A BAYESIAN ANALYSIS OF MULTIPLE INTERVAL-CENSORED FAILURE TIME EVENTS WITH APPLICATION TO AIDS DATA

LUCKY MOKGATLHE

A BAYESIAN ANALYSIS OF MULTIPLE INTERVAL-CENSORED FAILURE TIME EVENTS WITH APPLICATION TO AIDS DATA

By

LUCKY MOKGATLHE

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR

THE

DEGREE OF

DOCTOR OF PHILOSOPHY

in

MATHEMATICAL STATISTICS

THE FACULTY OF AGRICULTURE AND NATURAL SCIENCES

DEPARTMENT OF MATHEMATICAL STATISTICS

at the

UNIVERSITY OF FREE STATE

BLOEMFONTEIN

SOUTH AFRICA

MAY 2003

PROMOTER: PROFESSOR PIET C. N. GROENEWALD

THE DEPARTMENT OF MATHEMATICAL STATISTICS

UNIVERSITY OF FREE STATE

BLOEMFONTEIN, SOUTH AFRICA

CO-PROMOTER: PROFESSOR DANIEL J. DE WAAL

THE DEPARTMENT OF MATHEMATICAL STATISTICS

UNIVERSITY OF FREE STATE

BLOEMFONTEIN, SOUTH AFRICA

In memory of Grandma

Tite Nanesi Ndome

ACKNOWLEDGEMENT

Along the labyrinth that one treads in quest for knowledge, there are some amazing encounters whose impact on one's life is permanently engraved. I am immensely indebted to the following whose different contributions made my academic pursuit a success.

My profound thanks goes to Professor Piet C.N. Groenewald my promoter and mentor for his guidance and support throughout the entire duration of my research. His dedication and enthusiasm has been very inspiring to say the least. May God bless you Prof. Groenewald.

Special thanks goes to Professor Daniel J. De Waal, my co-promoter for his invaluable advice and comments. SANPAD for co-sponsoring my research through workshop presentations, Dr Pieter Van Gelder (SANPAD project leader – Delft University of Technology, Netherlands) for his encouraging discussions at the SANPAD workshops I participated at. Dr Kim Soyeon (Biostatistician at Botswana-Harvard Partnership) and Dr. Ronald Geskus (Biostastistician at Municipal Health Service in Amsterdam) for availing the main ACTG175 and Tri-Continental AIDS data sets respectively used in this research. I thank all staff members of Department of Mathematical Statistics (UFS) who made my studying at University of Free State worthwhile, especially Mrs Merriman, Carin, Delson, Martin and Prof. M. Finkelstein. To all UFS International Office staff, especially Hannalie Barnard, a big thank you guys.

To my lovely wife, Esther for her patient endurance of my long absence from home and words of encouragement each time things appeared to be going haywire. "The greatest gift one can get is education" are some of the words from my Mother to me once. I am very grateful to you Mom because today I realized that gift through your unrelenting support. To my two beautiful daughters Lerato and Thulisiwe, I say 'Daddy has finally made it'. To my brothers, sisters and Nkoko Nene, a big thank you for the support you offered my family while I was away.

In my long student life, there are teachers (the list is enormous) whose contribution towards the shaping of my academic destiny is amazing, those I wish to thank. Finally to all my friends while I was studying in Bloemfontein especially John, Andreas, Thembele (Ph.D) and Zawade, I say thank you guys, the hours we spent arguing over what may appeared to be trivial to you, meant a lot to me.

Lucky Mokgatlhe

CONTENTS Page

CHAPTER 1

INTRODUCTION AND BACKGROUND TO STUDY

1.1 Introduction

One of the aims of statistical modelling is to predict relationships among events in our surroundings. This is accomplished by finding predictive patterns that relates quantities in the real world. Linear modelling is one such statistical tool that is used to determine a linear function between the predicted and input attributes, say Y and Z. The investigation may seek to establish a linear relationship that exists between two phenomena that occurs naturally or by experimental design. This is generally denoted as

$$Y_i = \mu(z_i) + \varepsilon_i \quad (i = 1,2,...,n) \tag{1.1}$$

where $\mu(z_i)$ is a deterministic linear function of Z, while Y and ε are stochastic components. Y, also referred to as the response variable, is dependant on Z while ε is a latent variable whose distribution usually is assumed known. The most commonly used of which is the Normal distribution, thus rendering Y a continuous random variable. Y as a response variable evolves in any of the following scenarios: (1) As the actual response measure like weight in maize yield after dosage of fertilizer, or amount of viral load in a blood specimen taken from a patient who is undergoing anti-retroviral therapy. (2) Time-to-realization of event of interest, like time to recuperation after undergoing surgery. The two response situations, different as they are, address a common phenomenon using diverse analytical approaches. In ordinary linear modelling for instance, the response variable is a direct function of the linear combination

of explanatory variables. This method is well established and the abundance of literature that discusses this field of study vindicates this statement. Meanwhile analysis of response variable emanating from a lifetime requires the use of survival analysis techniques that implicitly relate the time response to the explanatory variables, resulting in departure from linear relationship. Since time is the actual measure, the response is assumed continuous and can only take positive values, thus normality assumptions no longer hold. Relaxing the continuity assumption for the response variable in both cases lead to a discrete Y as shall be elucidated in chapter 2, and when this happens, modifications of standard techniques is a necessity.

One aspect of survival data is that some experimental units do not realize the event of interest within the predetermined duration of study hence such observations are censored. This is a fundamental characteristic of survival data. A brief description of the kind of data and types of censoring prelude Chapter 2, with a subsequent review of literature that is related to the subject. It is important to mention that the main focus of this research shall be on interval right censoring for both grouped (disjoint) and overlapping intervals.

In Chapter 3 the likelihood functions are derived for both grouped and overlapping data types using distribution-free methods for a single lifetime. The results of using a non-parametric approach shall be compared to the parametric approach to be discussed in chapter 4. Dependence between

compliance by study units and censoring can have adverse effect on estimation of parameter values, thus as proposed by Finkelstein et al (2002), a likelihood that conditions, hence eliminating the effect of such a phenomena, is also derived. Still using non-parametric method, the univariate likelihood is extended to multiple failure time. Three methods are applied, these are by deriving likelihood functions to be used to estimate parameters under the independence working assumption method (IW), the Conditional Bivariate (CB) method by conditioning on the coordinate of one lifetime against the other, and use of Clayton Copula method (CC). In chapter 4 a Weibull distribution is assumed to be the underlying distribution for interval-censored data sets, and hence likelihood functions are derived for both univariate and multivariate distributed lifetimes situation using IW and CC methods.

A fundamental objective of this research is to explore the use of Bayesian methods by estimating posterior distributions for the unknown parameters, and where feasible, results from classical method of maximum likelihood estimation will also be computed. Since the type of prior distribution used for the parameters influences the final posterior function, our priority will be to derive non-informative priors where possible. Depending on the resultant posteriors, appropriate Monte Carlo Markov Chain methods like importance sampling, Gibbs sampling and Metropolis Hastings algorithm will be applied where possible. This topic is covered in Chapter 5. Illustrative examples, using several data sets, are given at the end of the chapter.

An alternative approach to survival analysis is explored in Chapter 6. The goal is to see if the same conclusions drawn using survival analysis results can be attained using other methods. The use of latent variables as illustrated by Albert and Chib (1993) is used, but with a logistic distributed latent variable. All statistical methods proposed for use undergo vigorous checking for estimation adequacy using simulated data. In Chapter 7 a Matlab computer program to simulate bivariate data using a Farlie-Morgenstern family of distributions, with exponentially distributed marginals, is used. The use of bivariate distributions is to depict the two lifetimes, and to address the question of multiple failure times and the inherent problem of correlated responses. The task therefore is to formulate stable parameter estimators in the presence of correlation on sparse data points. Finally, having established the adequacy of the aforementioned methods, they shall be applied on the estimation of parameters for explanatory variables in the Aids Clinical Trial Group (ACTG 175) data set.

Finally a brief summary of all major findings emanating from this research, are reported in Chapter 8.

1.2 Background Study and Variables of Interest

Human Immuno-deficiency Virus (HIV) and its related disease status Acquired Immune Deficiency Syndrome (AIDS), threaten to decimate human population from the face of the earth. Even though the pandemic is a universal tragedy, the

situation in Sub-Saharan Africa has reached genocide proportions, with some countries experiencing estimated national HIV prevalence rate of over 30% among the productive population (UNAIDS, (2002)), http://www.unaids.org. Researchers world-over are devoting time and massive resources to investigate the effect of AIDS on human kind. The recent development of antiretroviral (ARV) drugs offers hope, temporary as it may, in that sero-converted (HIV positive) patients' lifetime can be prolonged. Unfortunately the cost relating to acquisition of these drugs are prohibitive to the majority of third world countries. Thus to juggle the already over-stretched meagre resources to avail these drugs at affordable price to the ailing population, it is very important that the most potent and effective drugs are selected. Such information is not readily available, but applying methods mooted in this research on data solicited from a study conducted in the USA and described below, we partially address some of the aforementioned issues.

The ACTG 175 is a clinical trial study to assess the effectiveness of nucleosides on sero-converted or HIV positive patients, whose CD4 cell count just prior to enrolment into the study was measured to be between 200 and 500 per cubic millilitres. Patients were recruited from 43 Clinical Trials Units and 9 National Haemophilia Foundation sites in the United States and Puerto Rico. The study involved 2467 subjects whose time of enrolment varied between December 1991 and October 1992. Criterion for eligibility into the study were that subjects be of age 12 years or more, having laboratory documentation of HIV-1 type infection,

their CD4 cell count range between 200 and 500 per cubic millimetre within a month prior to the date of trial treatment, have no AIDS defining illnesses, a Karnofsky performance score of at least 70 and acceptable laboratory results. All patients were randomly assigned to any of the two single nucleosides (600 mg of zidovudine or 400 mg of didanosine) or a combination of nucleosides (600 mg of zidovudine plus 400 mg of didanosine or 600 mg of zidovudine plus 2.25 mg of zalcitabine). A monitoring and determination of CD4 cell levels were done at week 8 and every 12 weeks thereafter, with a primary study endpoint of 50% decline in CD4 cell count from the average of two pre-treatment counts, development of AIDS or death (Hammer et al 1996).

Prior studies show that plasma HIV Ribonucleic Acid (RNA) load (see section 7.3) is increasingly being used as a measure of viral replication in order to adequately evaluate the effect of antiretroviral drugs. Thus running concurrently with this study, a virology subgroup of 391 patients from among the main study patients were enrolled at 11 study sites and had their plasma HIV RNA concentrations also monitored. A primary study endpoint of 1 unit increase in the log base 10 of the number of copies per millimetre to the baseline concentrations of plasma HIV RNA was used. The copies of RNA per millimetre of blood were transformed in order to eliminate the variation between measurements. For instance some patients had values below the limit of detection (200 copies per millimetre), yet some had up to 1.45 million copies per millimetre. The first two monitoring period were at week 8 and 20, which

synchronized with CD4 cell determination periods, but for viral load, the monitoring was subsequently done every 36 weeks, provided the patients continued to receive assigned treatment. Thus this provided a bivariate measures of viral load and CD4 cell count as dependant variables, including the time period at which these measurements were assessed. Also recorded for each patient is the baseline demographic characteristics like age (years), race (white, black, Hispanic and Other race), weight (pounds) and gender. Also recorded were homosexual tendency, haemophilic, Karnofsky score, history of anti-retroviral use (ZDV), intravenous drug use (IDU), assigned treatment and presence or absence of syncytium-inducing phenotype. The study terminated in February 1995.

Some of the statistical methods employed towards the analysis of the data from ACTG 175 were univariate Cox's proportional hazard model for time-to either of the two variable end points, ANOVA with mean levels of CD4 cells and plasma concentration of HIV RNA, log-rank tests and two sample t-tests (Katzenstein et al 1996).

CHAPTER 2

LITERATURE REVIEW AND RESEARCH OBJECTIVES

2.1 Literature Review

Survival analysis is a special case of linear modelling which deals with time to occurrence of an event of interest. This may be time to death, failure, detection of some phenomena, etc. This inevitably renders a lifetime T non-negative and continuous random variable. Available literature on survival analysis like, Klein and Moeschberger (1997), Crowder (2001), Kalbfleisch and Prentice (2002), etc, have all shown important functions that describe the distributions of lifetime for both continuous and discrete variables. We shall only define the continuous case as follows; the unconditional probability of event (say failure) occurring at an infinitesimally small interval (t, t+Δt) gives a probability density function;

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t)}{\Delta t}. \qquad (2.1)$$

Probability of event occurring at or prior to time t is the probability distribution

$$F(t) = P(T \le t) = \int_0^t f(u)du\,. \qquad (2.2)$$

Probability that a subject survives beyond time t is the survival function:

$$S(t) = P(T > t) = \int_t^\infty f(u)du \qquad (2.3)$$

The conditional probability of failure or the chance that a subject who has survived to time t experiencing failure in the next instant, i.e. instantaneous rate of failure at time t given the subject survives up until t, is called the hazard rate.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}. \qquad (2.4)$$

Finally the cumulative hazard function is given by

$$H(t) = \int_0^t h(u)du\,. \qquad (2.5)$$

Survival analysis literature elucidates the relationship that exists between all the five functions. For instance, (2.6) shows the relation between all five functions.

$$H(t) = \int_0^t \frac{f(u)}{1-F(u)}du = -\ln[S(t)]. \qquad (2.6)$$

Time to event data could be easy to manipulate if data was complete for all subjects in the study, which is not the case in survival analysis. Instead, time-to-event maybe known to have occurred prior to the inception of the study, or certain subjects in the study may not have experienced failure at the time of termination of the study. Furthermore a design of study may militate that subjects be observed for failure at predetermined intervals, hence the occurrence of failure will only be known to have occurred between two time points, all of which are not exact lifetime. The above situations give rise to what is termed as censoring in survival analysis. There are three types of censoring. Right censoring, assumes a fixed termination point of study C_r and for each subject determine a lifetime T_A. These are independent and identically distributed with probability density function f(t) and survival function S(t). Thus exact lifetime is realized if $T_A \le C_r$, otherwise it is censored. This arises due to either subjects still surviving at termination period or some subjects moving

away from study for other reasons. Data obtained from each subject is represented as (t,δ) where δ is an indicator variable taking value 1 if the subject has an exact lifetime or 0 if the subject is censored, implying that $T=\min(T_A,C_r)$. A subject displaying an exact lifetime provides information that the probability of failure occurring is approximated by a density function of T at t, whilst a censored subject shows probability of survival evaluated at the termination of study.

$$P(T,\delta = 0) = S(C_r).$$

$$P(T,\delta = 1) = f(t).$$

Thus

$$P(T,\delta) = [f(t)]^{\delta} [S(t)]^{1-\delta}. \qquad (2.6)$$

A subject whose failure time t is known to have occurred prior to inception of study at time C_l is left censored. We observe that the exact event time is unknown, but that $T\in[0,t]$ and is analogous to right censoring, hence when contrasted with right censoring, $T=\max(T_A,C_l)$. The third type, the one that this research focuses on, is called interval censoring. It may not be feasible to observe the actual time of occurrence of event of interest, instead a time of last absence and first detection may be known, hence an interval. This happens in a longitudinal study like in clinical trial studies where treatment effect on study units are monitored and obtained periodically during clinic visits resulting in either complete or incomplete data. Turnbull (1976), in the estimation of distribution function for incomplete, censored and truncated data, proposed letting the checking times form a grid of time points which are completely

covered by end points $T \in [L_i, R_i]$ for all participants. Finkelstein (1986) discussed this scenario for fitting a proportional hazards regression model on a single lifetime where the intervals are disjoint, while Guo and Lin (1994) illustrated the same model for complete data situation.

Study design may vary, in that though predetermined clinic visits and study termination period is common to all units and is strictly enforced, study units may not commence study at the same time. A situation where study units start at varying times results in varying duration of study after rescaling. Thus for those units not having realized failure at termination point, will be censored, but the censoring may occur at any of the intervals. Yet if all units had commenced study simultaneously and no units lost (attrition), then any censoring will be at the last interval. A follow-up paper by Goggins and Finkelstein, (2000) discussed multiple failure lifetimes for interval-censored data with overlapping and non-disjoint intervals. In general, knowledge of type of censoring enables one to compute a likelihood function, which is of the form:

$$L \alpha \prod_{i \in E} f(t_i) \prod_{i \in G} S(C_r) \prod_{i \in H} (1 - S(C_i) \prod_{i \in I} (S(L_i) - S(R_i)) . \qquad (2.7)$$

E a subset of all individuals having exact lifetimes, G is all individuals whose lifetime is right censored, H individuals whose lifetime is left censored and all individuals whose lifetime is interval censored will belong in I, where L and R is the lower and upper end points of an interval, respectively.

There exist two reasons that impede the use of ordinary linear models. First, due to censoring, ordinary linear models will either omit those units that are censored or subdivide units into two groups for analysis. Efromovich (1999) illustrated the pitfalls of endeavouring to analyse survival data using the above approach. Secondly the distribution of lifetime data deviate from the accustomed normal distribution, hence conventional linear model results based on normal assumption cannot hold. Moreover, censoring invalidates the use of moments due to difficulty associated with estimation of right tail, which in reality may have significant influence on the mean. It is plausible that the distribution of a lifetime can be specified, resulting in parametric models as discussed by Lawless (1982), Cox and Oakes (1984) and Kalbfleisch and Prentice (2002). Not withstanding the difficulty associated with identifying a distribution that closely fit the data at hand, due to their restrictive distributional nature, lifetimes requires some transformations like logarithm, etc. Cox's (1972) proportional hazard model (PH), a distribution free method, is championed as robust in that it is able to handle survival data without having to resort to any of the afore-specified intervention. Its appeal is based on its avoidance of assuming an underlying distribution for the data, yet through the hazard function, is able to relate the response variable with the covariates.

$$\lambda(t \mid z) = \lambda_0(t)h(t,z). \tag{2.8}$$

Here $\lambda_0(t)$ is an arbitrary and unspecified baseline hazard function, and relative risk function $h(t,z)$ specifies the relationship between covariates and the hazard function. When the covariates in the model are fixed so that $Z(t) = Z$ for all t,

then the hazard function is independent of time, implying that the relative risk for any two individuals with different covariates are proportional, hence proportional hazard model. This model requires estimation of the baseline hazard. Cox's (1975) version of proportional hazards model is only partially parametric in that baseline parameters take arbitrary values and do not feature in the estimating equations, hence partial likelihood model. Satten (1996) also showed an approach that used marginal likelihood on interval-censored data to estimate parameters in the proportional hazard model without having to estimate the baseline hazard.

Motivated by Satten's paper, Pan and Chapell (2002) showed that the nonparametric MLE of the regression coefficient from the joint likelihood works well for the PH model with left truncated and interval censored data. If covariates vary with time, then there exist models that allows for the time variation in these variables, hence are called time-dependant covariates. It may happen that there exist unobservable heterogeneity among units, and to account for this random variability, frailty model discussed by among others, Hougaard (2001) and McGilchrist and Aisbett (1991), has proved to yield good results. Other suggested interventions to improve analysis of survival data include data imputation using auxiliary variables (Faucett et al; 2002). In their paper, Fleming and Lin (2000) outline what has been achieved in terms of research on survival analysis, summarily giving potential future research areas.

The dominance by a classical approach in survival analysis cannot be ignored, hence their call for contributions from a Bayesian perspective.

Yet there are other models which Fleming and Lin (2000) termed semi-parametric transformed models. These models depict lifetime as a function of an unspecified link function h(T), which in turn is a linear function of the covariates and random error term with a given distribution function F. If h(T) is a log-log transform, resulting in F being an extreme value distribution, this yields a proportional hazard model. Meanwhile a logit transformation h(T), with a logistic distribution F, will result in Collett's proportional odds (Collett, (1994)) model discussed by Colosimo et al (2000) and Cheng, Wei and Ying (1995). Lawless (1982) illustrated the use of both proportional hazard and proportional odds (PO) models for grouped interval censored data for a single lifetime. The methods presented are subsequently used to make comparisons between several independent cohorts. Even though the literature discusses both the continuous and discrete cases for both parametric and distribution free situation, we shall highlight the scenario for the interval data with both grouped and overlapping time intervals since it epitomizes this research's focal interest.

On the issue of multiple failure lifetimes with interval censoring experienced by an individual, the dependence that exists between the two measures cannot be wished away. The dependence structure varies with field of study, for instance

the dependence between competing risks will differ from recurrent events (Crowder, (2001)). Fleming and Lin once again mention the use of frailty models on bivariate survival data from a parametric perspective. The deficiency in exploring the use of non-parametric methods is apparent as amplified by their comment "No such results are available for general interval-censored data, although *ad hoc* methods (e.g. Finkelstein, (1986); Satten, (1996)) have been suggested". Goggins and Finkelstein (2000) used an independence assumption model (IW) approach to estimate the required parameters. The inherent dependence structure between lifetimes is thus unaccounted for except through the use of common covariates parameters between the marginal distributions of lifetimes. Hougaard (2001) terms it marginal modelling. The model seems to thrive in estimating the parameter values if the dependence between the failure types is not so strong and for relatively large sample sizes, though the same cannot be said with regard to variance estimation. Use of a 'sandwich estimator' to stabilize the variance is roped in as a mechanism to eliminate the inconsistencies. A complementary approach to the marginal modelling is the concept of using copulas. The problem of specifying a probability model for independent observations from a bivariate population with non-normal distribution function H(x,y) can be simplified by expressing H in terms of marginal distributions and its associated dependence function implicitly defined. (Genest and Rivest, (1993)) This assumes a uniform distribution on the unit square hence, if the distributions are continuous, they are transformed to the uniform case. The class of dependence functions or copulas are widely

available, as shown by Clayton (1978), Oakes (1982), Prentice and Cai (1992) and Gumbel (1960), to name a few. Betensky and Finkelstein (2002) made a follow-up on the question of non-compliance on a single failure time to illustrate how the dependence between failure time and visit compliance can affect the estimation of parameters. Sinha et al (1999) put together several Bayesian models which they compare using Bayes factors.

If all units realize the event of interest, i.e. in the absence of censoring, then familiar methods are available for analysing this type of data, one of which is the Generalized Linear models (GLM). The application of GLM, alongside with details on the estimation of parameters is given by McCullagh and Nelder (1989). The models involve a mean of observations given by the linear combination of unknown parameters and covariates on a link function transformed scale. Use of log-log, logit and probit link functions has been illustrated for both nominal and ordinal responses, as shown by Amemiya (1981), Agresti (1990) and Powers and Xie (2000). These models bear resemblance to the semi-parametric models on lifetime data. A cumulative logit on polychotomous responses is similar to proportional odds model on interval complete survival data, yet log-log is similar to proportional hazard on the same scale. These treat the ordinal responses as emanating from an unobservable latent lifetime variable. Mallick and Gelfand's (1994) approach is to treat the link function as an unknown, thus estimate it jointly with mean structure. Meanwhile Albert and Chib (1993) used data augmentation from a

Bayesian perspective to fit models on ordinal response variables. In their paper (Albert and Chib, (2001)), they apply the logit model (sequential ordinal modelling) to survival data. Earlier paper by Tanner and Wong (19987) described the data-augmentation algorithm for calculating marginal distributions. A method similar to the one applied by Albert and Chib (1993) will be applied to Tri-Continental AIDS data as an alternative to using survival methods.

2.2 Research Objectives

1 An endeavour to develop methods for analysing multiple lifetime data emanating from the same individual resulting in correlated failure time data. This give rise to multiple and correlated failure times.

2 A Longitudinal study results in interval data if patients are monitored at predetermined time periods. Overlapping intervals may arise and are difficult to handle. Methods that address this situation shall be presented.

3 To assess the impact of baseline predictors (covariates) of participating units on survival probabilities.

4 An exploratory investigation involving non-parametric methods with some Copulas like Farlie-Morgenstein and Clayton combined with either Cox's

Proportional Hazards or Proportional odds models will be closely scrutinized for their efficiency in parameter estimation.

5 Augmentation estimation techniques will be used since there is a tendency for methods to crash due to data scarcity. The method of Maximum Likelihood for instance, has shown to be highly sensitive to small sample situation. A Bayesian approach with good prior distributions for parameter and using Metropolis-Hastings, a branch of MCMC methods, is presented as an alternative. This takes centre stage in this research. Yet for large samples it will be shown that the two methods complement each other by using the MLE estimate of the covariance matrix for the covariance of the proposal distribution in posterior estimation.

6 The iterative process of cycling between parameters to simulate the next single parameter value in the MCMC method can be slow if the number of parameters involved is large. Suggested methods of alternating conditional sampling using blocks of parameters will be used.

7 Check the asymptotic traits of the parameter estimates by bootstrap methods for simulating pseudo samples and checking if the sampling distributions of the estimators converge to the true population values. This calls for the writing of appropriate computer programs to generate and analyse data.

8 Explore alternative methods to the survival ones that can be used to analyse data sets available. This will be tied to existing Generalized Linear Models techniques for categorical data, by adopting a Bayesian approach

9 Finally of profound interest is to assess the general applicability of the methods developed on real data situations using the following data sets: Aids Clinical Trials Group Study (ACTG 175) data, Mango data, Kidney data and Tri-continental Aids data.

CHAPTER 3

NONPARAMETRIC SURVIVAL MODELS

3.1 Introduction

Measure of time to event (failure) for observations cannot always be ascertained exactly. As indicated earlier, a clinical study where units are expected to be checked at predetermined checking times $0=t_0 < t_1 <....< t_{r+1}=\infty$ is a good example. Two scenarios arise when one views the regularity with which units adhere to the clinic monitoring times. If units observe and attend at all predetermined times, their failures will fall within two successive end points of an interval $I_j = (t_{j-1}, t_j]$, this results in Grouped failure times (complete data), with every unit described by a single interval within which failure/censoring occurs. Grouping of observed time into categories according to intervals results in discrete data. But any non-compliance results in failures stretching over several intervals, resulting in overlapping and non-disjoint intervals over individuals and are of varying lengths. This may be due to a subject missing several visits such that by the time they return, their response status has changed, hence their interval is now indexed by two end points L_i and R_i which may encompass several of the predetermined intervals I_j. Modification of methods used to analyse complete interval data is necessary since interval censoring of this nature is more intricate. To analyse this data, a distribution-free approach of Cox's proportional hazard and Collett's proportional odds is assumed, hence a non-parametric approach. Section 3.2 of this chapter shows

methods applied if a single lifetime is involved, which is extended in subsequent sections to address multiple failure time situations.

3.1.1 Proportional Hazards Models

Define the probability of an event occurring in time interval $(t_{j-1},t_j]$, $P(T \varepsilon (t_{j-1},t_j])$ as

$$\Delta_j = F(t_j) - F(t_{j-1}) \qquad j=1,2,\ldots,r+1$$

$$= S(t_{j-1}) - S(t_j). \tag{3.1}$$

Survival function at t_j, the probability of surviving interval $(t_{j-1},t_j]$ is given by

$$S(t_j) = P(T > t_j)$$

$$= \sum_{l>j} \Delta_l . \tag{3.2}$$

The conditional probability of failure in interval $(t_{j-1},t_j]$ is

$$h(t_j) = P(t_{j-1} < T \leq t_j \mid T>t_{j-1}) = \frac{\Delta_j}{S(t_{j-1})}$$

$$= 1 - \frac{S(t_j)}{S(t_{j-1})}. \tag{3.3}$$

So $\Delta_j = h(t_j)S(t_{j-1})$ is the unconditional probability of failure in interval $(t_{j-1},t_j]$. Survival function can therefore be modified to (Cox and Oakes (1984))

$$S(t_j) = \prod_{s=1}^{j} (1 - h(t_s)) . \tag{3.4}$$

Of fundamental importance is the conditional probability of survival beyond interval I_j given that one has survived to the interval. Let the conditional probability of being free of failure at the end of the j^{th} interval be

$$P_j = \frac{P(T > t_j)}{P(T > t_{j-1})} = \frac{\sum_{s=j+1}^{r} \Delta_s}{\sum_{s=j}^{r} \Delta_s}, \qquad (3.5)$$

where upon the survivor function (3.4), can be written in terms of P_j;

$$S(t_j) = \sum_{s=j+1}^{r} \Delta_s = P_j P_{j-1} \sum_{s=j-2}^{r} \Delta_s \qquad (s<j)$$

$$= \prod_{s=1}^{j} P_s \,. \qquad (3.6)$$

Let the unconditional probability of failure at I_j for a given unit rewritten in terms of conditional survival probability be

$$\Delta_j = (1 - P_j) \prod_{s=1}^{j-1} P_s. \qquad (3.7)$$

Under proportional hazards model, the probability that a person characterized by a vector of covariates **z**, survives beyond an interval is:

$$S(t_j) = \left(\prod_{s=1}^{j} P_j \right)^{e^{\beta z}} . \qquad (3.8)$$

Take a complementary log-log link function that relates the monotone differential function of the conditional survival probability P_j to the linear term composed of the explanatory variables, (Fahrmeir & Tutz (1994)). Then

$$P_j(\mathbf{z}) = \exp(-\exp(\gamma_j + \boldsymbol{\beta}\mathbf{z})) \qquad (3.9)$$

The transformation yields γ_j's, which are known as baseline survivor parameters. These parameters, unlike the conditional probability parameters, have a support that belongs to a real line. The unrestrictive nature of the parameters enhances easier estimation for any given likelihood function.

3.1.2 Proportional Odds Models

Collett's Proportional Odds model is defined as

$$\frac{F(I_j \mid \mathbf{z})}{1-F(I_{j-1} \mid \mathbf{z})} = \exp(\boldsymbol{\beta}\mathbf{z}). \tag{3.10}$$

The proportional odds model is the odds ratio of failure at interval I_j given survival to beginning of I_j.

Let

$$\tau_j(\mathbf{z}) = P(Y \leq I_j \mid Y > I_{j-1})$$

$$= 1 - P_j(\mathbf{z}).$$

By taking a logit transformation relating the conditional survival probability and the linear parameter function, we show that the resultant distribution is Logistic. For $\mathbf{z} = \mathbf{0}$, the baseline log odds of failure at I_j in terms of conditional survival probability,

$$\alpha_j = \log\left(\frac{1-P_j(\mathbf{0})}{P_j(\mathbf{0})}\right), \tag{3.11}$$

can be written in terms of conditional failure probability,

$$\frac{\tau_j(\mathbf{0})}{1-\tau_j(\mathbf{0})} = \exp(\alpha_j).$$

We note then that if the effect of explanatory variables is included, then

$$\frac{\tau_j(\mathbf{z})}{1-\tau_j(\mathbf{z})} = e^{\alpha_j} e^{\boldsymbol{\beta}\mathbf{z}}$$

$$= e^{\alpha_j + \boldsymbol{\beta}\mathbf{z}}, \tag{3.12}$$

hence

$$\tau_j(\mathbf{z}) = \frac{e^{\alpha_j+\boldsymbol{\beta}\mathbf{z}}}{1+e^{\alpha_j+\boldsymbol{\beta}\mathbf{z}}} \tag{3.13}$$

The conditional survival probability under proportional odds model is then given as

$$P_j(\mathbf{z}) = \frac{1}{1+e^{\alpha_j+\boldsymbol{\beta}\mathbf{z}}} \tag{3.14}$$

3.2 Univariate Failure Time

Suppose all the observations have a single lifetime with survival space sub-divided into r intervals (j=1,2,...,r) denoting the checking times, then depending on the choice of model we shall derive the appropriate likelihood. A univariate model is defined using (3.6 and 3.7) where the conditional survival probability P_j is replaced by (3.9) for a proportional hazard model or (3.14) for a proportional odds model. Define a dichotomous random variable φ_{ij} for each observation taking the value 1 if the failure occurs at interval I_j for i^{th} subject, and 0 otherwise, such that the contribution by i^{th} unit to the likelihood is

$$g(\boldsymbol{\theta} \mid \mathbf{z}_i) = \left(1-P_j(\mathbf{z}_i)\right)^{\varphi_{ij}} \left(P_j(\mathbf{z}_i)\right)^{1-\varphi_{ij}} \prod_{s=1}^{j-1} P_j(\mathbf{z}_i) \tag{3.15}$$

where $\boldsymbol{\theta}$ is a vector of parameters to be estimated for a specific model, for instance $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \gamma_1, \gamma_2, \ldots \gamma_r\}$ for a PH likelihood. The above model applies if the intervals at which failure occurs are disjoint, but a need for modification arises if the intervals are non-disjoint and overlapping as described in section 2.1. TurnBull (1976) proposed letting $0 = t_0 < t_1 < \ldots < t_r = \infty$ be a grid of time points

which includes all L_i and R_i for all participants that experience failure. Inability to know the single interval at which failure occurs, implies that we need to ascertain the failure probability by summing failure probabilities (3.7) over all intervals falling within the two end points. Meanwhile for censored observations, the potential intervals of failure are all intervals subsequent to the lower censoring endpoint L_i. Define an indicator $\omega_{ij} = 1$ if the interval $(t_{j-1}, t_j]$ is contained in the end points $(L_i, R_i]$ and 0 otherwise. Then the log likelihood for i^{th} unit is denoted as

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}_i) = \log \sum_{j=1}^{r} \omega_{ij} \Delta_{ij}, \tag{3.16}$$

where Δ_{ij} is from (3.7) and is a function of $P_j(\mathbf{z}_i)$ which can be derived from any of the transformation models. The overall log likelihood is then the sum of individual units' log likelihood.

3.3 Failure Time Data with Dependent Interval Censoring

The use of informative censoring and its effect on the estimated parameter values is important and feasible if the follow-up period is long enough. In this section we explore the effect any dependence that exist between clinic visiting times and failure intervals may have on the results. Generally it is assumed that the true failure time is independent of censoring mechanisms that controls visits. But that may not necessarily be the case in that for instance, if a study is on deadly diseases, time to detection of disease preceded by symptoms compels study units to see a doctor and have tests done without failure, resulting in

some dependence between failure time and interval. Likewise, detection of the disease may prompt a unit to strictly adhere to clinic checks thereafter for treatment. (Finkelstein et al, (2002)).

Let's assume that all units commence study at the same time with j=1,2,..,r clinic checks such that if any unit is censored at the termination of study, this will be at the r^{th} interval i.e. in the absence of units lost due to study attrition. To analyse this kind of data, it is essential that information on visit compliance before and after failure be taken into consideration. Let the unobservable continuous failure time be denoted by T. Since we only observe interval I_j within which failure occurred, all intervals preceding the j^{th} include units that have experienced failure, thus we can model the likelihood from the interval perspective. Let v be a vector of binary indicator variables taking value 1 if a visit is made and 0 otherwise. Let π_{Bj} be the probability of making the visit in interval I_j before the failure occurred, and π_{Aj} be the probability of making the visit in interval I_j after the failure occurred. The probability of failure in interval I_j is $\Delta_j = P(t_{j-1} < T < t_j)$. Let n_{Bj} be the number of patients who make the visit in interval I_j and for whom that visit was one before they failed and n_{Aj} be the number of units who make the visit in interval I_j and for whom that visit was one after they failed. Also let d_j be the number of units who failed at interval I_j. Finally, let r_{Bj} be the number of people who were under observation and had not failed at time j whether or not they made their visit, while r_{Aj} is the

number of people who were under observation and had already failed by time j whether or not they made their visit. By Bayes theorem, $P(\mathbf{v},T) = P(\mathbf{v} \mid T)P(T)$ is the joint likelihood of failure time T and visit schedule. The joint conditional probability of a unit making a visit is the product of individual conditional probabilities at interval I_j.

$$P(v \mid T) = P(v_1 \mid T)P(v_2 \mid T)....P(v_r \mid T).$$

Probability of failure for a unit at interval I_j is denoted as in (3.7) where $P_j(\mathbf{z})$ is as defined in (3.5). The probability of failure at interval I_j is based on all d_j units in that interval. The product of each unit's failure probability therefore will yield the necessary failure probability for that interval. If all units have complied with clinic visit times, then such data is complete or grouped. To derive likelihood for this data, we define the conditional probability of a unit making a visit at interval I_j before failure as:

$$\pi_{Bj}^{\,v_{Bj}}(1-\pi_{Bj})^{1-v_{Bj}}, \tag{3.17}$$

and the conditional probability of a unit making a visit at interval I_j after failure is:

$$\pi_{Aj}^{\,v_{ij}}(1-\pi_{Aj})^{1-v_{ij}} \tag{3.18}$$

where

$$v_{Bj} = \begin{cases} 1 & \text{if patient i makes } j^{th} \text{ visit before failure} \\ 0 & \text{otherwise,} \end{cases}$$

and similarly

$$v_{Aj} = \begin{cases} 1 & \text{if patient i makes } j^{th} \text{ visit after failure} \\ 0 & \text{otherwise.} \end{cases}$$

Hence at a given interval, the likelihood of a unit is described by the product of probability of failure, probability of a unit making a visit prior to failure and probability of making a visit after failure.

$$\ell_{ij} = \Delta_j \left\{ \pi_{Bj}^{\ v_{Bj}} (1-\pi_{Bj})^{1-v_{Bj}} \right\} \left\{ \pi_{Aj}^{\ v_{Aj}} (1-\pi_{Aj})^{1-v_{Aj}} \right\}. \tag{3.19}$$

For all patients in interval I_j, the likelihood is given by

$$\ell_j = \Delta_j^{\ d_j} \left\{ \pi_{Bj}^{\ n_{Bj}} (1-\pi_{Bj})^{r_{Bj}-n_{Bj}} \right\} \left\{ \pi_{Aj}^{\ n_{Aj}} (1-\pi_{Aj})^{r_{Aj}-n_{Aj}} \right\},$$

hence, the overall likelihood across all intervals is the product of individual interval's likelihood,

$$\ell = \prod_{j=1}^{r} \Delta_j^{\ d_j} \left\{ \pi_{Bj}^{\ n_{Bj}} (1-\pi_{Bj})^{r_{Bj}-n_{Bj}} \right\} \left\{ \pi_{Aj}^{\ n_{Aj}} (1-\pi_{Aj})^{r_{Aj}-n_{Aj}} \right\}. \tag{3.20}$$

However, if the intervals are non-disjoint and overlapping (incomplete) such that are defined by lower and upper endpoints, (L_i, R_i), modifications need be made on the likelihood. Define an indicator variable ω_{ij} as in section 3.2. Equally vital to observe is the fact that units may have varying numbers of intervals in the study due to study attrition, hence each unit will have its own r_i, the last checking interval. The likelihood then is

$$\ell = \prod_{i=1}^{n} \sum_{j=1}^{r} \omega_{ij} \Delta_j \prod_{s=1}^{j-1} \left\{ \pi_{Bs}^{\ v_{is}} (1-\pi_{Bs})^{1-v_{is}} \right\} \prod_{s=j}^{r_i} \left\{ \pi_{As}^{\ v_{is}} (1-\pi_{As})^{1-v_{is}} \right\}. \tag{3.21}$$

Under Cox's proportional hazard model with individual units' covariates, failure probability is denoted by

$$\Delta_{IJ} = \left(1 - e^{-e^{\gamma_j + \beta Z_I}} \right) \prod_{s=1}^{j-1} e^{-e^{\gamma_s + \beta Z_I}}, \tag{3.22a}$$

whereas proportional odds model has

$$\Delta_{IJ} = \left(\frac{e^{\gamma_j + \beta Z_I}}{1 + e^{\gamma_j + \beta Z_I}} \right) \prod_{s=1}^{j-1} \frac{1}{1 + e^{\gamma_s + \beta Z_I}} . \qquad (3.22b)$$

A logit transform for the uniformly distributed visit probabilities, yields a logistic distribution that enable the inclusion of individual unit covariates (Finkelstein et al 2002). Then, the probability of the i^{th} patient making the j^{th} visit at s^{th} failure interval can be written as

$$\pi_{ijs} = \frac{e^{\mu_j + \lambda \eta_{js} + \nu z_i}}{1 + e^{\mu_j + \lambda \eta_{js} + \nu z_i}} \qquad (3.23)$$

where μ_j is a constant for the j^{th} visit time irrespective of failure time (baseline or post-failure visits) and η_{js} is a binary variable taking value 1 if s<j and zero otherwise, with a coefficient λ. This coefficient therefore give a direction as to whether a unit is likely to make more visits prior or after a failure. The presence of such a coefficient in the model allows for the combination of the two Bernoulli components into one Bernoulli with two inbuilt indicator variables catering for visiting periods (before and after failure) and whether an interval is contained by the upper and lower endpoints. Meanwhile ν measures the effect that covariates may have on probability of visit. The modified likelihood is denoted by

$$\ell = \prod_{i=1}^{n} \sum_{j=1}^{r} \omega_{ij} \Delta_j \prod_{s=1}^{r_i} \pi_{ijs}^{\nu_{ij}} (1 - \pi_{ijs})^{1-\nu_{ij}} . \qquad (3.24)$$

Application for this method is shown in chapter 7 where both simulated data is generated and then an analysis of ACTG 175 AIDS data is used on CD4 failure times.

