. Engl. J. Med. 322, 352-358.

Moore, E.H. (1920). On the reciprocal of the general algebraic matrix. Bull. Am. Math. Soc. 26, 394-395.

Nelder, J.A. and Wedderburn, R.W.M. (1972). Generalised linear models. J. R. Statist. Soc. A 135, 370-384.

Nielsen, G.G., Gill, R.D., Andersen, P.K. and Sørensen, T.I.A. (1992). A counting process approach to maximum likelihood estimation in frailty models. Scand. J. of Statistics. 19, 25-44.

Oakes, D. (1982). A model for association in bivariate survival data. J. R. Statist. Soc. B 44, 414-422.

Oakes, D. (1989). Bivariate survival models induced by frailties. J. Amer. Statist. Assn. 84, 487-493.

Oakes, D. (1994). Use of frailty models for multivariate survival data. The XVIIth International Biometric Conference (invited paper), 275-286.

Ochi, Y. and Prentice, R.L. (1984). Likelihood inference in correlated probit regression. Biometrika 71, 531-543.

Patterson, H.D. and Thompson, R. (1971). Recovery of interblock information when the block sizes are unequal. *Biometrika* 58, 545-554.

Penrose, R.A. (1955). A generalised inverse for matrices. *Proc. Cambridge Philos. Soc.* 51, 406-413.

Pepe, M.S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Amer. Statistician* 88, 811-820.

Prentice, R.L., Williams, B. and Peterson, A.V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* 68, 373-379.

Pulver, A.E. and Liang, K.Y. (1991). Estimating effects of proband characteristics on familial risk: II. the association between age at onset and familial risk in the Maryland schizophrenia sample. *Genetic Epidemiology* 8, 339-350.

Robinson, G.K. (1991). The estimation of random effects (with discussion). *Statistical Science* 6, 15-51.

Schall, R. (1991). Estimation in generalised linear models with random effects. *Biometrika* 78, 719-727.

Schatzkin, A.L., Cupples, L.A., Heeren, T., Morelock, S. and Kannel, W.B. (1984). Sudden death in the Framingham Heart Study, differences in

incidence and risk factors by sex and coronary disease status. Amer. J. of Epidemiology 120, 888-899.

Self, S. (1993). A regression model for counting processes with a time dependent frailty. Scand. J. of Statistics

Simons, L.A., McCallum, J., Simons, J., Powell, I., Ruys, J., Heller, R. and Lerba, C. (1990). The Dubbo Study: An Australian prospective community study of the health of the elderly. Australian and New Zealand Journal of Medicine 20, 783-789.

Simons, L.A., McCallum, J., Friedlander, Y., Simons, J., Powell, I. and Heller, R. (1991). The Dubbo Study of the health of the elderly: Medical and sociological findings at entry. Australian and New Zealand Journal of Medicine 21, 701-709.

Solomon, P.J. and Cox D.R. (1992). Nonlinear components of variance models. Biometrika 79, 1-11.

Speed, T. (1991). Comment on Robinson: Estimation of random effects. Statistical Science 6, 42-44.

Stiratelli, R., Laird, N. and Ware, J. (1984). Random effect models for serial observations with binary response. Biometrics 40, 961-971.

Storb, R., Pepe, M., Anasetti, C., Appelbaum, F.R., Beatty, P., Doney, K., Martin, P., Stewart, P., Sullivan, K.M., Witherspoon, R., Bensinger, W.,

Buckner, C.D., Clift, R., Hansen, J., Longton, G., Loughran, T., Petersen, F.B., Singer, J., Sanders, J. and Thomas, E.D. (1990). What role for prednisone in prevention of acute graft-versus-host disease in patients undergoing marrow transplants? *Blood* 76, 1037-1045.

Thompson, R. (1980). Maximum likelihood estimation of variance components. *Math. Operationsforsch. Statist. Ser. Statistics* 11, 545-561.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of individual mortality. *Demography* 16, 439-454.

Vaupel, J.W. and Yashin, A.I. (1983). The deviant dynamics of death in heterogeneous populations, RR-83-1, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ware, J.H. (1985). Linear models for the analysis of longitudinal studies. *Amer. Statistician* 39, 95-101.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 61, 439-447.

Wei, L.J., Lin, D.Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *J. Amer. Statist. Assn.* 84, 1065-1073.

Wild, C.J. (1983). Failure time models with matched data. *Biometrika* 70, 633-641.

Williams, D.A. (1982). Extra-binomial variation in logistic linear models. *Appl. Statist.* 31, 144-148.

Wolf, P.A., Abbott, R.D. and Kannel, W.B. (1987). Atrial fibrillation: A major contributor to stroke in the elderly. *Archives of Internal Medicine* 147, 1561-1564.

Wolf, P.A., D'Agostino, R.B., Kannel, W.B. and Belanger, A.J. (1988). Cigarette smoking as a risk factor for stroke. *J. Amer. Medical Assn.* 259, 1025-1029.

Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika* 80, 791-796.

Wolfinger, R. and O'Connell, M. (1993). Generalised linear mixed models: a pseudolikelihood approach. *J. Statist. Comput. Simul.* 48, 233-244.

Yau, K.K.W. and McGilchrist, C.A. (1995a). ML and REML estimation in survival analysis with time dependent correlated frailty. *Biometrics*. (revising).

Yau, K.K.W. and McGilchrist, C.A. (1995b). Analysis of multivariate survival data using the GLMM approach. *Appl. Statist.* (submitted).

Zeger, S.L. and Liang, K.Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42, 121-130.

Zeger, S.L., Liang, K.Y. and Albert, P.S. (1988). Models for longitudinal data: A generalised estimating equation approach. *Biometrics* 44, 1049-1060.

Zeger, S.L., Liang, K.Y. and Self, S.G. (1985). The analysis of binary longitudinal data with time independent covariates. *Biometrika* 72, 31-38.

Zhaorong, J., Matawie, K.M. and McGilchrist, C.A. (1992). Variance components for discordances. *Math. Biosciences* 110, 119-124.

Zhaorong, J., McGilchrist, C.A. and Jorgensen, M.A. (1992). Mixed model discrete regression. *Biom. J.* 34, 691-700.

APPENDIX I

DATA SETS

This Appendix contains three data sets. They are the Chronic Granulomatous Disease (CGD) data, the litter matched tumorigenesis experiment data and the Dubbo study data. Background of these data sets are explained in chapter 2. Detailed format of these three data sets are described in chapters 5, 7 and 8 respectively.

The Chronic Granulomatous Disease (CGD) Data

1	294	0	1	0	38	152.2	66.7	0	1	2	0	1	0	0
2	256	0	0	1	14	144	32.8	0	1	1	0	1	0	0
3	214	0	0	1	26	81.25	55	0	1	1	0	1	0	0
4	204	0	0	1	26	178.5	69.3	0	1	1	0	1	0	0
5	220	1	1	0	12	147	62	0	0	2	0	1	0	0
5	154	1	1	0	12	147	62	0	0	2	0	1	0	0.6023
5	41	0	1	0	12	147	62	0	0	2	0	1	0	1.024
6	9	1	0	0	15	159	47.5	0	1	1	0	1	0	0
6	18	1	0	0	15	159	47.5	0	1	1	0	1	0	0.0246
6	126	1	0	0	15	159	47.5	0	1	1	0	1	0	0.0739
6	89	1	0	0	15	159	47.5	0	1	1	0	1	0	0.4189
6	8	1	0	0	15	159	47.5	0	1	1	0	1	0	0.6626
6	73	1	0	0	15	159	47.5	0	1	1	0	1	0	0.6845
6	28	1	0	0	15	159	47.5	0	1	1	0	1	0	0.8843
6	89	0	0	0	15	159	47.5	0	1	1	0	1	0	0.961
7	383	0	1	1	19	171	72.7	0	1	1	0	1	0	0
8	389	0	1	1	12	142	34	0	1	1	0	1	0	0
9	212	1	0	1	1	79	10.5	0	1	1	0	1	0	0
9	49	1	0	1	1	79	10.5	0	1	1	0	1	0	0.5804
9	51	0	1	1	1	79	10.5	0	1	1	0	1	0	0.7146
9	41	0	1	1	1	79	10.5	0	1	1	0	1	0	0.7283
9	38	1	0	1	1	79	10.5	0	1	1	0	1	0	0.7392
9	56	0	0	1	1	79	10.5	0	1	1	0	1	0	0.8433
10	83	1	1	0	9	134.5	32.7	0	1	1	0	1	0	0
10	32	1	1	0	9	134.5	32.7	0	1	1	0	1	0	0.2272
10	223	1	1	0	9	134.5	32.7	0	1	1	0	1	0	0.3149
10	30	0	1	0	9	134.5	32.7	0	1	1	0	1	0	0.9254
11	19	1	0	1	1	79	11.47	0	1	1	0	1	0	0
11	344	0	0	1	1	79	11.47	0	1	1	0	1	0	0.052
12	268	1	1	1	5	102	18	0	1	1	0	1	0	0
12	93	0	1	1	5	102	18	0	1	1	0	1	0	0.7337
13	338	0	1	0	22	169	52.2	0	1	1	0	1	0	0
14	271	0	0	0	19	159	46	0	0	2	0	1	0	0
15	275	0	0	1	7	115.5	19.5	0	1	1	0	1	0	0
16	272	0	1	1	25	185	58.4	0	1	1	0	1	0	0
17	253	0	0	1	31	170	80.5	0	1	1	0	1	0	0
18	244	0	1	1	37	155	67.5	0	1	2	0	1	0	0
19	105	1	0	0	6	130	21.6	0	1	2	0	1	0	0
19	123	0	0	0	6	130	21.6	0	1	2	0	1	0	0.2875
20	228	0	0	0	3	96	13.1	0	1	2	0	1	0	0
21	199	0	0	0	22	163	49.1	0	1	2	0	0	0	0
22	208	0	0	1	17	169	63.5	0	1	1	0	0	0	0
23	169	1	0	1	19	182	63.9	0	1	1	0	0	0	0
23	32	0	0	1	19	182	63.9	0	1	1	0	0	0	0.4627
24	198	0	1	0	36	167	60.8	0	1	2	0	0	0	0
25	247	1	0	1	17	162.5	52.7	0	1	1	1	0	0	0
25	7	1	0	1	17	162.5	52.7	0	1	1	1	0	0	0.6762
25	130	0	0	1	17	162.5	52.7	0	1	1	1	0	0	0.6954
26	295	1	0	0	27	176	82.8	0	1	1	1	0	0	0
26	55	0	0	0	27	176	82.8	0	1	1	1	0	0	0.8077
27	372	0	1	1	5	113	19.5	0	1	1	1	0	0	0

28	20	1	0	1	2	93	13.2	0	1	1	1	0	0	0
28	83	0	0	1	2	93	13.2	0	1	1	1	0	0	0.0548
29	374	1	1	1	8	124	25.4	0	1	1	1	0	0	0
29	15	0	1	1	8	124	25.4	0	1	1	1	0	0	1.024
30	389	0	1	1	12	144	36.9	0	1	1	1	0	0	0
31	366	0	0	1	27	174	67.8	0	1	1	1	0	0	0
32	335	1	0	1	14	143.5	33.4	0	1	1	1	0	0	0
32	36	1	0	1	14	143.5	33.4	0	1	1	1	0	0	0.9172
32	12	0	0	1	14	143.5	33.4	0	1	1	1	0	0	1.0157
33	374	0	1	0	11	149	50.9	0	1	1	1	0	0	0
34	281	1	0	0	29	175	73.1	0	1	1	1	0	0	0
34	105	0	0	0	29	175	73.1	0	1	1	1	0	0	0.7693
35	377	0	1	0	31	167	51.8	0	1	2	1	0	0	0
36	361	0	1	1	7	121	19.9	0	1	1	1	0	0	0
37	307	0	0	0	26	153	46.9	0	1	2	1	0	0	0
38	119	1	1	1	13	145.2	36.2	0	1	1	1	0	0	0
38	122	1	1	1	13	145.2	36.2	0	1	1	1	0	0	0.3258
38	11	0	1	1	13	145.2	36.2	0	1	1	1	0	0	0.6598
39	188	1	1	0	25	168	68.9	0	1	2	1	0	0	0
39	169	0	1	0	25	168	68.9	0	1	2	1	0	0	0.5147
40	340	0	1	0	9	140	36	0	1	2	1	0	0	0
41	7	1	0	1	28	174	63.7	0	1	1	1	0	0	0
41	295	0	0	1	28	174	63.7	0	1	1	1	0	0	0.0192
42	335	0	1	0	13	139	34.8	0	1	2	1	0	0	0
43	270	0	1	0	24	177	78.4	0	1	1	1	0	0	0
44	274	0	0	1	11	123	24.3	0	1	1	1	0	0	0
45	114	1	1	1	4	103.8	16.8	0	1	1	1	0	0	0
45	160	0	1	1	4	103.8	16.8	0	1	1	1	0	0	0.3121
46	100	1	0	1	19	170	71.2	0	1	1	1	0	0	0
46	207	1	0	1	19	170	71.2	0	1	1	1	0	0	0.2738
47	264	0	1	1	18	166	58.1	0	1	1	1	0	0	0
48	168	1	1	0	7	135	42.9	0	1	1	1	0	0	0
48	73	0	1	0	7	135	42.9	0	1	1	1	0	0	0.46
49	272	0	0	0	12	166	51.9	0	1	2	1	0	0	0
50	255	0	1	1	10	129	27.4	0	1	1	1	0	0	0
51	279	0	0	1	9	129.4	28.7	0	1	1	0	1	0	0
52	266	0	1	1	5	112.3	20.7	0	1	1	0	1	0	0
53	218	0	1	1	1	76.3	11.3	0	1	1	0	1	0	0
54	166	1	1	1	7	119	20.6	0	1	1	0	1	0	0
54	45	0	1	1	7	119	20.6	0	1	1	0	1	0	0.4545
55	193	0	0	0	11	137.5	40.3	0	1	1	0	1	0	0
56	12	1	0	0	4	98.3	14.4	0	1	1	0	1	0	0
56	11	1	0	0	4	98.3	14.4	0	1	1	0	1	0	0.0329
56	147	1	0	0	4	98.3	14.4	0	1	1	0	1	0	0.063
56	26	0	0	0	4	98.3	14.4	0	1	1	0	1	0	0.4654
57	121	1	0	0	7	113	20.4	0	1	2	0	1	0	0
57	83	0	0	0	7	113	20.4	0	1	2	0	1	0	0.3313
58	198	0	1	1	15	178.7	60.5	0	1	1	0	1	0	0
59	207	1	0	1	1	79	12.2	0	1	1	0	1	0	0
59	141	0	0	1	1	79	12.2	0	1	1	0	1	0	0.5667
60	336	0	1	1	7	116.6	23.3	0	1	1	0	1	0	0
61	275	1	1	1	17	170.2	47.9	0	1	1	0	1	0	0
61	87	1	1	1	17	170.2	47.9	0	1	1	0	1	0	0.7529
61	4	0	1	1	17	170.2	47.9	0	1	1	0	1	0	0.9911
62	337	0	1	1	8	125.4	27.7	0	1	1	0	1	0	0

63	53	1	0	1	20	173	68.4	0	1	1	0	1	0	0
63	13	1	0	1	20	173	68.4	0	1	1	0	1	0	0.1451
63	190	1	0	1	20	173	68.4	0	1	1	0	1	0	0.1807
63	15	0	0	1	20	173	68.4	0	1	1	0	1	0	0.7009
64	68	1	0	1	5	114.5	23	0	0	1	0	1	0	0
64	181	1	0	1	5	114.5	23	0	0	1	0	1	0	0.1862
64	2	1	0	1	5	114.5	23	0	0	1	0	1	0	0.6817
64	34	1	0	1	5	114.5	23	0	0	1	0	1	0	0.6872
64	63	0	0	1	5	114.5	23	0	0	1	0	1	0	0.7803
65	319	0	1	1	6	105.5	19.5	1	1	1	0	1	0	0
66	319	1	0	1	9	129.6	29.3	0	0	1	0	1	0	0
66	41	0	0	1	9	129.6	29.3	0	0	1	0	1	0	0.8734
67	260	0	1	1	5	110.4	19.1	0	1	1	0	1	0	0
68	365	0	1	0	44	153.3	45	0	0	2	0	1	0	0
69	293	1	0	1	22	175	59.7	0	1	1	0	1	0	0
69	72	0	0	1	22	175	59.7	0	1	1	0	1	0	0.8022
70	364	0	1	1	7	111	17.4	0	1	1	0	1	0	0
71	351	0	1	1	19	173.6	61.4	0	0	1	0	1	0	0
72	270	0	1	1	34	182.6	94.8	0	0	1	0	1	0	0
73	186	0	1	1	32	177.9	63.35	0	0	1	0	1	0	0
74	92	0	0	1	25	185	74.55	0	0	1	0	1	0	0
75	92	0	0	1	21	189	101.5	0	0	1	0	1	0	0
76	358	0	0	1	7	109	14.65	0	1	1	0	1	0	0
77	176	1	0	1	5	101	15.4	0	1	1	0	1	0	0
77	105	1	0	1	5	101	15.4	0	1	1	0	1	0	0.4819
77	73	0	0	1	5	101	15.4	0	1	1	0	1	0	0.7693
78	344	0	0	1	24	169.7	58.95	0	1	1	0	1	0	0
79	266	1	1	1	12	138	28.1	0	1	1	0	1	0	0
79	38	0	1	1	12	138	28.1	0	1	1	0	1	0	0.7283
80	227	1	0	1	5	97.1	15.3	0	1	1	0	1	0	0
80	96	0	0	1	5	97.1	15.3	0	1	1	0	1	0	0.6215
81	256	0	1	1	24	171	55.1	0	1	1	0	1	0	0
82	66	1	1	1	26	176.8	66	0	1	1	0	0	0	0
82	278	0	1	1	26	176.8	66	0	1	1	0	0	0	0.1807
83	295	0	0	1	6	104.4	13.8	0	1	1	0	0	0	0
84	304	0	1	0	9	122.8	20.1	0	1	2	0	0	0	0
85	24	1	0	1	25	176.9	73.5	0	0	1	0	0	0	0
85	247	0	0	1	25	176.9	73.5	0	0	1	0	0	0	0.0657
86	271	0	1	1	2	93.5	14.2	0	1	1	0	0	0	0
87	246	0	0	1	8	121.6	22.4	0	1	1	0	0	0	0
88	262	0	0	1	10	125.9	29.7	0	1	1	0	0	0	0
89	285	0	1	1	20	169.8	50.1	0	1	1	0	0	0	0
90	277	0	0	0	34	166.5	58.2	0	1	1	0	0	0	0
91	295	0	1	1	6	119.9	26.2	0	1	1	0	0	0	0
92	278	0	1	1	11	139.2	34.9	0	1	1	0	0	0	0
93	5	1	0	0	3	91.7	14.4	0	1	1	0	0	0	0
93	155	1	0	0	3	91.7	14.4	0	1	1	0	0	0	0.0137
93	54	1	0	0	3	91.7	14.4	0	1	1	0	0	0	0.4381
93	74	0	0	0	3	91.7	14.4	0	1	1	0	0	0	0.5859
94	332	0	1	1	9	131.3	24	0	1	1	0	0	0	0
95	289	0	0	0	11	138.6	36.1	0	1	2	0	0	0	0
96	270	0	0	1	17	156.7	36.8	0	1	1	0	0	0	0
97	270	0	0	1	10	143	31.3	0	1	1	0	0	0	0
98	331	0	0	0	17	171	46.1	0	1	2	0	1	0	0
99	58	1	0	0	8	115	19	0	1	2	0	1	0	0

99	64	1	0	0	8	115	19	0	1	2	0	1	0	0	0.1588
99	230	0	0	0	8	115	19	0	1	2	0	1	0	0	0.334
100	298	0	1	1	6	111.9	17.9	0	1	1	0	1	0	0	
101	282	0	0	0	7	116.2	22.2	0	1	1	0	1	0	0	
102	277	0	1	1	7	119.7	21.5	0	1	1	0	1	0	0	
103	280	0	1	1	8	116.8	20	0	1	1	0	1	0	0	
104	15	1	0	1	11	141.5	36	0	1	1	0	1	0	0	
104	264	0	0	1	11	141.5	36	0	1	1	0	1	0	0	0.0411
105	200	0	1	1	4	108.1	17.3	0	1	1	0	1	0	0	
106	309	0	0	0	18	179	67	0	1	1	0	0	1	0	
107	328	0	1	1	13	151	49	0	1	1	0	0	1	0	
108	330	0	0	1	11	136.5	31.2	0	0	1	0	0	1	0	
109	319	0	1	1	2	86	13.5	0	1	1	0	0	1	0	
110	305	0	1	1	17	180	68	0	1	1	0	0	1	0	
111	317	0	0	0	35	181.5	80	0	0	1	0	0	1	0	
112	301	0	0	1	25	172.5	50	0	1	1	0	0	1	0	
113	147	1	1	1	14	145	29	0	0	1	0	0	1	0	
113	42	1	1	1	14	145	29	0	0	1	0	0	1	0	0.4025
113	112	0	1	1	14	145	29	0	0	1	0	0	1	0	0.5175
114	305	1	0	0	25	187.5	74	0	1	1	0	0	1	0	
114	8	0	0	0	25	187.5	74	0	1	1	0	0	1	0	0.835
115	92	1	0	1	27	169.5	65	1	0	2	0	0	1	0	
115	30	1	0	1	27	169.5	65	1	0	2	0	0	1	0	0.2519
115	82	1	0	1	27	169.5	65	1	0	2	0	0	1	0	0.334
115	84	0	0	1	27	169.5	65	1	0	2	0	0	1	0	0.5585
116	294	0	0	0	32	185	95	0	1	1	0	0	1	0	
117	294	0	1	1	6	120	23	0	1	1	0	0	1	0	
118	265	1	0	0	8	138	31	0	0	1	0	0	1	0	
118	22	0	0	0	8	138	31	0	0	1	0	0	1	0	0.7255
119	287	0	1	0	9	144	34.4	0	1	1	0	0	1	0	
120	274	0	1	0	23	170	49	0	0	2	0	0	1	0	
121	237	1	0	0	23	122	31.2	1	1	2	0	0	1	0	
121	37	0	0	0	23	122	31.2	1	1	2	0	0	1	0	0.6489
122	274	0	1	0	17	127	27	0	1	2	0	0	1	0	
123	208	1	1	0	21	158	49	0	1	1	0	0	1	0	
123	66	0	1	0	21	158	49	0	1	1	0	0	1	0	0.5695
124	265	0	0	1	1	81	10.4	0	1	1	0	0	1	0	
125	161	0	1	1	12	136.5	30	0	1	1	0	1	0	0	
126	147	1	0	1	7	120	23	0	1	1	0	1	0	0	
126	170	0	0	1	7	120	23	0	1	1	0	1	0	0	0.4025
127	317	0	0	1	1	79	10.6	0	1	1	0	1	0	0	
128	316	0	1	1	3	97.8	13	0	1	1	0	1	0	0	