3.4 Marginal Likelihood Model for Multiple Failure Interval-Censored Data

A phenomenon can be described by several events, thus rendering it a multivariate type. For n observations, let there be M failure times (m=1,2,...,M) with each having survival space sub-divided into r_m intervals (j_m=1,2,...,r_m) representing the checking times. Consequently the failure times may be correlated to a reasonable degree since i^{th} subject's failure event at j^{th} interval for all lifetimes is defined as a hyperspace described by a vector $\{I_{1ij}, I_{2ij}, ..., I_{Mij}\}$ depicting the intervals at which each lifetime event occurred. We shall restrict our illustration to two lifetimes, (j=1,2,..r_1; q=1,2,..,r_2), hence a region $I_{jq} = \{I_j, I_q\}$.

Define for each subject an indicator variable

$$\varphi_{mij} = \begin{cases} 1 & \text{if } m^{th} \text{ failure lifetime occurs at } j^{th} \text{ interval for patient i} \\ 0 & \text{if censored} \end{cases}$$

The unconditional probability that a unit experiences both failure events in the intervals I_{1ij} and I_{2iq} assuming the failure times are independent and the intervals are disjoint, is a product of the marginal probabilities,

$$\Delta_{jq} = P(t_{1j-1} < T_1 \le t_{1j}, t_{2q-1} < T_2 \le t_{2q}). \qquad (3.25)$$

The conditional survival probability shown in (3.5) is subscripted by lifetimes and interval at which the event of interest occurs, hence the overall likelihood using appropriate conditional survival probabilities is denoted as

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}) = \sum_{m=1}^{M}\sum_{i=1}^{n}\left[\varphi_{mij}\log\left(1 - P_{mj}(\mathbf{z}_i)\right) + (1-\varphi_{mij})\log P_{mj}(\mathbf{z}_i) + \sum_{s=1}^{j-1}\log P_{ms}(\mathbf{z}_i)\right], \quad (3.26)$$

where φ_{mij}=1 if i^{th} patient's m^{th} failure occurs in the j^{th} interval and 0 otherwise. Where it not for the common covariates, hence common parameter estimates

for the marginal, the model would be equivalent to simply combining the marginal likelihood functions used in the individual univariate lifetime analysis. Thus the dependence, if any, is accounted for by the common effect of the covariates. This model is called Independence assumption model with proportional hazard (IWH) or proportional odds (IWO). In a similar fashion, if the intervals are overlapping and non-disjoint, define an indicator variable ω_{mij} (m=1,2) for each lifetime taking value 1 if the end points $(L_i,R_i]$ contain the i^{th} failure interval. The overall log likelihood model is denoted by

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}) = \sum_{m=1}^{M}\sum_{i=1}^{n}\left[\log\sum_{j=1}^{r_m}\omega_{mij}\left(1-P_{mj}(\mathbf{z}_i)\right)\prod_{s=1}^{j-1}P_{ms}(\mathbf{z}_i)\right]. \tag{3.27}$$

3.5 Conditional Bivariate Model

When we analyse data from a bivariate distribution, the data set depicts two failure times whose relation is brought about by a dependence parameter. An example is the bivariate normal distribution. The parameter ρ measures the dependence between the two variables involved. In the previous section we analysed survival data using marginal likelihood because we assumed that data would reveal its dependence through the explanatory variable parameter estimates. We know that if the failure types are independent, then we have two independent marginal distributions, hence their joint likelihood is a product of individual marginal distributions. A deviation from the above expectation can only be explained by existence of dependence between the data set. A weak dependence may not be discernable by using marginal parameter estimates

hence a need to apply a technique that would take cognisance of any prevailing dependence between two failure types in the computation of the parameter estimates. By conditioning on the coordinate of one lifetime, we compute the joint likelihood using conditional survival probabilities as before, simultaneously considering the position of failure of the other lifetime. This is presented as one option described as follows. Define P_{2qj} as a subject's conditional survival probability at interval $(t_{q-1}, t_q]$ for second lifetime, given the first lifetime's failure occurred at $(t_{j-1}, t_j]$, $j = 1,2,\ldots,r_1+1$; $q = 1,2,\ldots,r_2$.

Let

$$P_{2qj} = P[T_2 > t_q \mid T_2 > t_{q-1}, t_{j-1} < T_1 \leq t_j] \tag{3.28}$$

where for continuous T_1 and T_2,

$$P[T_2 > t_q, t_{j-1} < T_1 < t_j] = \int_{t_q}^{\infty} \int_{t_{j-1}}^{t_j} f(t_1, t_2) dt_1 dt_2 . \tag{3.29}$$

The conditional survival probabilities are easily extended to include the effect of covariates using any of the transformations described in section 3.1. The marginal failure probability at I_j is as defined as (3.7), while the conditional failure probability at interval $(t_{q-1}, t_q]$ for the second lifetime, given that failure for first lifetime occurred at $(t_{j-1}, t_j]$, is denoted by

$$\Delta_{2qj} = P[t_{q-1} < T_2 \leq t_q \mid t_{j-1} < T_1 \leq t_j]$$

$$= (1 - P_{2qj}) \prod_{l=1}^{q-1} P_{slj} . \tag{3.30}$$

Thus the unconditional joint failure probability at the intervals $(t_{j-1}, t_j]$ and $(t_{q-1}, t_q]$ is by Bayes theorem, the product of Δ_{2qj} and the marginal

unconditional failure probabilities for the conditioning lifetime, Δ_{1j}. We note that if the two lifetimes are independent, then $\Delta_{2qj} = \Delta_{2q}$ for all j, hence the joint unconditional probabilities will be the product of the marginal unconditional probabilities, $\Delta_{jq} = \Delta_{2q}\Delta_{1j}$, failure of which we conclude that the two lifetimes are dependent. If a unit's first lifetime is censored at the conditioning interval I_j, failure can only occur in one of the subsequent intervals, hence sum up all the joint failure probabilities of those intervals for known intervals of second lifetime. This gives the joint failure probability for this unit as

$$P\left[T_1 > t_j, t_{q-1} < T_2 < t_q\right] = \sum_{s=j+1}^{r_1+1} \Delta_{1s}\Delta_{2qs} \tag{3.31}$$

Then using (3.28), (3.30) and indicator variables in section 3.3, the likelihood for a subject whose event of interest occurs at interval I_j for first lifetime and interval I_q for second lifetime is given by

$$g(\boldsymbol{\theta} \mid \mathbf{z}_i) =$$

$$\left[(1-P_{1j}(z_i))\prod_{s=1}^{j-1}P_{1s}(z_i)(1-P_{2qj}(z_i))\prod_{l=1}^{q-1}P_{2lj}(z_i)\right]^{\varphi_{1ij}\varphi_{2iq}}\left[(1-P_{1j}(z_i))\prod_{s=1}^{j-1}P_{1s}(z_i)\prod_{l=1}^{q}P_{2lj}(z_i)\right]^{\varphi_{1ij}(1-\varphi_{2iq})}$$

$$\sum_{s=j+1}^{r+1}\left[(1-P_{1s}(z_i))\prod_{l=1}^{s-1}P_{1s}(z_i)(1-P_{2qs}(z_i))\prod_{l=1}^{q-1}P_{2ls}(z_i)\right]^{(1-\varphi_{1ij})\varphi_{2iq}}$$

$$\sum_{s=j+1}^{r+1}\left[(1-P_{is}(z_i))\prod_{l=1}^{s-1}P_{il}(z_i)\prod_{l=1}^{q}P_{2ls}(z_i)\right]^{(1-\varphi_{1ij})(1-\varphi_{2iq})} \tag{3.32}$$

This gives the overall likelihood over all units to be the product of all individual's likelihood. The number of baseline parameters to be estimated depends on the predetermined intervals involved, i.e. including the covariates parameters, there are $r_1(1+r_2) + p$ parameters to be estimated.

3.6 Use of Copulas for Bivariate Models

The use of a conditional bivariate model approach presented in section 3.4 poses problems due to the number of parameters involved. Therefore with small data sample, the method is bound to collapse, especially since some of the intervals may be empty. An alternative would be to use models that are built from the Copula distributions as per definition presented by Prentice & Cai (1992), based on the i^{th} subject's joint survival or failure function for two failure times. The method breaks away from the independence working assumption adopted for univariate likelihood in section 3.3, in that this method introduces dependence parameter between the two lifetimes. For example, a Clayton copula is depicted in terms of marginal survival functions. (Prentice and Cai, (1992))

$$S(t_{1j}, t_{2q}) = \left[S(t_{1j})^{\frac{-1}{\kappa}} + S(t_{2q})^{\frac{-1}{\kappa}} - 1 \right]^{-\kappa}, \qquad (3.33)$$

where $(0<\kappa<\infty)$ $\kappa \to 0$ implies a perfect correlation between the two failure times and absolute independence when $\kappa \to \infty$. $S(t_{1j})$ is the marginal survival probability for the first lifetime at interval I_j. For instance, under Cox's proportional hazard model, the i^{th} subject with a covariate $\mathbf{z}$ and discrete random variable T whose marginal survival at the j^{th} interval is as shown in (3.8) contributes the following component to the joint survival:

$$S(t_{i1})^{-\frac{1}{\kappa}} = \left[\prod_{s=1}^{j} P_{1s}(\mathbf{z}) \right]^{-\frac{1}{\kappa}} = \prod_{s=1}^{j} e^{\frac{1}{\kappa} h_s(\gamma, \boldsymbol{\beta})},$$

where $h_s(\gamma, \boldsymbol{\beta}) = e^{\gamma_{ms} + \boldsymbol{\beta} \mathbf{z}}$, yielding a joint survival function

$$S(t_{1j}, t_{2q}) = \left[\prod_{s=1}^{j} e^{\frac{h_s(\gamma,\beta)}{\kappa}} + \prod_{l=1}^{q} e^{\frac{h_l(\gamma,\beta)}{\kappa}} - 1 \right]^{-\kappa} \tag{3.34}$$

To write a joint likelihood, we use the same definition for the indicator variables φ_{1ij} and φ_{2iq} as in previous sections. The likelihood of a unit who has bivariate disjoint intervals for both lifetimes is given by

$$g(\gamma,\beta \mid \mathbf{z}_i) = \Big(\left[S(t_j,t_q) + S(t_{j-1},t_{q-1}) - S(t_{j-1},t_q) - S(t_j,t_{q-1}) \right]^{\varphi_{1ij}\varphi_{2iq}} \left[S(t_{j-1},t_q) - S(t_j,t_q) \right]^{\varphi_{1ij}(1-\varphi_{2iq})}$$

$$\left[S(t_j,t_{q-1}) - S(t_j,t_q) \right]^{(1-\varphi_{1ij})\varphi_{2iq}} \left[S(t_j,t_q) \right]^{(1-\varphi_{1ij})(1-\varphi_{2iq})} \Big) \tag{3.35}$$

If the regions are non-disjoint and overlapping, let there be two indicator variables ω_{mj}, for each lifetime as previously defined in section 3.3, also the probability of failure at region I_{jq} (3.25) is given by

$$\Delta_{jq} = S(t_j,t_q) + S(t_{j-1},t_{q-1}) - S(t_{j-1},t_q) - S(t_j,t_{q-1})$$

Hence the overall log likelihood is the sum of each unit's likelihood for each lifetime as in (3.16);

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}) = \sum_{m=1}^{M} \sum_{i=1}^{n} \log \sum_{s=1}^{r_m+1} \phi_{mij} \Delta_{ms} \,. \tag{3.36}$$

The Farlie- Morgenstein copula is in terms of marginal cdfs, hence of the form

$$F(t_1,t_2) = F(t_1)F(t_2)\left[1 + \kappa(1-F(t_1))(1-F(t_2))\right], \tag{3.37}$$

with κ the measure of association between the two failure times such that correlation between them is $\rho = \kappa/4$. For this model the maximum κ attainable is 1, representing a not so strong dependency. The failure probability (3.25) is then defined in terms of joint CDF, not joint survival probabilities, as is the case with

Clayton copula. Then use the function (3.37) to attain an overall likelihood (3.36) for observations drawn from Farlie-Morgenstein distribution that has overlapping intervals.

3.7 Chapter Summary

The standard distributions we could in most cases apply to other types of data are not necessarily relevant to survival data. This is true especially with interval-censored data with overlapping intervals. This chapter introduced two types of interval-censored data in the grouped (disjoint) intervals and the overlapping intervals data. The latter kind is complicated to handle and the parameters estimated under this situation are susceptible to regularity of visits. The Cox's method of using distribution-free hazard function, make the non-parametric method more preferable since the models have proved to be easier to evaluate and can adjust to any kind of data distribution. Three non-parametric models for survival data were introduced in this chapter. These are: Independence assumption model, conditional bivariate model and the use of copulas to allow for measure of association between failure times. For each of the models, a likelihood function was derived after transforming the hazard function by either a log-log transform to get a proportional hazard model or a logit transform to get a proportional odds model, so as to facilitate for the use of unrestricted parameters.

CHAPTER 4:

PARAMETRIC SURVIVAL MODELS

4.1 Introduction

The most extensively used methods in analysing lifetime data is the parametric methods. The difficulty associated with these methods is of ascertaining that the underlying distribution of the lifetime data is correctly specified. Failure to do so would result in erroneous conclusions. Since lifetime is a time measure, hence continuous, the distribution involved needs be continuous. Properties of continuous distributions are well documented and readily available such that their use in modelling, if the data does comply, makes them more eligible for use. Moreover, for parametric models all four functions characterizing lifetime data, namely probability density function, probability distribution, hazard or cumulative hazard rates and survivor functions, are of closed form for most of these distributions. This renders estimation of a single failure time event more feasible. These distributions include the Weibull, Exponential, Log-Logistic, etc., most of which emanate from the Generalized Gamma (GG) family of distributions denoted as

$$f(t) = \frac{\xi t^{\xi\omega-1}\exp(-\frac{t^{\xi}}{\lambda})}{\lambda^{\xi}\Gamma(\omega)} \qquad 0 < t < \infty. \tag{4.1}$$

Extensions to multiple failure times employ the use of copulas with marginal distributions from known univariate continuous distributions. Section 4.3 discusses the use of Farlie-Morgensten and Clayton distribution copulas the

same way as outlined in section 3.5 for distribution-free situation, with built-in relatedness measure to assess dependence between failure times, if it exists.

4.2 Weibull Distribution

One versatile distribution bearing some unique properties in the analysis of lifetime data T is the Weibull (λ,ξ) distribution (a special case of the GG when ω=1) whose density and probability distribution are given by respectively

$$f(t) = \frac{\xi}{\lambda} t^{\xi-1} \exp(-\frac{t^{\xi}}{\lambda}) \tag{4.2}$$

and

$$F(t) = 1 - \exp(-\frac{t^{\xi}}{\lambda}),$$

and its k^{th} moment is given by

$$E(T^k) = \lambda^{\frac{k}{\xi}} \Gamma\left(1 + \frac{k}{\xi}\right).$$

An Exponential (λ) distribution is a special case of the Weibull when ξ=1, hence most of the characteristics of a Webull distribution are easily extendable to the exponential situation. A Weibull distributed lifetime with ξ not equal to 1 gives a non-constant hazard function, a property better suited for certain lifetimes. Let Y=log T, then

$$f(y) = \frac{\xi}{\lambda} \exp\left\{\xi y - \frac{\exp(\xi xp}{\lambda}\right\} \tag{4.3}$$

Let $\mu = -\sigma \log \frac{1}{\lambda}$ and $\sigma = \frac{1}{\xi}$, then

$$f(y) = \frac{1}{\sigma}\left\{\exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\} \tag{4.4}$$

Furthermore, a unique property which makes this distribution attractive to use in survival analysis is that a log transformation of the original data distributed as Weibull with both λ and ξ equal to one, yields a standard Extreme Value distribution with a survivor function S(y), similar to the Cox's proportional hazard shown in previous chapters:

$$S(y) = \exp(-\exp(y)) \qquad -\infty < y < \infty.$$

Our focus is to model the impact of a vector of explanatory variables $\mathbf{z} = (z_1, z_2, \ldots, z_{p-1})$ on the lifetime. One way will be by fitting a linear model of the form (1.1) to the transformed data Y, where the error component ε, follows an Extreme Value distribution and the mean $\mu(\mathbf{z})$ is a function of explanatory variables, i.e.

$$\mu(\mathbf{z}) = \psi_0 + \boldsymbol{\psi}\mathbf{z}, \tag{4.5}$$

where the vector $\boldsymbol{\psi} = (\psi_1, \psi_2, \ldots, \psi_{p-1})$ consists of regression coefficients. Suppose the original untransformed data was used to fit a linear model with vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_{p-1})$ being the regression coefficients, then the two sets of regression coefficients relates through

$$\boldsymbol{\beta} = -\frac{\boldsymbol{\psi}}{\sigma}. \tag{4.6}$$

With a log transformation, the survival function of the i^{th} unit is written as

$$S(y_i \mid \sigma, \boldsymbol{\psi}, \mathbf{z}_i) = \exp\left\{-\exp\left(\frac{y_i - \mu(\mathbf{z}_i)}{\sigma}\right)\right\}, \quad i=1,2,\ldots,n. \tag{4.7}$$

The density is of the form defined in (4.4) with expected logarithm lifetime $\mu(\mathbf{z})$ and standard deviation σ. It is vital to note that the resulting survivor function has reduced number of parameters. Knowledge that some of the subjects haven't realized the event of interest and hence are censored has to be taken into consideration when analysing survival data. Maintaining the same indicator variable as in section 3.2 to depict the censoring status of an experimental unit characterized by a vector of covariates $\mathbf{z}$, the general form of the likelihood is

$$g(\boldsymbol{\theta} \mid \mathbf{z}) = \prod_{i=1}^{n} \left\{ \left(f(y_i)^{\varphi_i} \right) \left(S(y_i)^{1-\varphi_i} \right) \right\}.$$

Hence

$$g(\boldsymbol{\theta} \mid \mathbf{z}) = \prod_{i=1}^{n} \left[\frac{1}{\sigma} \left\{ \exp\left(\frac{y_i - \mu(\mathbf{z}_i)}{\sigma}\right) - \exp\left(\frac{y_i - \mu(\mathbf{z}_i)}{\sigma}\right) \right\}^{\varphi_i} \left\{ \exp\left(-\exp\left(\frac{y_i - \mu(\mathbf{z}_i)}{\sigma}\right)\right) \right\}^{1-\varphi_i} \right]. \qquad (4.8)$$

Suppose the experimental units are observed only at intervals $I_j = (t_{j-1}, t_j]$ such that the event of interest is only known to occur in I_j, then the likelihood for interval-censored data has to be applied. Let L_i and R_i be the lower and upper endpoints of the interval containing the i^{th} failure as previously defined, such that the difference between survival probabilities at the endpoints yields the failure probability for a unit in interval $(L_i, R_i]$. Then

$$P(y_i \in (L_i, R_i) \mid \mathbf{z}_i, \boldsymbol{\theta}) = \exp\left\{ -\exp\left(\frac{L_i - \mu(\mathbf{z}_i)}{\sigma}\right) \right\} - \exp\left\{ -\exp\left(\frac{R_i - \mu(\mathbf{z}_i)}{\sigma}\right) \right\}. \qquad (4.9)$$

Since the afore-specified interval is composed of several sub-intervals, Finkelstein (1986) showed that the contribution of the i^{th} observation to the

likelihood is the sum of all failure probabilities of sub-intervals enclosed within the endpoints. The conditional survival probability at I_j is given by

$$P_j(\mathbf{z}) = \exp\left\{-\left[\exp\left(\frac{y_{ij-1} - \mu(\mathbf{z}_i)}{\sigma}\right)\right] - \left[\exp\left(\frac{y_{ij} - \mu(\mathbf{z}_i)}{\sigma}\right)\right]\right\} \tag{4.10}$$

Groenewald & Mokgatlhe (2002) have shown that the log-likelihood can be written in terms of conditional survival probabilities hence write the unconditional failure probability as

$$P(y_i \in I_j \mid \boldsymbol{\theta}, \mathbf{z}) = P(T > t_{j-1})\left(1 - P_j(\mathbf{z}_i)\right)$$

$$= \exp\left(-e^{\xi\left(y_{ij-1} - \mu(z_i)\right)}\right)\left\{\left(1 - e^{-e^{-\xi\mu(z_i)}}\right)\exp\left(e^{\xi y_{ij-1}} - e^{\xi y_{ij}}\right)\right\},$$

where $\xi = \frac{1}{\sigma}$. Thus the log likelihood is of form

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}) = \sum_{i=1}^{n} \log \sum_{j=1}^{r} \omega_{ij}[P(y_i \in I_j \mid \boldsymbol{\theta}, \mathbf{z}_i)], \tag{4.11}$$

where ω_{ij} is as defined in section 3.3. This is a univariate failure time situation, and is easily extended to multiple failures by using a multivariate family of distributions, as shall be illustrated in the next section.

4.3 Use of Copulas with Weibull distribution

Assuming a Weibull distributed variable T, a log transformation results in a variable Y with density as shown in (4.4). Its probability distribution is denoted by

$$F(y) = 1 - \exp\left\{-\exp\left(\frac{y-\mu}{\sigma}\right)\right\}$$

Define $\mathbf{Y} = \{Y_{1i}, Y_{2i}\}$ as a matrix consisting of two vectors representing the two failure times, (i=1,2,...,n). Then the joint distribution may be represented by a Farlie-Morgenstern distribution, denoted by

$$F(y_{1i}, y_{2i}) = \left(1 - e^{-w_{1i}}\right)\left(1 - e^{-w_{2i}}\right)\left[1 + \kappa e^{-(w_{1i}+w_{2i})}\right], \tag{4.12}$$

where κ is the measure of association, and

$$w_{mi} = e^{\left(\frac{y_{mi}-\mu}{\sigma}\right)} \qquad m = 1,2.$$

To account for explanatory variables, the mean term is rewritten as a function of covariates and their regression coefficients as in (4.5). The interval is now described by four coordinates $I_{jq} = \{(t_{1j-1}, t_{1j}], (t_{2q-1}, t_{2q}]\}$, thus probability of failure in interval I_{jq} is (j=1,2,...,r_1; q=1,2,...r_2)

$$\begin{aligned}\Delta_{jq} &= P(y_i \in I_{jq} \mid \mathbf{z}) \\ &= P\left(t_{ij-1} < T_1 \le t_{1j}, t_{2q-1} < T_2 \le t_{iq} \mid \mathbf{z}\right) \\ &= F(t_{1j-1}, t_{2q-1}) - F(t_{1j-1}, t_{2q}) - F(t_{1j}, t_{2q-1}) + F(t_{1j}, t_{2q}) \end{aligned} \tag{4.13}$$

Note that $F(t_{10}, t_{20}) = 0$. Hence the log likelihood will be

$$\ell(\boldsymbol{\theta} \mid \mathbf{z}) = \sum_{i=1}^{n} \log \sum_{j=1}^{r_1} \sum_{q=1}^{r_2} \omega_{ijq} \left[P(\Delta_{jq} \mid \boldsymbol{\theta}, \mathbf{z}_i)\right] \tag{4.14}$$

Meanwhile the Clayton Copula (3.33), which is more tractable compared to the Farlie-Morgensten copula, utilizes marginal survivor functions. Assume a Weibull marginal density, and use the Extreme Value distribution marginal

survival function from the transformation, as in (4.7). Then the Clayton copula survivor function is given as

$$S(y_{1j}, y_{2q}) = [(S(y_1)^{-\frac{1}{\kappa}} + (S(y_2)^{-\frac{1}{\kappa}} - 1]^{-\kappa}, \qquad (4.15)$$

resulting in failure probability in interval I_{jq} in terms of joint survival probabilities as follows, while log likelihood is denoted as in (4.14).

$$\Delta_{jq} = S(y_{1j-1}, y_{2q-1}) - S(y_{1j-1}, y_{2q}) - S(y_{1j}, y_{2q-1}) + S(y_{1j}, y_{2q}).$$

4.4 Chapter Summary

A relatively easier to use and involving fewer parameters, is the parametric method. Its only weakness is that one can never be sure of which distribution to assume for the data. A goodness-of-fit criterion test on interval-censored data with overlapping intervals is not yet available. Likelihood functions derived for the parametric model are the Weibull distribution based, and use of copulas with either a Farlie-Morgenstern or a Clayton distributions for bivariate data, all having Weibull marginal distributions.

CHAPTER 5

PARAMETER ESTIMATION

5.1 Bayesian Approach

5.1.1 Introduction

Bayesian prior probability of an event x is a person's degree of belief in that event, based on cumulative evidence and evidence based on practice. Whereas a classical probability is a tangible property of the world, (e.g. the probability that a coin will land head), a Bayesian probability is a property left entirely to the person who assigns the probability (e.g. your degree of belief that the coin will land heads). Thus to sum up, the classical probability is based on the concept of repeated trials while the degree of belief in an event is a Bayesian or personal probability. One clear distinction between the two is that to measure personal probability we do not need to think of repeated trials, which is an acceptable notion in Bayesian, contrary to classical probability. The problem with Bayesian prior probability is that these probabilities seem arbitrary. This is further compounded by fact that whereas it would be satisfying to assign probability one (zero) to an event that will occur (not), it becomes problematic to assign probabilities to beliefs which are not on the extreme.

For illustration, take the tossing of a fair coin, which either rest on its head or tail. Suppose we toss the coin N+1 times, making sure that the conditions on each toss remain identical. Base on the first N outcomes, we want to determine the probability of head being the outcome on the N+1 toss. In the classical

analysis of this problem, we assert that there is some unknown but fixed probability Δ, of head. We estimate this probability from the N observations using criteria such as low bias and low variance. Then we use this estimate Δ as our probability for heads on the $(N+1)^{th}$ toss. In the classical statistics, estimation of an unknown parameter from sample data can be done, among other procedures, by least square estimates, maximum likelihood estimation, etc. In the estimation of these parameters, any information relating to the unknown parameter prior to data collection is not used. Moreover, the unknown parameter is considered fixed, while the sample data of the response variable to be used in the estimation of the parameter is treated as random. Nothing in terms of a probability distribution is mentioned in reference to the parameter.

From a Bayesian perspective, there is no distinction between observations and parameters of a statistical model in that they are all viewed as random quantities. Hence, Bayesians assume that there is some belief about the probability of heads and that the unknown parameter is a random variable. Any privy information regarding the parameter that is known prior to data sampling must be used in the estimation. So the beliefs is combined with the results from the N tosses of the coin to estimate the probability on the $(N+1)^{th}$ toss. Viewing the parameter as a random variable implies that a distribution that is a function of the parameter exists. Thus the prior information relating to parameter and the information from the sample data enables one to estimate

this function of parameter, known as the posterior distribution. Since it is a probability distribution, it need conform to all aspects of a probability density function.

Suppose we denote random variables $X_1,....,X_N$ and assign a true state or value to each variable in a given set as $x_1,...,x_N$. Also denote $P(X=x \mid \theta)$ or $P(x \mid \theta)$ as the probability density function for x given the parameter θ. Let Θ be the parameter space. We express the uncertainty about θ using the probability density function $P(\theta \mid \chi)$, where χ is any related information. Let $D=\{X_1=x_1,X_2=x_2,...,X_N=x_N)$ denote a set of observations. Bayes' rule obtains the probability distribution for θ given D and χ,

$$P(\theta \mid D, \chi) = \frac{P(\theta|\chi)g(D|\theta,\chi)}{\int_{\theta\in\Theta} g(D|\theta,\chi)P(\theta|\chi)d\theta}. \qquad (5.1)$$

Both Bayesian and classical statisticians agree on the likelihood function $g(D|\theta,\chi) = \prod_{i=1}^{N} P(x_i|\theta)$. The denominator represents a normalizing constant, which for complicated likelihood functions, is not easy to evaluate, Gilks et al (1996). Any distributional characteristics of a posterior distribution are legitimate for Bayesian inference, e.g. moments, highest posterior regions, etc. The k^{th} moment of a function $f(\theta)$ is

$$E\{f(\theta)^k \mid D\} = \frac{\int P(\theta)g(D \mid \theta)f(\theta)^k d\theta}{\int_{\theta\in\Theta} P(\theta)g(D \mid \theta)d\theta}. \qquad (5.2)$$

Just as the Newton-Raphson or Least Square estimation methods enables a classical statistician to estimate a parameter for a given model, Bayesians use

the idea of Markov chain simulation to generate a random walk in the parameter space of θ which converges to a stationary distribution called joint posterior distribution as defined in (5.1). This requires evaluation of a normalising constant, thus multiple integration would be impractical if a large number of parameters are involved. With a reasonably large sample from a density we can approximate the mathematical form of that density via curve estimation or kernel density method (Izenman, (1991)) as observed by Smith and Gelfand (1992). Samples from a density whose functional form is implicitly defined can also be simulated. This method hinges largely on the large sample theory, in that the larger the sample of θ simulated from a posterior function whose normalizing constant cannot be ascertained become the more accurately we can recreate the posterior density. This method is called Markov chain Monte Carlo (MCMC) method, and there are a variety of these situational methods in place.

5.1.2 Prior Distributions

In this section we will discuss priors for various likelihood functions discussed in the previous chapters specifically for interval censored data. The likelihood function for a Clayton model with multiple interval-censored lifetimes for instance, given in (3.35), where **z** are fixed observed covariates and $\theta = \{\beta, \gamma_m, \kappa,\}$is a vector of unknown parameters with $\gamma_m = \{\gamma_{m1}, \gamma_{m2}, \ldots, \gamma_{mr_m}\}$, i=1,2,...,n; s<j=1,2,...,r_m; m=1,2,...,M. The Bayesian approach postulates that the

joint posterior distribution for the unknown parameters given data is proportional to the product of the likelihood and prior distribution of the parameter of interest.

Prior 1

To derive a prior for the conditional survival probability under nonparametric proportional hazard model (PHM), let P_{mj} from (3.9), with **z=0** follow a Uniform(0,1) distribution for all m,j. Taking the log(-log) link transform, then γ_{mj} has an Extreme Value (EV) distribution;

$$f(\gamma_{mj}) = e^{-e^{\gamma_{mj}}} e^{\gamma_{mj}} \qquad -\infty < \gamma_{mj} < \infty$$

$$F(\gamma_{mj}) = e^{-e^{\gamma_{mj}}}.$$

Allow probability density function (5.3) to be prior distribution for γ_{mj}, while the rest of the parameters assume diffuse or non-informative prior:

$$\pi(\beta) \propto k$$

$$\pi(\gamma_{mj}) = e^{-e^{\gamma_{mj}}} e^{\gamma_{mj}}. \qquad (5.3)$$

In a similar fashion for the nonparametric proportional odds model (POM), in the absence of covariate effects, let P_{mj} from (3.14) be i.i.d. Uniform(0,1). By taking a logit transformation,

$$\log\left(\frac{P_{mj}(0)}{1-P_{mj}(0)}\right) = \alpha_{mj},$$

we have that α_{mj} has a logistic probability density function.

$$\pi(\alpha_{mj}) = \frac{e^{\alpha_{mj}}}{(1+e^{\alpha_{mj}})^2} \qquad -\infty < \alpha_{mj} < \infty \qquad (5.4)$$

Priors (5.3) and (5.4) assume that all the baseline conditional survival probabilities are uniformly and independently distributed. The resultant posterior function from combining a likelihood function like (3.36) with prior distributions (5.3) or (5.4) is proportional to the posterior distribution for PHM;

$$P(\theta \mid D,z) \;\alpha \prod_{j=1}^{r} e^{-e^{\gamma_j}} e^{\gamma_j} \prod_{i=1}^{n} \sum_{j=1}^{r} \left\{ \left(1 - e^{-e^{\gamma_i + z_i\beta}} \right) \prod_{s=1}^{j-1} e^{-e^{\gamma_s + z_i\beta}} \right\}, \qquad (5.5a)$$

and for POM

$$P(\theta \mid D,z) \;\alpha \prod_{j=1}^{r} \frac{e^{\alpha_j}}{\left(1+e^{\alpha_j}\right)^2} \prod_{i=1}^{n} \sum_{j=1}^{r} \left\{ \left(\frac{e^{\alpha_j + z_i\beta}}{1+e^{\alpha_j + z_i\beta}} \right) \prod_{s=1}^{j-1} \frac{1}{1+e^{\alpha_s + z_i\beta}} \right\}. \qquad (5.5b)$$

The above functions are of the form that is intractable to evaluate analytically.