The Litter Matched Tumorigenesis Experiment Data

1	101	0	1
1	49	1	0
1	104	0	0
2	104	0	1
2	102	0	0
2	104	0	0
3	104	0	1
3	104	0	0
3	104	0	0
4	77	0	1
4	97	0	0
4	79	0	0
5	89	0	1
5	104	0	0
5	104	0	0
6	88	1	1
6	96	1	0
6	104	0	0
7	104	1	1
7	94	0	0
7	77	1	0
8	95.9	1	1
8	104	0	0
8	104	0	0
9	82	0	1
9	77	0	0
9	104	0	0
10	70	1	1
10	104	0	0
10	77	0	0
11	88.9	1	1
11	91	0	0
11	90	0	0
12	91	0	1
12	70	0	0
12	92	0	0
13	39	1	1
13	45	0	0
13	50	1	0
14	102.9	1	1
14	69	0	0
14	91	0	0
15	93	0	1
15	104	0	0
15	103	0	0
16	85	0	1
16	72	0	0
16	104	0	0
17	104	0	1
17	63	0	0
17	104	0	0

18	104	0	1
18	104	0	0
18	74	0	0
19	81	0	1
19	104	0	0
19	69	0	0
20	67	1	1
20	104	0	0
20	68	1	0
21	104	0	1
21	104	0	0
21	104	0	0
22	104	0	1
22	104	0	0
22	104	0	0
23	104	0	1
23	83	0	0
23	40	1	0
24	87	0	1
24	104	0	0
24	104	0	0
25	104	0	1
25	104	0	0
25	104	0	0
26	89	0	1
26	104	0	0
26	104	0	0
27	78	0	1
27	104	0	0
27	104	0	0
28	104	0	1
28	81	1	0
28	64	1	0
29	86	1	1
29	55	1	0
29	94	0	0
30	34	1	1
30	104	0	0
30	54	1	0
31	76	0	1
31	87	0	0
31	74	0	0
32	102.8	1	1
32	73	1	0
32	83.9	1	0
33	101.9	1	1
33	104	0	0
33	80	0	0
34	79.9	1	1
34	104	0	0
34	73	0	0
35	45	1	1
35	79	0	0
35	104	0	0
36	94	1	1

36	104	0	0
36	104	0	0
37	104	0	1
37	104	0	0
37	104	0	0
38	104	0	1
38	101	1	0
38	94	0	0
39	76	0	1
39	84	1	0
39	78	1	0
40	80	1	1
40	80.9	1	0
40	76	0	0
41	72	1	1
41	95	0	0
41	104	0	0
42	72.9	1	1
42	104	0	0
42	66	1	0
43	92	1	1
43	104	0	0
43	102	1	0
44	104	0	1
44	98	0	0
44	73	0	0
45	55	0	1
45	104	0	0
45	104	0	0
46	49	0	1
46	83	0	0
46	77	0	0
47	89	1	1
47	104	0	0
47	104	0	0
48	88	0	1
48	79	0	0
48	99	0	0
49	103	1	1
49	91	0	0
49	104	0	0
50	104	0	1
50	104	0	0
50	79	1	0

The Dubbo Study Data (male individuals of age 70)

1	679	1	0	0	0	41.99	0	0	1	1	0	1
1	1389	0	0	0	0	41.99	0	0	1	1	0	1
2	1552	1	0	0	1	35.88	0	0	0	1	0	1
2	520	0	0	0	1	35.88	0	0	0	1	0	1
3	1447	1	0	0	0	28.3	0	0	0	1	0	0
3	565	1	0	0	0	28.3	0	0	0	1	0	0
3	45	0	0	0	0	28.3	0	0	0	1	0	0
4	282	1	0	1	0	40.83	0	1	0	1	1	0
4	1	0	1	1	0	40.83	0	1	0	1	1	0
5	2050	0	0	0	0	51.55	0	1	0	1	0	0
6	319	1	0	0	0	41.6	0	0	1	0	0	0
6	785	1	0	0	0	41.6	0	0	1	0	0	0
6	941	0	0	0	0	41.6	0	0	1	0	0	0
7	760	1	0	0	0	34.22	0	0	0	1	1	0
7	79	1	0	0	0	34.22	0	0	0	1	1	0
7	1211	0	0	0	0	34.22	0	0	0	1	1	0
8	567	1	0	0	0	36.36	0	0	1	1	1	0
8	114	1	0	0	0	36.36	0	0	1	1	1	0
8	483	1	0	0	0	36.36	0	0	1	1	1	0
8	5	1	0	0	0	36.36	0	0	1	1	1	0
8	133	1	0	0	0	36.36	0	0	1	1	1	0
8	54	1	0	0	0	36.36	0	0	1	1	1	0
8	42	1	0	0	0	36.36	0	0	1	1	1	0
8	42	1	0	0	0	36.36	0	0	1	1	1	0
8	119	1	0	0	0	36.36	0	0	1	1	1	0
8	84	1	0	0	0	36.36	0	0	1	1	1	0
8	56	1	0	0	0	36.36	0	0	1	1	1	0
8	332	0	0	0	0	36.36	0	0	1	1	1	0
9	1021	1	0	0	0	27.04	0	0	0	1	1	0
9	390	1	0	0	0	27.04	0	0	0	1	1	0
9	620	0	0	0	0	27.04	0	0	0	1	1	0
10	2029	0	0	1	0	42.38	0	0	0	0	1	0
11	42	1	0	1	0	25.81	1	0	0	0	0	1
11	1242	1	0	1	0	25.81	1	0	0	0	0	1
11	116	1	0	1	0	25.81	1	0	0	0	0	1
11	116	1	0	1	0	25.81	1	0	0	0	0	1
11	12	0	1	1	0	25.81	1	0	0	0	0	1
12	549	1	0	1	0	41.47	0	1	0	1	0	1
12	308	1	0	1	0	41.47	0	1	0	1	0	1
12	38	1	0	1	0	41.47	0	1	0	1	0	1
12	17	1	0	1	0	41.47	0	1	0	1	0	1
12	3	1	0	1	0	41.47	0	1	0	1	0	1
12	10	0	1	1	0	41.47	0	1	0	1	0	1
13	2018	0	0	0	0	55.95	0	0	1	1	0	0
14	263	1	0	1	0	36	0	0	0	0	1	0
14	418	1	0	1	0	36	0	0	0	0	1	0
14	608	1	0	1	0	36	0	0	0	0	1	0
14	42	1	0	1	0	36	0	0	0	0	1	0
14	2	0	1	1	0	36	0	0	0	0	1	0
15	531	1	0	0	1	31.7	0	0	0	1	1	0
15	491	1	0	0	1	31.7	0	0	0	1	1	0

15	101	1	0	0	1	31.7	0	0	0	1	1	0
15	875	0	0	0	1	31.7	0	0	0	1	1	0
16	412	1	0	1	1	29.38	0	0	0	1	0	0
16	561	1	0	1	1	29.38	0	0	0	1	0	0
16	922	1	0	1	1	29.38	0	0	0	1	0	0
16	21	1	0	1	1	29.38	0	0	0	1	0	0
16	65	1	0	1	1	29.38	0	0	0	1	0	0
16	5	1	0	1	1	29.38	0	0	0	1	0	0
16	88	0	0	1	1	29.38	0	0	0	1	0	0
17	146	1	0	1	1	42.25	1	0	1	0	1	0
17	273	1	0	1	1	42.25	1	0	1	0	1	0
17	31	1	0	1	1	42.25	1	0	1	0	1	0
17	80	1	0	1	1	42.25	1	0	1	0	1	0
17	909	1	0	1	1	42.25	1	0	1	0	1	0
17	27	1	0	1	1	42.25	1	0	1	0	1	0
17	486	0	0	1	1	42.25	1	0	1	0	1	0
18	512	1	0	1	0	42.77	0	0	0	1	1	0
18	1477	0	0	1	0	42.77	0	0	0	1	1	0
19	1959	0	0	0	0	45.29	0	1	0	1	1	0
20	1983	0	0	0	1	40.2	0	0	0	1	0	0
21	398	1	0	1	1	33.29	0	0	0	1	1	0
21	943	1	0	1	1	33.29	0	0	0	1	1	0
21	22	1	0	1	1	33.29	0	0	0	1	1	0
21	619	0	0	1	1	33.29	0	0	0	1	1	0
22	48	1	0	0	0	32.38	0	0	0	1	0	0
22	1912	0	0	0	0	32.38	0	0	0	1	0	0
23	1953	0	0	0	0	31.58	1	0	0	0	1	
24	1954	0	0	0	0	45.43	0	1	0	1	0	0
25	1745	1	0	0	0	75	0	0	0	1	0	1
25	280	0	0	0	0	75	0	0	0	1	0	1
26	662	1	0	0	0	22.94	0	0	0	1	1	0
26	243	1	0	0	0	22.94	0	0	0	1	1	0
26	1128	0	0	0	0	22.94	0	0	0	1	1	0
27	750	1	0	1	1	21.25	0	0	0	1	1	0
27	30	1	0	1	1	21.25	0	0	0	1	1	0
27	16	0	1	1	1	21.25	0	0	0	1	1	0
28	508	1	0	1	0	52.56	0	0	0	1	0	0
28	243	1	0	1	0	52.56	0	0	0	1	0	0
28	24	1	0	1	0	52.56	0	0	0	1	0	0
28	1150	0	0	1	0	52.56	0	0	0	1	0	0
29	511	1	0	0	0	55.95	0	1	0	1	0	0
29	618	1	0	0	0	55.95	0	1	0	1	0	0
29	789	0	0	0	0	55.95	0	1	0	1	0	0
30	41	1	0	0	1	36.24	0	0	1	1	1	0
30	34	1	0	0	1	36.24	0	0	1	1	1	0
30	20	1	0	0	1	36.24	0	0	1	1	1	0
30	178	1	0	0	1	36.24	0	0	1	1	1	0
30	321	1	0	0	1	36.24	0	0	1	1	1	0
30	590	1	0	0	1	36.24	0	0	1	1	1	0
30	21	1	0	0	1	36.24	0	0	1	1	1	0
30	3	0	1	0	1	36.24	0	0	1	1	1	0
31	1041	1	0	1	0	25.81	0	0	0	1	1	0
31	81	1	0	1	0	25.81	0	0	0	1	1	0
31	23	1	0	1	0	25.81	0	0	0	1	1	0
31	307	1	0	1	0	25.81	0	0	0	1	1	0