Prior II

In the previous section, non-informative priors were derived from the staring assumption that the conditional probabilities P_{mj}, are all independently Uniform(0,1) distributed. An Alternative would be, for the partitioned survival spaces of each lifetime, perceive the grouping of observations into intervals within which failure occurs, as similar to allotting n objects into r+1 bins with probabilities Δ_{mk} (k=1,2,...r+1) as defined in (3.1). Allotment probabilities in each lifetime are assumed independent. The distribution for a given lifetime is then Dirichlet with unknown hyper parameters δ_{mk}.

$$f(\Delta_{m1}, \Delta_{m2}, \ldots, \Delta_{mr+1}) = \frac{\Gamma\left(\sum_{k=1}^{r+1} \delta_{mk} \right)}{\prod_{k=1}^{r+1} \Gamma(\delta_{mk})} \Delta_{m1}^{\delta_{m1}-1} \Delta_{m2}^{\delta_{m2}-1} \ldots \Delta_{mr}^{\delta_{mr}-1} \left(1 - \sum_{k=1}^{r} \Delta_{mk}\right)^{\delta_{mr+1}-1}. \qquad (5.6)$$

For a Binomial distribution, the Beta distribution is a conjugate prior. Thus a multivariate generalization of Beta distribution or Dirichlet is a conjugate prior for a Multinomial, (Agresti (1990)). Defining the cumulative probability at j^{th} interval, $o_{mj} = \sum_{k=1}^{j} \Delta_{mk}$ $(o_{m1} < o_{m2} < \ldots < o_{mr} < 1)$, yields an ordered r-variate Dirichlet distribution (Wilks (1962)). The conditional probability of surviving beyond interval I_j given survival up to interval I_{j-1} in terms of cumulative probabilities is then

$$P_{mj} = \frac{1 - o_{mj}}{1 - o_{mj-1}}.$$

The marginal distribution of (o_{mj-1}, o_{mj}) is the ordered bivariate Dirichlet distribution:

$$f(o_{mj-1}, o_{mj}) = \frac{\Gamma(\sum_{k=1}^{r+1} \delta_{mk})}{\Gamma(\delta_{mk})\Gamma(\delta_{mj})\Gamma(\sum_{k=j+1}^{r+1} \delta_{mk})} o_{mj-1}^{\sum_{k=1}^{j-1} \delta_{mk} - 1} (o_{mj} - o_{mj-1})^{\delta_{mj}-1} (1 - o_{mj})^{\sum_{j=j+1}^{r+1} \delta_{mj} - 1}, \quad (5.7)$$

$$0 < o_{mj-1} < o_{mj} < 1.$$

Take a transform $P_{mj} = \frac{1 - o_{mj}}{1 - o_{mj-1}}$ and $W_{mj} = 1 - o_{mj}$, then the joint distribution of (P_{mj}, W_{mj}) is Dirichlet, while the marginal of P_{mj} reduces to a Beta$(\sum_{k=j+1}^{r+1} \delta_{mk}, \delta_{mj})$ distribution. Furthermore, the P_{mj}'s are independently distributed (j=1,2,...,r). The log-log transformation (3.9) maps the support of the parameters concerned from a probability space onto an open space, hence for multiple lifetimes, the prior for the baseline parameters is becomes

$$\pi(\gamma_{mj}) = \frac{1}{B(\sum_{k=j+1}^{r+1}\delta_{mk}, \delta_{mj})} e^{\gamma_{mj} - e^{\gamma_{mj}} \sum_{k=j+1}^{r+1}\delta_{mk}} (1 - e^{-e^{\gamma_{mj}}})^{\delta_{mj}-1}. \quad (5.8)$$

Assuming a Dirichlet prior distribution with $\delta_{mj} = 1$ for all m,j, then

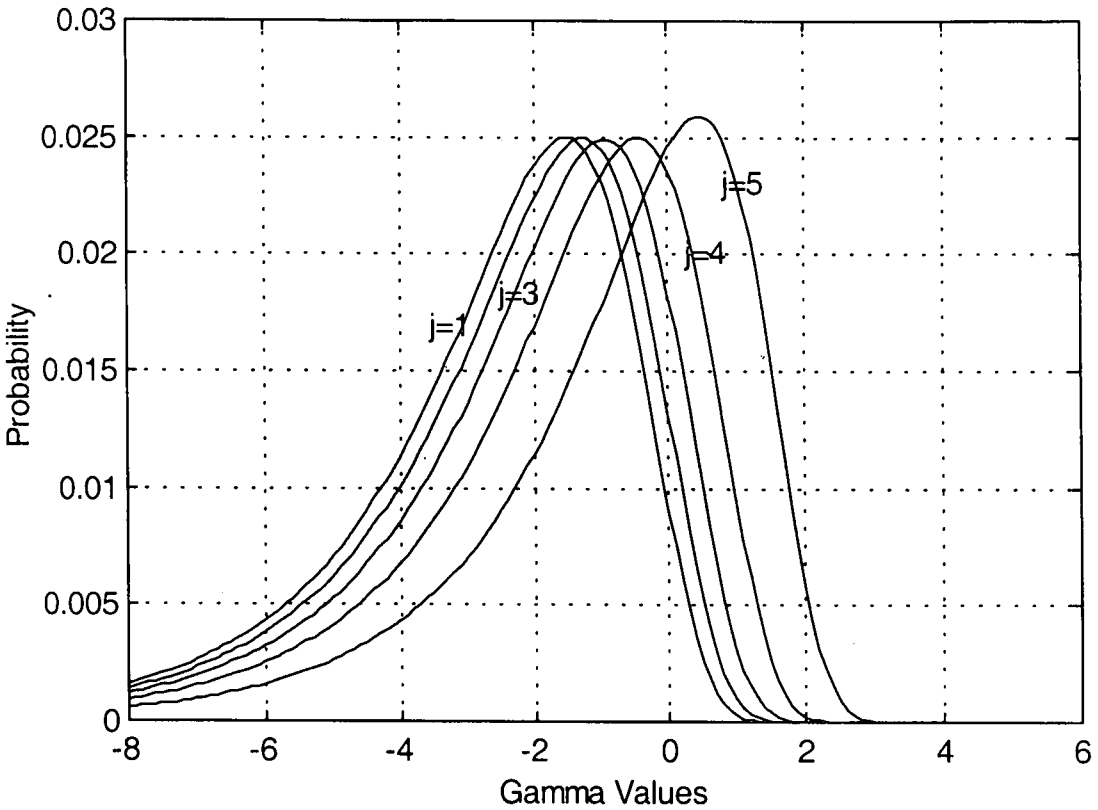
$$\pi(\gamma_{mj}) = (r+1-j)\ e^{\gamma_{mj} - (r+1-j)e^{\gamma_{mj}}}. \quad (5.9)$$

If $\delta_{mj} = 0.5$ for all m,j, a Jeffreys prior, then

$$\pi(\gamma_{mj}) = \frac{1}{B\left(\frac{r+1-j}{2}, \frac{1}{2}\right)} e^{\gamma_{mj} - 0.5(r+1-j)e^{\gamma_{mj}}} (1 - e^{-e^{\gamma_{mj}}})^{-0.5}. \quad (5.10)$$

The distribution of the r parameters is shown in Figure 5.1. All the distributions are negatively skewed, moving form left to right with increase in the interval number. From (5.9), the baseline parameter for the j^{th} interval has Extreme Value distribution. A transformation $U_{mj} = e^{\gamma_{mj}}$ results in an Exponential(r+1-j) distribution.

Figure 5.1: *The prior distributions for (r=5) baseline parameters under PHM with* $\delta_j=0.5$



In the same way, if a logit transformation of P_{mj} is effected for the POM model,

$$\pi(\alpha_{mj}) = \frac{1}{B(\sum_{k=j+1}^{r+1}\delta_{mk}, \delta_{mj})} \frac{e^{\delta_{mj}\alpha_{mj}}}{\left(1+e^{\alpha_{mj}}\right)^{\sum_{k=j}^{r+1}\delta_{mk}}}. \qquad (5.11)$$

If all hyper-parameters take the value 1, the prior distribution for the r^{th} interval reduces to a Logistic density. We used hyper-parameters of 0.5 for r intervals.

$$\pi(\alpha_{mj}) = \frac{1}{B\left(\frac{r+1-j}{2}, \frac{1}{2}\right)} \frac{e^{0.5\alpha_{mj}}}{\left(1+e^{\alpha_{mj}}\right)^{r\frac{r+2-j}{2}}}.$$

Figure 5.2: *The prior distributions for the (r=5) baseline parameters under POM with* $\delta_j = 0.5$

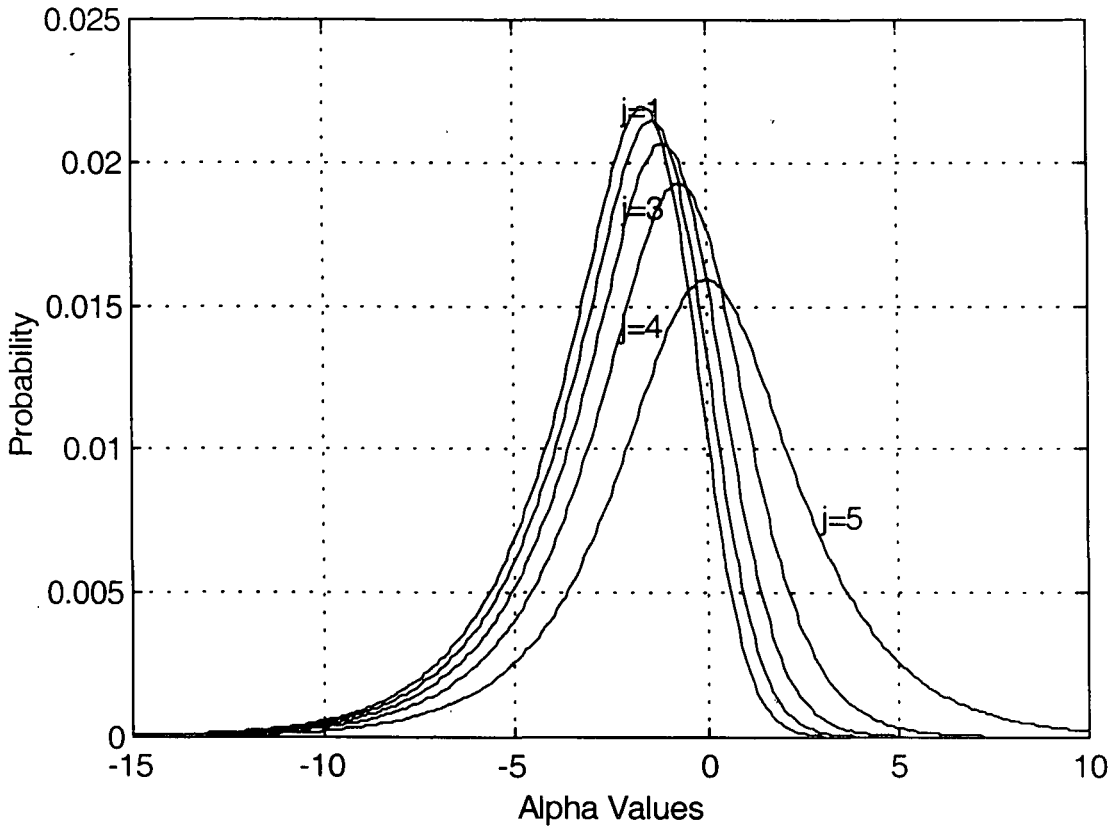


Figure 5.2 depicts the prior distribution from left to right for the smallest to the largest value of j respectively. The shape of the distributions for the first r-1 parameters is negatively skewed, but the distribution for the r^{th} parameter displays some symmetry around 0, thus the prior for the last interval yield a distribution similar to a Logistic density.

Importantly to note is that both Dirichlet priors assumes uniformly distributed failures over all intervals, resulting in a shift to the right in the means of the resultant baseline parameters as the interval number increases.

5.1.3 Metropolis-Hastings Algorithm

A Metropolis algorithm is an example of the MCMC methods introduced in section 5.1.1. Some posterior functions are of the form such that a normalizing constant of the joint posterior distribution cannot be analytically evaluated. Furthermore, if the conditional distribution function of the parameter involved is of non-standard form that does not enable drawing of sample estimates from it, a Metropolis type method is then applied to draw samples of parameter values. This method is premised on drawing a sequence of values of θ from a symmetric proposal density and comparing the posterior likelihood of the newly sampled parameter with that of the current parameter estimate. The new value is accepted as the updated parameter estimate if the posterior likelihood of the new value exceeds that of the current, otherwise is accepted with probability computed as the ratio of the posterior likelihoods. The following steps summarize the algorithm as follows: (Chib and Greenberg, (1995)).

- Start with a good approximate starting value θ^0 for which $P(\theta \mid y) > 0$
- Then sample a candidate point θ^* from a proposal distribution $f(\theta^* \mid \theta^{(t-1)})$ at time t. The jumping distribution must be symmetric about the current value of θ.
- Calculate the ratio

$$\varpi = \frac{P(\theta^* \mid \text{data})}{P(\theta^{t-1} \mid \text{data})} \tag{5.12}$$

and let

$$\theta^t = \begin{cases} \theta^* \text{ with probability } \min(\varpi, 1) \\ \theta^{t-1} \text{ otherwise} \end{cases} \tag{5.13}$$

The relative importance ratio ϖ will have to be computed for each (θ^*, θ).

- If $\varpi > 1$ then the estimate is automatically accepted. If $\varpi < 1$, then draw a U~Uniform(0,1) observation and accept θ^* if $U < \varpi$, thus jumping to higher posterior value. If the new value from proposal density, θ^* is such that $U > \varpi$, it will not be accepted hence set $\theta^* = \theta^{t-1}$. In summary, the importance ratio determines the direction of the jump, always accepting (jumping towards) parameter values that yield larger values of the posterior likelihood, yet accepting parameter values that yield smaller values of the posterior likelihood with a probability less than 1.

The rate at which new values are accepted is known as acceptance rate. A generalization of the Metropolis algorithm by doing away with symmetry in the jumping distribution results in the Metropolis-Hastings algorithm, Gelman et al (1995)). This increases the speed of the random walk. Combining the Metropolis-Hastings algorithm with the Gibbs sampling technique of cycling between sub-vectors (blocks) is another way of dealing with multidimensional problems. These blocks need be conditionally independent so that updates can be carried out simultaneously within a block, (Besag et al (1995)). One cycle of the MCMC algorithm consists of a visit to each block in random order.

The repeated sampling generates dependant sequence of random draws, which subjet to certain regularity conditions, eventually forgets the starting values, and converges to a stationary density, (Congdon, (2001)). This is called

convergence, and the rate at which this occurs is important. Some of the factors influencing the rate of convergence are:

- Sample size
- The way parameters are expressed.
- Complexity of the problem (number of parameters matters)
- Sampling scheme adopted (e.g. use of blocks helps).
- Closeness of the starting values to that of the stationary distribution.

For a joint posterior function that consists of several diversely distributed parameters, the use of a simultaneous simulation of the next parameter values in a given block, using a multivariate proposal distribution is recommended. Effective distributions suited for the proposal density are the ones that are symmetric about the current value of θ and have a spread similar to that of the marginal posterior for that variable. Rectangular and Gaussian distributions are therefore recommended. If the support of the parameter is restricted to an interval, a suitable transformation enables us to use the above-mentioned distributions. (Roberts et al, 1995).

5.1.4. Illustration with Example on Mango Data

Application I: Colosimo et al (2000) reported data on survival of mango trees in a complete randomised block design. By grafting scions on stocks, such that scions produce mango fruits while stocks serve as conductors of nutritional

requirements, the effect of scions and stocks on the resistance to disease was observed. Thus a randomised block experiment was conducted with treatments in a 6x7 factorial design involving the 6 varieties of scions and seven varieties of stocks. The duration of experiment was 20 years (1972 to 1992), wherein 12 visits were conducted to assess the status (dead/alive) of mango trees. Since all 12 visits were honoured, this resulted in interval-censored grouped data. Using conditional survival probability under PHM (3.9) and POM (3.14) models with a linear component of the form $\gamma + \beta_r \mathbf{z}_r + \delta_p \mathbf{z}_p + \eta_{rp} \mathbf{z}_{rp}$ were fitted for univariate lifetime of the mango trees. The $\mathbf{z}$'s are dummy variables taking value 1 or 0, the vectors γ is the baseline (interval) effect for either of the models, β_r, 1,2,...,5 is the scion effect, δ_p, 1,2,...,6 is the stock effect and η_{rp} is the interaction effect. Due to insignificant effect of stock and interaction on lifetime of mango trees, the results reported are on the covariate of scion varieties compared to a baseline variety Extrema .

Using the log likelihood of the form (3.15) for both PHM and POM models and logarithm of priors type I and II were applied resulting in log posterior function of form (5.5a and 5.5b). Since this is a non-standard model, a Metropolis-Hastings algorithm was invoked, and with Normal distribution as the proposal density. Samples of parameter values were drawn and the posterior likelihood of each drawn value compared with that of the current value. There were 12+5+6+30 parameters to be estimated. To speed up the process, simulation of parameter values was done in four blocks according to the effect. For a given

cycle, a multivariate normal density was used to simultaneously simulate all six parameters values for scion effect while the values of other parameters were held constant. 80 000 sets of values were generated.

Table 5.1: *Estimated posterior means of regression coefficients obtained by fitting PHM and POM models on mango data.*

			PHM		POM	
Prior	**Varieties**	**Parameter**	**Mean**	**95% HPD**	**Mean**	**95% HPD**
I	Oliviera	β_1	-0.8793	**-1.46, -0.31**	-0.9407	**-1.53, -0.36**
	Pahiri	β_2	-0.4133	-0.94, 0.10	-0.4061	-0.96, 0.16
	Imperial	β_3	-0.6044	**-1.13, -0.09**	-0.6361	**-1.20, -0.06**
	Carlota	β_4	-0.6176	**-1.16, -0.06**	-0.6555	**-1.25, -0.11**
	Bourbon	β_5	0.4635	-0.03, 0.96	0.4467	-0.07, 0.97
II	Oliviera	β_1	-05704	**-1.14, -0.06**	-0.7520	**-1.35, -0.21**
	Pahiri	β_2	-0.1136	-0.62, 0.38	-0.2380	-0.83, 0.32
	Imperial	β_3	-0.3141	-0.85, 0.18	-0.4606	-1.05, 0.11
	Carlota	β_4	-0.3286	-086, 0.18	-0.4784	-1.09, 0.07
	Bourbon	β_5	0.7277	**0.24, 1.21**	0.6275	**0.10, 1.15**

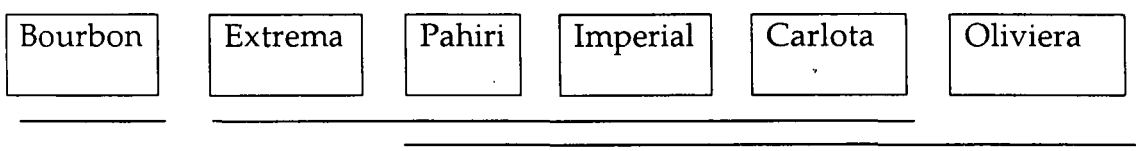
Results (Table 5.1) show that Extrema variety is not affected differently from Pahiri, but is significantly different from Oliviera variety for all models used. Whereas the POM and PHM methods do not yield different results, the type of prior used does matter. Using Prior II for both models shows that Oliviera variety has significantly higher resistance than Extrema, Bourbon has lower resistance, while the other varieties are not affected differently. If prior I is used, Oliviera, Imperial and Carlota varieties have higher resistance than Extrema

variety, while the rest are not affected differently from Extrema variety. This goes to show that the results are still sensitive to the prior.

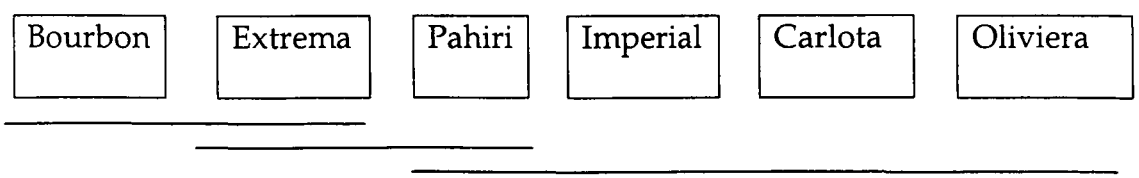
The varieties have varying resistance in the following sequence, Oliviera has the highest resistance, followed by Carlota, Imperial, Pahiri, Extrema and Bourbon is the most susceptible. Moreover, a 95% least significant difference (LSD) test on results obtained from PHM with prior II (Figure 5.3a) shows that Bourbon is different from the rest of the varieties. Whereas Oliviera and Extrema are significantly different on individual basis, these varieties are not different from Pahiri, Imperial and Carlota varieties. A POM with prior II return similar results. Meanwhile a 95% LSD test on the effect of prior I on PHM (Figure 5.3b) show that the mean difference between Bourbon and Extrema is not significant. Yet, Extrema is not different from Pahiri, which itself is different from Bourbon, but not from the rest of the varieties.

Figure 5.3: *Showing least significant difference between varieties for PHM with prior II (a) and with prior I (b).*

a)

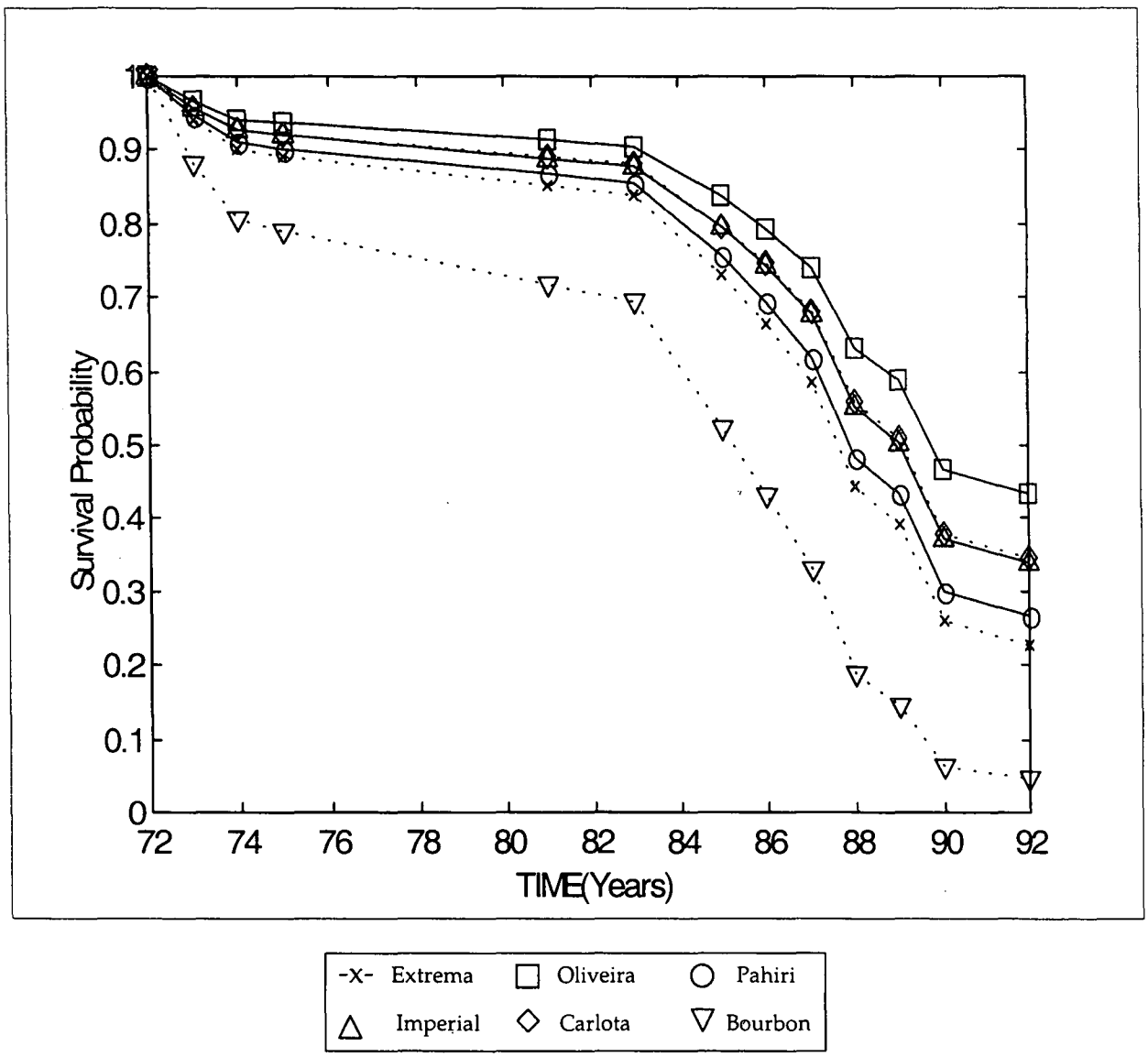


b)



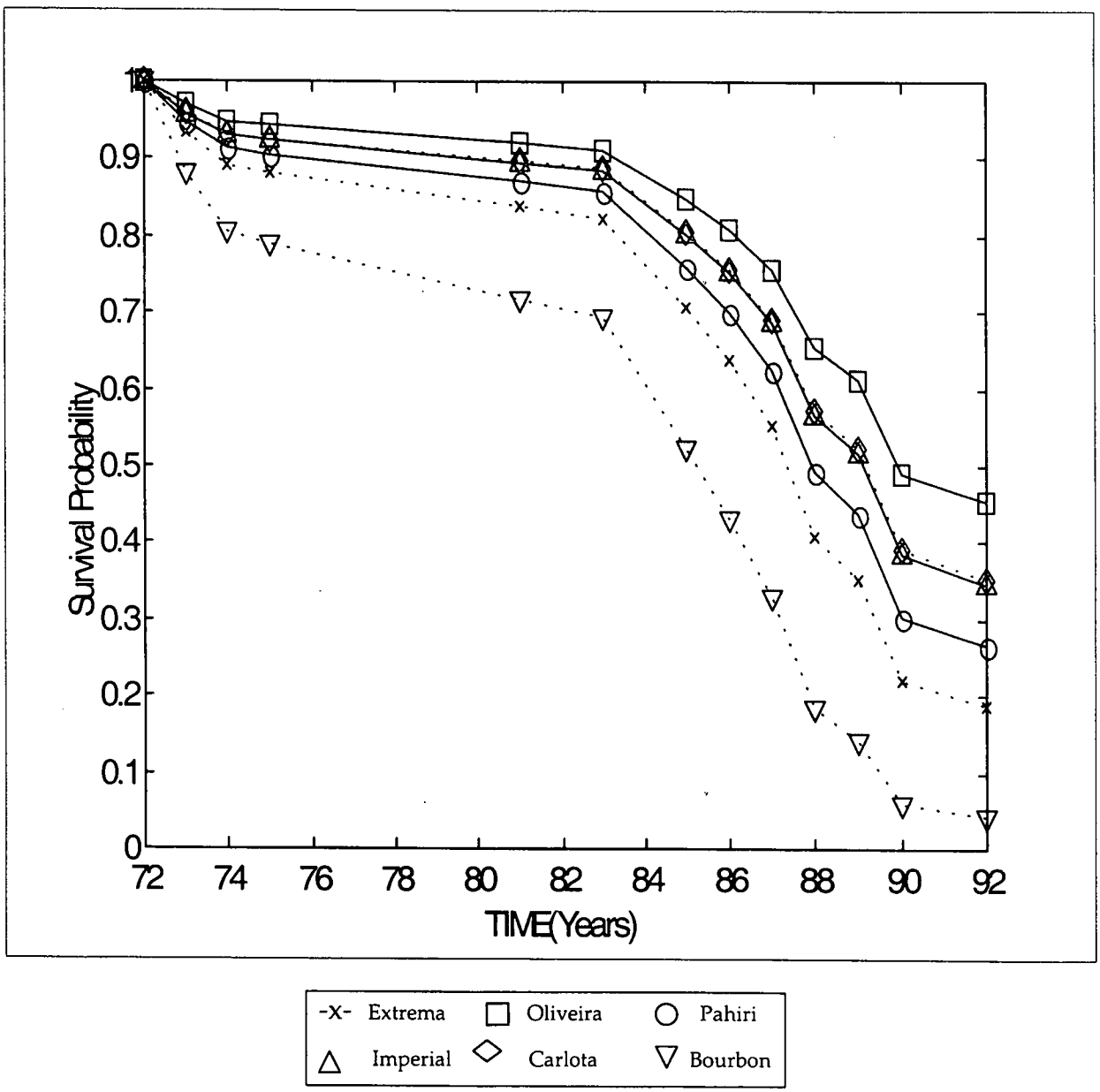
Curves for estimated survival functions at each time interval for each of the six varieties are shown in Figures 5.4 and 5.5. The survival probabilities for Bourbon rapidly decline to 0.80 in the first two years of study. By the end of study, it had reached 7% level, while that of the most resistant variety is around 40% at the end of study.

Figure 5.4. *Estimated survival functions for PHM: prior II*



Colosimo et al (2000), reported similar results obtained using the method of maximum likelihood estimation on the two models of PHM and POM. Likewise, in their results, the two models did not yield varying results.

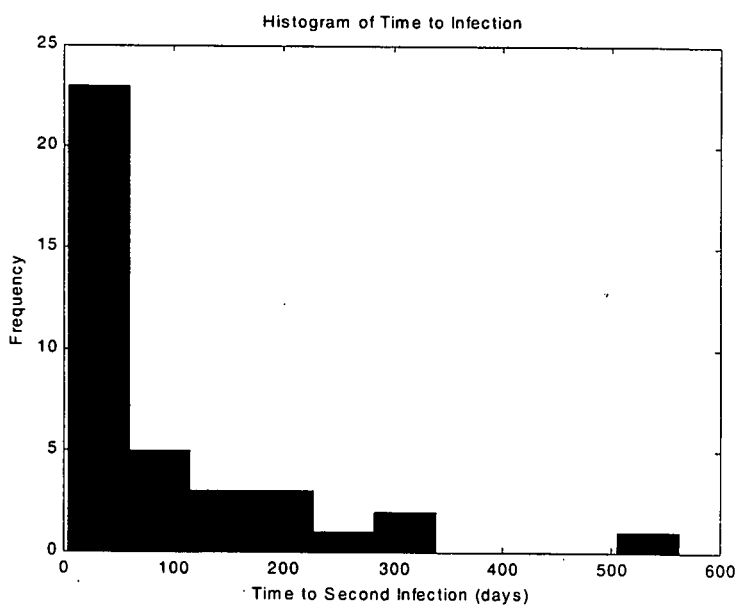
Figure 5.5 *Estimated survival functions for POM: prior II*



5.1.5 Illustration with Example on Bivariate Kidney Data

Application II: McGilchrist and Aisbett (1991) presented data on recurrence times to infections at point of insertion of the catheter (in days) on 38 kidney patients using a portable dialysis machine. Two exact right-censored times for each patient, their age and gender (0-male, 1-female) were reported. Also reported is whether a patient had disease type AN, GN or PKD. To analyse this data we assume pseudo-checking times at 30, 90, 180, 365 and 600 days for each of the infection times to create interval data. Thus data is analysed as bivariate grouped interval-censored. (see Klein and Moeschberger (1997) for data). Methods used to analyse this data included the Independence assumption (IW) model (3.26), Conditional Bivariate (CB) model (3.32) and Clayton copula (CC) model (3.34) and (3.35). These are all nonparametric methods, and for each both the PHM and POM transform were used.

Figure 5.6: *A frequency histogram of Time-to-Second infection*



Assuming a Weibull distribution for both infection lifetimes, parametric methods of Independence assumption (IWW) (4.12) and Clayton copula (4.15) and (4.14) giving (CCW) were used, (Weiclapos.m in Appendix B). For all the nonparametric models the linear component of the conditional survival probabilities is given as $\gamma_{mj} + \beta_1(age) + \beta_2(gender) + \beta_3(GN) + \beta_4(AN) + \beta_5(PKD)$ and for the parametric models, $\psi_{m0} + \psi_1(age) + \psi_2(gender) + \psi_3(GN) + \psi_4(AN) + \psi_5(PKD)$, m=1,2 and j=1,2,...,5.

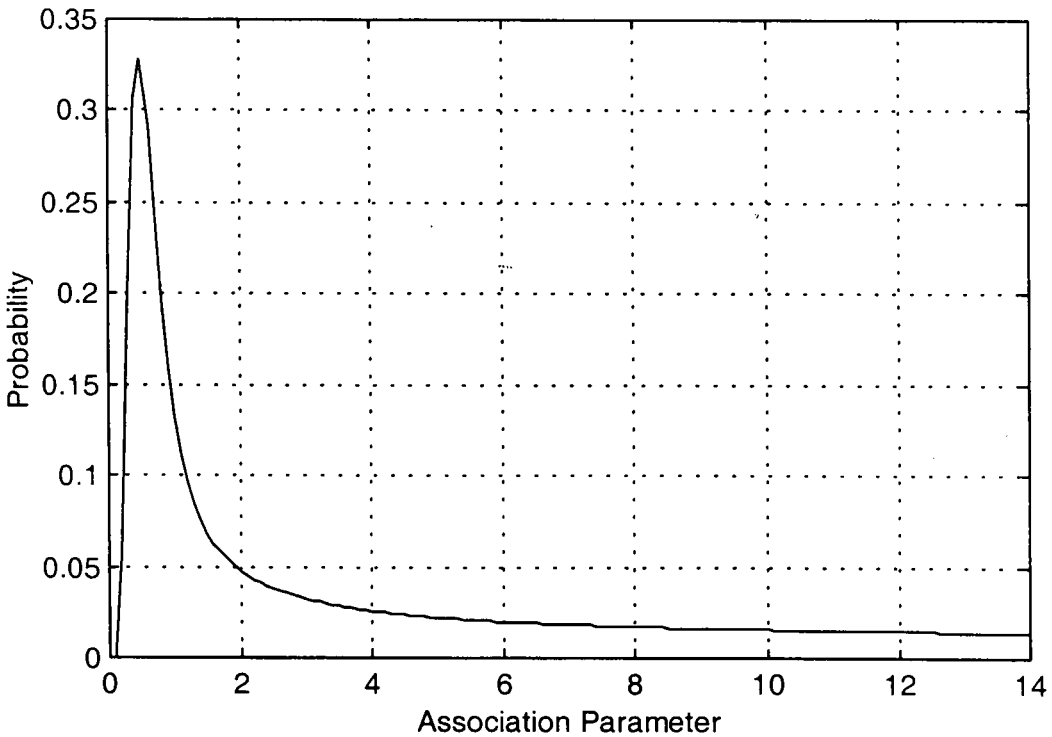
Table 5.2: *Estimated posterior means from fitting Independence assumption (IWW) and Clayton Copula (CCW) models assuming Weibull distributed marginals on kidney data.*

		CCW		IWW	
Variable	**Parameter**	**Estimate**	**95% HPD**	**Estimate**	**95% HPD**
Intercept1	ψ_{01}	6.1978	5.25, 7.38	4.3578	3.53, 5.98
Intercept2	ψ_{02}	6.7344	3.76, 6.19	6.8932	6.20, 7.47
Age	ψ_1	0.0124	-0.01, 0.05	-0.0037	-0.01, 0.00
Gender	ψ_2	-3.5991	**-5.62, -1.44**	-2.8035	**-3.24, -2.39**
GN	ψ_3	-0.8007	-3.31, 1.77	0.0254	-0.17, 0.16
AN	ψ_4	-1.5674	-4.24, 0.61	-0.4798	-1.11, 1.16
PKD	ψ_5	0.7646	-2.43, 3.71	1.7917	**1.23, 2.44**
Sigma	σ	2.9511	1.62, 4.55	1.5411	1.19, 2.03
Assoc	κ	2.9018	0.69, 4.55		

The parametric method was executed after doing a crude goodness of fit test on the interval data to determine if data was Weibull distributed. For first and second times to infection, χ^2 values of 52.13 and 5.13 respectively with 35 degrees of freedom were obtained. The values show that whereas data for time-to-second infection is distributed as Weibull, the first one is not. For analysis we

assumed that both infection times follow a Weibull distribution. The next characteristic of these two failure-time survival data is the degree of dependence.

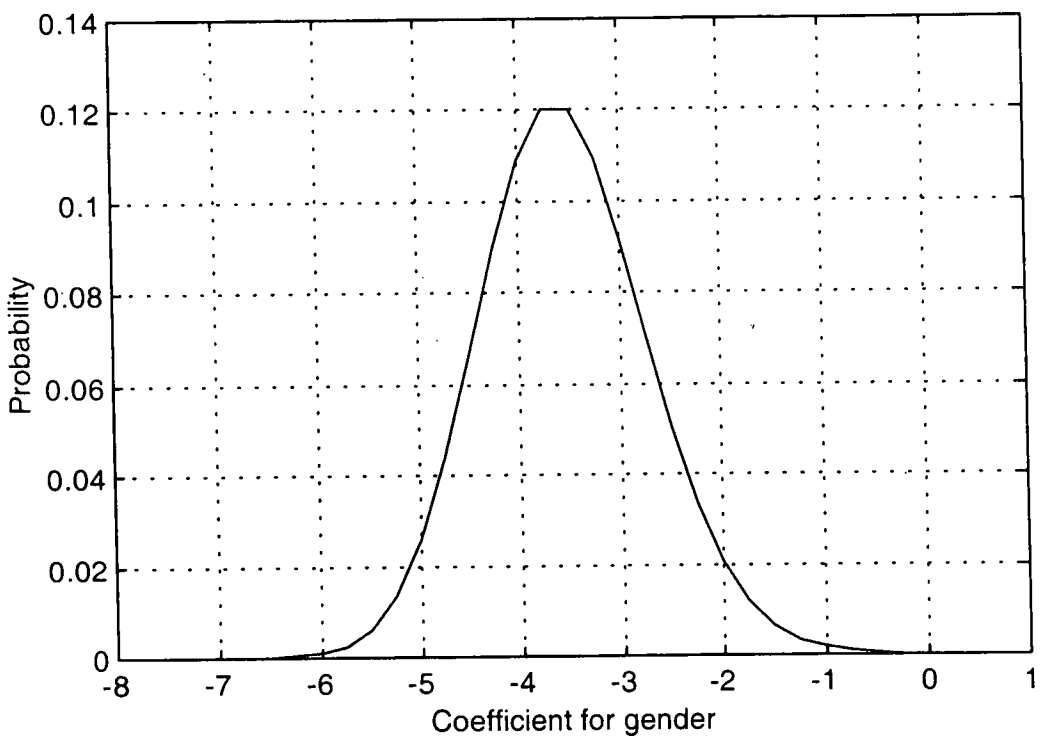
Figure 5.7: *Posterior distribution of association parameter κ in the Clayton model with Weibull marginals.*



The measure of association obtained from assuming a Weibull distribution is ($\hat{\kappa}$=2.90), while the non-parametric Clayton copula methods reveal a dependency measure of about 2.9 (Tables 5.2a and 5.2b) between the infection times. This show a not so strong relation, hence we expect the independence working assumption method to also do well in estimating the parameters. The crude estimated mean failure time ($\hat{\mu} = \exp(\psi_{m0} + \boldsymbol{\psi}\mathbf{z})$) for a 50 years old female patient with PKD disease, is $\hat{\mu}_1$ = 178 days to first infection and $\hat{\mu}_2$ = 305 days to

second infection. The delta method using $\hat{\sigma} = 2.95$ as shown by Worku (1997) and Rice (1995) is more precise, giving expected failure times to be 208 and 353 days for first and second infection respectively. The two models shows that gender has a significant effect on time to infection.

Figure 5.8: *Posterior distribution of parameter for gender in the Clayton model with Weibull marginals.*



When using a Clayton copula model, only gender is significant. Meanwhile, assuming independence between time to first and second infection, variables gender and disease type PKD, are significant, apparently due to unstable nature of the method when the sample is small, and ignoring the dependence effect. Using separate coefficients for the parameter for each failure time, the effect of gender on the two infection times varies, (table not included) with females

($\hat{\psi}_2$=-3.18) at higher risk towards first infection and male ($\hat{\psi}_2$=0.1669) towards time to second infection, yet the combined effect shows females are at higher risk for both times, lowering the expected infection times, (see Figure 5.9).

Table 5.2a: *Estimated posterior means from fitting Independence working (IWH), Bivariate Conditional (CBH) and Clayton Copula (CCH) models using proportional hazard on kidney data.*

		IWH		CBH		CCH	
Prior	**Var**	**Mean**	**95% HPD**	**Mean**	**95% HPD**	**Mean**	**95% HPD**
I	Age	-0.003	-0.01, 0.00	0.0003	-0.003, 0.003	0.003	-0.004, 0.01
	Gender	2.228	**1.90, 2.61**	1.793	**1.43, 2.01**	2.032	**1.51, 2.37**
	GN	0.185	**0.02, 0.28**	0.125	-0.06, 0.26	0.105	-0.04, 0.20
	AN	-0.004	-0.04, 0.03	0.101	**-0.03, 0.32**	-0.185	-0.33, 0.00
	PKD	-0.088	-0.15, 0.00	0.017	-0.11, 0.12	0.016	-0.06, 0.10
	Assoc					2.70	**2.55, 2.83**
II	Age	0.004	-0.01, 0.01	0.001	-0.001, 0.01	0.013	-0.00, 0.02
	Gender	1.870	**1.70, 2.06**	1.082	**0.76, 1.42**	2.041	**1.82, 2.22**
	GN	-0.434	**-0.75,-0.01**	0.272	**0.00, 0.50**	0.100	-0.02, 0.23
	AN	0.050	-0.00, 0.10	-0.123	-0.26, 0.07	0.103	-0.02, 0.19
	PKD	0.044	-0.08, 0.21	0.056	-0.04, 0.14	0.031	-0.12, 0.14
	Assoc					2.97	**2.74, 3.20**

Results in Table 5.2a and 5.2b shows that the estimated values of variable gender are significant for all three non-parametric methods used with both proportional odds and proportional hazard models. Hougaard (1987) made similar conclusions on this data, analysed using a frailty model. Under the independence assumption model (IW), an extra variable is significant apart from gender. For instance, both priors I and II used with PH, shows that disease type GN is significant, though the effects are inversely related. Meanwhile, with the same model using proportional odds, Age, gender and disease type PKD are significant for prior II. The Bivariate Conditional model (conditioned on

failure interval of first time to infection) shows that disease type GN has effect on time when a PH is used with prior I.

Table 5.2b: *Estimated posterior means from fitting Independence working (IWO), a Bivariate Conditional (CBO) and Clayton Copula (CCO) models using proportional odds on kidney data.*

		IWO		CBO		CCO	
Prior	**Var**	**Mean**	**95% HPD**	**Mean**	**95% HPD**	**Mean**	**95% HPD**
I	Age	-0.005	-0.02, 0.00	-0.002	-0.01,0.00	0.002	-0.00, 0.01
	Gender	1.987	**1.69, 2.42**	1.380	**1.26,1.53**	1.812	**1.72, 1.96**
	GN	-0.017	-0.13, 0.16	0.197	-0.02,0.43	-0.039	-0.31, 0.24
	AN	-0.083	-0.20, 0.04	0.051	-0.08,0.21	0.154	-0.21, 0.71
	PKD	0.095	-0.00, 0.21	0.154	-0.01,0.31	-0.243	-0.86, 0.29
	Assoc					2.78	**2.62, 2.89**
II	Age	0.016	**0.00, 0.03**	-0.0004	-0.01,0.00	-0.005	-0.00, 0.003
	Gender	2.033	**1.77, 2.36**	1.680	**1.45,1.85**	1.682	**1.63, 1.75**
	GN	-0.061	-0.33, 0.20	-0.019	-0.38,0.22	0.060	-0.26, 0.52
	AN	0.001	-0.07, 0.07	0.099	-0.01,0.21	-0.637	-1.64, 0.09
	PKD	0.280	**0.05, 0.45**	-0.116	-0.21,0.03	0.133	-0.38, 0.76
	Assoc					2.91	**2.77, 3.06**

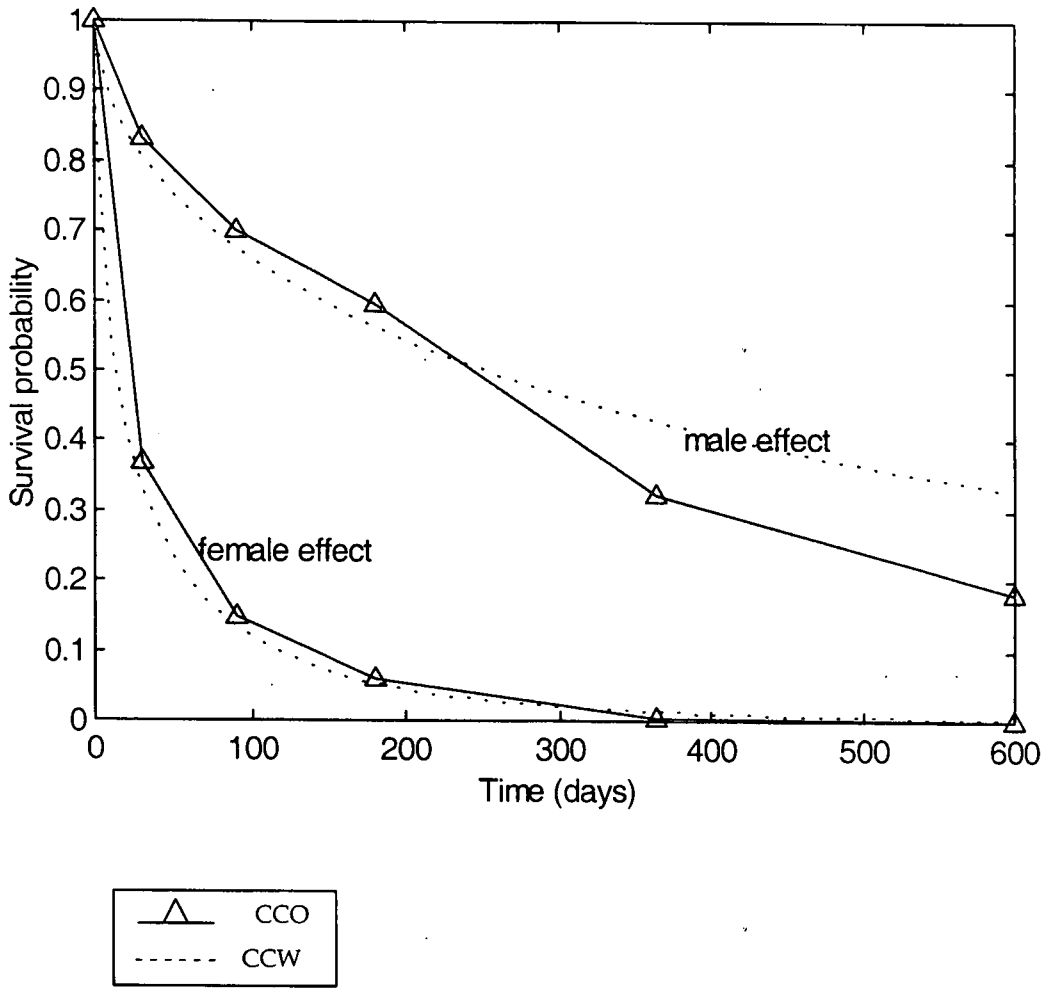
Table 5.3 below depicts the baseline conditional survival probabilities for the five intervals of first failure time. The results in table 5.3 show that IW yields higher survival probabilities compared to CB and CC, which seem to have comparable results. The quick drop in survival for time-to-first infection is apparent and is rapid in the first interval, at 68% and 53% for the CB and CC models respectively. These decline rates in survivals are slow compared to the rate for time-to-second infection.

Table 5.3: *Posterior mean estimates of baseline conditional survival probabilities of the first lifetime: Prior II*

		CB		*IW*		*CC*	
	Prob.	*Mean*	*95% HPD*	*Mean*	*95%PHD*	*Mean*	*95% HPD*
PH	P_{11}	0.6829	0.19,0.95	0.8350	0.71,0.92	0.5299	0.05,0.89
	P_{12}	0.6697	0.46,0.58	0.8458	0.77,0.92	0.8749	0.73,0.95
	P_{13}	0.7623	0.65,0.84	0.8090	0.68,0.90	0.8657	0.74,0.93
	P_{14}	0.2410	0.04,0.64	0.5927	0.39,0.71	0.7433	0.42,0.93
	P_{15}	0.0013	0.00,0.02	0.0915	0.00,0.24	0.0308	0.00,0.16
PO	P_{11}	0.5694	0.38,0.71	0.9019	0.84,0.95	0.8309	0.74,0.88
	P_{12}	0.8598	0.69,0.93	0.9117	0.88,0.93	0.8445	0.80,0.88
	P_{13}	0.6462	0.44,0.80	0.8414	0.74,0.89	0.8463	0.79,0.89
	P_{14}	0.6034	0.45,0.82	0.6436	0.54,0.71	0.5449	0.33,0.67
	P_{15}	0.1044	0.05,0.17	0.3497	0.16,0.54	0.5561	0.29,0.76

The non-parametric methods appear to return good estimates of survival probabilities for time to second infection (see figure 5.7), whereas assuming a Weibull distribution for the time to infection seem to only do well at the end-points. This is explained by the fact that a single baseline parameter is used to estimate survival probabilities at all intervals, whereas interval specific baseline parameter estimates are used in non-parametric estimation, hence more precision. For both models, the gender effect is in the same direction. For example, after surviving for 90 days without infection, males have more that 60% chance of being free from infection during the next 90 days, while females have less than 20% chance.

Figure 5.9: *Estimated survival probabilities obtained using CCW (dotted) and CCO methods for second infection.*



5.2 Classical Approach: MLE for Interval-Censored Data

Given a likelihood function $g(D \mid \theta)$, which is a function of θ, we need to determine a vector (p+1) of estimators $\hat{\theta}$ of θ taking on values in Θ, the parameter space. The values $\hat{\theta}$ which assigns the largest possible value to $g(D \mid \theta)$, so that they provide the best explanation of the observed values and thus are a natural estimator of θ are called Maximum Likelihood Estimators

(MLE's) if they exist. Instead of maximizing the likelihood $g(D \mid \boldsymbol{\theta})$, it is convenient (yielding same results) to maximize the log-likelihood, $\ell(\boldsymbol{\theta})$. Under some mild regularity conditions, $\hat{\boldsymbol{\theta}}$ maximizes $\ell(\boldsymbol{\theta})$ if the expectation of first derivative, also called the score equation, $E[\ell'(\boldsymbol{\theta})]=0$, and the second derivative $\ell''(\boldsymbol{\theta}) < 0$. Furthermore if the regularity conditions hold, then there exists a sequence of $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}_n(x_1, x_2, \ldots x_n)$ of local maxima of the log-likelihood function which is consistent, such that

$$\hat{\boldsymbol{\theta}}_n \xrightarrow{p} \boldsymbol{\theta} \text{ for all } \theta \in \Theta$$

and with probability tending to 1 as $n \to \infty$, is the MLE, (proof Lehmann (1999)). The local maxima are determined by setting the score function equal to 0.

$$q = \left[\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_0} \quad \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1} \ldots \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_{p+1}} \right] = \mathbf{0} \tag{5.13}$$

The difficulty arises in the choosing a local maximum when there exists several local maximum. One of the remedies to this dilemma is the use of the Newton-Raphson iterative method. It leads to estimators which are not exact roots of the log-likelihood but which have the same asymptotic behaviour as $\hat{\boldsymbol{\theta}}_n$. In this case the score equation of the log-likelihood $\ell'_n(\boldsymbol{\theta})$ is replaced by the linear term of its Taylor expansion about a starting value θ_n^0, and therefore replaces the log-likelihood score equation with the equation

$$\ell'_n(\theta_n^0) + (\theta - \theta_n^0)\ell''_n(\theta_n^0) = 0, \tag{5.14}$$

where

$$\ell''(\boldsymbol{\theta}) = \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_v^2} \right]$$

and

$$H(\boldsymbol{\theta}) = \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\theta_v \partial\theta_{v'}}\right], \qquad v',v=1,2,\ldots p \tag{5.15}$$

This suggests the solution of the equation (5.14) for θ to be

$$\tilde{\theta}_n = \theta_n^0 - H^{-1}(\theta_n^0) q_n(\theta_n^0), \tag{5.16}$$

as a first approximation to the solution of the log-likelihood equation. The solution in (5.16), constitute the Newton-Raphson method. The procedure is then iterated by replacing θ_n^0 by $\tilde{\theta}_n$ and continued until a convergence is achieved. It can be shown that the starting point as an estimator of θ is both consistent and $\sqrt{n}\left(\theta_n^0 - \theta\right)$ is bounded in probability.

Due to the complicated nature of the log-likelihood function for the CB, CC or CWM models, the MLE method could not be used for models seeking to encompass the dependence structure. For a single lifetime with overlapping intervals, the log-likelihood under PHM as denoted by (3.16), with a p explanatory variables, let $\boldsymbol{\theta}=\{\boldsymbol{\beta}, \boldsymbol{\gamma}\}$ be a r+p vector of unknown parameters and $H(\boldsymbol{\theta})$ be the negative definite r+p Hessian matrix of mixed second partial derivatives of $\ell(\boldsymbol{\beta},\boldsymbol{\gamma})$ defined by,

$$H\left(\tilde{\beta}_v\right) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial\beta_v \partial\beta_{v'}}\Big|_{\beta=\tilde{\beta}}, \qquad v',v= 1,2,\ldots,r+p;$$

and a vector of first derivatives

$$q(\tilde{\beta}_v) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial\beta_v}\Big|_{\beta_v=\tilde{\beta}_v}.$$

The MLE's are iteratively estimated as

$$\tilde{\beta}_{t+1} = \tilde{\beta}_t - H^{-1}(\tilde{\beta}_t)q(\tilde{\beta}_t),$$

the idea being to search for a global maximum point which is approximated when

$$\sum_{i=1}^{n} \frac{\partial}{\partial \beta_v} \log\left(g(\mathbf{z}_i, \boldsymbol{\beta})\right) = 0 \qquad \text{for all } \beta_v \quad v=1,2,\ldots,r+p+1$$

(Lehmann, (1999)). The above method derives the parameter estimates under the assumption of independent lifetimes. Independence is a necessary condition for estimation of MLE's, and the inverse of H, the Fisher information evaluated at the estimated $\tilde{\beta}$ value of parameter value is the estimate of variance of the parameter estimate, (Casella and Berger, (1990)). $\tilde{\beta}_{(t+1)}$ converge to the ML estimates $\hat{\beta}$ and $H(\tilde{\beta})$ converges to the matrix $\hat{H}$ which, when inverted will be the covariance matrix of $\hat{\beta}$ Agresti (1990).

Generally the log likelihood i^{th} unit for IWH model with disjoint intervals is given by (3.15). Guo and Lin (1994) presented the first and second derivatives for this model, (see appendix A1). If the intervals are overlapping the log-likelihood is written as (3.16). The derivatives will be dependent on the size of the endpoints (L_i, R_i), so let u_i = number of intervals contained in (L_i, R_i) and $h_j(\mathbf{z}) = e^{\gamma_j + \boldsymbol{\beta}\mathbf{z}}$. If a unit fails within the endpoints then, $\{ j \in (L_i, R_i) \}$ and s<j. Note that if a unit's observation is censored, then $R_i = r$. The first and second derivatives for model (3.27) are then as follows:

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_s} = -\sum_{i=1}^{n} h_s(z_i)$$

$$\frac{\partial^2 \ell(\beta,\gamma)}{\partial \gamma_s^2} = -\sum_{i=1}^{n} h_{ms}(z_i)$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_j} = \sum_{i=1}^{n} \frac{h_l(z_i)\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}}{\left(1-\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\right)}$$

$$\frac{\partial^2 \ell(\beta,\gamma)}{\partial \gamma_j^2} = \sum_{i=1}^{n} \frac{h_j(z_i)\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\left(1-\prod_{j=L_i}^{R_i} e^{-h_j(z_i)} - h_j(z_i)\right)}{\left(1-\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\right)^2}$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \beta_l} = \sum_{i=1}^{n} z_i \left\{ \frac{\prod_{j=L_i}^{R_i} h_j(z_i)\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}}{\left(1-\prod_{j=L_i}^{R_i} e^{-h_v(z_i)}\right)} - \sum_{s=1}^{L_i-1} h_s(z_i) \right\}$$

$$\frac{\partial^2 \ell(\beta,\gamma)}{\partial \beta_l^{\,2}} = \sum_{i=1}^{n} z_i \grave{} z_i \left[\frac{\prod_{L_i}^{R_i} h_j(z_i)\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\left\{u_i\left(1-\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\right) - \sum_{j=L_i}^{R_i} h_j(z_i)\right\}}{\left(1-\prod_{j=L_i}^{R_i} e^{-h_j(z_i)}\right)^2} - \sum_{s=1}^{L_i-1} h_j(z_i) \right]$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_j \partial \gamma_s} = 0$$

$$\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_i \partial \gamma_s} = -\sum_{i=1}^{n} z_i h_s(z_i)$$

$$\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_i \partial \gamma_j} = \sum_{i=1}^{n} z_i \left\{ \frac{h_j(z_i) \prod_{j=L_i}^{R_i} e^{-h_j(z_i)} \left(1 - \prod_{j=L_i}^{R_i} e^{-h_j(z_i)} - h_j(z_i) \right)}{\left(1 - \prod_{j=L_i}^{R_i} e^{-h_j(z)} \right)^2} \right\}$$

The above result is applied if there is one lifetime. Under the independence assumption, the derivatives can be extended to multiple failure times situation by summing over the m lifetimes involved.

5.2.1 Univariate lifetime with Illustrative Example: MLE

Application III: Using Kidney failure data example, MLE's for each univariate lifetime were computed assuming that data is Weibull distributed. A SAS program was used, and the following are the results obtained. Note: SAS can only handle one lifetime.

Table 5.4: *Results of MLE's obtained from Weibull distributed univariate kidney data.*

	First Infection Time			Second Infection Time		
Variable	**MLE**	**STD**	**P-val**	**MLE**	**STD**	**P-val**
Intercept	5.77	0.555	0.0001	4.832	0.657	0.0001
Age	-0.0047	0.013	0.7203	0.0007	0.016	0.9964
Gender	-2.9617	0.587	0.0001	-1.028	0.582	0.0771
GN	-0.2315	0.489	0.6357	0.069	0.653	0.9161
AN	-0.8898	0.449	0.0474	0.436	0.622	0.4833
PKD	1.3938	0.771	0.0708	1.242	0.850	0.1499
Assoc.	0.86	0.23	0.0001	1.06	0.44	0.008

In this case intercepts are significant for both infection times. Also significant are variables gender and disease AN, for time to first infection and none for second infection. These results compares well with the ones obtained using posterior means from Table 5.2, with minor variation in that gender is not significant for second infection. Apparently MLE method has conservative confidence intervals compared to the posterior mean's HPD.

5.2.2 Bivariate Lifetimes with Illustrative Example: MLE .

Ignoring the dependence component between the lifetimes, results in unstable variance estimates for the parameter estimates, which only stabilize when the sample size is large (see Chapter 7). A robust estimator of variance of parameters, which is an attempt to address the correlated effect between failure types, is by introducing a matrix **D** composed of products of first derivatives whose off-diagonal terms are nonzero, unlike the Information matrix. D is used to build an estimator called sandwich estimator as follows: (Guo and Lin, (1994)),

$$D_n = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \log f(\mathbf{z}_i,\boldsymbol{\beta})}{\partial \beta_v} \cdot \frac{\partial \log f(\mathbf{z}_i,\boldsymbol{\beta})}{\partial \beta_{v'}}, \qquad v,v'=1,2,\ldots,p \tag{5.18}$$

White (1982) showed that for large n, $D_n \xrightarrow{as} D$, where

$$D = E\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial \log f(\mathbf{z}_i,\boldsymbol{\beta})}{\partial \beta_v} \frac{\partial \log f(\mathbf{z}_i,\boldsymbol{\beta})}{\partial \beta_{v'}} \right). \tag{5.19}$$

An improved robust estimator of variance for the estimated parameters which accounts for the correlated failure type outcomes is given by

$$A_n = H_n^{-1} D_n H_n^{-1}. \tag{5.20}$$

where H_n is as in (5.15)

Application IV: Using Kidney data, maximum likelihood estimates are computed using the two lifetimes under the independence working assumption. To adjust for dependence between lifetimes, the variance estimate is computed using a sandwich estimator. Only results on gender are reported. Gender is the only significant factor towards infection, with female $\hat{\beta}_2 = 1.0050$ (p-value=0.0012) hence a shorter lifetime towards infection for both Naïve and Sandwich variance estimator. The sandwich estimator improves and stabilizes the variance estimation for baseline parameters (Table 5.5). Note that, for second infection at interval 5, the variance estimate is 4.93, which is improved to be 0.4786 by sandwich estimator, (Indmle.m in Appendix B).

Table 5.5: *Showing the MLE's for baseline parameters using a IWH model.*

		Naïve		Sandwich	
Parameter	**Estimate**	**STD**	**P-val**	**STD**	**P-val**
γ_{11}	-1.1663	0.3107	< 0.01	0.2778	< 0.01
γ_{12}	-1.4964	0.4505	< 0.01	0.4220	<0.01
γ_{13}	-0.6707	0.3863	>0.05	0.3785	>0.05
γ_{14}	-1.1367	0.4468	<0.01	0.5984	>0.05
γ_{15}	1.8931	1.8351	> 0.05	0.0995	< 0.01
γ_{21}	-1.9097	0.3990	< 0.01	0.3219	< 0.01
γ_{22}	-0.9814	0.3591	< 0.05	0.3772	<0.01
γ_{23}	-0.8005	0.3984	< 0.01	0.4050	0.05
γ_{24}	0.2689	0.4481	>0.05	0.5896	>0.05
γ_{25}	1.0099	4.9267	>0.05	0.4786	<0.05

5.3 Chapter Summary

This chapter introduced the approaches to be used for estimating parameters. There is the classical approach of maximising the likelihood. Thus for likelihoods derived in previous chapter, derivatives were computed. For the Bayesian approach, two proper prior distributions were derived based on the kind of transformations that involved the baseline parameters. Combining the likelihood with priors, resulting in a posterior function, determines the kind of estimation algorithm to be used. The Metropolis-Hastings algorithm was preferred.

Finally the Bayesian approach was applied on Mango data while both approaches were applied on the Kidney Infection data set. For the mango data set, results show that it does not matter which transformation to use between a

logit and log-log. Meanwhile the type of prior used does matter, with a more informative prior II preferable. As for Kidney infection data, the MLE method, which only could be applied under the independence assumption, was less preferred due to sample size and the visible dependence between failure times. Thus giving the Bayesian approach superiority over the Classical approach.

CHAPTER 6

OTHER MODELS WITH CATEGORICAL RESPONSE DATA

6.1 Introduction

A situation may present itself whereby we know the response outcome is neither time related nor continuous. An example is a study that observes the final response of a subject with regard to whether an event of interest has taken place or not, disregarding the time at which it occurs. To illustrate methods for analysing data of this nature, we start by describing the Linear Models (LM). The classical linear model assumes a response random variable Y_i (i=1,2,...,n) following a normal distribution, with expectation μ and variance σ^2, and is denoted as

$$Y_i = \mathbf{z}_i\boldsymbol{\beta} + \varepsilon_i \qquad i = 1,2,..,n,$$

where

$$\varepsilon_i \sim N(0, \sigma^2).$$

The expectation of the response variable $\mu_i = E(y_i \mid \mathbf{z}_i) = \mathbf{z}_i\boldsymbol{\beta}$ is a linear combination of observable covariates and unknown parameters $\beta_1,\beta_2,.....,\beta_p$. The task then is to estimate the unknown parameters. A specialized field of linear models called Generalized Linear Models (GLM), introduced by Nelder and Wedderburn (1972) handles such kind of data. The above characteristics can be relaxed if the assumptions mentioned below are adhered to so that data not normally distributed can also be handled.

A distribution belongs to a class of family called exponential family exponential family if its functional form can be expressed as: (Fahrmeir & Tutz (1994))

$$f(y \mid \theta,\phi) = \exp\left(\frac{y\,\theta - b(\theta($}{a(\phi)} + c(y;\theta) \right) \qquad (6.1)$$

where $c(y;\theta) \geq 0$ and is measurable, and θ is called the natural parameter which is a function of the mean, i.e. $\theta = \theta(\mu_i)$. The component ϕ is a dispersion parameter. Several distributions belong to this class, including among others, the Normal, Bernoulli, Poisson, etc.

GLM has been found to adequately handle qualitative (ordinal and nominal) response data. When there is ordering between successive intervals, then data is ordinal. Over and above naturally arising ordinal data, it can be created from continuous data by means of thresholds. Conversion of continuous data by dichotomising the entire continuous measurable space using a single threshold t_0, such that $T \geq t_0$ or $T < t_0$, results in ordinal data with dichotomous categories. Use of multiple thresholds to get interval data arises if the support space of a continuous response variable is subdivided into multiple intervals of interest and then labelled by numeric values, polychotomous response categories. Use of power family transformations like logit, probit, log-log links etc, to relate the expectation of the response variable with the predictor variable has been found to yield good estimation results. Thus

$$\eta = g(\mu) = \mathbf{z}\beta$$

where g(.) is the link function. There exist link functions like logit link, which has an error term with a Logistic distribution. The complementary log-log link, also used in survival analysis, has error term with Extreme Value (Gumbel) distribution, while probit link has a Normal distributed error term, and finally the t-link utilizes the Student t-distributed error term. This chapter briefly address the logit and probit links in GLM as alternative models for interval data.

6.2 Generalized Linear Models for Binary Responses.

In carrying out an experiment, the response from experimental units may consist of only two mutually exclusive outcomes. The outcome of a patient undergoing a surgical operation may either be complete recovery or not, or the outcome of an insect exposed to a lethal insecticide may be dead or alive. These are merely a classification without any ordering, hence nominal responses. For convenience, we let the two possible outcomes take numerical values, i.e. let Y be a Bernoulli random variable such that:

$$Y = \begin{cases} 1 & \text{with probability } \Delta \\ 0 & \text{with probability } (1-\Delta) \end{cases}.$$

Yet we can also use a threshold if we know that every insect could resist death (tolerance) up to a certain amount of insecticide used, but beyond that amount the insecticide will be lethal. We define T as the random variable denoting the amount of insecticide used, and any amount $T>t_0$ is lethal to the insect, once again resulting in a Bernoulli random variable. For $Y_1, Y_2, \ldots, Y_n$

identically and independently distributed random variables representing the n experimental units involved, we write

$$\Pr(Y_i = 1) = P(T_i \leq t_0) = \Delta$$

The first two moments of a Bernoulli distributed random variable are given by

$$E(Y_i) = \Delta \tag{6.2}$$

$$\mathrm{Var}(Y_i) = \Delta(1-\Delta).$$

As a linear model, the objective is to express the dichotomous responses as a linear combination of unknown parameters and covariate $\mathbf{z}=\{z_1, z_2,\ldots,z_p\}$, which in this instance are observable. These are measurable variables, which in their variation have an effect on the probability of a particular response. The expectation of the response variables given the covariates relates to linear predictor in that the dependence of Δ on $\mathbf{z}$ occurs through $\eta = \mathbf{z}\beta$, $(-\infty<\eta<\infty)$. Meanwhile the random component, which is the distributional function of Y also relates to Δ. Thus to reconcile the two disproportionate components, a transformation that map the linear predictor η from an unbounded linear space to a probability space (0,1) is necessary, hence conforming to probability rules. Such a function h is the link function

$$h(\mu_i) = \mathbf{z}_i\boldsymbol{\beta} \tag{6.3}$$

$$\mu_i = h^{-1}(\mathbf{z}_i\boldsymbol{\beta}) = \Delta_i .$$

Letting $h^{-1}=F$ be monotone differentiable, F is called a response function. The class of distributions belonging to the exponential family have natural link functions that relate the natural parameter to the linear predictor. The natural

link function for a Bernoulli distribution is called logit of Δ. Generalized linear models that use the logit link function are called logit models, (Hosmer & Lemeshow, (1989)). The function F is usually taken as the CDF of some continuous distribution, for the reason given in the following theorem.

Theorem *Let Y have a continuous cdf $F_Y(y)$ and define a random variable W such that $W = F_Y(y)$. Then W is uniformly distributed on (0,1), i.e. $P(W \leq w) = w$; $(0<w<1)$.*

Proof: Casella and Berger (1990).

Since F is a cumulative distribution function, by the Theorem let W=F, thus W~U(0,1). Taking a logit transformation

$$\eta = \log\left(\frac{w}{1-w}\right),$$

then

$$f(\eta) = \frac{\exp(\eta)}{(1+\exp(\eta))^2}. \qquad (6.4)$$

This is a logistic probability density for η $(-\infty<\eta<\infty)$ and

$$P(y_i = 1) = \Delta_i = \int_{-\infty}^{z_i\beta} \frac{\exp(\eta)}{(1+\exp(\eta))^2} d\eta$$

$$= \frac{1}{1+e^{-z_i\beta}}. \qquad (6.5)$$

There exist other link functions like probit, log-log, etc, for different distributions within the exponential class of distributions. Albert & Chib (1993) discuss the probit.

In the example of insects exposed to an insecticide, suppose T is distributed as N(0,1), then the probability of an insect surviving the exposure is

$$P(Y=1)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t_0} e^{-\frac{t^2}{2}}dt = \Phi(t_0), \tag{6.6}$$

where $F=\Phi$ is a standard normal distribution function relating the expectation to the linear predictor, and

$$\Delta_i = \Phi(z_i\beta) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z_i\beta} e^{-\frac{t^2}{2}}dt\,. \tag{6.7}$$

Assuming that all experimental units are i.i.d, we can derive the likelihood, making it possible to estimate the unknown parameters. By utilizing the known distribution of the response variable, which in this case is Bernoulli, the likelihood for all n units is denoted by

$$g(\boldsymbol{\beta} \mid \mathbf{z}_i, y_i) = \prod_{i=1}^{n}\left[F(\mathbf{z}_i\boldsymbol{\beta})\right]^{y_i}\left[1-F(\mathbf{z}_i\boldsymbol{\beta})\right]^{1-y_i}\,. \tag{6.8}$$

6.3 Generalized Linear Models for Polychotomous Responses

For independent binary response variables, the outcome follows a Bernoulli distribution. For N repeated trials of this experiment the number of successes or failures can be grouped together resulting in a Binomial distribution. Now suppose instead of dealing with dichotomous responses, we have more than two mutually exclusive response categories. The response variable can be categorical (nominal or ordinal) or continuous. If continuous, then the multivariate linear model suffices for the analysis of that kind of data, but our

interest here is to address the categorical data. Example is the rolling a six-faced fair die, where the mutually exclusive possible outcomes in a roll are $\{1,2,\ldots,6\}$ occurring with probability $\Delta_1, \Delta_2,\ldots, \Delta_6$ respectively, $0<\Delta_j<1$. This is a single trial multinomial distribution and is a generalization of the Bernoulli distribution for more than two categories.

If a response is nominal there is no ordering, instead it is a classification or discrete choice modelling of a subject's choice of one of several response options. Then comparison can be made between a baseline category and the rest of the r-1 categories. To analyse this data, let random variable X take any of the values $j=\{1,2,\ldots,r\}$, where

$$\Pr(X_i = j) = \Delta_j \text{ and } \sum_{j=1}^{r} \Delta_j = 1. \qquad (6.9)$$

The last condition implies that the probabilities of r categories are not independent, but are independent only if r-1 categories are considered. The r^{th} category that is omitted is known as the 'reference' category.

In exactly the same way that a Bernoulli distributed trial is extended to a Binomial situation by repeated experiments, if N repeated independent trials are executed and grouping together similar outcomes that belong to the j^{th} category such that Y_j is the number of X_i's taking value j, then Y_j is a multinomial random variable with a distribution:

$$\Pr(Y_1=y_1, Y_2=y_2,\ldots,Y_r=y_r) = \frac{N!}{y_1!\ldots y_r!}\Delta_1^{y_1}\ldots\Delta_r^{y_r}. \qquad y_j = 0,1,\ldots,N \qquad (6.10)$$

$$\Delta_r = 1 - \sum_{j=1}^{r-1} \Delta_j \text{ and } \sum_{j=1}^{r} y_j = N.$$

A multinomial distribution belongs to the exponential family, thus can be written as

$$\Pr(Y_r = y_r) = \exp\left\{ y_r \log\left(\frac{\Delta_r}{1-\Delta_j} \right) + N\log\left(1 - \sum_{j=1}^{r-1} \Delta_j \right) + C \right\} \quad j=1,2,...r-1. \tag{6.11}$$

Since the model in use treat variable Y as a response, the expected probability given the p global covariates is given by $\mu = F(\mathbf{z}_i'\boldsymbol{\beta})$, where F is the CDF related to the link function of our choice. Whereas the baseline parameters are category specific, the covariates can either be category-specific or global. In this case the probabilities are a function of a vector of covariates. Hence

$$\Pr(Y_i = j) = \Delta_{ij} \quad \text{for } j=1,2,...,r-1$$

and

$$\log\left(\frac{\Pr(Y_i = j)}{\Pr(Y_i = r)} \right) = \log \frac{\Delta_{ij}}{\Delta_{ir}} = \beta_{0j} + \mathbf{z}_i\boldsymbol{\beta}_j \tag{6.12}$$

Therefore

$$\Pr(Y_i = r) = 1 - \Delta_{i1} - ... - \Delta_{ir-1}$$

$$= \frac{1}{1 + \sum_{j=1}^{r-1} \exp(\beta_{0j} + \mathbf{z}_i\boldsymbol{\beta}_j)} \tag{6.13}$$

while

$$\Pr(Y_i = j) = \frac{\exp(\beta_{0j} + \mathbf{z}_i\boldsymbol{\beta}_j)}{1 + \sum_{j=1}^{r-1} \exp(\beta_{0j} + \mathbf{z}_i\boldsymbol{\beta}_j)}. \tag{6.14}$$

(See appendix A3)

Lack of independence between the response categories dictates that if r categories are involved, only r(r-1)/2 pairs of responses can have their logits formulated. Furthermore, the above model requires 2(r-1) parameter estimates for a single covariate of which each is category specific. This model therefore, would be difficult to estimate if several explanatory variables are used with multiple categories.

Ordinal response data arises in situation where there is some ordering of response categories, but the magnitude between orders is not of any importance. There are several logits that can be used on ordinal data as shown below:

1 *Adjacent Categorical logits*

$$\log \frac{\Pr(Y_i = j)}{\Pr(Y_i = j+1)} = \mathbf{z}_i\boldsymbol{\beta} \tag{6.15}$$

2 *Cumulative logits*

$$\log \frac{\Pr(Y_i \leq j)}{\Pr(Y_i > j)} = \mathbf{z}_i\boldsymbol{\beta} \tag{6.16}$$

3 *Continuation-ratio logits*

$$\log \frac{\Pr(Y_i = j)}{\sum_{s=j+1}^{r} P(Y_i = s)} = \mathbf{z}_i\boldsymbol{\beta} \tag{6.17}$$

Of interest is that the cumulative logit is equivalent to Collett's proportional odds (3.14) defined in terms of conditional failure probability, $\tau_j(\mathbf{z}) = 1 - P_j(\mathbf{z})$ in survival analysis, (Albert and Chib (2001)). Also to be noted, is that if a log-log transform is used in (6.16) with conditional probability of success instead

of failure, then this result s in proportional hazard model (3.10). Ignoring the covariate effect, the cumulative logit model is denoted as

$$\log \frac{\Pr(Y_i \leq j)}{\Pr(Y_i > j)} = \nu_j. \qquad (6.18)$$

The ν_j is called the 'cut-off' parameters. Agresti(1990) explain this model as assuming a non-observable latent response variable Y^* which is continuous and dependent on covariates through $\eta(\mathbf{z})=\mathbf{z}\beta$. Suppose then that the cut-off points, $-\infty=\nu_0<\nu_1<\ldots<\nu_r=\infty$, is such that the ordinal response Y is

$$Y_i = j \quad \text{if} \quad \nu_{j-1}<Y^*<\nu_j.$$

The likelihood is then formulated for the event as in form (6.8), using appropriate probabilities, and finally parameters relating to the covariates and baseline cut-offs are estimated.

6.4 Gibbs Sampling for Nominal Responses using a Latent Variable

The Gibbs sampler is a special case of MCMC, and with some modification, can be applied in a multidimensional problem it is also called alternating conditional sampling (Geman et al (1995)). There are sub-vectors representing the dimensions and a complete iteration of a Gibbs Sampler cycle through the sub-vectors of θ draws each subset conditional on the value of all the others. The idea is to generate random estimates of parameters from the conditional distribution of parameters in question, while holding constant other related estimates of parameters. This procedure is a method for generating a sample from the marginal in an indirect way, by sampling instead from the

conditional distributions that are known in statistical models, (Casella and George, (1996)). This is done in the following steps:

1 Begin with some initial guess or estimates of parameters, $\theta=\{\theta_1,\theta_2,\ldots\ \theta_r\}$ and denote their initial values by $\theta_1^0,\theta_2^0,\ldots,\theta_r^0$.

2 Generate random draws in sequence from the conditional posterior distributions:

$$\theta_j^{(p+1)} \sim P(\theta_j \mid \text{data}, \theta_{-j}^p) \qquad j=2,\ldots,r.$$

3 Repeat step 2 many times, say L, $L\rightarrow\infty$, conditioning at each iteration on the most recently generated parameter vector for the other partitions. Discard the first P sets of θ to avoid dependence on the initial values. Keep the next L-P sets.

4 Check for convergence to ensure optimality. For large enough L, the sequence of realized sets of parameter estimates would then approximate the random sample emanating from $f(\theta_j \mid \text{data})$, (Gelfand et al, (1990)).

A simulated approach using data augmentation method assumes that the observable variable Y is merely a categorized version of a latent continuous variable W. In the case of a grouped response variable, W may be considered as the unobserved underlying continuous variable. Thus W, the latent variable is primarily used in this type of approach, called data augmentation

or threshold approach (Fahrmeir & Tutz (1994)). The algorithm is expressed as follows:

Start with values $\theta^{(0)}$ generated from either the prior distribution for θ or other unbiased estimators of θ. Then iterate as follows:

Choose θ^{i+1} of θ from the conditional density $P(\theta \mid W^{(i)}, \text{data})$

Choose W^{i+1} of W from the density $P(W \mid \theta^{i+1}, \text{data})$

If the values (θ, W) given the values up to now depends only on the present ones, then it is a Gibbs sampler (Lee, (1997)). After a large number of iterations the resulting values of θ and W have a joint density close to the true density $P(\theta, W \mid \text{data})$. Successive θ^i , θ^{i+1} will in general not be independent since each depend on the previous value of the parameter, but the ideal is to obtain an independently and identically distributed set of observations of W and θ. To attain the objective, one can run the process through k successive iterations, retaining only the final values obtained on m different replications.