31	200	1	0	1	0	25.81	0	0	0	1	1	0
31	51	1	0	1	0	25.81	0	0	0	1	1	0
31	266	0	0	1	0	25.81	0	0	0	1	1	0
32	1669	1	0	0	1	38.81	0	0	0	1	1	0
32	214	0	0	0	1	38.81	0	0	0	1	1	0
33	77	1	0	1	0	70.73	0	0	0	1	1	0
33	744	1	0	1	0	70.73	0	0	0	1	1	0
33	1175	0	0	1	0	70.73	0	0	0	1	1	0
34	1834	1	0	1	0	36.24	0	0	0	1	1	0
34	9	1	0	1	0	36.24	0	0	0	1	1	0
34	47	0	0	1	0	36.24	0	0	0	1	1	0
35	1883	0	0	0	0	28.41	0	0	0	1	1	0
36	14	1	0	1	1	37.45	1	1	0	0	0	0
36	42	1	0	1	1	37.45	1	1	0	0	0	0
36	134	1	0	1	1	37.45	1	1	0	0	0	0
36	35	1	0	1	1	37.45	1	1	0	0	0	0
36	148	1	0	1	1	37.45	1	1	0	0	0	0
36	1161	0	1	1	1	37.45	1	1	0	0	0	0
37	928	1	0	1	1	31.14	0	0	0	1	0	0
37	689	1	0	1	1	31.14	0	0	0	1	0	0
37	318	0	0	1	1	31.14	0	0	0	1	0	0
38	205	1	0	0	1	54.76	0	0	1	0	1	0
38	7	1	0	0	1	54.76	0	0	1	0	1	0
38	406	0	1	0	1	54.76	0	0	1	0	1	0
39	1365	1	0	1	0	60.68	0	0	0	1	0	0
39	490	0	0	1	0	60.68	0	0	0	1	0	0
40	303	1	0	0	0	77.79	0	0	0	1	1	0
40	741	1	0	0	0	77.79	0	0	0	1	1	0
40	9	1	0	0	0	77.79	0	0	0	1	1	0
40	3	1	0	0	0	77.79	0	0	0	1	1	0
40	2	0	1	0	0	77.79	0	0	0	1	1	0
41	1008	1	0	0	0	27.98	0	0	0	1	0	0
41	933	0	0	0	0	27.98	0	0	0	1	0	0
42	1669	1	0	1	0	75.52	0	0	0	1	0	1
42	161	0	0	1	0	75.52	0	0	0	1	0	1
43	1764	0	0	0	0	29.81	0	0	1	1	1	0
44	549	1	0	0	0	7.02	1	1	0	0	1	0
44	40	0	1	0	0	7.02	1	1	0	0	1	0
45	1017	1	0	1	0	36	0	1	0	1	0	0
45	804	0	0	1	0	36	0	1	0	1	0	0
46	1391	1	0	0	0	26.11	0	1	0	1	0	1
46	46	1	0	0	0	26.11	0	1	0	1	0	1
46	18	1	0	0	0	26.11	0	1	0	1	0	1
46	28	1	0	0	0	26.11	0	1	0	1	0	1
46	308	0	0	0	0	26.11	0	1	0	1	0	1
47	1017	1	0	0	0	26.52	0	0	0	0	0	1
47	105	1	0	0	0	26.52	0	0	0	0	0	1
47	20	1	0	0	0	26.52	0	0	0	0	0	1
47	637	0	0	0	0	26.52	0	0	0	0	0	1
48	196	1	0	1	1	40.32	0	0	0	1	1	0
48	347	1	0	1	1	40.32	0	0	0	1	1	0
48	262	1	0	1	1	40.32	0	0	0	1	1	0
48	41	1	0	1	1	40.32	0	0	0	1	1	0
48	5	1	0	1	1	40.32	0	0	0	1	1	0
48	78	1	0	1	1	40.32	0	0	0	1	1	0

48	7 1 0 1 1	40.32 0 0 0 1 1 0
48	7 1 0 1 1	40.32 0 0 0 1 1 0
48	7 1 0 1 1	40.32 0 0 0 1 1 0
48	15 1 0 1 1	40.32 0 0 0 1 1 0
48	25 1 0 1 1	40.32 0 0 0 1 1 0
48	3 1 0 1 1	40.32 0 0 0 1 1 0
48	28 1 0 1 1	40.32 0 0 0 1 1 0
48	14 1 0 1 1	40.32 0 0 0 1 1 0
48	61 1 0 1 1	40.32 0 0 0 1 1 0
48	37 1 0 1 1	40.32 0 0 0 1 1 0
48	29 1 0 1 1	40.32 0 0 0 1 1 0
48	7 1 0 1 1	40.32 0 0 0 1 1 0
48	22 1 0 1 1	40.32 0 0 0 1 1 0
48	23 1 0 1 1	40.32 0 0 0 1 1 0
48	23 1 0 1 1	40.32 0 0 0 1 1 0
48	8 1 0 1 1	40.32 0 0 0 1 1 0
48	25 1 0 1 1	40.32 0 0 0 1 1 0
48	9 1 0 1 1	40.32 0 0 0 1 1 0
48	6 1 0 1 1	40.32 0 0 0 1 1 0
48	45 1 0 1 1	40.32 0 0 0 1 1 0
48	63 1 0 1 1	40.32 0 0 0 1 1 0
48	1 1 0 1 1	40.32 0 0 0 1 1 0
48	40 1 0 1 1	40.32 0 0 0 1 1 0
48	38 1 0 1 1	40.32 0 0 0 1 1 0
48	36 1 0 1 1	40.32 0 0 0 1 1 0
48	39 1 0 1 1	40.32 0 0 0 1 1 0
48	45 1 0 1 1	40.32 0 0 0 1 1 0
48	30 1 0 1 1	40.32 0 0 0 1 1 0
48	6 1 0 1 1	40.32 0 0 0 1 1 0
48	35 1 0 1 1	40.32 0 0 0 1 1 0
48	28 1 0 1 1	40.32 0 0 0 1 1 0
48	28 1 0 1 1	40.32 0 0 0 1 1 0
48	34 1 0 1 1	40.32 0 0 0 1 1 0
48	25 1 0 1 1	40.32 0 0 0 1 1 0
48	7 0 0 1 1	40.32 0 0 0 1 1 0
49	1757 0 0 1 0	31.25 0 0 0 1 1 0
50	705 1 0 0 0	95.06 0 0 0 1 1 0
50	22 0 1 0 0	95.06 0 0 0 1 1 0
51	1008 1 0 1 1	78.32 1 0 1 1 0 0
51	982 0 0 1 1	78.32 1 0 1 1 0 0
52	653 1 0 0 0	112.36 0 1 0 1 1 0
52	119 1 0 0 0	112.36 0 1 0 1 1 0
52	245 1 0 0 0	112.36 0 1 0 1 1 0
52	84 1 0 0 0	112.36 0 1 0 1 1 0
52	210 1 0 0 0	112.36 0 1 0 1 1 0
52	336 1 0 0 0	112.36 0 1 0 1 1 0
52	61 1 0 0 0	112.36 0 1 0 1 1 0
52	43 1 0 0 0	112.36 0 1 0 1 1 0
52	239 0 0 0 0	112.36 0 1 0 1 1 0
53	102 1 0 0 1	35.05 0 0 1 1 1 0
53	20 1 0 0 1	35.05 0 0 1 1 1 0
53	42 1 0 0 1	35.05 0 0 1 1 1 0
53	31 1 0 0 1	35.05 0 0 1 1 1 0
53	1045 1 0 0 1	35.05 0 0 1 1 1 0
53	88 1 0 0 1	35.05 0 0 1 1 1 0

53	10	1	0	0	1	35.05	0	0	1	1	1	0
53	27	1	0	0	1	35.05	0	0	1	1	1	0
53	371	1	0	0	1	35.05	0	0	1	1	1	0
53	234	0	0	0	1	35.05	0	0	1	1	1	0
54	1754	1	0	0	0	30.58	0	1	0	1	1	0
54	222	0	0	0	0	30.58	0	1	0	1	1	0
55	1464	1	0	0	0	32.6	0	0	0	0	0	0
55	489	0	0	0	0	32.6	0	0	0	0	0	0
56	1204	1	0	0	0	35.88	0	0	0	1	0	1
56	282	1	0	0	0	35.88	0	0	0	1	0	1
56	434	0	0	0	0	35.88	0	0	0	1	0	1
57	1309	1	0	1	1	46.51	1	1	0	1	0	0
57	289	1	0	1	1	46.51	1	1	0	1	0	0
57	273	0	0	1	1	46.51	1	1	0	1	0	0
58	1875	0	0	0	0	57.76	0	0	0	1	1	0
59	560	1	0	0	0	35.76	1	1	0	1	0	0
59	1247	0	0	0	0	35.76	1	1	0	1	0	0
60	4	1	0	1	0	39.06	0	0	0	1	0	0
60	1824	0	0	1	0	39.06	0	0	0	1	0	0
61	308	1	0	1	1	46.24	1	0	0	1	1	0
61	1526	0	0	1	1	46.24	1	0	0	1	1	0
62	264	1	0	1	1	58.52	0	0	0	0	0	1
62	1559	0	0	1	1	58.52	0	0	0	0	0	1

APPENDIX II

APL PROGRAMS

This Appendix contains those APL programs that has been used to analyse the three data sets provided in Appendix I. Main APL programs and their corresponding models are listed in the following table.

Chapter	Model	APL program
Five	Time independent frailty model	GROWTH
	Model 1	GROWTH1
	Model 2	GROWTH2
	Model 3	GROWTH3
Six	AR(1) frailty model	AR
Seven	Baseline frailty model	GROWTH
	Justify the exp. relative risk function	GENGROWTH
	Random block frailty model	RANBLKB
	Simulation (baseline frailty model)	SIMBASELINE
	Simulation (random block frailty model)	SIMRANBLK3
Eight	Model 1	COMRSK1
	Model 2	COMRSK2

The rest are the APL subprograms that appear in one or more of the main APL programs.