For the Bayesian estimation of the parameters for binary responses, let $P(Y_i=1)$ be a Logistic function and introduce the latent variable $W_1, W_2, \ldots W_N$ into the problem, where W_i has a Uniform(0,1) distribution. For a given matrix Z (Nxp) of predictor variable and unknown vector of parameters β (px1), the threshold approach postulates that the response variable $Y \in [0,1]$ and the unobservable latent variable W relate through

$$Y_i = 1 \Leftrightarrow W_i \geq c_i; \qquad 0 < c_i < 1,$$

$$\Pr(Y_i = 1) \Leftrightarrow \Pr(W_i \geq c_i),$$

$$c_i = \frac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{1+\exp(\mathbf{z}_i\boldsymbol{\beta})}.$$

Here c_i is the threshold or cutoff point. The joint posterior distribution of β and W for known prior distribution, $\pi(\beta)$, observed values of Y and given covariates Z is

$$\pi(\beta, W \mid y, z) = \pi(\beta)\prod_{i=1}^{N}\left\{I_{(w_i \geq \frac{\exp(z_i\beta)}{1+\exp(z_i\beta)})}I_{(Y_i = 1)} + I_{(w_i < \frac{\exp(z_i\beta)}{1+\exp(z_i\beta)})}I_{(Y_i = 0)}\right\}I_{(0,1)}(w_i) \quad (6.19)$$

$$= \pi(\beta)\prod_{i=1}^{N}\left\{I_{(z_i\beta \geq -\log\frac{w_i}{1-w_i})}\, I_{(Y_i = 1)} + I_{(z_i\beta < -\log\frac{w_i}{1-w_i})}\, I_{(Y_i = 0)}\right\}I_{(0,1)}(w_i), \quad (6.20)$$

where I is the indicator function. Assuming a non-informative prior for β, then the marginal conditional distribution of β_l given the rest of the β's is a Uniform($A_L(l)$,$A_U(l)$) distribution for l = 1,2,...,p where

$$A_L(l) = \max_i\left[-\frac{1}{z_{il}}(\mathbf{z}\boldsymbol{\beta} - z_{il}\beta_l + \ln(\frac{1-w_i}{w_i}))I_{(Y_i = 1,\, z_{il} > 0)}\right]$$

$$A_U(l) = \min_i\left[-\frac{1}{z_{il}}(\mathbf{z}\boldsymbol{\beta} - z_{il}\beta_l + \ln(\frac{1-w_i}{w_i}))I_{(Y_i = 0,\, z_{il} < 0)}\right]. \quad (6.21)$$

The conditional distribution of the latent variable W_i is given by

$$\pi(W_i \mid \boldsymbol{\beta}, y, z) \sim \begin{cases} \text{Uniform}\left(0, \dfrac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{1+\exp(\mathbf{z}_i\boldsymbol{\beta})}\right) & \text{if} \quad Y_i = 0 \\ \text{Uniform}\left(\dfrac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{1+\exp(\mathbf{z}_i\boldsymbol{\beta})}, 1\right) & \text{if} \quad Y_i = 1 \end{cases} \quad (6.22)$$

Given starting values of β that can be any of the unbiased estimators of β like the MLE or least square estimate, a cycle of the Gibbs algorithm (see

Logitpol.m in Appendix B) will yield a W and β. The cycles are permitted to run k times where k is large to obtain large sample from the marginal posterior distributions.

Consider now the case where the response variable Y has more than two (polychotomous) response categories and assume independence among the repeated trials. This results in a data set of N response observations whose distribution is Multinomial. Assuming a logit link function (see Appendix A4), the probability of observing a j^{th} category response follows a logistic distribution given as:

$$\Pr(Y_i = j) = \frac{\exp(\mathbf{z}_i\boldsymbol{\beta}_j)}{1+\sum_{j=1}^{r-1}\exp(\mathbf{z}_i\boldsymbol{\beta}_j)}. \qquad \begin{matrix} i = 1,2,\ldots,N \\ j = 1,2,\ldots,r-1. \end{matrix} \tag{6.23}$$

For easier illustration, let there be one explanatory variable, p=1. The response probability matrix Δ, is a (N x k-1).

Let $a = \sum_{j\neq j'}^{r-1}\exp(\mathbf{z}_i\boldsymbol{\beta}_j)$ for $j\neq j'$, then

$$\Delta_{ij} = \frac{\exp(\mathbf{z}_i\boldsymbol{\beta}_{j'})}{1+a+\exp(\mathbf{z}_i\boldsymbol{\beta}_{j'})}.$$

Let

$$c_{ij'} = \frac{\exp(\mathbf{z}_i\boldsymbol{\beta}_{j'})}{1+a+\exp(\mathbf{z}_i\boldsymbol{\beta}_{j'})}$$

be the cut-off points. The transformation variable W_{ij} has a uniform distribution on interval (0,1), and the resulting distribution is a CDF of a

logistic distribution with r response categories. The joint posterior distribution of β and **W** (Nxr-1) given observed data is:

$$\pi(\beta,\mathbf{W} \mid y) = \pi(\beta) \prod_{i=1}^{N} \sum_{j=1}^{r-1} \left[I_{(Y_i=j)} I_{\left(w_{ij} < \frac{\exp(z_i\beta_j)}{1+\sum_{j=1}^{r-1}\exp(z_i\beta_j)}\right)} \right] I_{(0,1)}(w_{ij}) \qquad (6.24)$$

Letting the prior distribution of β be proportional to 1, then the conditional distribution of the i,jth element of W is:

$$\pi(w_{ij} \mid \beta, Y_i) = \text{Uniform}\left(0, \frac{\exp(z_i\beta_j)}{1+\sum_{j=1}^{r-1}\exp(z_i\beta_j)}\right) \qquad \text{if } Y_i = j$$

$$= \text{Uniform}\left(\frac{\exp(z_i\beta_j)}{1+\sum_{j=1}^{r-1}\exp(z_i\beta_j)}, 1\right) \qquad \text{if } Y_i \neq j \qquad (6.25)$$

To define the conditional distribution of β_j, let i(j) be the set of all i for which $Y_i = j$, then

$$a_j = \max_{i(j)}\left[\frac{1}{z_i}\ln\left\{\frac{w_{ij}}{1-w_{ij}}\left(1+\sum_{j=1}^{r-1}\exp(z_i\beta_j)\right)\right\}\right]$$

is the lower limit of β_j if $z_{i(j)} > 0$ and

$$b_j = \min_{i(j)}\left[\frac{1}{z_i}\ln\left\{\frac{w_{ij}}{1-w_{ij}}\left(1+\sum_{j=1}^{r-1}\exp(z_i\beta_j)\right)\right\}\right] \qquad (6.26)$$

is the upper limit of β_j if $z_{i(j)} > 0$. The reverse would be true if $z_{i(j)} < 0$ for the upper and lower limits, and

$$\pi(\beta_j \mid \mathbf{W}, y_i, \beta) = \text{Uniform}(a_j, b_j) \quad j=1,2,...r-1.$$

An extension of the above result to encompass multiple explanatory variables is such that j = 1,2,...,r-1; i=1,2,...,N; l=0,1,2,...p; β is a (p+1xr-1) matrix, $\beta_j = [\beta_{0j},\beta_{1j},...,\beta_{pj}]'$, $\mathbf{z}_i = [1\ z_{i1},z_{i2},..,.z_{ip}]$. In exactly the same way as we derived the above results, assuming a uniformly distributed latent variable **W**, the joint posterior distribution of β and **W** given observed categorical data $y_1,....,y_N$ and explanatory variables **z**, is given by:

$$\pi(\beta, \mathbf{W} \mid y) = \pi(\beta) \prod_{i=1}^{N} \sum_{j=1}^{r-1} I_{(Y_i = j)} I_{\left(w_{ij} < \frac{\exp \sum_{l=1}^{p} (z_i \beta_{lj})}{1 + \sum_{j=1}^{r-1} \exp \sum_{l=1}^{p} (z_i \beta_{lj})}\right)} I_{(0,1)}(w_{ij}) \qquad (6.27)$$

Letting the prior distribution of β be proportional to 1, then conditional distribution of latent variable W_{ij} is:

$$\pi(w_{ij} \mid \beta, Y_i) = \text{Uniform}\left(0,\ \frac{\exp \sum_{l=1}^{p} (z_i \beta_{lj})}{1 + \sum_{j=1}^{r-1} \exp \sum_{l=1}^{p} (z_i \beta_{lj})}\right) \quad \text{if } Y_i = j$$

$$= \text{Uniform}\left(\frac{\exp \sum_{l=1}^{p} (z_i \beta_{lj})}{1 + \sum_{j=1}^{r-1} \exp \sum_{l=1}^{p} (z_i \beta_{lj})},\ 1\right) \quad \text{if } Y_i \neq j \qquad (6.28)$$

If i(j) is the set of all i for which $Y_i = j$, then

$$a_{l'j'} = \max_{i(j)} \left[\frac{1}{z_{il'}} \left\{\ln\left[\frac{w_{ij}}{1 - w_{ij}} \left(1 + \sum_{j \neq j'}^{r-1} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)\right)\right] + \sum_{l \neq l'}^{p} z_i \beta_{lj}\right\}\right] \qquad (6.29)$$

is the lower limit of $\beta_{j'j'}$ if $z_{i(j)} > 0$ and if i(j) is the set of all i for which $Y_i \neq j$, then

$$b_{l'j'} = \min_{i(j)} \left[\frac{1}{z_{il'}} \left\{\ln\left[\frac{w_{ij}}{1 - w_{ij}} \left(1 + \sum_{j \neq j'}^{r-1} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)\right)\right] + \sum_{l \neq l'}^{p} z_i \beta_{lj}\right\}\right] \qquad (6.30)$$

is the upper limit of $\beta_{l'j'}$ if $Z_{i(j)} > 0$ (Appendix A5), and

$$\pi(\beta_{lj} \mid \mathbf{w}_{ij}, y_i, \beta) = \text{Uniform}(\mathbf{a}_{lj}, \mathbf{b}_{lj}) \quad j=1,2,...r-1.$$

6.5 Gibbs Sampling for Ordinal Responses Using a Latent Variable

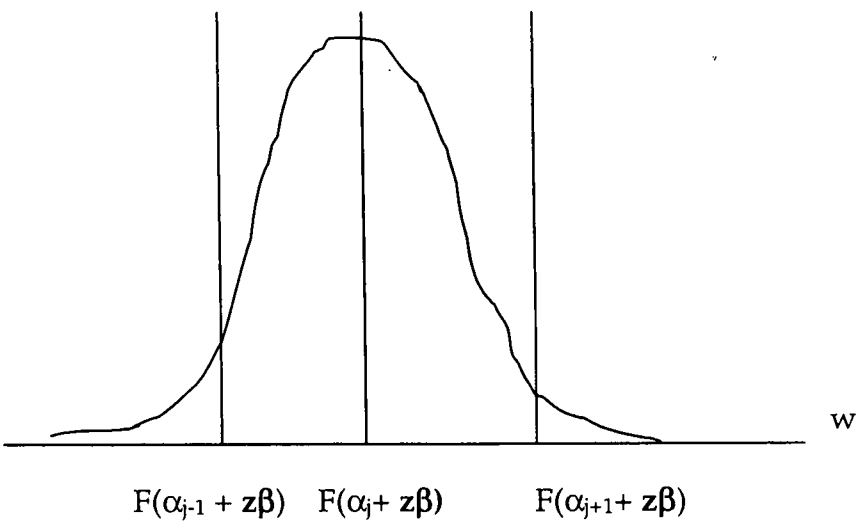
For an explanatory variables **z**, p unknown parameters β and the observable variable Y falling in any of the r-ordered categories, the idea is to utilize the ordering of response categories. Take a continuous latent variable W_i uniformly distributed over [0,1] interval such that the following would be true:

$$j\text{-}1 < Y_i \leq j \Leftrightarrow F(\alpha_{j-1} + \mathbf{z}_i\beta) < W_i < F(\alpha_j + \mathbf{z}_i\beta) \qquad j = 1,2,\ldots,r; \quad i = 1,2,\ldots,N$$

$$\Pr(Y_i \leq j \mid \mathbf{z}, \beta) = F(\alpha_j + \mathbf{z}_i\beta) = \frac{\exp(\alpha_j + \mathbf{z}_i\beta)}{1 + \exp(\alpha_j + \mathbf{z}_i\beta)}.$$

The α_j are the ordered cut-off points for the continuous latent variable W.

Figure 6.1: *Example of cut-off points for the distribution of a latent variable*



The joint distribution of β, α, W, given the response variable Y and covariates **z** is given by

$$\pi(\beta, \alpha, W \mid y, z) \propto \pi(\beta, \alpha)$$

$$\prod_{i=1}^{N}\left\{\sum_{j=1}^{r-1} I_{(Y_i=j)} I_{\left(\frac{\exp(\alpha_{j-1}+z_i\beta)}{1+\exp(\alpha_{j-1}+z_i\beta)}<w_i<\frac{\exp(\alpha_j+z_i\beta)}{1+\exp(\alpha_j+z_i\beta)}\right)}\right\} I_{(0,1)}(w_i). \qquad (6.31)$$

Assuming a non-informative prior distribution for β and α, then the conditional posterior distribution of latent variable W_i is given by

$$\pi(w_i \mid \beta,\alpha, y, z) = \text{Uniform}\left(\frac{\exp(\alpha_{j-1}+z_i\beta)}{1+\exp(\alpha_{j-1}+z_i\beta)}, \frac{\exp(\alpha_j+z_i\beta)}{1+\exp(\alpha_j+z_i\beta)}\right), \ \forall i \text{ which } Y_i = j.$$

The conditional posterior distribution for parameter $\beta_{l'}$ given α, w, y, z is

$$\Pi(\beta_{l'} \mid \alpha, w,\beta\, y, z) = U(A_L, A_U)$$

where

$$A_L = \max_i\{\max_j\{\frac{1}{z_{il}}[\ln(\frac{w_i}{1-w_i}) - \alpha_j - \sum_{l\neq l'}^{r-1} z_{il}\beta_l]\}\},$$

and

$$A_U = \min_i\{\min_j\{\frac{1}{z_{il}}[\ln(\frac{w_i}{1-w_i}) - \alpha_j - \sum_{l\neq l'}^{r-1} z_{il}\beta_l]\}\}. \qquad (6.32)$$

The posterior distribution for the cut-point α_j is derived under the following conditions:

i) If $Y_i = j$ then $F(\alpha_{j-1}+z_i\beta) < W_i < F(\alpha_j + z_i\beta)$

ii) if $Y_i = j+1$then $F(\alpha_j + z_i\beta) < W_i < F(\alpha_{j+1} + z_i\beta)$.

Thus $W_i < \frac{\exp(\alpha_j+z_i\beta)}{1+\exp(\alpha_j+z_i\beta)}$ implies that

$$\alpha_{j-1} < \ln(\frac{W_i}{1-W_i}) - z_i\beta < \alpha_j \quad \text{set of i's for which } Y_i = j$$

and

$$\alpha_j < \ln(\frac{W_i}{1-W_i}) - z_i\beta < \alpha_{j+1} \quad \text{set of i's for which } Y_i = j+1,$$

giving the posterior distribution of the cut-off points to be uniformly distributed,

$$\pi(\alpha_j \mid \beta, W, y, \mathbf{z}) = \text{Uniform}(C_L, C_U)$$

with

$$C_L = \max_i [\max\{ \ln(\frac{W_i}{1-W_i}) - \mathbf{z}_i \beta, \alpha_{j-1} \}]$$

$$C_u = \min_i [\min\{ \ln(\frac{W_i}{1-W_i}) - \mathbf{z}_i \beta, \alpha_{j+1} \}]. \qquad (6.33)$$

Once again the maximisation (minimisation) to find the intervals CL (Cu) is only over the set of i's for which $Y_i = j$ ($Y_i = j+1$).

To carry out a Gibbs sampler in this case, we start with good estimators of β and α like the MLE or least square estimate, and then simulate from conditional posterior distribution of α_j in (6.33), followed by simulating the latent variable and finally use the conditional posterior distribution of β in (6.32).

6.6 Gibbs Sampling on Tri-Continental AIDS Data

Illustration V: Data is from an AIDS study on sero-converted men in the Amsterdam, Netherlands from 1985 to 1996, when the study terminated. This was a section of the study termed *The European Seroconverter Study among Injecting Drug Users* whose data was analysed by Prins and Veugelers (1997). 117 subjects entered the study at different times, a year after sero-conversion

with earliest entrant being followed for a period of a little over 10 years. Data is only on subjects who were AIDS free one year after sero-conversion. Measured on each patient are the following variables:

1 **Entry Time**: Date of study entry a year after sero-conversion

2 **CD4**: The number of CD4 cells per micro-litre of blood at entry

3 **AIDS**: Time (years) to diagnosis of AIDS if subject is HIV positive otherwise it is termination period.

4 **AIDS Status**: Indicator for subject's AIDS status (0=negative; 1=positive)

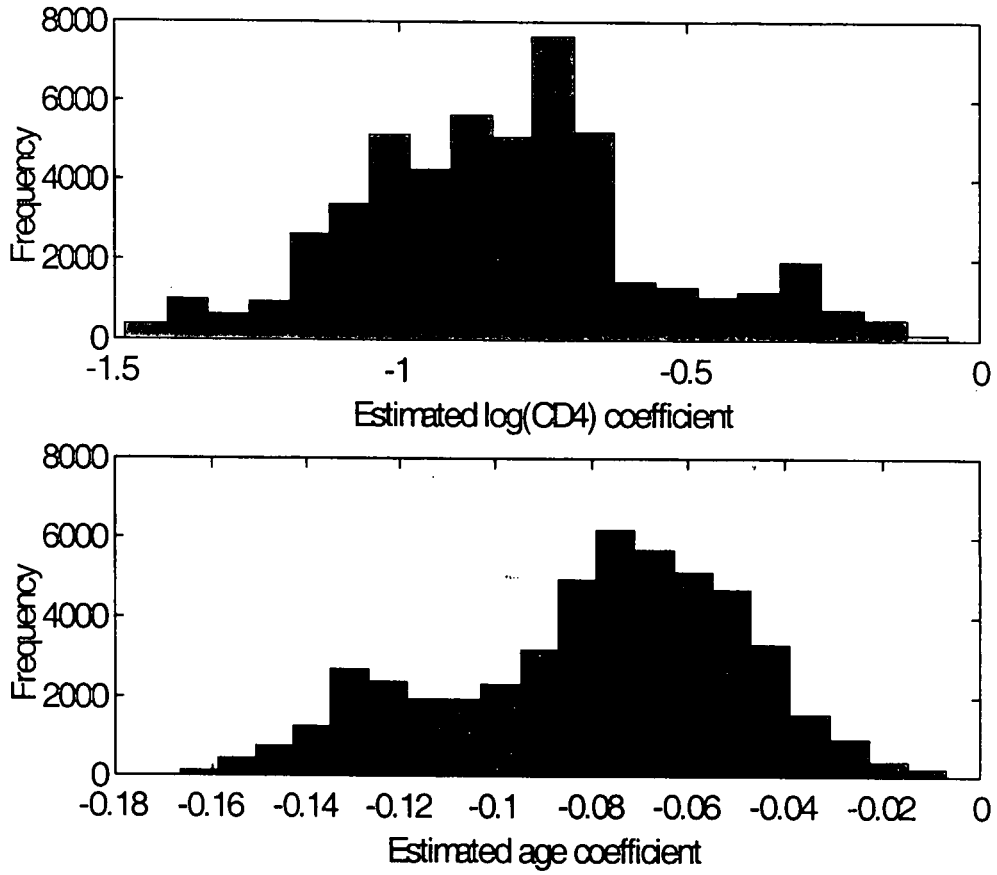
5 **Death**: Time (years) to death if AIDS related, otherwise its termination period study for those subjects alive.

6 **Death Status**: Indicator for death status (0=dead 1=alive)

7 **Age**: age of participant at entry point

Using the joint posterior (6.20) with AIDS status (1 or 0) as a nominal response variable, we used a Gibbs Sampler (Logitbi.m in appendix B) to simulate the parameters from conditional distributions given in (6.21) and (6.22) with a latent variable method to estimate the probability of developing AIDS given the explanatory variables age and CD4 cell counts of a patient at entry point. 50000 Gibbs cycles was run. The parameter estimate for age is –0.08 (-0.14, -0.03) and for the log of CD4 levels at entry is -0.8211 (-1.35, -0.26), which implies that probability of developing AIDS is lowered for high CD4 levels at entry and older males. The odds in favour of developing AIDS decreases by $e^{(-0.8211)}=0.44$ for every unit increase in the log of CD4 cell counts.

Figure 6.1: *A frequency histogram to show the posterior distributions of log(CD4) and age explanatory variable parameters.*



The posterior distributions in terms of frequency histograms of the two coefficients of explanatory variables in the AIDS data are shown in Figure 6.1. The two parameters show signs of being negatively skewed.

6.7 Chapter Summary

A computational method for the Bayesian analysis for the Logistic regression is given in this chapter snd applied to AIDS data. The method handles both nominal and ordinal response variables with multiple categories. Thus for

time related variables whose exact values cannot be ascertained, hence converting it to discrete variable, this method can be applied. Application of this method to AIDS data reveal that the levels of CD4 and age of a patient are factors influencing the probability of a change in status from AIDS-free to AIDS. Albert and Chib (1993) applied a similar method with probit link and latent variable on some binary and polychotomous response data.

CHAPTER 7

DATA ANALYSIS AND RESULTS

7.1 Simulation Results

7.1.1 Introduction

A Matlab (1992) computer program was used to generate and analyse both the simulated and ACTG 175 data sets (see Appendix B7 for programs). For the MLE method convergence for small samples occurred after a few iterative searches, but got slower (8 seconds) for large samples of size 200. Thus to draw m repeated samples posed no problem. For the MLE approach, 2000 samples were generated for samples sizes of 50 and 100. Meanwhile the Bayesian approach was affected more by a slight increase in sample size. The method took 6 minutes to make 2000 cycles for a sample of size 50. This not only constrained the number of samples that could be generated, but also curtailed the use of large sample sizes. At most 300 samples were generated with maximum size of 50 observations. The univariate parameter estimates when assuming Weibull distributed lifetimes were computed using SAS (1986).

7.1.2 Results from Simulated Bivariate Data

To ascertain the validity of methods suggested in this research, they were applied to some simulated data, and the sampling distribution of the estimated parameters determined. Some m pseudo samples of size 50 were

simulated from the Farlie-Morgenstern (F-M) distribution (3.37) with some specified measures of association and covariate parameters. Using these samples, parameter estimates were computed using each of the methods derived in Chapter 3. An array of marginal distributions can be used, including among others, the Exponential, Weibull, etc. For simulation, the Exponential marginal distributions for each of the failure times were preferred. (Simfm.m in Appendix B).

The joint probability distribution function for the Farlie-Morgenstern copula is denoted by

$$F(t_1,t_2) = \left(1-e^{-t_1}\right)\left(1-e^{-t_2}\right)\left[1+\kappa e^{-(t_1+t_2)}\right], \qquad (7.1)$$

with κ the measure of association between the failure times such that the correlation is $\rho=\kappa/4$. For this model the maximum κ attainable is 1, representing a not so strong dependency. With an exponential marginal, the hazard function is of the form $\lambda=e^{z\beta}$. To simulate the two random variables T_1 and T_2, first generate T_1 from the marginal distribution $F(t_1)=1-e^{-\lambda t_1}$ and thereafter for given values of T_1 simulate T_2 from the conditional distribution $F(t_2 \mid t_1)$,

$$F(t_2 \mid t_1) = (1-e^{-t_2})[1+\kappa(e^{-t_2})(2e^{-t_1}-1)] \qquad (7.2)$$

This is done by setting

$$T_1 = -\lambda^{-1}\log(1-u_1) \text{ and } T_2=-\lambda^{-1}\log(1-v),$$

where u_1 and u_2 are Uniform(0,1), and

$$v = \frac{(1-\kappa-2\kappa a)-\sqrt{(1+\kappa-2\kappa a)^2-4u_2(1-2a)}}{2\kappa(1-2a)} \text{ with } a=\exp(-\lambda t_1).$$

Steps towards generating T_1 and T_2 are as follows;

- A sample of size n of a single normally distributed covariate, Z~N(0,1), is generated, and a vector of hazard functions obtained where, $\lambda_i = e^{\beta z_i}$. This assumes common hazard functions for both failure times, unless the β's are made different.
- For given λ_i, generate T_1 from an Exponential distribution by the inversion method, (Devroye (1986)).
- Finally using the values of T_1, generate T_2 using the inversion method on the conditional distribution (7.2) by calculating v.
- To address the issue of intervals, the simulated values for both lifetimes are then grouped into non-overlapping intervals. The following intervals were used: {0-0.5; 0.5-1.0; 1.0-2.0; 3-∞} resulting in 3 intervals with observations belonging to the open interval brought to the 3rd as censored.

An alternative to a Farlie-Mogenstern copula would be to use the Clayton Copula distribution, which allows for a stronger dependence between T_1 and T_2, to generate the data. The survival function by assuming Exponential marginal distribution is

$$S(t_1,t_2) = (e^{\frac{t_1}{\kappa}} + e^{\frac{t_2}{\kappa}} - 1)^{-\kappa}.$$

Hence setting $\lambda_i = e^{\beta z_i}$ to account for the covariate effect, then $T_1=-\lambda^{-1}\log(1-u_1)$ and $T_2=-\lambda^{-1}\kappa\log(v)$ where $v=(1-a)+a(1-u_2)^{-(1+\kappa)^{-1}}$ with $a=\exp(-\lambda t_1)$, (Prentice and Cai, (1992)).

Using the Exponential marginal distributions with a Farlie-Morgenstern copula, the baseline conditional survival probabilities P_{mj} for the three intervals from each failure type were computed as {0.606, 0.606, 0.368} which are transformed by log-log to give γ_{mj}={-0.6914, -0.6914, 0} for PHM, or a logit transform with α_{mj}={-0.4305, -0.4305, 0.5408} for the POM. Of interest to note is that these conditional probabilities are based on the independent marginal distributions. Therefore if some association between the variables exist, it would be revealed by the data. The baseline parameter values from the two transformations were taken as true values. To assess the performance of the independence assumption with proportional hazard model IWH (3.26), data with two failure times simulated using either the $\kappa=0$ ($\rho=0$) or $\kappa=1$ ($\rho=0.25$) on the F-M copula was used, (Indmlesim.m in Appendix B). Results from the method of Maximum Likelihood and posterior distribution (Indposim.m in Appendix B) for the parameter estimates are given on Tables 7.1.1-7.1.4. SE and CP are respectively the sampling averages of standard errors (from diagonal of H^{-1} or $H^{-1}DH^{-1}$ of estimator) and the coverage probability of the 95% confidence interval or 95% posterior density region. SSE is square-root of the sampling variance, $\frac{\sum_{l=1}^{m}(\tilde{\beta}_l-\beta)^2}{m-1}$, of the estimator $\hat{\beta}$ where, $\tilde{\beta}_l$ is the

estimate from the l^{th} sample. Estimate (Est.) is the value of $\hat{\beta}$ obtained by averaging over all samples, while bias is the difference between the true value and the estimated value.

Table 7.1.1: *Maximum Likelihood mean estimates of parameters computed under IWH model, for data simulated from a F-M copula using 2000 samples. n=50, β=0.25, κ=0, and κ=1.*

					Naïve		*Robust*	
Par.	ρ	Mean	Bias	SSE	SE	CP	SE	CP
$\hat{\beta}$	0	0.2592	0.0092	0.1238	0.1197	0.938	0.1155	0.929
	0.25	0.2528	0.0028	0.1313	0.1183	0.922	0.1216	0.930
$\hat{\gamma}_{11}$	0	-0.7173	-0.0260	0.2374	0.2334	0.949	0.2337	0.949
	0.25	-0.7098	-0.0185	0.2329	0.2327	0.944	0.2329	0.944
$\hat{\gamma}_{12}$	0	-0.7143	-0.0230	0.3219	0.3079	0.932	0.3078	0.932
	0.25	-0.7207	-0.0294	0.3170	0.3091	0.941	0.3088	0.941
$\hat{\gamma}_{13}$	0	0.0169	0.0169	0.3496	16.230	1.000	0.3251	0.938
	0.25	0.0082	0.0082	0.3515	28.420	1.000	0.3245	0.935
$\hat{\gamma}_{21}$	0	-0.7126	-0.0213	0.2371	0.2329	0.947	0.2334	0.948
	0.25	-0.3333	0.3580	0.4138	0.2058	0.587	0.2064	0.588
$\hat{\gamma}_{22}$	0	-0.7164	0.0251	0.3242	0.3090	0.944	0.3086	0.943
	0.25	-0.4469	0.2444	0.4130	0.3188	0.870	0.3184	0.868
$\hat{\gamma}_{23}$	0	0.0208	0.0208	0.3395	0.3271	0.940	0.3240	0.939
	0.25	0.2427	0.2427	0.5796	97.609	1.000	0.3828	0.894

The MLE approach (Table 7.1.1) is sensitive to small sample sizes as evidenced by its tendency to either crush or return biased estimates with highly inflated standard errors, when estimating parameters for samples of 30 or less observations. The effect of small sample is shown by the standard errors of γ_{13} and γ_{23} which are inflated. The bias in estimating the baseline parameter for the second failure when there is dependence (ρ=0.25) is conspicuously evident compared to bias when failures times are independent.

The covariate parameter is less affected. The standard errors by naïve estimator are highly erratic. The robust estimator's endeavour to correct the anomaly succeeds quite well.

Table 7.1.2: *Maximum likelihood mean estimates of parameters computed under IWH, for data simulated from F-M copula using 2000 samples. n=100,* β*=0.25 and* κ*=0 and 1.*

					Naïve		*Robust*	
Par	ρ	Est.	Bias	SSE	SE	CP	SE	CP
$\hat{\beta}$	0	0.2553	0.0053	0.0838	0.0820	0.940	0.0808	0.939
	0.25	0.2426	-0.0074	0.0877	0.0813	0.934	0.0855	0.944
$\hat{\gamma}_{11}$	0	-0.7078	-0.0165	0.1653	0.1632	0.9450	0.1633	0.949
	0.25	-0.6960	-0.0047	0.1609	0.1625	0.950	0.1625	0.950
$\hat{\gamma}_{12}$	0	-0.7045	-0.0132	0.2153	0.2131	0.950	0.2131	0.950
	0.25	-0.6925	-0.0012	0.2150	0.2126	0.941	0.2125	0.941
$\hat{\gamma}_{13}$	0	0.0014	0.0014	0.2311	0.2257	0.940	0.2246	0.939
	0.25	0.0048	0.0048	0.2319	0.2269	0.949	0.2253	0.947
$\hat{\gamma}_{21}$	0	-0.7002	-0.0089	0.1665	0.1627	0.946	0.1629	0.946
	0.25	-0.3292	0.3621	0.1444	0.1443	0.950	0.1445	0.950
$\hat{\gamma}_{22}$	0	-0.7014	-0.0101	0.2141	0.2133	0.949	0.2132	0.949
	0.25	-0.4239	0.2694	0.2213	0.2180	0.950	0.2181	0.950
$\hat{\gamma}_{23}$	0	0.0028	0.0028	0.2343	0.2264	0.939	0.2251	0.938
	0.25	0.1892	0.1892	0.2743	0.2665	0.942	0.2651	0.942

Finally for samples of size 100 the MLE method begins to display stability, (see Table 7.1.2). Increasing the sample size to 100 greatly improves the standard errors of the parameter estimates resulting in more precise coverage. However the parameter estimates for the second failure remains different from the assumed true parameter values if the association measure is not zero.

Bayesian methods of using posterior distributions with Priors I and II were also applied. 274 samples were used in the Bayesian approach with a Metropolis-Hastings algorithm, compared to 2000 samples for the MLE method. Prior II return large bias for values of $\hat{\beta}$, compared to both prior I and MLE estimates, (see Table 7.1.3).

Table 7.1.3: *Posterior means estimated using IWH for samples of size 50 generated from F-M copula with* $\beta = 0.25$ *and 2000 cycles. 274 samples were generated.*

		Prior I				Prior II			
Par.	ρ	Mean	Bias	SSE	CP	Est.	Bias	SSE	CP
$\hat{\beta}$	0	0.2618	0.0118	0.1269	0.987	0.2697	0.0197	0.1265	0.993
	.25	0.2518	0.0018	0.1352	0.971	0.2601	0.0101	0.1343	0.981
$\hat{\gamma}_{11}$	0	-0.7232	-0.0319	0.2205	0.983	-0.7696	-0.0780	0.2205	0.990
	.25	-0.7196	-0.0283	0.2249	0.971	-0.7777	-0.0860	0.2403	0.985
$\hat{\gamma}_{12}$	0	-0.7235	-0.0322	0.3167	0.990	-0.8002	-0.1090	0.3128	0.987
	.25	-0.7325	-0.0412	0.3344	0.986	-0.7751	-0.0850	0.3250	0.992
$\hat{\gamma}_{13}$	0	-0.0064	-0.0064	0.3158	0.980	-0.0669	-0.0669	0.3432	0.973
	.25	-0.0046	-0.0046	0.3344	0.914	-0.0094	-0.0092	0.3433	0.974
$\hat{\gamma}_{21}$	0	-0.7346	-0.0433	0.2510	0.997	-0.7616	-0.0702	0.2535	0.987
	.25	-0.3554	0.3359	0.3859	0.686	-0.4069	0.2844	0.3492	0.883
$\hat{\gamma}_{22}$	0	-0.7300	-0.0387	0.3125	0.967	-0.7259	-0.0353	0.2830	0.987
	0.25	-0.4623	0.2290	0.3917	0.857	-0.5152	0.1761	0.3497	0.947
$\hat{\gamma}_{23}$	0	-0.0614	-0.0614	0.3252	0.990	-0.0523	-0.0523	0.3327	0.980
	.25	0.1436	0.1436	0.4062	0.903	0.1265	0.1265	0.3697	0.911

The use of posterior measures does well in the estimation of standard errors, hence a good coverage, but the issue of deviating estimates of the second failure when dependence is not accounted for, persists. It is vital to acknowledge that since data is correlated, we do not expect the baseline

parameters for the second failure to be estimated as well as those for the first failure time. Interestingly, the sampling standard errors using both priors are consistently small, comparable to the sandwich standard error estimates, actually remains so even when correlation is taken aboard. In the absence of dependence between failure types, the standard errors of parameters relating to conditional probabilities are also well estimated. Furthermore, data simulated using a Farlie-Morgenstein copula with a fairly weak association of κ=0.6 and same Exponential distribution for failure times, hence same true values as above, was generated. One again a Metropolis-Hastings MCMC algorithm on a nonparametric Clayton Copula with proportional hazard model (CCH) was used to analyse the data and results given in Table 7.1.4.

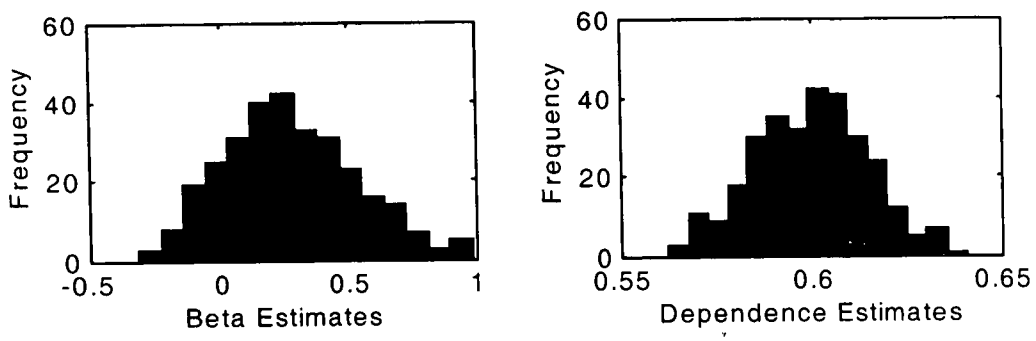
Table 7.1.4: *Posterior means estimated using a CCH with prior I for samples of size 50 generated from a F-M copula with $\beta = 0.25$ and 2000 iterations. 274 samples were generated.*

Para	$\hat{\beta}$	$\hat{\gamma}_{11}$	$\hat{\gamma}_{12}$	$\hat{\gamma}_{13}$	$\hat{\gamma}_{21}$	$\hat{\gamma}_{22}$	$\hat{\gamma}_{23}$
Est	0.2715	-0.6949	-0.6884	-0.0017	-0.6679	-0.7009	-0.0136
Bias	0.0215	-0.0036	0.0029	-0.0017	0.0234	-0.0096	-0.0136
SSE	0.0667	0.0670	0.0713	0.0688	0.0644	0.0811	0.0649
CP	0.9533	0.9533	0.9567	0.9500	0.9500	0.9500	0.9500

The CCH model returns good estimates with small SSE. The baseline parameters also have good coverage probabilities. The fact that we can also ascertain the measure of association between the failure times augurs well for the method. The graph below (Figure 7.1.1) shows posterior distributions parameter coefficients β and the measure of association.

Once again it has been demonstrated that the ability to estimate using a classical approach is affected by the small sample size hence an improvement is brought about by increasing the simulated data observations to 100. Meanwhile the more stable results from Bayesian approach come at price, since this requires a large number of cycles to converge, explaining why a small number (2000) of cycles was used. Due to similarity in results obtained from either PH or PO transforms, only PH was used in analysing simulated data. Prior type II tends to perform better. Finally, the results show that if dependence exist between lifetimes, then model that account for the measure must be used.

Figure 7.1.1: *Histograms showing the posterior distributions of simulated covariate and dependence parameters from 300 samples using a CCH model with prior I.* (β=0.25; κ=0.6)



7.1.3 Simulation Data with Dependent Visiting Times

In this section, simulated data is used to assess the effect non-compliance to predetermined visiting times has on the estimation of parameters values. The model (3.24) shall be applied to the simulated data. To simulate pseudo data, assume a lifetime data that follows an Exponential distribution, X~Exp.(0.3), (0<X<12) with each integer value representing an interval. The assumption is that the actual lifetimes are not observable (latent) but we can observe the intervals, $(t_{j-1}, t_j]$. With intervals and distribution now known, failure probabilities for each of the 12 intervals plus the open interval involved were computed using

$$g_j = e^{-0.3t_{j-1}} - e^{-0.3t_j} \qquad j = 1,2,\ldots,13, \text{ where } t_{13}=\infty.$$

Let P_j j=1,2,...,12, again be the conditional probability of surviving beyond j^{th} interval given survival to its beginning, these are computed to be 0.7408 for all intervals. For explanatory variable effect, a single normally distributed variable with 50 observations with mean and variance of 0 and 1 respectively (Vispos.m in Appendix B), were simulated. Also of importance is knowledge of each unit's frequency of visits prior and post failure time, since we are interested in the relationship between visiting compliance and failure. For each unit, a binary random variable η_{js} taking value 1 when a unit makes visit before failure at j and 0 otherwise, thus the parameter λ allows for test of compliance before or after failure was used. Also included is the constant μ_j, specific to interval j and the effect of covariates on the compliance probabilities (3.23) is measured by parameter ν.

For true values of μ_j=1.25-0.25(j-1), β=0.30, λ=1, ν=-0.5, and for each sequence a failure interval was generated, and then the visiting status for each interval was subsequently generated. Using a Metropolis-Hastings algorithm in blocks on model (3.24) with a PH transform, one complete cycle accomplished by:

- Simulating values of P_j simultaneously from a proposal distribution Multivariate Normal MV($\mathbf{P}$,Σ). Σ is a diagonal matrix with all its elements equal 0.4. A logit transformation on the hazard was applied hence prior II was used. Meanwhile the rest of the parameters in other blocks are held constant.
- Next values of μ_j, λ, and ν were also simulated from a MV proposal distribution with current values of P and β.
- Finally the value of β was simulated to complete a cycle. This was iterated 2000 times.

The results obtained from simulated data are given as $\hat{\beta}$=0.3138 (0.015; 0.630), $\hat{\lambda}$=1.4962 (0.917; 2.000) and $\hat{\nu}$=-0.8573 (-1.281; -0.467). The plot in Figure 7.1.2 shows the estimated values of explanatory variable returned by the jumping distribution. The plot reveal some stabilization and hence convergence in the 2000 cycles. A plot of the visiting probabilities at each of the 12 simulated intervals is shown in Figure 7.1.3. The simulated data reveals decreasing compliance probabilities for increasing interval numbers with prior-to-failure visiting probabilities higher than after failure for both true and estimated values as indicated by the positive value of the estimated parameter.

Figure 7.1.2: *2000 iterations from a Metropolis-Hastings algorithm for the explanatory variable coefficient with a proposal distribution of* $N(\beta^{(t-1)},0.3)$.

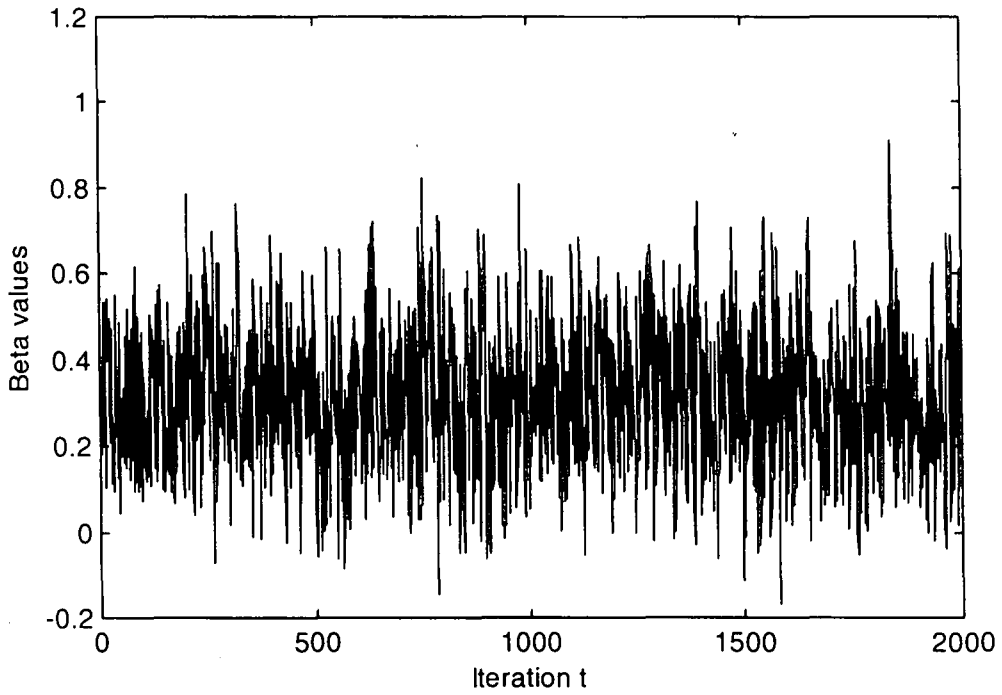
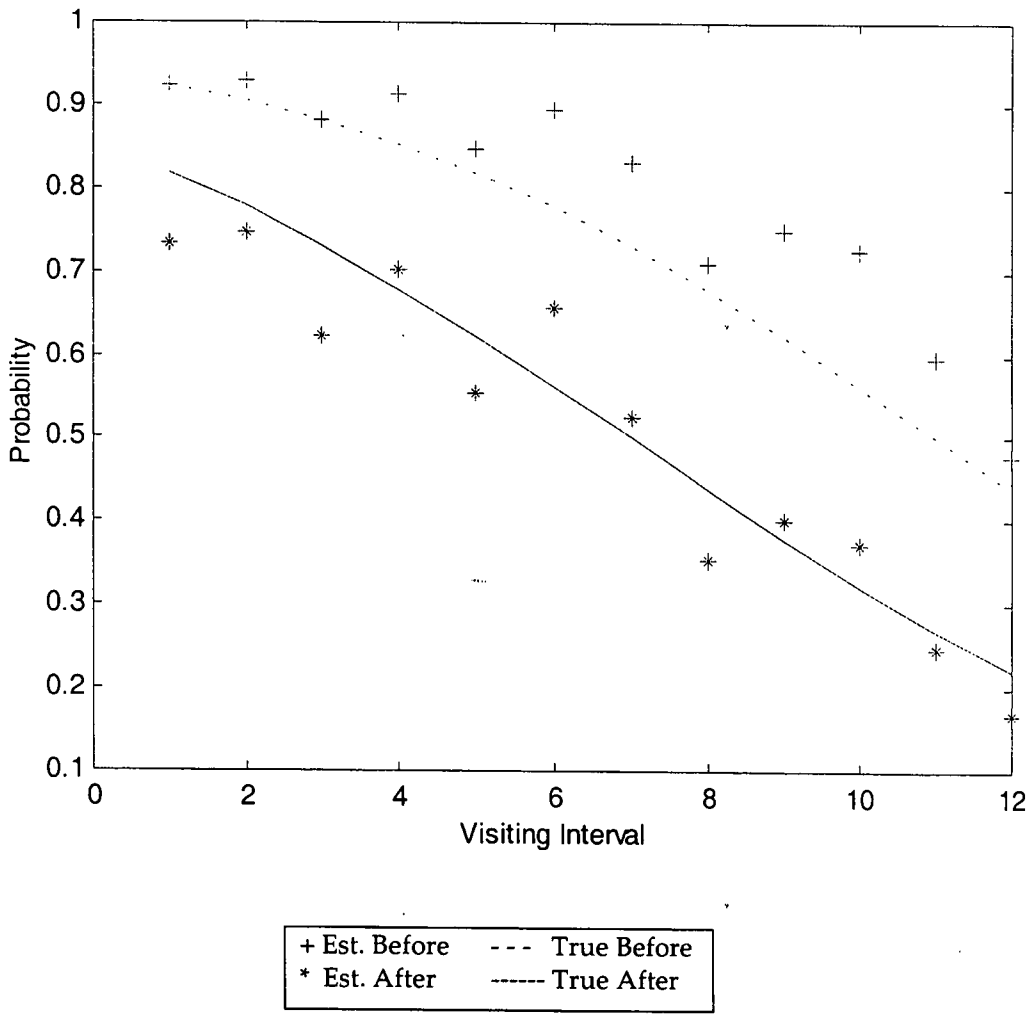


Table 7.1.5: *Results of parameters estimates obtained from simulated visiting data*

Interval	μ_j	$\hat{\mu}_j$	$\hat{\pi}_{Bj}$	$\hat{\pi}_{Aj}$	π_{Bj}	π_{Aj}
1	1.5	1.01 (-0.05, 2.21)	0.9246	0.7330	0.9241	0.8176
2	1.25	1.08 (0.29, 1.88)	0.9293	0.7465	0.9047	0.7773
3	1	0.50 (-0.24, 1.25)	0.8804	0.6225	0.8808	0.7311
4	0.75	0.86 (0.07, 1.56)	0.9134	0.7027	0.8520	0.6792
5	0.5	0.22 (-0.42,0.80)	0.8476	0.5548	0.8176	0.6223
6	0.25	0.65 (-0.04,1.35)	0.8976	0.6570	0.7773	0.5622
7	0	0.10 (-0.48, 0.67)	0.8315	0.5250	0.7311	0.5000
8	-0.25	-0.60 (-1.27, 0.10)	0.7102	0.3543	0.6792	0.4378
9	-0.50	-0.40 (-1.03, 0.18)	0.7495	0.4013	0.6225	0.3775
10	-0.75	-0.52 (-1.27, 0.12)	0.7264	0.3729	0.5622	0.3208
11	-1.00	-1.11 (-1.91,-0.23)	0.5954	0.2479	0.5000	0.2689
12	-1.25	-1.59 (-2.44,-0.70)	0.4766	0.1694	0.4378	0.2227

The estimated coefficient for the parameter λ of the visiting compliance variable (before/after) is overestimated, (See Figure 7.1.3).

Figure 7.1.3: *True and estimated probabilities of making a visit before and after failure.*



The effect of compliance cannot be ignored since it affects the estimation of baseline survival probabilities. (See Finkelstein et al (2002)).

7.2 Results from ACTG 175 AIDS Data

Background information on the design and analysis of ACTG 175 study data was described in Chapter 1. In this section we apply some of the models derived in Chapters 3, 4 and 5. Specialized medical terms have been used to describe the study, hence brief information on HIV and the meaning of some of the related key words used are explained in the next section.

7.2.1 Information on HIV and Treatment therapies

Human immunodeficiency virus (HIV) was isolated almost at the same time at three different places by three independent groups of scientists. A French cancer specialist Luc Montaigner who is a scientist at Pasteur Institute in Paris isolated a human retrovirus from the lymph node of a man at risk of Aids. Almost simultaneously, two separate American scientists in Robert Gallo of the National Cancer Institute in Bethesda, Maryland and virologist Jay Levy of the University of California in San Francisco also isolated the same retrovirus from people with Aids. The virus was subsequently named HIV, a virus that plays a pivotal role in progression towards AIDS. HIV is spread through exchange of body fluids like semen, blood and blood products. Since 1984, its been known that HIV enters human cells by binding with a receptor protein called CD4 located on human immune cells surfaces. A person infected with HIV gradually loses immune function, thus rendering him/her fatally susceptible to opportunistic infections (infection by organism that do

not under normal circumstances cause disease except in people whose immune system has been greatly compromised).

Ribonucleic Acid (RNA) has the same structural composition as Deoxyribonucleis Acid (DNA), a genetic material that carries information that determines protein structure in every cellular organism. HIV RNA belongs to the single strand type of viruses called retroviruses. On entering a host, HIV binds to specific receptors (CD4 and Chemokine receptors) found on the surface cells of a host. CD4 T-4 cells are responsible for every human body's immunity. On release into the cell, a viral protein called *Enzyme Reverse Transcriptase* (RT) converts the single strand RNA into double strand DNA and is subsequently joined into the cellular DNA chain for multiplication. After multiplication, a *protease* protein facilitate the formation of new single strand HIV RNA, and this exit the cell ready to infect other cells. During the process of replication, CD4 cells get destroyed directly. The destruction of CD4 occurs gradually in stages that determine the progression of a seroconverted patient, culminating in AIDS when the cell counts fall beyond 200 per millimetre of blood.

The only available intervention on the deteriorating status of an AIDS patient is through the use of Anti-retroviral (ARV) drugs, which are therapeutic in nature. Available ARV exploits the characteristic behaviour of the virus. The nucleosides for instance, are viral DNA-chain terminator in that if the virus

mistakes the drug for viral DNA nucleotide, then it gets attached to the DNA chain, terminating any replication. Some of the available nucleosides are Zidovudine (AZT), Didanosine (DDI), Zalcitabine (DDC) and Stavudine (d4T), among others. Some of these treatment therapies were administered to patients in the ACTG175 study.

7.2.2 Univariate Cox's models on Failure Times

The preliminary descriptive statistics on actual measure of CD4 cell count and plasma concentrations of HIV RNA for the sample data from ACTG175 is given by Katzenstein et al (1996). Results are based on the 348 patients having their follow-up data for both CD4 cells count and HIV RNA levels over the three and half years of study. Table 7.2.1 show the distribution of patients according to failure/censored for each of the CD4 and RNA intervals.

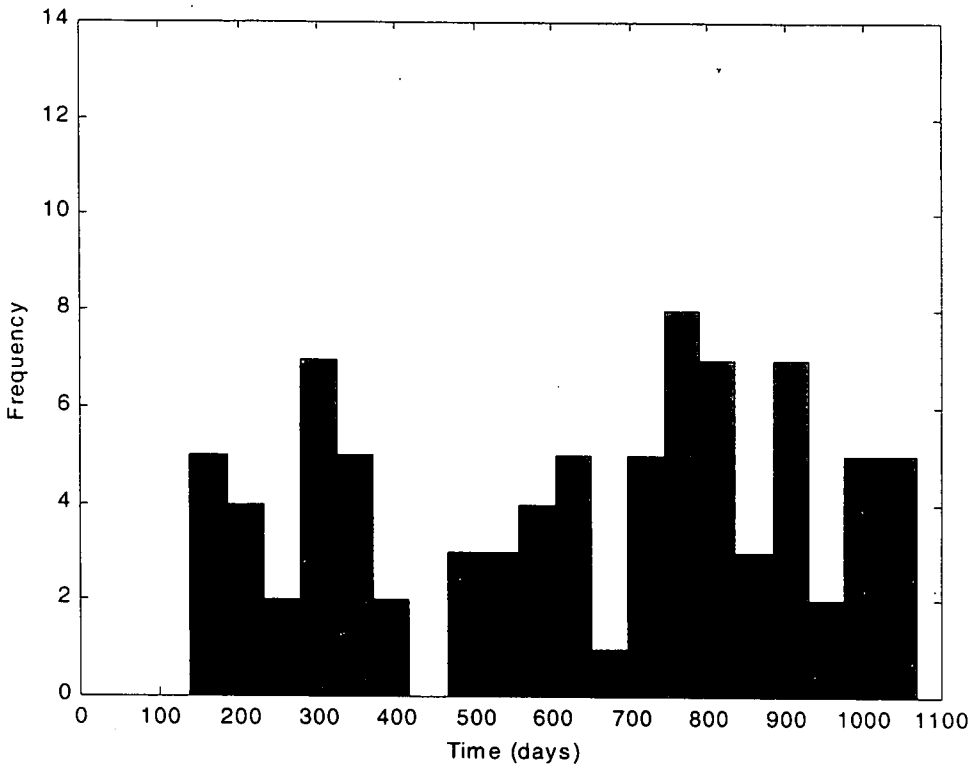
The failure for CD4 variable is a 50% decline in CD4 cells from the baseline entry count, while a unit increase of log base 10 measure of plasma concentration of HIV RNA was used as an endpoint for the RNA variable. Attention should be paid to the fact that this data has overlapping intervals and data reported are the observation's endpoints. Inevitably, some units have their times transcend over several intervals. This will influence the computation of interval conditional probabilities, since any unit's contribution to an interval is taken aboard even if its endpoints are not in that interval.

Table 7.2.1: *The number of observations in each interval of the two failure days.*

	Failure Time				
	CD4		RNA		
Time Interval	Failure (%)	Censored (%)	Time Interval	Failure (%)	Censored (%)
0-56	0 (0)	3(100)	0-56	0 (0)	22 (100)
56-140	1 (8)	12 (92)	56-140	5 (12)	37 (88)
140-224	8 (38)	13 (62)	140-392	26 (38)	42 (62)
224-308	7 (44)	9 (56)	392-694	29 (40)	43 (60)
308-392	8 (50)	8 (50)	694-1164	21 (15)	123 (85)
392-476	2 (25)	6 (75)			
476-560	6 (46)	7 (54)			
560-644	7 (70)	3 (30)			
644-728	6 (43)	8 (57)			
728-812	13 (43)	17 (57)			
812-896	8 (14)	50 (86)			
896-980	7 (14)	45 (86)			
980-1064	9 (11)	70 (89)			
1064-1148	1 (7)	14 (93)			
TOTAL	83	265		81	267

To show the distribution of patients who experienced failure in CD4 cell counts during the study period, Figure 7.2.1 a histogram depict this information. Since patients' exact failure time is not known, shown in the histogram are the upper-end points of a given patient. The failures are clustered into two groups, with the first early group's failure occurring between 140 and 425 days. A temporary absence in of failure, then failures resume at 500 days, peaking at 800 days, and then declines gradually.

Figure 7.2.1: *A frequency histogram to show distribution of time to failure for CD4 cell counts. (failed units only)*



A decline by 50%, seemed to occur mostly between 700 and 1100 days after assignment to treatment as shown on Figure 7.2.1, whereas increase in HIV RNA peaked between 140 and 700 days. A decline in CD4 cells is generally superseded by increase in RNA levels, as indicated in Table 7.2.1. At 140 days, 5 patients under treatment had experienced a unit increase in viral load, yet only one patient had CD4 cells below 50% of the entry value at that time.The time to 50% decline in CD4 cells count does not follow a Weibull distribution, hence parametric inference on this data will be kept at minimum.

The departure point of analysis of ACTG 175 data was through the use of a univariate proportional hazards and odds models on individual failure times, using model (3.16) with all 14 explanatory variables described in Chapter 1 included in the regression. The Metropolis-Hastings algorithm was used with explanatory variables coefficients estimated simultaneously in a one block and the interval specific baseline parameters simultaneously simulated in another. For each block of parameters, a Multivariate Normal was used as a proposal distribution with a vector of either $\tilde{\beta}$ or $\tilde{\gamma}$ (updating values), and variance matrix from information matrix of maximum likelihood estimates. (See Indmleaid.m in Appendix B). Due to the size of the data set hence was slow to make a complete cycle, 5000 iterations where done.

Table 7.2.2: *Estimated posterior means of explanatory variable coefficients for univariate failure times using the PHM model.*

	CD4 Failure Time		**RNA Failure Time**			
	Prior II		**Prior I**		**Prior II**	
Variable	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD
Age	0.012	-0.01, 0.04	0.0009	-0.02,0.02	0.004	-0.02,0.03
Gender	-0.183	-0.44, 0.01	0.2986	-0.06,0.49	0.435	-0.05,0.73
Karnofsky	-0.025	**-0.04, -0.01**	-0.0257	**-0.04,-0.00**	-0.018	-0.02,0.00
ZDV	0.573	**0.43, 0.70**	0.6587	**0.53,0.84**	0.511	**0.41,0.70**
Weight	-0.008	-0.03, 0.01	-0.0131	-0.03,0.00	-0.012	-0.03,0.00
Homosexual	0.341	**0.16, 0.57**	0.0701	-0.27,0.36	0.054	-0.32,0.31
Symptomatic	0.958	**0.79, 1.06**	0.5417	**0.41,0.66**	0.515	**0.45,0.59**
ID use	-0.360	-0.85, 0.08	-0.2965	-0.98,0.23	-0.320	-0.91,0.12
Ethnic white	0.951	**0.54, 1.28**	0.5491	**0.12,0.93**	0.405	-0.08,0.80
Ethnic black	1.162	**0.59, 1.66**	0.6844	-0.03,1.28	0.494	-0.10,0.93
Hispanic	0.501	-0.10, 0.95	1.0021	**0.57,1.44**	0.692	**0.37,1.15**
DdI	-0.419	**-0.51, -0.30**	-0.6322	**-0.78,-0.50**	-0.725	**-0.88,-0.49**
AZT + ddI	-1.051	**-1.13, -0.94**	-0.9234	**-1.16,-0.92**	-1.074	**-1.20,-0.99**
AZT + ddC	-0.713	**-0.78, -0.65**	-0.5438	**-0.72,-0.34**	-0.702	**-0.90,-0.60**

The results (Tables 7.2.2, 7.2.3 and 7.2.4) emanating from univariate analysis of the sample data suggest that neither age, gender and weight has significant effect on both time to 50% decline in CD4 cell counts and time to increase in viral load measured by HIV RNA for all models applied. The two transform models PHM and POM applied under Bayesian method gave similar results, but the PHM under maximum likelihood estimation method gave out differing results.

The Bayesian approach has the following variables featuring in all transforms irrespective of prior type, as having significant effect on both failure times: History of antiretroviral use (ZDV), having AIDS symptoms at entry and using DDI, AZT+DDI or AZT+DDC. ZDV and symptomatic hasten time to failure. The effect of all ARV's is slowing progression towards failure relative to AZT alone, as shown by negative estimated values even though it has slightly varying effect on both failure times and among themselves. The effect is in the following order: AZT+DDI is more potent, followed by AZT+DDC, DDI and AZT in that sequence. The rest of the variables fluctuate between having effect for one model, and not for the other. What is apparent is that some covariates have differing effect on the two failure times. A high Karnofsky score at the commencement of the study signify a lower risk in time to CD4 decline, yet doesn't have effect on time to increase in viral load. This is also true with homosexual tendencies, which increases the risk of

decline in CD4 cells but doesn't have effect on time to unit increase in log of viral RNA.

Using the classical approach by estimating MLE's using univariate PH models, results show (Table 7.2.4) that only ZDV, symptomatic signs, and all ARV affect the progression time to decline in CD4 cell count. Meanwhile time to a unit increase in $\log_{10}$ of viral load is affected by ZDV, symptomatic signs and AZT+DDI combination only. The rest of the variables are not significant. This begins to show the difference between the two approaches.

Table 7.2.3: *Estimated posterior means of explanatory variable coefficient for univariate failure times using the POM model.*

	CD4 Failure Time				**RNA Failure Time**	
	Prior I		**Prior II**		**Prior I**	
Variables	*Mean*	*95% HPD*	*Mean*	*95% HPD*	*Mean*	*95% HPD*
Age	0.0107	-0.01,0.03	0.008	-0.01,0.03	0.0012	-0.03,0.03
Gender	-0.2209	-0.54,0.11	0.024	-0.16,0.37	0.3537	**0.02,0.60**
Karnofsky	-0.0397	**-0.05,-0.02**	-0.027	**-0.04,-0.01**	-0.0242	**-0.04,-0.01**
ZDV	0.6394	**0.43,0.82**	0.680	**0.51,0.80**	0.6088	**0.52,0.73**
Weight	-0.0100	-0.02,0.01	-0.007	-0.02,0.01	-0.0116	-0.03,0.01
Homosexual	0.1143	-0.23,0.42	0.096	-0.11,0.29	0.1684	-0.14,0.53
Symptomatic	0.8893	**0.70,1.07**	0.905	**0.82,0.97**	0.4844	**0.31,0.89**
ID use	-0.7625	-1.63,0.21	-0.441	**-0.78,-0.11**	-0.0473	-0.49,0.41
Ethnic white	1.8027	**1.20,2.33**	0.871	**0.52,1.28**	0.7449	-0.10,1.35
Ethnic black	2.0732	**1.20,2.66**	1.049	**0.62,1.81**	1.0667	**0.41,1.61**
Hispanic	1.0088	**0.31,1.79**	0.200	-0.24,0.78	1.2733	**0.64,1.87**
Ddi	-0.4967	**-0.69,-0.33**	-0.506	**-0.62,-0.43**	-0.6017	**-0.79.-0.49**
AZT + ddi	-1.2413	**-1.52,-1.09**	-1.046	**-1.16,-0.89**	-1.3951	**-0.79,-0.49**
AZT + ddc	-0.7500	**-1.07,-0.54**	-0.577	**-0.72,-0.49**	-0.6064	**-0.75,-0.49**

The other variable that has a significant effect on time to decline in CD4 cells counts is ethnicity (white and black), which has a high risk to decline in CD4

cells. Thus both white and black ethnicity has increased risk of time to decline in CD4 in comparison to Other races.

The results mirror the ANOVA findings by Katzeintein et al (1996), using the means of CD4 cell counts and log of HIV RNA concentration levels found per cubic millimetre of blood extracted from patients. A univariate analysis of time-to-increase in HIV RNA using the same covariates reveals a similar trend with some minor deviations.

Figure 7.2.3: *Frequency histograms from 5000 MCMC simulations for some of the significant covariates using POM with Prior I on RNA time to failure.*

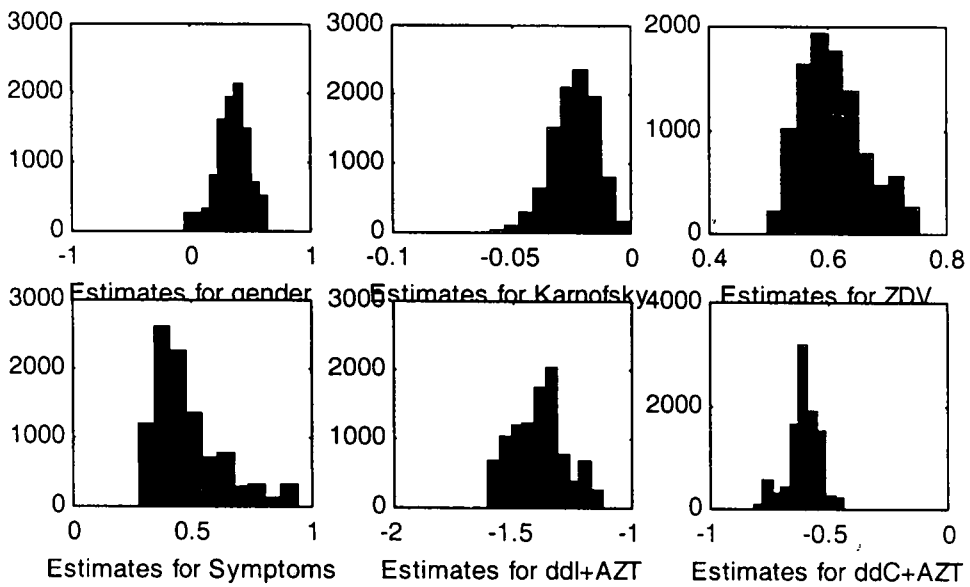


Table 7.2.4: *Maximum Likelihood estimates of explanatory variable coefficients using univariate PH models.*

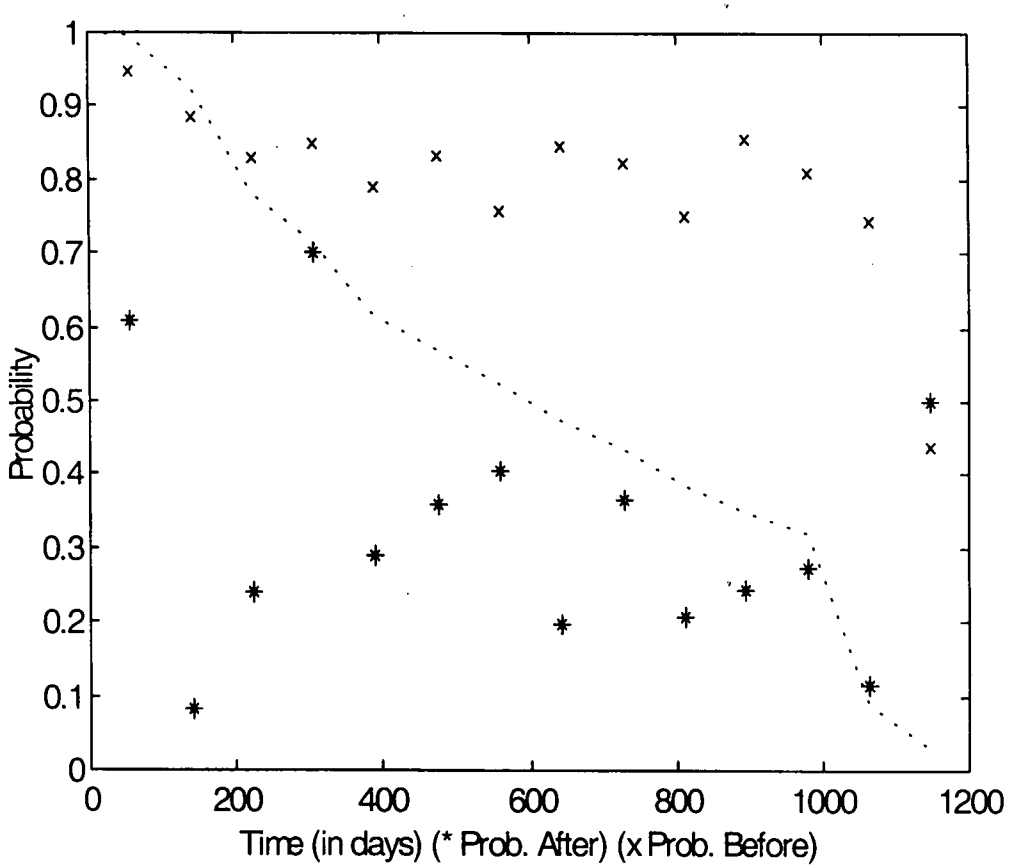
	CD4 Failure Time			RNA Failure Time		
Variable	**Estimate**	**STD**	**p-value.**	**Estimate**	**STD**	**p-value**
Age	0.0187	0.0125	0.1336	0.0112	0.0128	0.3844
Gender	0.1027	0.4648	0.8258	-0.2091	0.5302	0.7040
Karnofsky	-0.0121	0.0103	0.2420	0.0044	0.0109	0.6892
ZDV	0.7351	0.2364	**0.0020**	0.6523	0.2363	**0.0058**
Weight	-0.0026	0.0099	0.7948	-0.0069	0.0099	0.4902
Homosexual	-0.3774	0.4116	0.3576	0.3236	0.4916	0.5092
Symptomatic	1.0400	0.2857	**0.0000**	0.7100	0.2866	**0.0136**
ID use	-0.4419	0.3699	0.2340	-0.2779	0.3955	0.4840
white	-0.1157	0.8647	0.8966	0.1473	0.8916	0.8650
Black	-0.0529	0.9092	0.9522	0.2800	0.9409	0.7718
Hispanic	-0.7715	0.9153	0.4010	0.6839	0.8979	0.4472
DdI	-0.6389	0.3003	**0.0340**	-0.5771	0.3039	0.0574
AZT + ddI	-1.2693	0.3652	**0.0006**	-1.1190	0.3687	**0.0024**
AZT +ddC	-0.7965	0.2991	**0.0078**	-0.5692	0.2944	0.0536

7.2.3 Non-Compliance Effect on Parameter Estimation: ACTG 175 Data

To assess the impact of non-compliance to visits by study patients on the parameter estimation, a univariate analysis of the CD4 visits was carried out. CD4 failure time was preferred because patients had their CD4 levels monitored more regularly than for the RNA failure time. The only problem was that for ethical reasons, patients who experienced failure were removed from the study hence their after-failure visits could not be ascertained. An MCMC method on the joint posterior using a non-parametric approach with PH transform was used, (Visposaid.m in Appendix B).

The probability of visiting is high (all larger than 0.70) for all intervals except the last one, before the occurrence of a 50% decline in CD4 cells. The lowest visiting probability prior to CD4 decline is at the last interval estimated to be 0.44. Meanwhile the probabilities of continuing to visit after 50% decline are low, all except for the 1st and 4th intervals are below 0.5. This may be explained by the fact that some patients had their treatment changed once there was a 50% decline in CD4 cells, while some had died, hence the measure for post-failure visits is not observed by a lot of patients.

Figure 7.2.4: *Visiting probabilities prior and after 50%decline in CD4 cells and interval survival probabilities estimated with compliance effect.*



The high prior-failure probabilities imply that the effect of compliance can be ignored. Thus we expect the method that accounts for visiting chances not to differ much from the one that excludes the visiting effect.

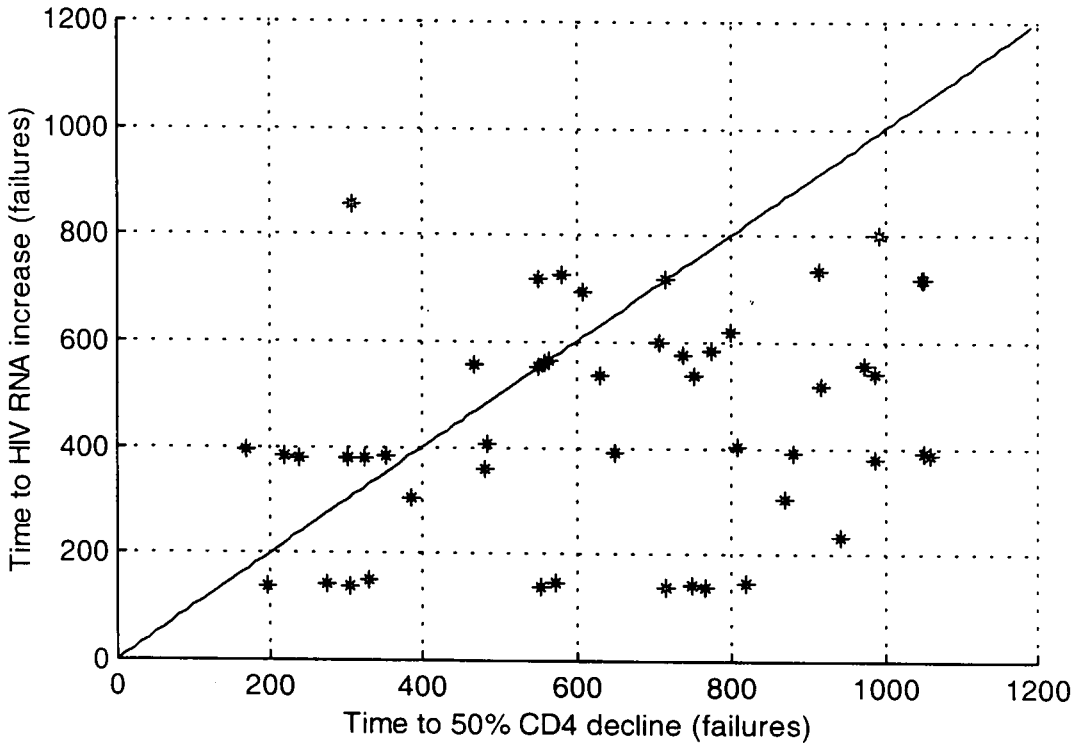
7.2.3 Bivariate Models on ACTG 175 Data

As elucidated in early chapters, the bivariate models endeavour to illustrate the joint effect of explanatory variables on the two failure times. In general we envisage the parameter estimates from individual analysis of failure times to be averaged out with some weighting dependent on the number of observations experiencing that phenomenon for the given explanatory variable. For instance, with time to increase in RNA, 25% of females had experienced one unit ($\log_{10}$) increase in viral load by end of the study, compared to 23% of males. Thus for this phenomenon, males may have a higher conditional survival probability. Meanwhile with time to 50% decline in CD4 cells counts, 25% of males experience the decline against 22% of females. The issue then is, given the gender of a patient can we determine the joint risk of time-to 50% decline in CD4 cell counts and a unit increase in log of HIV RNA concentrations after undergoing treatment?

For those patients who experienced both failures in shown on Figure 7.2.5, majority of them experienced a failure in RNA before CD4 failure. Only one patient's RNA failure occurred after 800 days, with the majority occurring before 2 years. Meanwhile CD4 failures took longer than that. Some patients

experienced failure in only one of the times (Figure 7.2.5), these are points depicted on the axis.

Figure 7.2.5: *A scatter diagram of time to failure for CD4 and HIV RNA (failures only), the 45degree line shows equal times.*



Four models, the independence assumption (IW) model, bivariate conditional (CB) model, non-parametric Clayton Copula (CC) and use of Weibull marginal distributions with the Clayton copula (CCW) will be applied to the ACTG 175 bivariate data. For each of the non-parametric methods, either a proportional hazard or proportional odds model is used in a combination with priors I and II and, a Metropolis-Hastings algorithm applied to estimate the posterior distributions for the parameters.

Figure 7.2.6: *A scatter plot of CD4 follow-up time against RNA follow-up time in days.*

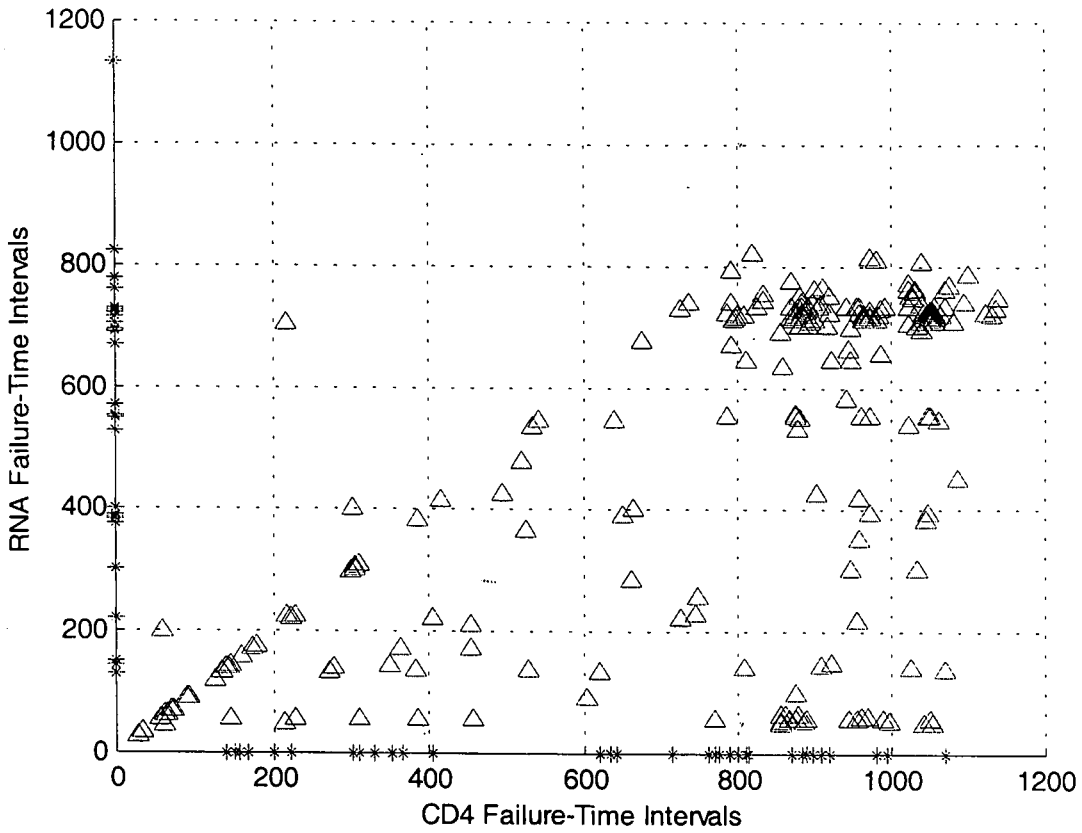


Table 7.2.5 summarizes the bivariate failure time data showing the frequencies of failure on time to decline in CD4 for a given time of RNA follow ups.