GROWTH

=====

v BETA<-GROWTH X;M;N;NV;MA;THETA;RHO;ML;L;I;A;W;V
[1] ^FITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2] ^OF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[3] ^FOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES. COX MODEL
[4] ^IS FITTED WITH AN INCLUSION OF A RANDOM FRAILTY TERM FOR EACH PATIENT.
[5] M<-T[X[,1]
[6] N<-+1+pX
[7] NV<-+/-3+-1+pX
[8] MA<-T/+/(,M)=X[,1]
[9] X<-X,X[,1]=,M
[10] BETA<-(+/-3+1+pX)p0
[11] THETA<0.41
[12] ML<[],0,p0[]<'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[13] X<-X[X[,2];]
[14] LBL1:L<-X LL02(0 3+X)+.x*BETA
[15] I<-,M=,M
[16] A<-(M+NV)NVp0,(NV Mp0),[1]@I*THETA
[17] W<-[((,0 3+X)+.x(0 1+L)+.x0 3+X)+A
[18] BETA<-(V*BETA)+W+.x((,0 3+X)+.xL[,1])-A+.x*BETA
'BETA ESTIMATE',6 RND NV+BETA
[20] +(0.01<|V-BETA)/LBL1
[21] +(ML=0)/LBL3
[22] +(ML#1)/LBL2
[23] W<-(NV NV+W),(NV Mp0)),[1](M NVp0),@NV NV+@W
LBL2:A<-(+/-/(T+NV NV+W)x(,M)=,M)+(+/NV+BETA)*2
[25] +L4
LBL3:A<-/(NV+BETA)*2
L4:V<-THETA
[27] THETA<-A+M
'THETA ESTIMATE',THETA
[29] +(0.0001<|THETA-V)/LBL1
+(ML=0)/L5
[31] 'SE OF BETA ESTIMATE',(+/NV NV+W)x(,NV)=,NV)*0.5
[32] A<-M+((+/-/(T+.xT)x(,M)=,M)+THETA*2)-(2*x+/-/T*(,M)=,M)+THETA
[33] 'SE OF THETA',((2*THETA*2)+A)*0.5
[34] 'ESTIMATION OF FRAILTY'
[35] +(M 1;,M),M 1;NV+BETA
[36] L5:'THE END'

v

GROWTH1

=====

▼ BETA←GROWTH1 X;M;N;NV;MA;THETA;RHO;ML;L;I;A;W;V
[1] □FITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2] □OF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[3] □FOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES, TIME TO
[4] □FAILURE (MEASURES FROM THE FIRST FAILURE EPISODE). COX MODEL IS
[5] □FITTED WITH AN INCLUSION OF A RANDOM FRAILTY TERM FOR EACH PATIENT.
[6] □THE LONGITUDINAL PARAMETER IS ASSUMED TO BE FIXED.

M←[/X[,1]
N←+/1+pX
NV←+/^3+^-1+pX
MA←[/+/(,1M) . . = ,1M
X←X,X[,1] . . = ,1M
BETA←(+/^3+1+pX)p0
THETA←0.41
ML←[],0p[]←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
X←X,[AX[,2];]
LBL1:L←X LL02(0 3+X)+.×BETA
I←(,1M) . . = ,1M
A←((M+NV)NVp0),(NV Mp0),[1]⊗(I×THETA)
W←⊗((,0 3+X)+.×(0 1+L)+.×0 3+X)+A
BETA←(V-BETA)+W+.×((,0 3+X)+.×L[,1])-A+.×BETA
'BETA ESTIMATE',6 RND NV+BETA
+(0.01z[/|V-BETA)/LBL1
+(ML=0)/LBL3
+(ML≠1)/LBL2
W←((NV NV+W),(NV Mp0)),[1](M NVp0),⊗NV NV+⊗W
LBL2:A←(+/+/(T+NV NV+W)×(,1M) . . = ,1M)+(+/NV+BETA)*2
→L4
LBL3:A←+/(NV+BETA)*2
L4:V←THETA
THETA←A+M
'THETA ESTIMATE',THETA
+(0.0001<|THETA-V)/LBL1
+(ML=0)/LS
'SE OF BETA ESTIMATE',(+/NV NV+W)×(,1NV) . . = ,1NV)*0.5
A←M+((+/+/(T+.×T)×(,1M) . . = ,1M)+THETA*2)-(2×+/+/(T×(,1M) . . = ,1M)+THETA
'SE OF THETA',((2×THETA*2)+A)*0.5
'ESTIMATION OF FRAILTY'
(M 1,pM),M 1,pNV+BETA
LS:'THE END'

GROWTH2

```

v GROWTH2 X;M;N;NV;BETA;TA1;TA2;ML;L;A;W;A1;A2
[1]   aFITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2]   aOF MATRIX X AS PATIENT NUMBER (NUMBER CONSECUTIVELY STARTING AT 1)
[3]   aFOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES, TIME TO
[4]   aFAILURE (MEASURES FROM THE FIRST FAILURE EPISODE). COX MODEL IS
[5]   aFITTED WITH AN INCLUSION OF A RANDOM FRAILTY TERM FOR EACH PATIENT.
[6]   aTHE LONGITUDINAL PARAMETER IS ASSUMED TO BE RANDOM.
[7]   M←[X[,1]
[8]   N←+/1+pX
[9]   NV←+/^3+^-1+pX
[10]  X←X,W,(X[,;(NV+3)]..×Mp0)×W+X[,1]=:M
[11]  BETA←(+/-3+1+pX)p0
[12]  TA1←0.2198
[13]  TA2←0.05
[14]  ML←[],0p[]←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[15]  X←X[4X[,2];]
[16]  LBL1:L←X LL02(0 3+X)+.×BETA
[17]  A←(((((1M)=:1M)+TA1),M Mp0),[1](M Mp0),((1M)=:1M)+TA2
[18]  A←(((2×M)+NV)NVp0),(NV(2×M)p0),[1]A
[19]  W←@((@0 3+X)+.×(0 1+L)+.×0 3+X)+A
[20]  BETA←(V←BETA)+W+.×((@0 3+X)+.×L[,1])-A+.×BETA
[21]  'BETA ESTIMATE',6 RND NV+BETA
[22]  +(0.01s/[|V-BETA])/LBL1
[23]  +(ML=0)/LBL3
[24]  +(ML=1)/LBL2
[25]  W←((NV NV+W),(NV(2×M)p0)),[1]((2×M)NVp0),@NV NV+@W
[26]  LBL2:
[27]  A1←(+/+/(M M+NV NV+W)×(1M)=:1M)+(+/(M+NV+BETA)*2)
[28]  A2←(+/+/(M M+NV NV+W)×(1M)=:1M)+(+/(M+NV+BETA)*2)
[29]  +L4
[30]  LBL3:A1←+/(M+NV+BETA)*2
[31]  A2←+/(M+NV+BETA)*2
[32]  L4:L←TA1
[33]  A←TA2
[34]  TA1←A1+M
[35]  TA2←A2+M
[36]  'THETA1 ESTIMATE',TA1
[37]  'THETA2 ESTIMATE',TA2
[38]  +(0.0001<|TA1-L)∨(0.0001<|TA2-A))/LBL1
[39]  +(ML=0)/L5
[40]  'SE OF BETA ESTIMATE',(+/(NV NV+W)×(1NV)=:1NV)*0.5
[41]  LBL4:L←M M+NV NV+W
[42]  A←M M+NV NV+W
[43]  W←M(-M)+NV NV+W
[44]  L←(M+((+/-/(L+.*L)×(1M)=:1M)+TA1*2)-(2×+/-/L×(1M)=:1M)+TA1)+(2×TA1*2)
[45]  A←(M+((+/-/(A+.*A)×(1M)=:1M)+TA2*2)-(2×+/-/A×(1M)=:1M)+TA2)+(2×TA2*2)
[46]  W←(+/-/(W+.*W)×(1M)=:1M)+(2×(TA1*2)×(TA2*2))
[47]  'SE OF THETA1',(A÷(A×L)-W*2)*0.5
[48]  'SE OF THETA2',(L÷(A×L)-W*2)*0.5
[49]  L5:'THE END'

```

GROWTH3

=====

```

v GROWTH3 X;M;N;V;NV;I;BETA;TA1;TA2;ML;L;A;W;A1;A2
[1]   a FITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2]   a OF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[3]   a FOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES, TIME TO
[4]   a FAILURE (MEASURES FROM THE FIRST FAILURE EPISODE). COX MODEL IS
[5]   a FITTED WITH AN INCLUSION OF A RANDOM FRAILTY TERM FOR EACH PATIENT.
[6]   a THE LONGITUDINAL PARAMETER IS ASSUMED TO BE RANDOM AND RELATES TO THE
[7]   a PATIENT CHARACTERISTICS.
[8]   M-[/X[,1]
[9]   N-+/-1+pX
[10]  V-+/-4+-1+pX
[11]  NV-@0 -1+0 3+X
[12]  L-+/(+(,M),=X[,1]
[13]  A-V 1@NV[,L[1]]
[14]  I-2
[15]  L1:A-A,NV[,L[I]]
[16]  I-I+1
[17]  -(I,M)/L1
[18]  W-W,V-(X[,,(V+4)]*M@1)*W-X[,]=,M
[19]  X-X,V+.*@A
[20]  NV-+/-3+-1+pX
[21]  X-X,W
[22]  BETA-+/-3+1+pX)@0
[23]  TA1-0.225
[24]  TA2-0.043
[25]  ML-0,0@[]-0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[26]  X-X[,@X[,2];]
[27]  LBL1:L-X LL02(0 3+X)+.*BETA
[28]  A-(((,(,M),=,(M)+TA1),M M@0),[1](M M@0),((,(M),=,(M)+TA2
[29]  A-(((2*M)+NV)NV@0),(NV(2*M)@0),[1]A
[30]  W-0
[31]  W-@((@0 3+X)+.*(0 1+L)+.*0 3+X)+A
[32]  BETA-(V-BETA)+W+.*((@0 3+X)+.*L[,1])-A+.*BETA
[33]  'BETA ESTIMATE',6 RND NV+BETA
[34]  A-0
[35]  L-0
[36]  +(0.015/[|V-BETA])/LBL1
[37]  +(ML=0)/LBL3
[38]  +(ML#1)/LBL2
[39]  W-((NV NV+W),(NV(2*M)@0)),[1]((2*M)NV@0),@NV NV+@W
[40]  LBL2:A1-+(/+/(M M+NV NV+W)*(,(M),=,(M)+(+/(M+NV+BETA)*2
[41]  A2-+(/+/(M M+NV NV+W)*(,(M),=,(M)+(+/(M+NV+BETA)*2
[42]  ->L4
[43]  LBL3:A1-+/(M+NV+BETA)*2
[44]  A2-+/(M+NV+BETA)*2
[45]  L4:L-TA1
[46]  V-TA2
[47]  TA1-A1+M
[48]  TA2-A2+M
[49]  'THETAT1 ESTIMATE',TA1
[50]  'THETAT2 ESTIMATE',TA2
[51]  +((0.0001<|TA1-L)<(0.0001<|TA2-V))/LBL1
[52]  +(ML=0)/L5
[53]  'SE OF BETA ESTIMATE',(+/(NV NV+W)*(,(NV),=,(NV))*0.5
[54]  L-M M+NV NV+W
[55]  A-M M+NV NV+W
[56]  W-M(-M)+NV NV+W
[57]  L-((M+((/+/(L+,*L)*(,(M),=,(M)+TA1*2)-(2*+/*L*(,(M),=,(M)+TA1)+(2*TA1*2
[58]  A-((M+((/+/(A+,*A)*(,(M),=,(M)+TA2*2)-(2*+/*A*(,(M),=,(M)+TA2)+(2*TA2*2
[59]  W-((/+/(W+,*@W)*(,(M),=,(M)+((2*(TA1*2)*(TA2*2))
[60]  'SE OF THETAT1',(A+(A*L)-W*2)*0.5
[61]  'SE OF THETAT2',(L+(A*L)-W*2)*0.5
[62]  L5:'THE END'
```