Table 7.2.5: *Frequencies of failure /censored and estimated survival probabilities of CD4 time for given failure time interval of RNA time, computed using CBH with prior II.*

	RNA Time (days)					
CD4 Time (days)	**0-56**	**56-140**	**140-392**	**392-694**	**694-1164**	**Total**
0-56	0(3) 0.6009	0(0) 0.8717	0(0) 0.8663	0(0) 0.8978	0(0) 0.9711	0(3) 0.9887
56-140	1(1) 0.3605	0(9) 0.5732	0(2) 0.6255	0(0) 0.7622	0(0) 0.7798	1(12) 0.8999
140-224	2(1) 0.2346	2(1) 0.3545	3(10) 0.3570	1(0) 0.5788	0(1) 0.7253	8(13) 0.8401
224-308	0(1) 0.0886	1(1) 0.2074	5(6) 0.2361	0(1) 0.3947	1(0) 0.5899	7(9) 0.7217
308-392	0(0) 0.0533	2(3) 0.1186	5(5) 0.1628	1(0) 0.3156	0(0) 0.5622	8(8) 0.6942
392-476	1(0) 0.0412	0(1) 0.0351	0(3) 0.0843	1(2) 0.1649	0(0) 0.5411	2(6) 0.6451
476-560	0(0) 0.0161	1(1) 0.0195	1(1) 0.0524	3(5) 0.0764	1(0) 0.4663	6(7) 0.5776
560-644	0(0) 0.0075	0(2) 0.0167	1(0) 0.0308	3(1) 0.0687	3(0) 0.3817	7(3) 0.4805
644-728	0(0) 0.0022	1(0) 0.0015	0(3) 0.0111	3(4) 0.0355	2(1) 0.2694	6(8) 0.3978
728-812	1(0) 0.0005	2(1) 0.0001	1(3) 0.0038	7(2) 0.0165	2(11) 0.1563	13(17) 0.3060
812-896	0(4) 0.0001	3(6) 0.0000	2(1) 0.0012	2(11) 0.0131	1(28) 0.1388	8(50) 0.2522
896-980	0(2) 0.0001	0(2) 0.0000	2(6) 0.0002	3(10) 0.0060	2(25) 0.0855	7(45) 0.1954
980-1064	0(5) 0.0000	0(2) 0.0000	2(5) 0.0000	3(8) 0.0023	4(50) 0.0786	9(70) 0.1603
1064-1148	0(0) 0.0000	0(1) 0.0000	0(1) 0.0000	0(1) 0.0007	1(12) 0.0485	1(14) 0.1243
Total	3(16)	12(30)	22(46)	27(35)	17(128)	83(265)

Also given are the corresponding interval survival probabilities for given RNA failure time. These are results obtained from the CBH model (3.32), computed by using (3.28). The results suggest that if a patient's viral load increases shortly after taking treatment then his/her survival probability for time to decline in CD4 cell count will be lowered.

Fitting a Clayton model (Claposaid.m in Appendix B), the estimated measure of association between time to 50% decline in CD4 cells and time to a unit ($\log_{10}$) increase in viral load is 0.25 and 0.30 for CCH with prior I and II respectively, while CCO gives 0.70 and 0.53 for prior I and II respectively. This represents a fairly strong positive dependence between the two failure times. This implies that if the suppression of viral multiplication by Anti-retroviral drugs takes longer, then we expect commensurately longer time towards decline in CD4 cells.

Table 7.2.6a: *Estimated posterior means of covariate parameters from fitting Independence assumption model, Conditional bivariate and Clayton copula using proportional hazard with prior I.*

	IWH		*CBH*		*CCH*	
Variables	*Mean*	*95% HPD*	*Mean*	*95% HPD*	*Mean*	*95% HPD*
Age	0.012	-0.01, 0.03	-0.0004	-0.02,0.01	0.002	-0.02,0.02
Gender	0.259	-0.14, 0.68	-0.069	-0.13,0.00	0.153	-0.07,0.33
Karnofsky	-0.030	**-0.05, -0.02**	-0.022	**-0.03,-0.01**	-0.018	-0.03,0.00
ZDV	0.872	**0.73, 0.98**	0.566	**0.55,0.59**	0.561	**0.52,0.59**
Weight	-0.006	-0.02, 0.02	-0.001	-0.01,0.01	-0.010	-0.03,0.00
Homosex	-0.038	-0.31, 0.13	0.013	-0.15,0.20	-0.405	-0.73,0.05
Symptom	0.954	**0.84, 1.10**	0.583	**0.49,0.67**	0.870	**0.78,0.97**
ID use	-0.569	-1.55, 0.27	-0.080	-0.40,0.25	-0.811	**-1.19,-0.19**
white	0.419	**0.06, 0.86**	0.089	-0.26,0.33	0.446	-0.11,1.16
black	0.803	**0.06, 1.13**	0.071	-0.25,0.32	0.237	-0.42,1.16
Hispanic	-0.475	-0.83, 0.11	-0.007	-0.47,0.27	0.431	-0.13,0.98
DDI	-0.345	**-0.47,-0.27**	-0.633	**-0.68,-0.57**	-0.642	**-0.71,-0.60**
AZT + DDI	-1.042	**-1.18,-0.86**	-0.860	**-0.92,-0.80**	-1.150	**-1.21,-1.07**
AZT + DDC	-0.720	**-0.77,-0.67**	-0.680	**-0.71,-0.65**	-0.629	**-0.67,-0.58**
Assoc.					0.25	**0.18,0.36**

The presence of association renders the independence assumption results inefficient as shown in Table 7.2.6a and 7.2.6b. The estimates returned by the method tend to be larger with large standard errors, as shown by its HPD

intervals. Meanwhile the Conditional bivariate model performs well and its results are comparable to the ones from the Clayton model.

Table 7.2.6b: *Estimated posterior means of covariate parameters from fitting Independence assumption model, Conditional bivariate and Clayton copula using proportional hazard with prior II.*

	IWH		*CBH*		*CCH*	
Variables	*Mean*	*95% HPD*	*Mean*	*95% HPD*	*Mean*	*95% HPD*
Age	0.009	-0.02, 0.03	0.002	-0.01, 0.01	0.002	-0.02, 0.02
Gender	0.089	-0.07, 0.24	-0.246	-0.45, 0.02	0.036	-0.24, 0.28
Karnofsky	-0.029	**-0.05,-0.01**	-0.024	**-0.03,-0.01**	-0.019	**-0.03, -0.01**
ZDV	0.738	**0.65,0.85**	0.558	**0.53, 0.59**	0.528	**0.45, 0.58**
Weight	-0.004	-0.02,0.02	-0.001	-0.01, 0.01	-0.009	-0.02, 0.01
Homosex	0.207	**0.03,0.38**	0.1657	-0.08,0.34	-0.030	-0.26, 0.18
Symptom	0.915	**0.80, 1.15**	0.6676	**0.51, 0.78**	0.9216	**0.82, 1.02**
ID use	-0.046	-0.16,0.07	-0.124	-0.51, 0.20	-0.297	-0.69, 0.27
white	0.729	**0.14,1.26**	0.276	**0.01, 0.52**	0.341	-0.56, 1.32
black	0.547	**0.03, 0.93**	0.274	**0.06, 0.50**	0.239	-0.89,1.23
Hispanic	0.233	-0.19,0.75	0.145	**-0.17, 0.45**	0.386	-0.59, 1.47
DDI	-0.534	**-0.61, -0.46**	-0.596	**-0.68, -0.54**	-0.649	**-0.74, -0.57**
AZT + DDI	-1.059	**-1.17, -0.95**	-0.856	**-0.91, -0.81**	-1.18	**-1.24,-1.12**
AZT + DDC	-0.511	**-0.68, -0.37**	-0.704	**-0.75, -0.66**	-0.692	**-0.73, -0.64**
Assoc.					0.30	**0.23,0.38**

Results show that before study use of anti-retroviral therapy (ZDV) by patients, having HIV related symptoms at entry and treatment therapies are variables that consistently and significantly affect time to progression towards a 50% decline in CD4 cells and a unit increase (log 10) in viral load for patients undergoing ARV treatment, irrespective of model used. Thereafter different variables' effect on time-to-event depends on which model is used. For instance, level of Karnofsky score at entry significantly affect progression time if we assume independence between failure times (IW), or analysing

data for a known interval of the other failure time (CB model), yet it is not so if we use a Clayton copula with non-parametric approach. Race and intravenous drug use, are the other variables with effects dependent on the model used.

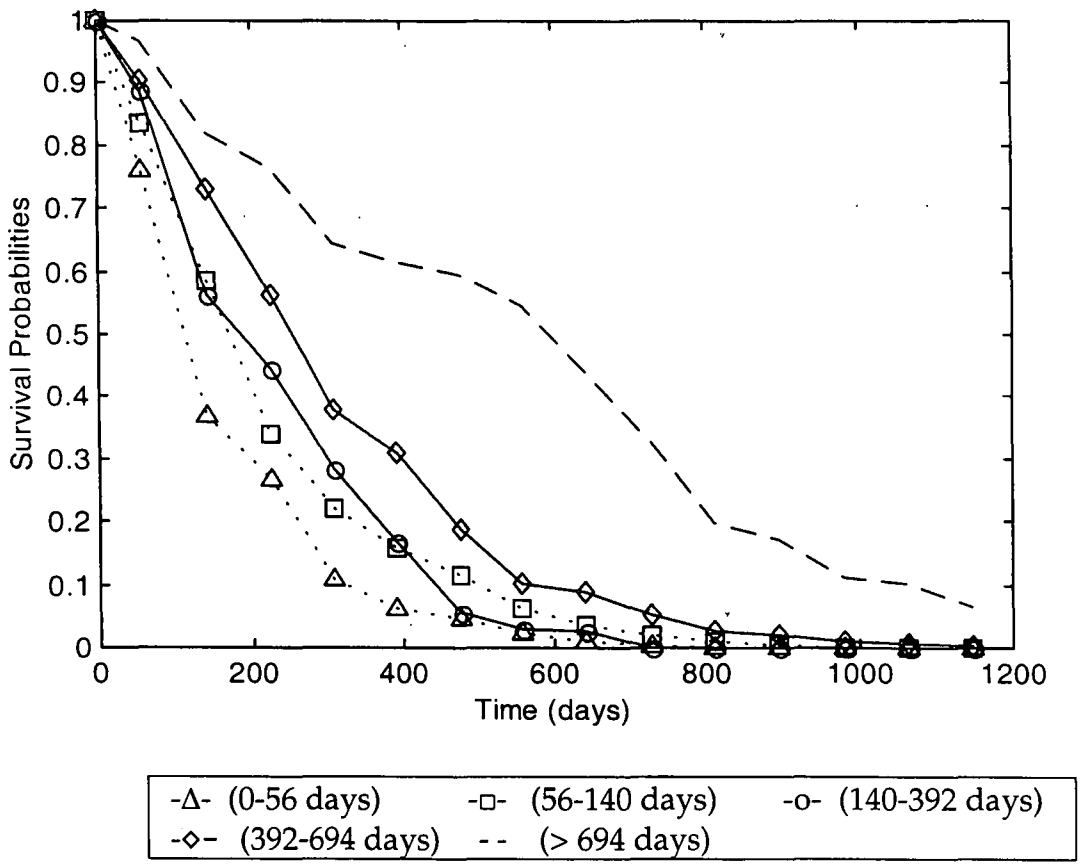
The baseline survival probability estimates $\hat{\alpha}_j$ generated using a CB model with baseline parameters having prior II (5.11) applied to a proportional odds model are shown on Figure (7.2.7). Results obtained from proportional hazard and proportional odds models do not differ much.

Table 7.2.7: *The baseline survival probabilities estimated using CB and CC for proportional hazards and odd models at RNA time intervals.*

	Model	CC		CB	
Prior	**Interval**	POM	PHM	POM	PHM
I	0	1	1	1	1
	0-56	0.9924	0.9881	0.8609	0.8892
	56-140	0.5437	0.4357	0.4872	0.3727
	140-392	0.3799	0.2433	0.2112	0.0945
	392-694	0.2824	0.2103	0.0742	0.0190
	694-1164	0.1170	0.0690	0.0010	0.0000
II	0	1	1	1	1
	0-56	0.9927	0.9822	0.8911	0.8533
	56-140	0.5109	0.5015	0.5000	0.3726
	140-392	0.3185	0.1652	0.2096	0.1229
	392-694	0.1941	0.0177	0.0741	0.0176
	694-1164	0.0108	0.0005	0.0008	0

Both indicate that if a patient's viral load increases at early stage of therapy, then their progression to 50% decline in CD4 cells will also occur early, hence low survival probabilities. For instance, a patient whose RNA level failure time occurs at interval between 0 and 56 days, will only have a 10% chance of surviving beyond 300 days before 50% decline in CD4 cells occurs.

Figure 7.2.7: *Baseline survival probabilities of CD4 time for given RNA intervals obtained using Conditional Bivariate with prior II on proportional odds model*



Meanwhile if a patient's viral load increase occurs nearly two years after treatment, there is 30% chance that their CD4 cell count will not have declined by 50% in the same period. It is easy to establish that a viral load increase

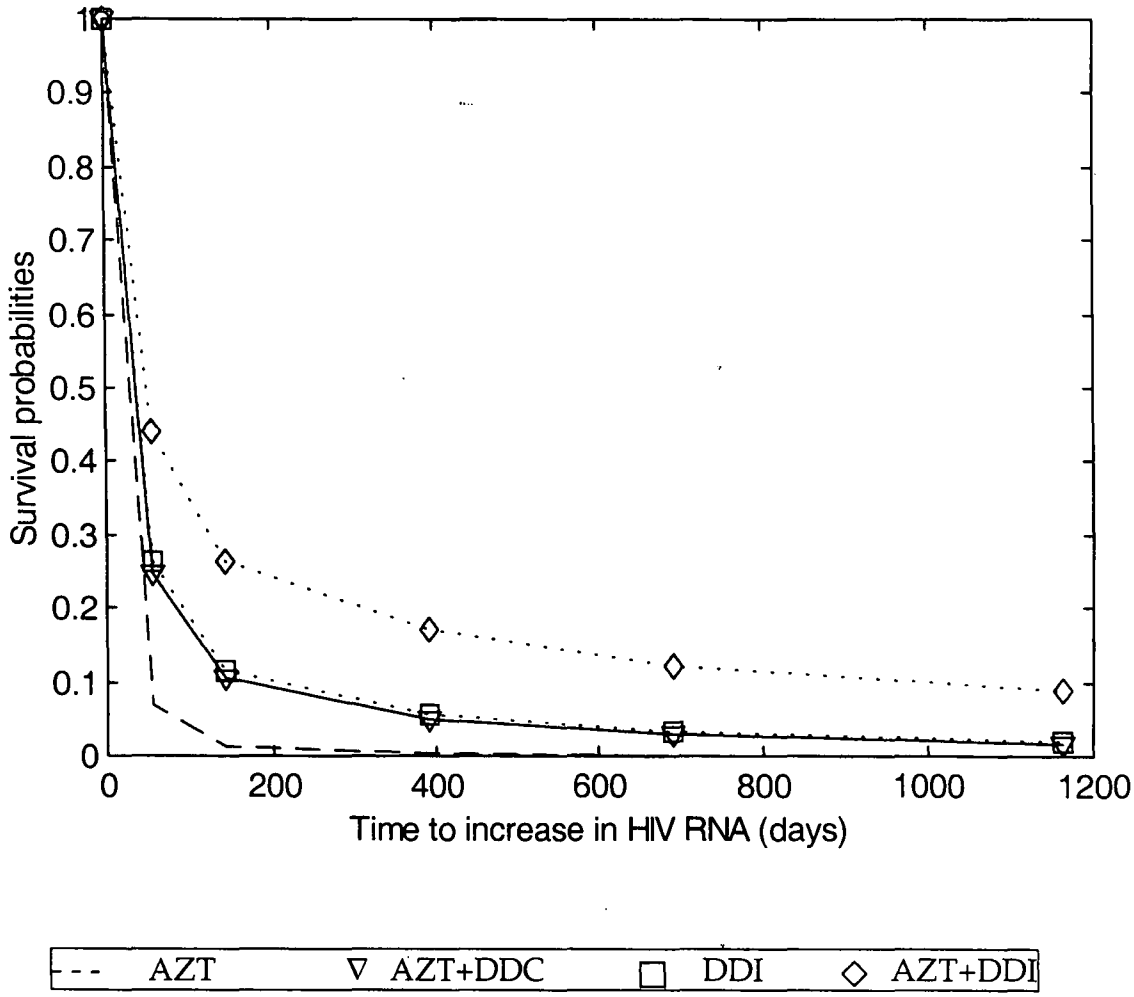
precede a decline in CD4 cells with low CD4 survival probabilities as duration of study increases. If a patient's viral HIV RNA increase of 1 unit takes place at 0-56 days interval, then there is between 75%-100% chance for CBO model, and 60%-100% chance for CBH model, that this patient's CD4 cell count will not decreased by 50% at this time.

Also, if viral increase occurs at 694-1164 days, then there is between 70% and 90% chance that 50% decline in CD4 cell count has occurred. Thus for monitoring, it may be advisable to regularly monitor the viral load of a patient undergoing treatment for early warning of adverse effects, than relying on CD4 cells count, despite costs involved in using the former. To illustrate the impact of treatment therapy variable on survival times, we plot the baseline survival probabilities of time to increase in viral load (equivalent to AZT mono-therapy) in Figure 7.2.8, and then show the effect of each of the treatment therapies under a Clayton Copula with prior II for baseline parameters and proportional hazard model.

The superior effect of combination AZT and DDI is apparent. The mono-therapy of AZT results in very low survival rate for patients on that therapy. The effect of DDI alone and AZT + DDC is significantly superior to that of AZT alone, but the distinction between the two is negligible. Survival from viral increase is low for AZT patients, at a mere 5% at 56 days, yet AZT+DDI boost its patients to a 45% chance for the same period. Whereas patients have no chance of survival after 3 years of AZT therapy, those on combination

therapy, still have a 10% chance of survival. Also given that one's survival interval from increase in viral load is known, a history of prior use of anti-retroviral at the inception of the study renders one having high risk towards a decline in CD4 cells. The same applies to homosexuals, while ethnic blacks and Hispanics are at high risk compared to Other races.

Figure 7.2.8: *Showing the effect of four treatment therapies on estimated survival for RNA probabilities using CCH model with prior II for baseline parameters.*



7.3 Chapter Summary

On simulated data, bivariate data generated using a Farlie-Morgenstern copula with a maximum allowable dependence measure under this copula, was analysed using independence assumption model. The model proved to be inefficient for small sample size and presence of dependence between failure times. An improvement in using the method of maximum likelihood was only realized after increasing the sample size, and use of sandwich estimator on variance. A nonparametric Bayesian method was also applied with PH transform. A type II prior that distinguishes baseline parameters belonging to different intervals was preferred for use.

The focus of the study culminates in this chapter when ACTG175 AIDS data is analyzed. A strong association between CD4 and HIV RNA failure times was established. The use of either Clayton copula or Conditional bivariate method with Prior II was found more appealing. Results further showed that certain factors play a significant role in the progression of the failure times. These are history prior of ARV's and showing AIDS symptoms at entry. These have adverse effects. Meanwhile the treatment therapies ranked according to effectiveness as AZT+DDI, DDI and AZT+DDC, while AZT alone is the lease effective of the four.

CHAPTER 8

RECOMMENDATIONS AND CONCLUSIONS

The effect of dependence among multiple failure times is critical and cannot be ignored. Its effect on small sample sizes when MLE are used can yield biased results in parameter estimation. The use of robust estimators was found to work well when sample sizes are large enough. The Bayesian approach of using MCMC methods on copulas is robust and mathematically less intense, but computationally intense. Its usefulness is evidenced by its ability to estimate for small samples, as long as proper priors like type I and II are used. In all applications prior II proved to perform better since it make a distinction between different baseline parameters in each failure interval.

Methods for testing for goodness of fit on survival data with overlapping intervals are not so developed. It is therefore difficult to make conclusive decisions on the use of parametric methods. This limits the use of parametric methods, which otherwise involves fewer parameters. On comparable basis, the non-parametric method of either Cox' proportional hazard or proportional odds applies and combines well with copulas (Clayton) when using MCMC algorithms. Once again the MLE method is confined in application by its mathematical intensity.

Methods that estimate the measure of association between failure times are highly recommended. It remains to be seen if these copulas can be extended to handle multiple failures exceeding two, since this would require multiple correlation matrix. Through the Clayton method the presence of strong association between the markers of the two failure times in the ACTG 175 data was determined which enables prediction of one variable using the other. To determine the level of HIV RNA in a patient is expensive, yet early detection of viral increase is important. Thus knowledge of levels of CD4 can be used to augment sparse data on time to viral load increase, which would prove to be cost effective.

The issue of assessing how well the Cox's model of using the hazard function fit interval data with overlapping intervals need to be explored further. In this study, an ad-hoc approach of assessing the behaviour of the sampling distributions of the estimated parameter values was used, this hinges on the large sample theory. Ibrahim et al (2001) suggested some Bayesian model diagnostics that can be used on simpler interval-censored data when the intervals are not overlapping.

APPENDIX A

Appendix A1: Derivatives for Grouped Interval Data: PHM

$$\frac{\partial \ell(\beta,\gamma)}{\partial \beta_l} = \sum_{i=1}^{n}\sum_{m=1}^{M} z_i \left[\varphi_{mij} \frac{(e^{-e^{\gamma_{mj}+z_i\beta}})(e^{\gamma_{mj}+z_i\beta})}{(1-e^{-e^{\gamma_{mj}+z_i\beta}})} - \sum_{s=1}^{j-1} e^{\gamma_{ms}+z_i\beta} \right], \qquad \text{for} \qquad s<j$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_{ms}} = -\sum_{i=1}^{n} e^{\gamma_{ms}+z_i\beta} \qquad \text{for } s<j$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_{mj}} = \sum_{i=1}^{n} \varphi_{mij} \frac{(e^{-e^{\gamma_{mj}+z_i\beta}})(e^{\gamma_{mj}+z_i\beta})}{(1-e^{-e^{\gamma_{mj}+z_i\beta}})} \qquad \text{for } s = j$$

For $m \neq m'$, the second derivatives are

$$\frac{\partial \ell(\beta,\gamma)}{\partial \beta_{ml}\beta_{ml'}} = 0$$

$$\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_{m}\partial \gamma_{m'}} = 0$$

$$-\frac{\partial^2 \ell(\beta,\gamma)}{\partial \beta_l \partial \beta_{l'}} = \sum_{i=1}^{n}\sum_{m=1}^{M} z_i z_i' \left[\varphi_{mij} \frac{(e^{-e^{\gamma_{mjt}+z_t\beta}})(e^{\gamma_{mjt}+z_t\beta})\left[e^{-e^{\gamma_{mjt}+z_t\beta}} + e^{\gamma_{mjt}+z_t\beta} - 1\right]}{(1-e^{-e^{\gamma_{mjt}+z_t\beta}})(1-e^{-e^{\gamma_{mjt}+z_t\beta}})} - \sum_{s=1}^{j-1} e^{\gamma_{ms}+z_t\beta} \right]$$

$$-\frac{\partial \ell(\beta,\gamma)}{\partial \gamma_{ms}^2} = \sum_{i=1}^{n} e^{\gamma_{ms}+z_i\beta} \qquad \text{for } s<j$$

$$-\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{mj}^2} = \sum_{i=1}^{n}\left[\varphi_{mij}\frac{(e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})(e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})\left[e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}-1\right]}{(1-e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})(1-e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})}\right] \quad \text{for s=j}$$

$$-\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{mj}\gamma_{mv}} = 0 \qquad \text{if } v \neq j$$

$$\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \boldsymbol{\beta}\partial \gamma_{mj}} = \sum_{i=1}^{n}\mathbf{z}_i\left[\varphi_{mij}\frac{(e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})(e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})(e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}-1)}{(1-e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})(1-e^{-e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}})}\right]$$

Appendix A2: Derivatives for Grouped Interval Data: POM

Generally the log likelihood for proportional odds model is written in (3.16):

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{m=1}^{M}\sum_{i=1}^{n}\log\left[\sum_{j=1}^{r}\omega_{mij}\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}\prod_{s=1}^{j-1}\frac{1}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}\right]$$

If an observation is known to be censored in the j^{th} interval, then

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{m=1}^{M}\sum_{i=1}^{n}\log\left[\prod_{s=1}^{j}\frac{1}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}\right]$$

and

$$\frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_{ms}} = -\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}$$

$$\frac{\partial\,\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_l} = -\mathbf{z}_i\sum_{s=1}^{j}\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}$$

$$-\frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_{ms}^2} = \frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

$$\frac{\partial\,\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_l\partial\beta_{l'}} = -\mathbf{z}_i'\mathbf{z}_i\sum_{s=1}^{j}\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_l \partial \gamma_{ms}} = -\mathbf{z}_i \frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

If unit fails during the j^{th} interval

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{m=1}^{M}\sum_{i=1}^{n} \log\left[\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}\prod_{s=1}^{j-1}\frac{1}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}\right]$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{ms}} = -\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{mj}} = \frac{1}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_l} = -\mathbf{z}_i\left[\frac{1}{1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}} - \sum_{s=1}^{j-1}\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}\right]$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{ms}^2} = -\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_{mj}^2} = \frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_l \partial \beta_{l'}} = -\mathbf{z}_i'\mathbf{z}_i\left[\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})^2} - \sum_{s=1}^{j-1}\frac{e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{ms}+\mathbf{z}_i\boldsymbol{\beta}})^2}\right]$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_l \partial \gamma_{ms}} = -\mathbf{z}_i\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \beta_l \partial \gamma_{mj}} = \mathbf{z}_i\frac{e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}}}{(1+e^{\gamma_{mj}+\mathbf{z}_i\boldsymbol{\beta}})^2}$$

Appendix A3: Derivation of Category-Specific Probabilities for (6.12) and (6.13).

$$\log \frac{Pr(Y_i = j)}{Pr(Y_i = r)} = \log\left(\frac{\Pi_{ij}}{\Pi_{ir}}\right) = \mathbf{z}_i \boldsymbol{\beta}_j \qquad j = 1,2,\ldots,r\text{-}1$$

$$\pi_{ij} = \pi_{ir} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)$$

$$\sum_{j=1}^{r-1} \pi_{ij} = \sum_{j=1}^{r-1} \Pi_{ir} \exp(\mathbf{z}_i \boldsymbol{\beta}_j) \qquad \left(\sum_{j=1}^{r-1} \pi_{ij} = 1 - \pi_{ir}\right)$$

$$1\text{-}\pi_{ir} = \pi_{ir} \sum_{j=1}^{r-1} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)$$

$$\pi_{ir} = \frac{1}{1 + \sum_{j=1}^{r-1} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)}$$

and $$\pi_{ij} = \frac{\exp(\mathbf{z}_i \boldsymbol{\beta}_j)}{1 + \sum_{j=1}^{r-1} \exp(\mathbf{z}_i \boldsymbol{\beta}_j)}$$

Appendix A4: Deriving Conditional Posterior Distributions for r Nominal Categories with One Covariate used in (6.26).

For Y=j

$$w_{ij} < \frac{\exp(z_i\beta_j)}{1+\sum_{s=1}^{k-1}\exp(z_i\beta_s)}$$

$$w_{ij}(1+\sum_{s=1}^{r-1}\exp(z_i\beta_s)) < \exp(z_i\beta_j)$$

$$w_{ij}(1+\sum_{s\neq j}^{r-1}\exp(z_i\beta_s)) < \exp(z_i\beta_j) - w_{ij}\exp(z_i\beta_j)$$

$$\frac{w_{ij}}{1-w_{ij}}(1+\sum_{s\neq j}^{r-1}\exp(z_i\beta_s)) < \exp(z_i\beta_j)$$

$$\ln\left[\frac{w_{ij}}{1-w_{ij}}(1+\sum_{s\neq j}^{r-1}\exp(z_i\beta_s))\right] < z_i\beta_j$$

$$\frac{1}{z_i}\ln\left[\frac{w_{ij}}{1-w_{ij}}(1+\sum_{s\neq j}^{r-1}\exp(z_i\beta_s))\right] < \beta_j \qquad \text{for } z_i > 0.$$

Appendix A5: Deriving Conditional Posterior Distributions for p-Covariate Parameters on r Nominal Categories used in (6.27).

$$w_{ij} < \frac{\exp(\beta_{1j} + z_{i2}\beta_{2j} + \ldots + z_{ip}\beta_{pj})}{1 + \sum_{s=1}^{k-1} \exp(\beta_{1s} + z_{i2}\beta_{2s} + \ldots + z_{ip}\beta_{ps})}$$

$$w_{ij} < \frac{\exp \sum_{l=1}^{p} z_{il}\beta_{lj}}{1 + \sum_{s=1}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls})} \qquad \begin{cases} l = 1,2,\ldots,p \\ i = 1,2,\ldots,N \end{cases} \qquad s = 1,2,\ldots,k$$

$$w_{ij}\,(1 + \sum_{s=1}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls})) < \exp(\sum_{l=1}^{p} z_{i}\beta_{lj})$$

$$w_{ij}\,(1 + \sum_{s\neq j}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls})) < \exp(\sum_{l=1}^{p} z_{il}\beta_{lj}) - w_{ij}\exp(\sum_{l=1}^{p} z_{il}\beta_{lj})$$

$$\frac{w_{ij}}{1 - w_{ij}}\,(1 + \sum_{s\neq j}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls})) < \exp(\sum_{l=1}^{p} z_{il}\beta_{lj})$$

$$\ln\left[\ \frac{w_{ij}}{1 - w_{ij}}\,(1 + \sum_{s\neq j}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls}))\ \right] < \beta_{1j} + z_{i2}\beta_{2j} + \ldots + z_{ip}\beta_{pj}$$

$$\frac{1}{x_{il'}}\left\{\ln\left[\frac{w_{ij}}{1 - w_{ij}}(1 + \sum_{s\neq j}^{k-1} \exp(\sum_{l=1}^{p} z_{il}\beta_{ls}))\right] - \sum_{l\neq l'}^{p} z_{il}\beta_{lj}\right\} < \beta_{l'j}$$

This must hold for all $x_{il'} > 0$, i = 1,2,...,n, so equation (6.29) follows. Similarly for equation (6.30).

APPENDIX B

Computer Programs

Logitbi.m: A Matlab program for simulating parameter values using latent variable with a Gibbs Sampler applied on binary nominal responses of Tri-continental AIDS data withAIDS status as response variable and two explanatory variables.

```
**********************************************************************************************************
clear all
F=Tridata;        %Data file
F(:,4);           %CD4 cells counts
F(:,9)            %AGE
r=3;
a=5.4429; b=-0.4712; c=-0.156;  b0=a; b1=b; b2=c;
B=[8.4429 –0.4712 –0.156]';
n=length(F);
one=ones(n,1);
x=[one log(F(:,4)) F(:,9)];
y=F(:,5);
c=size(x);
p=c(2);
Lo=[]; Up=[]; BB=[]; P2=[];
for k=1:80000
  k
    e=exp(-x*B)./(1+exp(-x*B));
    for i=1:n
   if y(i)==1
     z(i)=e(i)+rand(1,1)*(1-e(i));
   else
     z(i)=e(i)*rand(1,1);
        end
    end;
    for j=1:p
        l1 = find(y==1 & x(:,j)>0);
        l2 = find(y==0 & x(:,j)>0);
        l3 = find(y==1 & x(:,j)<=0);
        l4 = find(y==0 & x(:,j)<=0);
        v=-(1./x(:,j)').*(B'*x'-B(j)*x(:,j)'+log(z./(1-z)));
        v1 = v([l1;l4]);
        v2 = v([l2;l3]);
        lo(j)=max(v1);
        up(j)=min(v2);
        B(j)=lo(j)+rand(1,1)*(up(j)-lo(j));
    end;
    BB=[BB;B'];
    Lo=[Lo;lo];
    Up=[Up;up];
end;
W=1./(Up-Lo);
FB=[]; M=[]; C=[]; H=[];
for j=1:p
```

```
    m=min(Lo(:,j)); c=max(Up(:,j));
    h=(c-m)/100;
    SS=[];
    for q=m:h:c
        J=find(Lo(:,j)<=q & Up(:,j)>=q);
        S=sum((W(J,j)));
        SS=[SS;S];
    end;
    FB=[FB SS]; M=[M m]; C=[C c]; H=[H h];
end;
FB=FB/k;
for j=1:p
    d=FB(:,j);
    q=M(j):H(j):C(j);
    figure (j)
    plot(q,d)
    TITLE('Posterior Distribution of Beta(j)')
    Grid
end;
```

Logitpol.m: A Matlab program for simulating a polychotomous nominal response data and analysing using latent variable approach .An MCMC method of using a Gibbs Sampler is applied on the data which has p explanatory variables.

```
**********************************************************************************************************
format compact
n=100;
J=3;
r=2;
b01=0.4; b02=.6; b03=0.21; b04=.002; b05=2; b06=.9; b07=.4;
b11=0.3; b12=.2; b13=0.54; b14=.002; b15=2; b16=.9; b17=.4;
b21=0.1; b22=.35; b23=0.34; b24=.002; b25=2; b26=.9; b27=.4;
b31=0.8; b32=.25; b33=0.2; b34=.002; b35=2; b36=.9; b37=.4;
Bs=[b01 b02 b03 b04 b05 b06 b07;b11 b12 b13 b14 b15 b16 b17;b21 b22 b23 b24 b25 b26 b27];
x1=3*rand(n,1);
Y=[];
B=[]; Py=[ ];
P=[];
for j=1:J-1
  b=Bs(:,j);
  B=[B b];
end;
B1=B(1,:);
  for i=1:n
     e=exp(x1*B1);
     t=1+(sum(e')');
     py=e(i,:)./t(i);
     Ps=sum(py);
     PL=1-Ps;
     Py=[py PL];
     Fy=cumsum(Py);
     z=rand(1,1);
     y=min(find(Fy>=z));
     Y=[Y;y];
  end;
  x=[ones(n,1) 3*rand(n,r)];
  c=size(x);
  p=c(2);
  BB=[]; LL=[];
  for k=1:50
    k
    U=[];
    e=exp(x*B);
    t=1+sum(e,2);
    Es=sum(e,2);
   for j=1:J-1
     uu=[ ];
     for i=1:n
       if Y(i)==j;
         u=rand(1,1)*exp(x(i,:)*B(:,j))/t(i);
       else
         u=(1-exp(x(i,:)*B(:,j))/t(i))*rand(1,1)+exp(x(i,:)*B(:,j))/t(i);
       end
```

```
      uu=[uu;u];
      end
    U=[U uu];
  end;
  D=[Y x U];
  LIM=[ ];
for j=1:J-1
  I1=find(Y==j); J1=find(Y~=j);
  Lim=[ ];
    for i=1:p
    e=exp(x*B);
    t=1+sum(e,2);
    Es=sum(e,2);
    Es1=1+(Es-e(:,j));
    M1=x(I1,:)*B(:,j)-x(I1,i)*B(i,j);
    M2=x(J1,:)*B(:,j)-x(J1,i)*B(i,j);
    v1=(1./x(I1,i)).*(log(Es1(I1).*(U(I1,j)./(1-U(I1,j))))-M1);
    v2=((1./x(J1,i)).*(log(Es1(J1).*(U(J1,j)./(1-U(J1,j))))-M2));
    Lo=max(v1);
    Up=min(v2);
    lim=[Lo Up];
    B(i,j)=(Up-Lo)*rand(1,1)+Lo;
    Lim=[Lim;lim];
    end
  LIM=[LIM;Lim];
 end
BB=[BB;B];
LL=[LL;LIM];
end;
W=1./LL(:,2)-LL(:,1);
for i=1:p*(J-1)
  I=i:p*(J-1):length(LL);
  a=min(LL(I,1)); c=max(LL(I,2));
  T=[];
  h=(c-a)/100;
  N=LL(I,:); NN=W(I);
  for q=a:h:c
    K=find(N(:,1)<=q & N(:,2)>=q);
    t=sum(NN(K));
    T=[T;t];
  end
  F(:,i)=T/k;
  q1=a:h:c;
  figure(i)
  plot(q1,F(:,i))
  grid
end
```