v

▼ AR X;M;N;NV;MA;BETA;THETA;RHO;ML;L;A;V;U2;U3;U4;C1;C2;C3;C4
 [1] *FITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
 [2] *OF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
 [3] *FOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES, TIME TO
 [4] *FAILURE (MEASURES FROM THE FIRST FAILURE EPISODE). COX MODEL IS
 [5] *FITTED WITH AN INCLUSION OF A RANDOM FRAILTY TERM AT EACH FAILURE.
 [6] *SUCH RANDOM FRAILTY TERM IS ASSUMED TO FOLLOW AN AR(1) PROCESS.
 [7] M←[X[,1]
 [8] N←+/1+pX
 [9] NV←+/^3+1+pX
 [10] MA←+//(,M)=X[,1]
 [11] X←X,(,N)=,N
 [12] BETA←(+/-3+1+pX)p0
 [13] THETA←1
 [14] RHO←0
 [15] ML←[],Op[]←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
 [16] X←X[,2];
 [17] LBL1:L←X LL02(0 3+X)+.xBETA
 [18] A←(,N)=,N
 [19] V←1
 [20] L3:W←(RHO*V)×(bW)+W←(X[,1]=X[,1])×(V+,N)=,N
 [21] A←A+W
 [22] V←V+1
 [23] +(V<MA)/L3
 [24] A←((N+NV)NVp0),(NV Np0),[1]@(A×THETA+1-RHO*2)
 [25] W←@((,0 3+X)+.x(0 1+L)+.x0 3+X)+A
 [26] BETA←(V-BETA)+W+.x((,0 3+X)+.xL[,1])-A+.xBETA
 [27] 'BETA ESTIMATE',6 RND NV+BETA
 [28] +(0.01<|V-BETA)/LBL1
 [29] +(ML=0)/LBL3
 [30] +(ML=1)/LBL2
 [31] W←((NV NV+W),(NV Np0)),[1](N NVp0),@NV NV+@W
 [32] LBL2:A←(NV NV+W)+(NV+BETA)+.xNV+BETA
 [33] +L4
 [34] LBL3:A←(NV+BETA)+.xNV+BETA
 [35] L4:U2←+/A×(,N)=,N
 [36] U3←+/+A×(X[,1]=X[,1])×((,N)=,1+;N)
 [37] V←+//(,M)=X[,1]
 [38] L←1,-1+1+V
 [39] U4←+/+(A[V;V]+A[L;L])×(,M)=,M
 [40] A←THETA
 [41] THETA←((U2×1+RHO*2)-(2×U3×RHO)+U4×RHO*2)+N
 [42] 'THETA ESTIMATE',THETA
 [43] C1←(N-M)×U2-U4
 [44] C2←((2×M)-N)×U3
 [45] C3←(N×U4)-(N+M)×U2
 [46] C4←N×U3
 [47] L←RHO
 [48] U3←(C1×RHO*3)+(C2×RHO*2)+(C3×RHO)+C4
 [49] U4←(3×C1×RHO*2)+(2×C2×RHO)+C3
 [50] RHO←L-U3+U4
 [51] 'RHO ESTIMATE',RHO
 [52] +(0.01<|THETA-A)v(0.01<|RHO-L))/LBL1
 [53] +(ML=0)/L5
 [54] L←X LL02(0 3+X)+.xBETA
 [55] A←(,N)=,N
 [56] V←1
 [57] L6:W←(RHO*V)×(bW)+W←(X[,1]=X[,1])×(V+,N)=,N
 [58] A←A+W
 [59] V←V+1
 [60] +(V<MA)/L6
 [61] A←A+1-RHO*2
 [62] L←@((,0 1+L)+THETA×(0(3+NV)+X)+.xA+.xb0(3+NV)+X)
 [63] W←N NV+0 3+X
 [64] 'SE OF BETA ESTIMATE',(+/(@(bW)+.xL+.xW)×(,N)=,N)*0.5
 [65] +(ML=1)/L7
 [66] L←L-L+.xW+.x(@(bW)+.xL+.xW)+.x(bW)+.xL
 [67] L7:W←L+.x(0(3+NV)+X)+.xA+.xb0(3+NV)+X
 [68] U2←2×RHO×(,N)=,N
 [69] U3←(bU3)+U3-((,N)=,1+;N)×(X[,1]=X[,1])
 [70] C1←+//(,M)=X[,1]
 [71] U4←2×RHO×(bU4)+.xU4-C1=,N
 [72] C1←1,-1+1+C1
 [73] U4←U4+2×RHO×(bC1)+.xC1-C1=,N
 [74] V←U2-U3-U4
 [75] V←-THETA×L+.x(0(3+NV)+X)+.xA+.xV+.xA+.xb0(3+NV)+X
 [76] U2←(0.5)x+/.+(W+.xW)×(,N)=,N
 [77] U3←(0.5)x+/.+(W+.xV)×(,N)=,N
 [78] U4←(0.5)x+/.+(V+.xV)×(,N)=,N
 [79] 'SE OF THETA AND RHO',((U4+C1),U2+C1-(U2×U4)-U3*2)*0.5
 [80] L5:'ESTIMATE OF FRAILTY',NV+BETA

▼

GENGROWTH

```

v GENGROWTH X;M;N;NV;MA;BETA;THETA;RHO;ML;L;I;A;W;V;K;K1
[1]   aFITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2]   aOF MATRIX X AS CLUSTER NUMBER (NUMBER CONSECUTIVELY STARTING AT 1)
[3]   aFOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES. A FAMILY
[4]   aOF RELATIVE RISK FUNCTIONS IS CONSIDERED IN WHICH COX MODEL IS
[5]   aINCLUDED AS A SPECIAL CASE WHEN K TENDS TO 0. K IS CHOSEN BY
[6]   aMAXIMIZING THE BLUP LIKELIHOOD.
[7]   M←Γ/X[,1]
[8]   N←+/-1+pX
[9]   NV←+/-3+-1+pX
[10]  MA←Γ/+/(,M)←X[,1]
[11]  X←X,X[,1]←,M
[12]  K1←[],0,p0←'CHOOSE THE STARTING VALUE OF K FROM 0 TO 1, 0 FOR LOGLINEAR, 1 FOR LINEAR'
[13]  BETA←(+/-3+1+pX)p0
[14]  THETA←0.43
[15]  ML←[],0,p0←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[16]  X←X[,X[,2],]
[17]  LBL1:K←K1
[18]  →(K=0)/L6
[19]  L←X LL03(0 3+X)+.,xBETA
[20]  →L7
[21]  L6:L←X LL02(0 3+X)+.,xBETA
[22]  L7:I←(,M)←,M
[23]  A←((M+NV)NVp0),(NV Mp0),[1]⊗(I×THETA)
[24]  W←⊗((,M 3+X)+.,x(0 1+L)+.,x0 3+X)+A
[25]  BETA←(V+BETA)+W+.x((,M 3+X)+.,xL[,1])-A+.xBETA
[26]  'BETA ESTIMATE',6 RND NV+BETA
[27]  →(0.01<Γ/|V-BETA|)/LBL1
[28]  K1←L01(0 2+N 3+X),(0 3+X)+.,xBETA
[29]  →(K1>0.1)/LBL8
[30]  aK1←0
[31]  LBL8:'ESTIMATE OF K',K1
[32]  →(K=K1)/LBL1
[33]  →(ML=0)/LBL3
[34]  →(ML=1)/LBL2
[35]  W←((NV NV+W),(NV Mp0)),[1](M NVp0),⊗NV NV+⊗W
[36]  LBL2:A←(+/+/(T-NV NV+W)×(,M)←,M)+(+/((NV+BETA)*2)
[37]  →L4
[38]  LBL3:A++/(NV+BETA)*2
[39]  L4:V←THETA
[40]  THETA←A+M
[41]  'THETA ESTIMATE',THETA
[42]  →(0.0001<|THETA-V|)/LBL1
[43]  →(ML=0)/L5
[44]  'SE OF BETA ESTIMATE',(+/(NV NV+W)×(,M)←,M)×(,M)←,NV)*0.5
[45]  A←M+(+/+(T+.,xT)×(,M)←,M)+THETA*2)-(2×+/+T×(,M)←,M)+THETA
[46]  'SE OF THETA',((2×THETA*2)+A)*0.5
[47]  L5:'ESTIMATE OF FRAILTY',NV+BETA

```

RANBLKB

=====

v RANBLKB X;BETA;M;N;NV;TA1;TA2;ML;L;A;W;A1;A2;V
[1] ^FITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2] ^OF MATRIX X AS CLUSTER NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[3] ^FOLLOWED BY TIME TO EVENT, FAILURE/CENSOR, RISK VARIABLES. COX MODEL
[4] ^IS FITTED WITH AN INCLUSION OF A RANDOM CLUSTER EFFECT AS WELL AS AN
[5] ^RANDOM INDIVIDUAL EFFECT.
[6] M←r/X[,1]
[7] N←+/-1+pX
[8] NV←+/-3+-1+pX
[9] X←X,((,:N)=,:N),X[,1]=,:M
[10] BETA←(+/-3+1+pX)p0
[11] TA1←0.1041
[12] TA2←0.4458
[13] ML←[],0,p[]←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[14] X←X[,X[,2];]
[15] LBL1:L←X LL02(0 3+X)+.x*BETA
[16] A←(((N,:N)=,:N)+TA1),N Mp0,[1](M Np0),((,:M)=,:M)+TA2
[17] A←((N+M+NV)NVp0),(NV(N+M)p0),[1]A
[18] W←((N+M+NV)+.x(0 1+L)+.x(0 3+X))+A
[19] BETA←((V-BETA)+W+.x((N+M+NV)+.xL[,1]))-A+.x*BETA
[20] 'BETA ESTIMATE',6 RND NV+BETA
[21] →(0.01<|V-BETA|)/LBL1
[22] →(ML=0)/LBL3
[23] →(ML=1)/LBL2
[24] W←((NV NV+W),(NV(N+M)p0)),[1]((N+M)NVp0),BNV NV+BW
[25] LBL2:
[26] A1←(+/+/(N N+NV NV+W)x(:N)=,:N)+(+/(N+NV+BETA)*2)
[27] A2←(+/+/(N N+NV NV+W)x(:M)=,:M)+(+/(N+NV+BETA)*2)
→L4
[28] LBL3:A1←+/(N+NV+BETA)*2
[29] A2←+/(N+NV+BETA)*2
[30] L4:L+TA1
[31] V+TA2
[32] TA1←A1+N
[33] TA2←A2+M
[34] 'THETA1 ESTIMATE',TA1
[35] 'THETA2 ESTIMATE',TA2
[36] →((0.0001<|TA1-L)∨(0.0001<|TA2-V))/LBL1
[37] →(ML=0)/L5
[38] 'SE OF BETA ESTIMATE',(+/(NV NV+W)x(:N)=,:N)*0.5
[39] L←N N+NV NV+W
[40] A←N N+NV NV+W
[41] W←N(-M)+NV NV+W
[42] L←N(-M)+NV NV+W
[43] L←(N+((+/-/(L+X*L)x(:N)=,:N)+TA1*2)-(2x+/-/Lx(:N)=,:N)+TA1)+(2xTA1*2)
[44] A←(M+((+/-/(A+X*A)x(:M)=,:M)+TA2*2)-(2x+/-/Ax(:M)=,:M)+TA2)+(2xTA2*2)
[45] W←((+/-/(W+X*W)x(:M)=,:M)+(2x(TA1*2)x(TA2*2)))
[46] 'SE OF THETA1',(A+(A*L)-W*2)*0.5
[47] 'SE OF THETA2',(L+(A*L)-W*2)*0.5
[48] L5:'THE END'

v

SIMBASELINE

```
=====
  v PR SIMBASELINE MVB;I;OUT;FR;W;C;V;M;U;X;K;EST;SE;SQE;SS;J
[1]  a RUNNING SIMULATIONS OF THE BASELINE FRAILTY MODEL
[2]  J+1
[3]  LL:I++/1+pPR
[4]  OUT+PR
[5]  'SIMULATION (ML)'
[6]  LOOP:FR=30 NORMALR(2+MVB[;J])
[7]  W=90 0p0
[8]  C=0
[9]  L1:V+=(-1+?30p2)*.x3p1
[10]  +(0=[/V]/L1
[11]  +(1=[/V])/L1
[12]  W=W,V
[13]  C=C+1
[14]  +(C#+/p2+MVB[;J])/L1
[15]  M=W+.x2+MVB[;J]
[16]  U=(-0.000001*790p999999)+0.1*x*M+,FR*.x3p1
[17]  X+=(.130)*.x3p1,U,(90p1),W
[18]  FR=((MVB[;J])[2]),FR
[19]  K=FR BASELINE X
[20]  +(0z+/-1+((+/^-2+pMVVB[;J])+K))/LL
[21]  +(0z+/-1+((+/^-1+pMVVB[;J])+K))/LL
[22]  OUT=OUT,K
[23]  I+=/1+pOUT
[24]  'NO. OF SIMULATION : ',I
[25]  +(I#100)/LOOP
[26]  EST=(+/-pMVB[;J])0+OUT
[27]  SE=(+/-1+pMVB[;J])0+(-1 0+OUT)
[28]  SQE=-1 I+OUT
[29]  a 'PARAMETER ESTIMATE'
[30]  a EST
[31]  a 'STANDARD ERROR'
[32]  a SE
[33]  a 'SQUARE ERROR IN FRAILTY'
[34]  a SQE
[35]  'ASSIGNED BETA',2+MVB[;J]
[36]  'ASSIGNED THETA',(MVB[;J])[2]
[37]  'AVERAGE BIAS OF ESTIMATE',((+/EST)+I)-E+(2+MVB[;J]),((MVB[;J])[2])
[38]  'AVERAGE OF SE OF ESTIMATE',(+/SE)+I
[39]  'SE OF ESTIMATE OVER SIMULATION',SS-((+/(EST-((+/EST)+I)*.xI*p1)*2)+I)*0.5
[40]  'SE IN BRACKET',SS+I*0.5
[41]  'MEAN SQUARE ERROR IN FRAILTY',(+/SQE)+I
[42]  J<-J+1
[43]  +(J<+/-1+pMVB)/LL
v
```

SIMRANBLK3

```
    v PR SIMRANBLK3 MVVB;I;OUT;FR;W;C;V;M;U;X;K;EST;SE;SQE;SS;J
[1]  a RUNNING SIMULATIONS OF THE RANDOM BLOCK FRAILTY MODEL
[2]  J<-1
[3]  LL:I<-+/1+pPR
[4]  OUT<-PR
[5]  'SIMULATION (REML)'
[6]  LOOP:FR<-(90 NORMALR(2+MVVB[,J]))+,(30 NORMALR(0,(MVVB[,J])[3]))..x3p1
[7]  W<-90 0p0
[8]  C<-0
[9]  L1:V<,-1+?30p2)..x3p1
[10]  +(0=I/V)/L1
[11]  +(1=L/V)/L1
[12]  W<-W,V
[13]  C<-C+1
[14]  +(C*+/p3+MVVB[,J])/L1
[15]  M<-W+.x3+MVVB[,J]
[16]  U<-(-0.000001x790p999999)+0.1*x*M+FR
[17]  X<-,(:30)..x3p1,U,(90p1),W
[18]  FR<-(MVVB[,J])[2],((MVVB[,J])[3]),FR
[19]  K<-FR RANBLK3 X
[20]  +(0z+/-1+((+/-2+pMVVB[,J])+K))/LL
[21]  +(0z+/-1+((+/-1+pMVVB[,J])+K))/LL
[22]  OUT<-OUT,K
[23]  I<-+/1+pOUT
[24]  'NO. OF SIMULATION : ',I
[25]  +(I*100)/LOOP
[26]  EST<-(+/-pMVVB[,J])0+OUT
[27]  SE<-(+/-1+pMVVB[,J])0+(-1 0+OUT)
[28]  SQE<-1 I+OUT
[29]  'PARAMETER ESTIMATE'
[30]  EST
[31]  a 'STANDARD ERROR'
[32]  a SE
[33]  a 'SQUARE ERROR IN FRAILTY'
[34]  a SQE
[35]  'ASSIGNED BETA',3+MVVB[,J]
[36]  'ASSIGNED THETAI',(MVVB[,J])[2]
[37]  'ASSIGNED THETA2',(MVVB[,J])[3]
[38]  'AVERAGE BIAS OF ESTIMATE',((+/EST)+I)-E<-(3+MVVB[,J]).((MVVB[,J])[2]).(MVVB[,J])[3]
[39]  'AVERAGE OF SE OF ESTIMATE',(+/SE)+I
[40]  'SE OF ESTIMATE OVER SIMULATION',SS<-(+/(EST-((+/EST)+I)..xIp1)*2)+I)*0.5
[41]  'SE IN BRACKET',SS+I*0.5
[42]  'MEAN SQUARE ERROR IN FRAILTY',(+/SQE)+I
[43]  J<-J+1
[44]  +(Js+/-1+pMVVB)/LL
```

COMRSK1

```

[1]   v U1~COMRSK1 X;M;N;NV;BETA;TA1;ML;Y;Z;X1;X2;XX;L1;L2;L;A;W;V;A1;TT;Y1;Y2
[2]   aFITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[3]   aOF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[4]   aFOLLOWED BY TIME TO EVENT, HOSPITALISED/CENSOR, DEATH/CENSOR, RISK
[5]   aVARIABLES. COX MODEL IS FITTED WITH AN INCLUSION OF A RANDOM FRAILTY
[6]   aTERM FOR EACH PATIENT. SUCH RANDOM FRAILTY TERM FOR EACH PATIENT IS
[7]   aASSUMED TO BE EQUAL FOR HOSPITALISATION AND DEATH.
[8]   M←[ /X[ ;1]
[9]   N←+/1+pX
[10]  NV←+/^4+^-1+pX
[11]  ML←0,0p0'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[12]  TT←PROG4 N 2+X
[13]  Y←N 4+X
[14]  Y1←N 3+Y
[15]  Y2←TT,N 1+pY[ ;4]
[16]  X←0 4+X
[17]  Z←Y[ ;1]..=1M
[18]  BETA←(M+2*NV)p0
[19]  TA1←0.51
[20]  X2←(N NVp0),X,Z
[21]  X2←X2[ ^Y2[ ;2]; ]
[22]  X1←X,(N NVp0),Z
[23]  X1←X1[ ^Y1[ ;2]; ]
[24]  XX←X1,[1]X2
[25]  Y1←Y1[ ^Y1[ ;2]; ]
[26]  Y2←Y2[ ^Y2[ ;2]; ]
[27]  LBL1:L1←(Y1,X1)LL02 X1+.×BETA
[28]  L2←(Y2,X2)LL02 X2+.×BETA
[29]  L←((2×N)1p(L1[ ;1],L2[ ;1])),((0 1+L1),N Np0),[1](N Np0),0 1+L2
[30]  A←((1M)..=1M)+TA1
[31]  A←((M+2*NV)(2×NV)p0),((2×NV)Mp0),[1]A
[32]  W←((bXX)+..×(0 1+L)+..×XX)+A
[33]  BETA←(V-BETA)+W+.×((bXX)+..×L[ ;1])-A+.×BETA
[34]  'BETA ESTIMATE',6 RND(2×NV)+BETA
[35]  →(0.01<|V-BETA|)/LBL1
[36]  →(ML=0)/LBL3
[37]  →(ML≠1)/LBL2
[38]  W←(((2×NV)(2×NV)+W),((2×NV)Mp0)),[1](M(2×NV)p0),((2×NV)(2×NV)+W)
[39]  LBL2:
[40]  A1←(+/+/(M M+(2×NV)(2×NV)+W)×(1M)..=1M)+(+/(M+(2×NV)+BETA)*2)
[41]  →L4
[42]  LBL3:A1←+/(M+(2×NV)+BETA)*2
[43]  L4:L←TA1
[44]  TA1←A1+M
[45]  'THETTA1 ESTIMATE',TA1
[46]  →(0.01<|TA1-L|)/LBL1
[47]  →(ML=0)/L5
[48]  'SE OF BETA ESTIMATE',(+/((2×NV)(2×NV)+W)×(12×NV)..=1(2×NV))*0.5
[49]  L←M M+(2×NV)(2×NV)+W
[50]  L←M+((+/+(L+.×L)×(1M)..=1M)÷TA1*2)-(2×+/+L×(1M)..=1M)÷TA1
[51]  'SE OF THETTA1',((2×TA1*2)+L)*0.5
[52]  L5:U1←4 RND M+(2×NV)+BETA

```

v

COMRSK2

```

v U1+COMRSK2 X;M;N;NV;BETA;TA1;TA2;GAMMA;DG;ML;Y;Z;X1;X2;XX;L1;L2;L;A;W;V;A1;A2;U2;TT;
    Y1;Y2;YY1
[1]   aFITS A GENERALISED MIXED MODEL TO SURVIVAL DATA WITH COLUMNS
[2]   aOF MATRIX X AS PATIENT NUMBER (NUMBERED CONSECUTIVELY STARTING AT 1)
[3]   aFOLLOWED BY TIME TO EVENT, HOSPITALISED/CENSOR, DEATH/CENSOR, RISK
[4]   aVARIABLES. COX MODEL IS FITTED WITH AN INCLUSION OF A RANDOM FRAILTY
[5]   aTERM FOR EACH PATIENT. THE FRAILTY OF HOSPITALISATION IS RELATED TO
[6]   aTHE FRAILTY OF DEATH BY A PARAMETER GAMMA FOR EACH PATIENT.
[7]   M←[ /X[ ;1]
[8]   N←+/1+pX
[9]   NV←+/^4+-1+pX
[10]  ML←[],0,p[]←'0 FOR BLUP, 1 FOR ML, 2 FOR REML'
[11]  TT←PROG4 N 2+X
[12]  Y←N 4+X
[13]  YY1←N 3+Y
[14]  Y2←TT,N 1+pY[ ;4]
[15]  X←0 4+X
[16]  Z←Y[ ;1]◦.=1M
[17]  BETA←((2×NV)+M)◦0
[18]  TA1←0.6
[19]  GAMMA←1.5
[20]  X2←(N NV◦0),X,Z
[21]  X2←X2[ ^Y2[ ;2];]
[22]  LBL4:X1←X,(N NV◦0),(GAMMA×Z)
[23]  X1←X1[ ^YY1[ ;2];]
[24]  XX←X1,[1]X2
[25]  Y1←YY1[ ^YY1[ ;2];]
[26]  Y2←Y2[ ^Y2[ ;2];]
[27]  LBL1:L1←(Y1,X1)LBL02 X1+.×BETA
[28]  L2←(Y2,X2)LBL02 X2+.×BETA
[29]  L←((2×N)1◦(L1[ ;1],L2[ ;1])),((0 1+L1),N NV◦0),[1](N NV◦0),0 1+L2
[30]  A←((1M)◦.=1M)×(1+GAMMA×2)+TA1
[31]  A←(((2×NV)+M)(2×NV)◦0),((2×NV)M◦0),[1]A
[32]  W←B((4XX)+.×(0 1+L)+.×XX)+A
[33]  BETA←(V+BETA)+W+.×((4XX)+.×L[ ;1])-A+.×BETA
[34]  'BETA ESTIMATE',4 RND(2×NV)+BETA
[35]  +(0.01s[ /|V-BETA])/LBL1
[36]  +(ML=0)/LBL3
[37]  +(ML≠1)/LBL2
[38]  W←(((2×NV)(2×NV)+W),((2×NV)M◦0)),[1](M(2×NV)◦0),B(2×NV)(2×NV)+B
[39]  LBL2:
[40]  A1←(1+GAMMA×2)×(+/+((2×NV)(2×NV)+W)×(1M)◦.=1M)+(+/((2×NV)+BETA)*2)
[41]  +L4
[42]  LBL3:A1←(1+GAMMA×2)×+/(((2×NV)+BETA)*2
[43]  L4:L←TA1
[44]  TA1←A1+M
[45]  'THETA ESTIMATE',4 RND TA1
[46]  +(0.01<|TA1-L)/LBL1
[47]  +(ML=0)/L5
[48]  'SE OF BETA ESTIMATE',4 RND(+/(((2×NV)(2×NV)+W)×(12×NV)◦.=12×NV)*0.5
[49]  L←(2×NV)(2×NV)+W
[50]  L-M+(+/+(L+.×L)×(1M)◦.=1M)×((1+GAMMA×2)+TA1)*2)-2×(+/+L×(1M)◦.=1M)×(1+GAMMA×2)+TA
    1
[51]  'SE OF THETA',4 RND((2×TA1*2)+L×(1+GAMMA×2)*2)*0.5
[52]  L5:V=GAMMA
[53]  GAMMA←Y1[ ;3]PROG5 BETA
[54]  'GAMMA ESTIMATE',GAMMA
[55]  +(GAMMA×V)/LBL4
[56]  U1←4 RND(2×NV)+BETA

```

L01

====

```
v K1-L01 W;I;K;L1;L;L0
[1] K<-(t100)+100
[2] I<-1
[3] L0<+/W[;1]*W[;2]-•φ+\\φ*W[;2])
[4] L1<-2 1o(L0,0)
[5] LBL1:->(0>|/1+K[I]*W[;2])/LBL2
[6] L<+/W[;1]*((1+K[I])*W[;2])-•φ+\\φ(1+K[I]*W[;2])*1+K[I])
[7] L1<-L1,L,K[I]
[8] LBL2:I<-I+1
[9] -(I<100)/LBL1
[10] L<-L1[;L1[1;1]]
[11] 'NO. OF LIKELIHOOD EXIST',+/1+pL
[12] -(0=+/1+pL)/L3
[13] K1<+/1+L[;1]
[14] 'LIKELIHOOD',+/1+L[;1]
[15] -L4
[16] L3:K1<-0
[17] L4:I<-1
v
```

LL02

====

```
v L-X LL02 W;N;A;B;C;I
[1] B<+\A-X[;3]+φ+\\φW-+W
[2] L<(N,0)o(N-pB),(I-1),C<+\A*2
[3] LBL1:L<-L,(W×B×I=tN)-W×W[I]×(C×Iz1N)+C[I]×I<1N
[4] -(NzI-I+1)/LBL1
[5] L<(X[;3]-W×B),L
v
```

LL03

====

```
v L-X LL03 W;N;A;B;C;I;F;G
[1] F<1+1+K×W
[2] B<+\A-X[;3]+φ+\\φW-(1+K×W)*1+K
[3] G<F×W
[4] L<(N,0)o(N-pB),(I-1),C<+\A*2
[5] LBL1:L<-L,(((K×X[;3])+(1-K)×F×G×B)×I=tN)-G×G[I]×(C×Iz1N)+C[I]×I<1N
[6] -(NzI-I+1)/LBL1
[7] L<(F×X[;3]-W×B),L
v
```

NORMALR

```

      v R←N NORMALR PA;V
[1]   R←(?(2,[N+2)ρ2147483647)÷2147483647
[2]   R←N†(.1 2°.◦R[2;]×◦2)×V,V←(‐2×◦R[1;])×0.5
[3]   →(^/PA=0 1)/0
[4]   R←PA[1]+R×PA[2]×0.5
[5]   v

```

RND

```

      v R←N RND X
[1]   R←(10×N)×[0.5+(10×N)×X
v

```

BASELINE

```

      v K←FR BASELINE X;M;N;NV;MA;THETA;RHO;ML;L;I;A;W;V;CC
[1]   M←[;/X[;1]
[2]   N←+/1+ρX
[3]   NV←+/3+‐1+ρX
[4]   MA←[/+/(‐M)◦.=X[;1]
[5]   X←X,X[;1]◦.=‐M
[6]   BETA←(+/‐3+1+ρX)ρ0
[7]   THETA←FR[1]
[8]   CC←0
[9]   ML←1
[10]  X←X[ΔX[;2];]
[11]  LBL1:L←X LL02(0 3+X)+.×BETA
[12]  I←(‐M)◦.=‐M
[13]  A←((‐M+NV)NVρ0),(NV Mρ0),[1]⊕(I×THETA)
[14]  W←⊕((‐0 3+X)+.×(0 1+L)+.×0 3+X)+A
[15]  BETA←(V‐BETA)+W+.×((‐0 3+X)+.×L[;1])-A+.×BETA
[16]  →(25×CC+CC+1)/L7
[17]  →(0.01<[;/V‐BETA])/LBL1
[18]  →(ML=0)/LBL3
[19]  →(ML≠1)/LBL2
[20]  W←((NV NV+W),(NV Mρ0)),[1](M NVρ0),⊕NV NV+⊕W
[21]  LBL2:A←(+/+/(T+NV NV+W)×(‐M)◦.=‐M)+(+/NV+BETA)×2
[22]  →L4
[23]  LBL3:A←+/(NV+BETA)×2
[24]  L4:V←THETA
[25]  THETA←A+M
[26]  →(0.01<|THETA‐V)/LBL1
[27]  C2←(+/(NV NV+W)×(‐NV)◦.=‐NV)×0.5
[28]  A←M+((+/+/(T+×T)×(‐M)◦.=‐M)+THETA×2)-(2×+/+/T×(‐M)◦.=‐M)+THETA
[29]  C3←((2×THETA×2)+A)×0.5
[30]  C4←(+/((NV+BETA)‐1+FR)×2)+M
[31]  K←(NV+BETA),THETA,C2,C3,C4
[32]  →L8
[33]  L7:K←(3+2×NV)0ρ0
[34]  L8:K←K
v

```

RANBLK3

=====

```

    v K←FR RANBLK3 X;M;N;NV;BETA;TA1;TA2;ML;L;A;W;A1;A2;V;C2;C3;C4;CC
[1]   M←Γ/X[,1]
[2]   N←+/1+ρX
[3]   NV←+/3+1+ρX
[4]   X←X,((,:N)=,:N),X[,1]=,:M
[5]   BETA←(+/-3+1+ρX)ρ0
[6]   TA1←FR[1]
[7]   TA2←FR[2]
[8]   CC←0
[9]   ML←2
[10]  X←X[↑X[,2];]
[11]  LBL1:L←X LL02(0 3+X)+.×BETA
[12]  A←(((,:N)=,:N)+TA1),N Mρ0,[1](M Nρ0),((,:M)=,:M)+TA2
[13]  A←(N+M+NV)NVρ0,(NV(N+M)ρ0,[1]A
[14]  W←((,:N+3+X)+.×(0 1+L)+.×0 3+X)+A
[15]  BETA←(V←BETA)+W+.×((,:N+3+X)+.×L[,1])-A+.×BETA
[16]  →(5×CC←CC+1)/L7
[17]  →(0.01<Γ/|V-BETA)/LBL1
[18]  →(ML=0)/LBL3
[19]  →(ML=1)/LBL2
[20]  W←((NV NV+W),(NV(N+M)ρ0)),[1]((N+M)NVρ0),BNV NV+BW
[21]  LBL2:
[22]  A1←(+/+/(N N+NV NV+W)×(:N)=,:N)+(+/(N+NV+BETA)*2)
[23]  A2←(+/+/(N N+NV NV+W)×(:M)=,:M)+(+/(N+NV+BETA)*2)
[24]  →L4
[25]  LBL3:A1←+/(N+NV+BETA)*2
[26]  A2←+/(N+NV+BETA)*2
[27]  L←L-TA1
[28]  V←TA2
[29]  TA1←A1+N
[30]  TA2←A2+M
[31]  →((0.01<|TA1-L)∨(0.01<|TA2-V)))/LBL1
[32]  C2←(+/(NV NV+W)×(:NV)=,:NV)*0.5
[33]  L←N N+NV NV+W
[34]  A←N N+NV NV+W
[35]  W←N(-M)+NV NV+W
[36]  L←(N+((+/+/(L+.×L)×(:N)=,:N)+TA1*2)-(2×+/-L×(:N)=,:N)+TA1)+(2×TA1*2)
[37]  A←(M+((+/+/(A+.×A)×(:M)=,:M)+TA2*2)-(2×+/-A×(:M)=,:M)+TA2)+(2×TA2*2)
[38]  W←(+/+/(W+.×B)×(:M)=,:M)+(2×(TA1*2)×(TA2*2))
[39]  C3←((A+(A×L)-W*2)*0.5),((L+(A×L)-W*2)*0.5)
[40]  C4←(+/(((N+NV+BETA)+,(N+NV+BETA)×.×3ρ1)-2+FR)*2)÷N
[41]  K←(NV+BETA),TA1,TA2,C2,C3,C4
[42]  →L8
[43]  L7:K←(5+2×NV)0ρ0
[44]  L8: '
    v

```

PROG5

=====

```

    v C←A PROG5 D;L;W;H;B;I;K;G
[1]   G←200
[2]   W←(V-1),(V-1)+(:G)+100
[3]   I←1
[4]   K←0ρ0
[5]   LBL1:H←X,(N NVρ0),(W[I]×Z)
[6]   H←H[↑YY1[,2];]
[7]   B←H+.×D
[8]   L←(+/A×B-•φ+\\φ*B)+0.5×(M×•1+W[I]*2)-(+/(NV+BETA)*2)×(1+W[I]*2)+TA1
[9]   K←K,L
[10]  I←I+1
[11]  →(I×G+2)/LBL1
[12]  C←2(G+1)ρW,K
[13]  C←1+,2 1+C[,;G[2;]]
    v

```