Logitord.m: A Matlab program for simulating a polychotomous ordinal response data and and analysing it using a latent variable approach. A Gibbs Sampler is applied on the data with p explanatory variables.

```
*****************************************************************************************************
clear
n=20;
x=2*rand(n,2);
B=[-2;1];
A=[1.5;2.5;30]; J=3;
for j=1:J
  N(:,j)=exp(A(j)+x*B)./(1+exp(A(j)+x*B));
end
z=rand(n,1);
for i=1:n
  Y(i)=min(find(z(i)<=N(i,:)));
end
I1=find(Y==1);
I2=find(Y==2);
I3=find(Y==3);
m=100; AA=[ ]; BB=[ ]; UU=[ ];
for k=1:m
  Gup=zeros(n,1); Glo=zeros(n,1);
  Gup(I1)=A(1); Gup(I2)=A(2); Gup(I3)=A(3);
  Glo(I2)=A(1); Glo(I3)=A(2);
  Lo=exp(Glo+x*B)./(1+exp(Glo+x*B));
  Up=exp(Gup+x*B)./(1+exp(Gup+x*B));
  U=Lo+rand(n,1).*(Up-Lo);
  UU=[UU U];
  log1=max([log(U(I1)./(1-U(I1)))+x(I1,:)*B;0]);
  upg1=min([log(U(I2)./(1-U(I2)))+x(I2,:)*B;A(2)]);
  log2=max([log(U(I2)./(1-U(I2)))+x(I2,:)*B;A(1)]);
  upg2=min([log(U(I3)./(1-U(I3)))+x(I3,:)*B;A(3)]);
  Lg=[log1 upg1;log2 upg2];
  A=[Lg(:,1)+rand(J-1,1).*(Lg(:,2)-Lg(:,1));30];
  lob11=max((log(U(I1)./(1-U(I1)))-A(1)-x(I1,2)*B(2))./x(I1,1));
  upb11=min((log(U(I2)./(1-U(I2)))-A(1)-x(I2,2)*B(2))./x(I2,1));
  lob12=max((log(U(I2)./(1-U(I2)))-A(2)-x(I2,2)*B(2))./x(I2,1));
  upb12=min((log(U(I3)./(1-U(I3)))-A(2)-x(I3,2)*B(2))./x(I3,1));
  lob1=max([lob11 lob12]); upb1=min([upb11 upb12]);
  b1=lob1+rand(1,1)*(upb1-lob1);
  lob21=max((log(U(I1)./(1-U(I1)))-A(1)-x(I1,2)*B(1))./x(I1,2));
  upb21=min((log(U(I2)./(1-U(I2)))-A(1)-x(I2,2)*B(1))./x(I2,2));
  lob22=max((log(U(I2)./(1-U(I2)))-A(2)-x(I2,2)*B(1))./x(I2,2));
  upb22=min((log(U(I3)./(1-U(I3)))-A(2)-x(I3,2)*B(1))./x(I3,2));
  lob2=max([lob21 lob22]); upb2=min([upb21 upb22]);
  b2=lob2+rand(1,1)*(upb2-lob2);
  B=[b1;b2];
  AA=[AA A]; BB=[BB B];
end
```

Simfm.m: A Matlab program used to simulate dependent bivariate data from a Farlie-Morgenstern distribution, with Exponential marginal distributions for each failure time.

```
*****************************************************************************************************
clear all
n=100;
r=1;
B=.25;
Int=0.5;
M=2;
κ=1;
u1=rand(n,1);
Z=[randn(n,1)-2 randn(n,1)];
Lam=exp(B*Z);
x1=-(1./Lam(:,1)).*log(1-u1);
u2=rand(n,1);
u3=rand(n,1);
d=κ*((2*Lam(:,1).*exp(-x1.*Lam(:,1)))-1);
a=d.*Lam(:,2);
b=1-a;
c=u2-1;
Rt=(b.^2-(4*a.*c)).^.5;
y1=(-b-Rt)./(2*a);
y2=(-b+Rt)./(2*a);
x2=-log(y2);
X=[x1;x2];
figure (1)
hist(x2,10)
xlabel('x2')
figure (2)
scatter(X(1:n),X(n+1:n*M))
xlabel('x1')
ylabel('x2')
grid
```

Indmlesim: A Matlab program for estimating MLE under independence assumption model from simulated data

```
**********************************************************************************************************************
clear all
simfm.m                    %simulated data
n=100;
r=3;
B1=.25;
g1=-.6931; g2=-.6931; g3=0;
g4=-.6931; g5=-.6931; g6=0;
delta=[g1 g2 g3;g4 g5 g6];
beta=[B1 B2 g1 g2 g3 g4 g5 g6];
beta=beta';
Int=[0.5 1 2];
M=2;
Z1=[Z(:,1);Z(:,2)];
for i=1:n*M
  if X(i)>=Int(r)
    alp(i,r+1)=1;
  elseif X(i)>=Int(r-1) & X(i)<Int(r)
    alp(i,r)=1;
  elseif X(i)>=Int(r-2) & X(i)<Int(r-1)
    alp(i,r-1)=1;
  elseif X(i)>=0 & X(i)<Int(r-2)
    alp(i,r-2)=1;
  end
end
% iteration should start here
BB1=[]; BB2=[]; Likel=[-337]; LLike=[]; diff=0.1; count=[0];
k=1
while diff>=.000000001
count=[count;k];
h=exp(delta+B*Z(i:,));
co=exp(-h);
rem=1-co;
data=[X alp];
data1=[rem h X];
for i=1:n
  if alp(i,4)==1
    li1(i)=-sum(h(i,1:r));
    fd1(i,:)=[-Z1(i)*sum(h(i,1:r)) 0 -h(i,1) -h(i,2) -h(i,3) 0 0 0];
    sdd1(i,:)=[-Z1(i)^2*sum(h(i,1:r)) 0 -h(i,1) -h(i,2) -h(i,3) 0 0 0];
    sdc1(i,:)=[0 0 -Z1(i)*h(i,1) -Z1(i)*h(i,2) -Z1(i)*h(i,3) 0 0 0];
  elseif alp(i,3)==1
    li1(i)=log(rem(i,r))-sum(h(i,1:r-1));
    fd1(i,:)=[Z1(i)*((co(i,r)*h(i,r)/rem(i,r))-sum(h(i,1:r-1))) 0 -h(i,1)
                -h(i,2) (co(i,r)*h(i,r))/rem(i,r) 0 0 0];
    sdd1(i,:)=[Z1(i)^2*(((co(i,r)*h(i,r)*(1-co(i,r)-h(i,r)))/rem(i,r)^2)-sum(h(i,1:r-1))) 0 -h(i,1) -h(i,2)
co(i,r)*h(i,r)*((1-co(i,r)
    -h(i,r)))/(rem(i,r))^2 0 0 0];
    sdc1(i,:)=[0 0 -Z1(i)*h(i,1) -Z1(i)*h(i,2) Z1(i)*(co(i,r)*h(i,r)*(1-co(i,r)-h(i,r)))/(rem(i,r))^2 0 0 0];
  elseif alp(i,2)==1
    li1(i)=log(rem(i,2))-h(i,1);
    fd1(i,:)=[Z1(i)*((co(i,r-1)*h(i,r-1)/rem(i,r-1))-h(i,1)) 0 -h(i,1) (co(i,r-1)*h(i,r-1))/rem(i,r-1) 0 0 0 0];
    sdd1(i,:)=[Z1(i)^2*(((co(i,r-1)*h(i,r-1)*(1-co(i,r-1)-h(i,r-1)))/rem(i,r-1)^2)-h(i,1)) 0 -h(i,1) co(i,r-1)*h(i,r-1)*(1-
co(i,r-1)
    -h(i,r-1))/(rem(i,r-1))^2 0 0 0 0];
```

```
    sdc1(i,:)=[0 0 -Z1(i)*h(i,1) Z1(i)*(co(i,r-1)*h(i,r-1)*(1-co(i,r-1)-h(i,r-1)))/(rem(i,r-1))^2 0 0 0 0];
  elseif alp(i,1)==1
    li1(i)=log(rem(i,1));
    fd1(i,:)=[Z1(i)*(co(i,1)*h(i,1)/rem(i,1)) 0 (co(i,1)*h(i,1))/rem(i,1) 0 0 0 0 0];
    sdd1(i,:)=[Z1(i)^2*((co(i,1)*h(i,1)*(1-co(i,1)-h(i,1)))/rem(i,1)^2) 0 (co(i,1)*h(i,1)*(1-co(i,1)-h(i,1)))/(rem(i,1))^2
0 0 0 0 0];
    sdc1(i,:)=[0 0 Z1(i)*(co(i,1)*h(i,1)*(1-co(i,1)-h(i,1)))/(rem(i,1))^2 0 0 0 0 0];
  end
end
Li=li1';
for i=n+1:n*M
  if alp(i,4)==1
    li2(i)=-sum(h(i,1:r));
    fd2(i,:)=[0 -Z1(i)*sum(h(i,1:r)) 0 0 0 -h(i,1) -h(i,2) -h(i,3)];
    sdd2(i,:)=[0 -Z1(i)^2*sum(h(i,1:r)) 0 0 0 -h(i,1) -h(i,2) -h(i,3)];
    sdc2(i,:)=[0 0 0 0 0 -Z1(i)*h(i,1) -Z1(i)*h(i,2) -Z1(i)*h(i,3)];
elseif alp(i,3)==1
    li2(i)=log(rem(i,r))-sum(h(i,1:r-1));
    fd2(i,:)=[0  Z1(i)*((co(i,r)*h(i,r)/rem(i,r))-sum(h(i,1:r-1))) 0 0
        0  -h(i,1) -h(i,2) (co(i,r)*h(i,r))/rem(i,r)];
    sdd2(i,:)=[0 Z1(i)^2*(((co(i,r)*h(i,r)*(1-co(i,r)
        -h(i,r)))/rem(i,r)^2)-sum(h(i,1:r-1))) 0 0 0 -h(i,1)
        -h(i,2) (co(i,r)*h(i,r)*(1-co(i,r)- h(i,r)))/(rem(i,r))^2];
    sdc2(i,:)=[0 0 0 0 0 -Z1(i)*h(i,1) -Z1(i)*h(i,2)
          Z1(i)*(co(i,r)*h(i,r)*(1-co(i,r)-h(i,r)))/(rem(i,r))^2];
 elseif alp(i,2)==1
    li2(i)=log(rem(i,2))-h(i,1);
    fd2(i,:)=[0 Z1(i)*((co(i,r-1)*h(i,r-1)/rem(i,r-1))-(h(i,1))) 0 0 0
          -h(i,1) (co(i,r-1)*h(i,r-1))/rem(i,r-1) 0];
    sdd2(i,:)=[0 Z1(i)^2*(((co(i,r-1)*h(i,r-1)*(1-co(i,r-1)
        -h(i,r-1)))/rem(i,r-1)^2)-h(i,1)) 0 0 0 -h(i,1)
        (co(i,r-1)*h(i,r-1)*(1-co(i,r-1)-h(i,r-1)))/(rem(i,r-1))^2 0];
    sdc2(i,:)=[0 0 0 0 0 -Z1(i)*h(i,1) Z1(i)*(co(i,r-1)*h(i,r-1)
        *(1-co(i,r-1)-h(i,r-1)))/(rem(i,r-1))^2 0];
 elseif alp(i,1)==1
    li2(i)=log(rem(i,1));
    fd2(i,:)=[0 Z1(i)*(co(i,1)*h(i,1)/rem(i,1)) 0 0 0
        (co(i,1)*h(i,1))/rem(i,1) 0 0];
    sdd2(i,:)=[0 Z1(i)^2*((co(i,1)*h(i,1)*(1-co(i,1)-
        h(i,1)))/rem(i,1)^2) 0 0 0
        (co(i,1)*h(i,1)*(1-co(i,1)-h(i,1)))/(rem(i,1))^2 0 0];
    sdc2(i,:)=[0 0 0 0 0 Z1(i)*(co(i,1)*h(i,1)*(1-co(i,1)-
        h(i,1)))/(rem(i,1))^2 0 0];
  end
end
li2=li2(n+1:n*M);
fd2=fd2(n+1:n*M,:);
sdd2=sdd2(n+1:n*M,:);
sdc2=sdc2(n+1:n*M,:);
Li=[Li;li2'];
fd=[fd1;fd2];
q=(sum(fd))';
sdd=[sdd1;sdd2];
sdc=[sdc1;sdc2];
sufd=sum(fd);
sudd=sum(sdd);
sudc1=sum(sdc1);
```

```
sudc2=sum(sdc2);
sudc=sum(sdc);
isudc=sudc';
diadd=diag(sudd);
for j=1:r*M+2
  for i=1:r*M+2
    if i==j
      H1(i,j)=diadd(i,j);
    elseif i==1 & j>1
      H1(i,j)=sudc1(i,j);
    elseif j==1 & i>1
      H1(i,j)=sudc1(j,i);
    elseif i==2 & j>M
      H1(i,j)=sudc2(1,j);
    elseif j==2 & i>M
      H1(i,j)=sudc2(1,i);
    else
      H1(i,j)=0;
    end
  end
end
H=inv(H1);
likel=sum(Li);
beta=beta-H*q/2;
delta=[beta(3) beta(4) beta(5); beta(6) beta(7) beta(8)];
B1=beta(1);
B2=beta(2);
BB1=[BB1;B1];
BB2=[BB2;B2];
LLike=[LLike;likel];
Likel=[Likel;likel]
diff=Likel(k+1)-Likel(k)
k=k+1
end
```

Indphmle: A Matlab program for estimating MLE under independence assumption model using a bivariate data with overlapiing intervals.

```
**********************************************************************************************
clear all
% Data file
% Specify starting values for all parameters.
beta=[B rdelta cdelta]';
v=length(B);
vr=v+rnr;
vrc=v+rnr+cdr;
% iteration should start here
BB=[]; Likel=[-2000]; Q=[];  LLike=[]; diff=0.1;
BETA=[];
  k=1;  %abs(diff)>=.0001 &
  while  diff>=0.1
    k
    LI=[];  SBB=[]; SBR=[]; SBCC=[];SBCR=[];
    for q=1:rnr
              for j=1:cdr
        for i=1:n
          rnh(i,q)=exp(rdelta(q)+Z(:,i)*B'
          cdh(i,j)=exp(cdelta(j)+Z(:,i)*B'
        end
      end
    end
    cdco=exp(-cdh);
    cdrem=1-cdco;
    rnco=exp(-rnh);
    rnrem=1-rnco;
    for i=1:n
      cl=cdalp(i,1);
      cu=cdalp(i,2);
      if cl==0
        cdli(i)=-sum(cdh(i,1:cu));
      elseif cl~=0 & cl==cu
        cdli(i)=log(cdrem(i,cu))-sum(cdh(i,1:cu-1));
      elseif cl~=0 & cl==cu-1
        cdli(i)=log(cdrem(i,cu)*exp(-sum(cdh(i,1:cu-1)))+cdrem(i,cu-1)*exp(-sum(cdh(i,1:cl-2))));
      elseif cl~=0 & cl==cu-2
        cdli(i)=log(cdrem(i,cu)*exp(-sum(cdh(i,1:cu-1)))+cdrem(i,cu-1)*exp(-sum(cdh(i,1:cu-
2)))+cdrem(i,cu-2)*exp(-sum(cdh(i,1:cl-3))));
      end
    end
    for i=1:n
      rl=rnalp(i,1);
      ru=rnalp(i,2);
      if rl==0
        rnli(i)=-sum(rnh(i,1:ru));
      elseif rl~=0 & rl==ru
        rnli(i)=log(rnrem(i,ru))-sum(rnh(i,1:ru-1));
      elseif rl~=0 & rl==ru-1
```

```
      rnli(i)=log(rnrem(i,ru)*exp(-sum(rnh(i,1:ru-1)))+rnrem(i,ru-1)*exp(-sum(rnh(i,1:ru-2))));
    elseif rl~=0 & rl==ru-2
      rnli(i)=log(rnrem(i,ru)*exp(-sum(rnh(i,1:ru-1)))+rnrem(i,ru-1)*exp(-sum(rnh(i,1:ru-
2)))+rnrem(i,ru-2)*exp(-sum(rnh(i,1:ru-3))));
    end
  end
  li=cdli'+rnli';
  for t=1:v
    for i=1:n
      cl=cdalp(i,1);
      cu=cdalp(i,2);
      if cl==0
        fc(i,1:cu)=-cdh(i,1:cu);
        fb(i,t)=-Z(i,t)*sum(cdh(i,1:cu));
        sc(i,1:cu)=-cdh(i,1:cu);
        sb(i,1:v)=-Z(i,t)*Z(i,1:v)*sum(cdh(i,1:cu));
        sbc(i,1:cu)=-Z(i,t)*cdh(i,1:cu);
      elseif cl~=0 & cl==cu
        fc(i,1:cu-1)=-cdh(i,1:cu-1);
        fc(i,cu)=cdco(i,cu)*cdh(i,cu)/cdrem(i,cu);
        fbc(i,t)=Z(i,t)*((cdco(i,cu)*cdh(i,cu)/cdrem(i,cu))-sum(cdh(i,1:cu-1)));
        sc(i,1:cu-1)=-cdh(i,1:cu-1);
        sc(i,cu)=cdco(i,cu)*cdh(i,cu)*(1-cdh(i,cu)-cdco(i,cu))/cdrem(i,cu)^2;
        sb(i,1:v)=Z(i,t)*Z(i,1:v)*((cdco(i,cu)*cdh(i,cu)*(1-cdh(i,cu)-cdco(i,cu))/cdrem(i,cu)^2)-
sum(cdh(i,1:cu-1)));
        sbc(i,1:cu-1)=-Z(i,t)*cdh(i,1:cu-1);
        sbc(i,cu)=Z(i,t)*cdco(i,cu)*cdh(i,cu)*(1-cdh(i,cu)-cdco(i,cu))/cdrem(i,cu)^2;
      elseif cl~=0 & cl==cu-1
        fc(i,1:cu-2)=-cdh(i,1:cu-2);
        fc(i,cu-1)=cdco(i,cu-1)*cdco(i,cu)*cdh(i,cu-1)/(1-cdco(i,cu)*cdco(i,cu-1));
        fc(i,cu)=cdco(i,cu)*cdco(i,cu-1)*cdh(i,cu)/(1-cdco(i,cu)*cdco(i,cu-1));
        fb(i,t)=(Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*(cdh(i,cu)+cdh(i,cu-1))/(1-cdco(i,cu)*cdco(i,cu-1)))-
Z(i,t)*sum(cdh(i,1:cu-2));
        sc(i,1:cu-2)=-cdh(i,1:cu-2);
        sc(i,cu-1)=cdco(i,cu)*cdco(i,cu-1)*cdh(i,cu-1)*(1-cdco(i,cu)*cdco(i,cu-1)-cdh(i,cu-1))/(1-
cdco(i,cu)*cdco(i,cu-1))^2;
        sc(i,cu)=cdco(i,cu)*cdco(i,cu-1)*cdh(i,cu)*(1-cdco(i,cu)*cdco(i,cu-1)-cdh(i,cu))/(1-
cdco(i,cu)*cdco(i,cu-1))^2;
        sb1=cdco(i,cu)*cdco(i,cu-1)*(cdh(i,cu)+cdh(i,cu-1))*(1-cdco(i,cu)*cdco(i,cu-1)-cdh(i,cu)-
cdh(i,cu-1))/(1-cdco(i,cu)*cdco(i,cu-1))^2;
        sb(i,1:v)=Z(i,t)*Z(i,1:v)*(sb1-sum(cdh(i,1:cu-2)));
        sbc(i,1:cu-2)=-Z(i,t)*cdh(i,1:cu-2);
        sbc(i,cu-1)=Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdh(i,cu-1)*(1-cdco(i,cu)*cdco(i,cu-1)-cdh(i,cu-
1)+cdh(i,cu))/(1-cdco(i,cu)*cdco(i,cu-1))^2;
        sbc(i,cu)=Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdh(i,cu)*(1-cdco(i,cu)*cdco(i,cu-1)-
cdh(i,cu)+cdh(i,cu-1))/(1-cdco(i,cu)*cdco(i,cu-1))^2;
      elseif cl~=0 & cl==cu-2
        fc(i,1:cu-3)=-cdh(i,1:cu-3);
        fc(i,cu-2)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-2)/(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2));
        fc(i,cu-1)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-1)/(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2));
```

```
        fc(i,cu)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu)/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-
2));
        fb(i,t)=(Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*(cdh(i,cu)+cdh(i,cu-1)+cdh(i,cu-2))/(1-
cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)))-Z(i,t)*sum(cdh(i,1:cu-3));
        sc(i,1:cu-3)=-cdh(i,1:1:cu-3);
        sc(i,cu-2)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-2)*(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2)-cdh(i,cu-2))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
        sc(i,cu-1)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-1)*(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2)-cdh(i,cu-1))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
        sc(i,cu)=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu)*(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-
2)-cdh(i,cu))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
        sb1=cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)-cdh(i,cu)-
cdh(i,cu-1)-cdh(i,cu-2))*(cdh(i,cu)+cdh(i,cu-1)+cdh(i,cu-2))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-
2))^2;
        sb(i,1:v)=Z(i,t)*Z(i,1:v)*(sb1-sum(cdh(i,1:cu-3)));
        sbc(i,1:cu-3)=-Z(i,t)*cdh(i,1:cu-3);
        sbc(i,cu-2)=Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-2)*(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2)-cdh(i,cu-2)+cdh(i,cu)+cdh(i,cu-1))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
        sbc(i,cu-1)=Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu-1)*(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2)-cdh(i,cu-1)+cdh(i,cu)+cdh(i,cu-2))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
        sbc(i,cu)=Z(i,t)*cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2)*cdh(i,cu)*(1-cdco(i,cu)*cdco(i,cu-
1)*cdco(i,cu-2)-cdh(i,cu)+cdh(i,cu-1)+cdh(i,cu-2))/(1-cdco(i,cu)*cdco(i,cu-1)*cdco(i,cu-2))^2;
      end
    end
    ssbc=sum(sbc);
    ssb=sum(sb);
    SBB=[SBB;ssb];
    SBCC=[SBCC;ssbc];
    for i=1:n
      rl=rnalp(i,1);
      ru=rnalp(i,2);
      if rl==0
        fr(i,1:ru)=-rnh(i,1:ru);
        fbr(i,t)=-Z(i,t)*sum(rnh(i,1:ru));
        scr(i,1:ru)=-rnh(i,1:ru);
        sbr(i,1:v)=-Z(i,t)*Z(i,1:v)*sum(rnh(i,1:ru));
        sbcr(i,1:ru)=-Z(i,t)*rnh(i,1:ru);
      elseif rl~=0 & rl==ru
        fr(i,1:ru-1)=-rnh(i,1:ru-1);
        fr(i,ru)=rnco(i,ru)*rnh(i,ru)/rnrem(i,ru);
        fbr(i,t)=Z(i,t)*((rnco(i,ru)*rnh(i,ru)/rnrem(i,ru))-sum(rnh(i,1:ru-1)));
        scr(i,1:ru-1)=-rnh(i,1:ru-1);
        scr(i,ru)=rnco(i,ru)*rnh(i,ru)*(1-rnh(i,ru)-rnco(i,ru))/rnrem(i,ru)^2;
        sbr(i,1:v)=Z(i,t)*Z(i,1:v)*((rnco(i,ru)*rnh(i,ru)*(1-rnh(i,ru)-rnco(i,ru))/rnrem(i,ru)^2)-
 sum(rnh(i,1:ru-1)));
        sbcr(i,1:ru-1)=-Z(i,t)*rnh(i,1:ru-1);
        sbcr(i,ru)=Z(i,t)*rnco(i,ru)*rnh(i,ru)*(1-rnh(i,ru)-rnco(i,ru))/rnrem(i,ru)^2;
      elseif rl~=0 & rl==ru-1
        fr(i,1:ru-2)=-rnh(i,1:ru-2);
        fr(i,ru-1)=rnco(i,ru-1)*rnco(i,ru)*rnh(i,ru-1)/(1-rnco(i,ru)*rnco(i,ru-1));
        fr(i,ru)=rnco(i,ru-1)*rnco(i,ru)*rnh(i,ru)/(1-rnco(i,ru)*rnco(i,ru-1));
```

```
        fbr(i,t)=(Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*(rnh(i,ru)+rnh(i,ru-1))/(1-rnco(i,ru)*rnco(i,ru-1)))-
Z(i,t)*sum(rnh(i,1:ru-2));
        scr(i,1:ru-2)=-rnh(i,1:ru-2);
        scr(i,ru-1)=rnco(i,ru)*rnco(i,ru-1)*rnh(i,ru-1)*(1-rnco(i,ru)*rnco(i,ru-1)-rnh(i,ru-1))/(1-
rnco(i,ru)*rnco(i,ru-1))^2;
        scr(i,ru)=rnco(i,ru)*rnco(i,ru-1)*rnh(i,ru)*(1-rnco(i,ru)*rnco(i,ru-1)-rnh(i,ru))/(1-
rnco(i,ru)*rnco(i,ru-1))^2;
        sb1=rnco(i,ru)*rnco(i,ru-1)*(rnh(i,ru)+rnh(i,ru-1))*(1-rnco(i,ru)*rnco(i,ru-1)-rnh(i,ru)-
rnh(i,ru-1))/(1-rnco(i,ru)*rnco(i,ru-1))^2;
        sbr(i,1:v)=Z(i,t)*Z(i,1:v)*(sb1-sum(rnh(i,1:ru-2)));
        sbcr(i,1:ru-2)=-Z(i,t)*rnh(i,1:ru-2);
        sbcr(i,ru-1)=Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnh(i,ru-1)*(1-rnco(i,ru)*rnco(i,ru-1)-rnh(i,ru-
1)+rnh(i,ru))/(1-rnco(i,ru)*rnco(i,ru-1))^2;
        sbcr(i,ru)=Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnh(i,ru)*(1-rnco(i,ru)*rnco(i,ru-1)-
rnh(i,ru)+rnh(i,ru-1))/(1-rnco(i,ru)*rnco(i,ru-1))^2;
      elseif rl~=0 & rl==ru-2
        fr(i,1:ru-3)=-rnh(i,1:ru-3);
        fr(i,ru-2)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-2)/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-
2));
        fr(i,ru-1)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-1)/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-
2));
        fr(i,ru)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru)/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2));
        fbr(i,t)=(Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*(rnh(i,ru)+rnh(i,ru-1)+rnh(i,ru-2))/(1-
rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)))-Z(i,t)*sum(rnh(i,1:ru-3));
        scr(i,1:ru-3)=-rnh(i,1:1:ru-3);
        scr(i,ru-2)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-2)*(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2)-rnh(i,ru-2))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
        scr(i,ru-1)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-1)*(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2)-rnh(i,ru-1))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
        scr(i,ru)=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru)*(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)-
rnh(i,ru))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
        sb1=rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*(rnh(i,ru)+rnh(i,ru-1)+rnh(i,ru-2))*(1-
rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)-rnh(i,ru)-rnh(i,ru-1)-rnh(i,ru-2))/(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2))^2;
        sbr(i,1:v)=Z(i,t)*Z(i,1:v)*(sb1-sum(rnh(i,1:ru-3)));
        sbcr(i,1:ru-3)=-Z(i,t)*rnh(i,1:ru-3);
        sbcr(i,ru-2)=Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-2)*(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2)+rnh(i,ru)+rnh(i,ru-1)-rnh(i,ru-2))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
        sbcr(i,ru-1)=Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru-1)*(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2)+rnh(i,ru)-rnh(i,ru-1)+rnh(i,ru-1))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
        sbcr(i,ru)=Z(i,t)*rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2)*rnh(i,ru)*(1-rnco(i,ru)*rnco(i,ru-
1)*rnco(i,ru-2)-rnh(i,ru)+rnh(i,ru-1)+rnh(i,ru-2))/(1-rnco(i,ru)*rnco(i,ru-1)*rnco(i,ru-2))^2;
      end
    end
    ssbcr=sum(sbcr);
    ssbr=sum(sbr);
    SBR=[SBB;ssbr];
    SBCR=[SBCR;ssbcr];
  end
  sfc=sum(fc);
  sfr=sum(fr);
  sfb=sum(fb);
```

```
sfbr=sum(fbr);
SC=[sum(scr) sum(sc)];
SBC=[SBCR SBCC];
GB1=[]; GB2=[]; DIA1=[]; DIA2=[]; CR1=[]; Ub=[];
for i=n
  for l=1:v
    ub(i,l)=fbr(i,l)*(fbr(i,l)+fb(i,l))+fb(i,l)*(fbr(i,l)+fb(i,l));
    crb1(i,1:rnr)=(fr(i,1:rnr)*fbr(i,l)+fr(i,1:rnr)*fb(i,l));
    crb2(i,1:cdr)=(fc(i,1:cdr)*fbr(i,l)+fc(i,1:cdr)*fb(i,l));
    GB1=[GB1 crb1]; GB2=[GB2 crb2];
  end
end
UB=sum(ub);
SGB1=sum(GB1); SGB2=sum(GB2);
DIA1=zeros(rnr,rnr);
DIA2=zeros(cdr,cdr);
CR1=zeros(rnr,cdr);
for i=1:n
  diA1=(fr(i,1:rnr)'*fr(i,1:rnr));
  diA2=(fc(i,1:cdr)'*fc(i,1:cdr));
  cr1=(fr(i,1:rnr)'*fc(i,1:cdr));
  DIA1=DIA1+diA1; DIA2=DIA2+diA2;
  CR1=CR1+cr1;
end
CRR=[];
CRC=[];
for j=1:v
  crr=SGB1((j-1)*rnr+1:j*rnr);
  crc=SGB2((j-1)*cdr+1:j*cdr);
  CRR=[CRR;crr];
  CRC=[CRC;crc];
end
D=[diag(UB) CRR CRC;CRR' DIA1 CR1;CRC' CR1' DIA2];
likel=sum(li)
q=[sfbr+sfb sfr sfc]';
Q=[Q q];
H1=[SBB SBC;SBC' diag(SC)];
H=inv(H1);
NH=H*D*H;
beta=beta+NH*q/10;
BETA=[BETA beta];
B=beta(1:v)';
B1=beta(1)';
    rdelta=beta(v+1:vr)';
    cdelta=beta(vr+1:vrc)';
    BB=[BB;B1];
    LLike=[LLike;likel];
    Likel=[Likel;likel];
    diff=Likel(k+1)-Likel(k);
     k=k+1;
end
plot(BB,LLike)
```

Datasim.m: A simulation of dependant visiting times, one covariate for failure times and covariate for visiting probabilities which vary over intervals

```
**********************************************************************************************************
clear
N=50;
J=12; t=0.3; z=randn(1,N); bet=0.3; x=round(rand(N,1));
X=0:J-1;
muj=1.5-0.25*X; nu=1; thet=-0.5;
g=exp(-t*X)-exp(-t*(X+1));
g=[g 1-sum(g)]';
G=[0;cumsum(g)];
Gi=1-G;
r=rand(N,1);
V=[]; A=[]; Gamjk=[];
for i=1:N
  G=1-Gi.^exp(bet*z(i));
  I(i)=min(find(G>=r(i)))-1;
  K=I(i);
  if I(i)==J+1;
    I(i)=J;
  end
  gamjk=[ones(1,I(i)) zeros(1,J-I(i))];
  vv=[];
  for j=1:J
    pijk=exp(muj(j)+nu*gamjk(j)+thet*x(i))/(1+exp(muj(j)+nu*gamjk(j)+thet*x(i)));
    r2=rand(1,1);
    if r2>=pijk
      v=0;
    else v=1;
    end
    vv=[vv;v];
  end
  vv=[vv;0];
  V=[V;vv'];
  a=zeros(1,J+1);
  if K<=J
    c=min(find(vv(I(i):J)==1))+I(i)-1;
    if isempty(c)
      c=J+1;
    end
    d=max(find(vv(1:I(i)-1)==1))+1;
    if isempty(d)
      d=1;
    end
  a(d:c)=ones(1,length(a(d:c)));
  else
  f=max(find(vv==1));
  a=[zeros(1,f) ones(1,J-f+1)];
end
A=[A;a];
Gamjk=[Gamjk;gamjk];
End
```

Vispos.m: A Gibbs samplere for Dependent Visits, data form DATASIM.M with one covariate for failure times and covariate for visiting probabilities that vary over intervals.

```
**********************************************************************************************************
clear
datasim.m
%STARTING VALUES
%---------------
muj=0.125*ones(J,1); nu=0; thet=0; bett=0;
for j=1:J+1
  P(j)=(1-sum(g(1:j)))/(1-sum(g(1:j-1)));
end
PP=[]; MUj=[]; NU=[]; THET=[]; tel1=0; tel2=0; tel3=0; BET=[];
for k=1:2000
  k
%SIMULATING P
%------------
L=[];
for i=1:N
  pijk=[exp(muj+nu*Gamjk(i,:)'+thet*x(i))./(1+exp(muj+nu*Gamjk(i,:)'+thet*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(P(j).^exp(bett*z(i))))*prod(P(1:j-1).^exp(bett*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:J
  pr=(P(j)^(0.5*(J-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pr);
end
lnpost=LL+lnpr;
Z1=min(P,0.9999); Z2=max(Z1(1:12),0.0001);
P=[Z2 0];
r1=1; a1=0;
while r1>=a1;
  r=0.4*randn(1,J)+log(P(1:J)./(1-P(1:J)));
  Po=[1./(1+exp(-r)) 0];
  L=[]; tel1=tel1+1;
  for i=1:N
  pijk=[exp(muj+nu*Gamjk(i,:)'+thet*x(i))./(1+exp(muj+nu*Gamjk(i,:)'+thet*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(Po(j).^exp(bett*z(i))))*prod(Po(1:j-1).^exp(bett*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
```

```
LL=sum(log(L));
lnpr=0;
for j=1:J
  pro=(Po(j)^(0.5*(J-j-1)))*((1-Po(j))^(-0.5));
  lnpr=lnpr+log(pro);
end
lnposto=LL+lnpr;
a1=exp(lnposto-lnpost);
r1=rand(1,1);
end
P=Po;
PP=[PP;P];
%SIMULATING MUj, NU AND THET
%---------------------------
L=[];
for i=1:N
  pijk=[exp(muj+nu*Gamjk(i,:)'+thet*x(i))./(1+exp(muj+nu*Gamjk(i,:)'+thet*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(P(j).^exp(bett*z(i))))*prod(P(1:j-1).^exp(bett*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:J
  pr=(P(j)^(0.5*(J-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pr);
end
lnpost=LL+lnpr;
r1=2; a2=0;
while r2>=a2;
  mujo=muj+randn(J,1)*0.15;
  nuo=nu+randn(1,1)*0.1;
  theto=thet+randn(1,1)*0.1;
  L=[]; tel2=tel2+1;
  for i=1:N
  pijk=[exp(mujo+nuo*Gamjk(i,:)'+theto*x(i))./(1+exp(mujo+nuo*Gamjk(i,:)'+theto*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(P(j).^exp(bett*z(i))))*prod(P(1:j-1).^exp(bett*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:J
  pro=(P(j)^(0.5*(J-j-1)))*((1-P(j))^(-0.5));
```

```
  lnpr=lnpr+log(pro);
end
lnposto=LL+lnpr;
a2=exp(lnposto-lnpost);
r2=rand(1,1);
end
muj=mujo; nu=nuo; thet=theto;
MUj=[MUj;muj']; NU=[NU;nu]; THET=[THET;thet];
%SIMULATING BETA
%---------------
L=[];
for i=1:N
  pijk=[exp(muj+nu*Gamjk(i,:)'+thet*x(i))./(1+exp(muj+nu*Gamjk(i,:)'+thet*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(P(j).^exp(bett*z(i))))*prod(P(1:j-1).^exp(bett*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:J
  pr=(P(j)^(0.5*(J-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pr);
end
lnpost=LL+lnpr;
r3=1; a3=0;
while r3>=a3;
  beto=bett+randn(1,1)*0.3;
  L=[]; tel3=tel3+1;
  for i=1:N
  pijk=[exp(muj+nu*Gamjk(i,:)'+thet*x(i))./(1+exp(muj+nu*Gamjk(i,:)'+thet*x(i)));0];
  B=prod((pijk'.^V(i,:)).*((1-pijk)'.^(1-V(i,:))));
  ll=0;
  for j=1:J+1
    l=A(i,j)*(1-(P(j).^exp(beto*z(i))))*prod(P(1:j-1).^exp(beto*z(i)))*B;
    ll=ll+l;
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:J
  pro=(P(j)^(0.5*(J-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pro);
end
lnposto=LL+lnpr;
a3=exp(lnposto-lnpost);
r3=rand(1,1);
end
```

```
bett=beto;
BET=[BET;bett];
%pause
end
save c:\matlabr11\deptsim1.mat PP MUj NU THET BET J g A V Gamjk x z
```

Weiclapos: A program for computing posterior distributions with a bivariate CC assuming Weibull Marginal distributions, this is applied to kidney infection data using Metropolis-Hastings algorithm.

```
**************************************************************************
clear all
P=kiddata;      %Kidney data file
x=P(:,2:3);
x1=x(:,1);
n=length(x);
stat=P(:,4:5);
Z0=ones(n,1);
Z1=P(:,6);
Z2=P(:,7);
Z3=P(:,9);
Z4=P(:,10);
Z5=P(:,11);
Z=[Z0 Z1 Z2 Z3 Z4 Z5];
tho=2.3781;
M=2;
cut=[0 30 90 180 365 600];
Int=[30 90 180 365 600];
LInt=log(Int);
cat=length(cut);
r=cat-1;
for m=1:M
  for j=1:r
    for i=1:n
      if x(i,m)>=cut(j) & x(i,m)<cut(j+1)
        alp(i,m)=j;
      end
    end
  end
end
%LInt1=log(Int(:,1));
%LInt2=log(Int(:,2));
sigma=2.033;
%sigma2=20.2789;
B0=[6.24 6.76];
B1=[0.1423  -1.3290  -0.6706  -0.1599   1.3630];
Al=0.5*ones(2,r+1);
K=5000; BB0=[]; BB1=[]; Fd=[]; Sd=[]; THO=[]; SIGMA=[];  tel1=0;
for It=1:K
  It
  Itho=1/tho;
  L=[];
  for i=1:n
    PS1=[1 cumprod(exp(-exp((LInt-Z0(i)*B0(1,1)-Z(i,:)*B1')/sigma)))].^-Itho;
    PS2=[1 cumprod(exp(-exp((LInt-Z0(i)*B0(1,2)-Z(i,:)*B1')/sigma)))].^-Itho;
    j=alp(i,1); q=alp(i,2);
    if stat(i,1)==1 & stat(i,2)==1
      l=((PS1(j+1)+PS2(q+1)-1)^-tho)+((PS1(j)+PS2(q)-1)^-tho)-((PS1(j)+PS2(q+1)-1)^-tho)-
((PS1(j+1)+PS2(q)-1)^-tho);
```

```
    elseif stat(i,1)==1 & stat(i,2)==0
      l=((PS1(j)+PS2(q+1)-1)^-tho)-((PS1(j+1)+PS2(q+1)-1)^-tho);
    elseif stat(i,1)==0 & stat(i,2)==1
      l=((PS1(j+1)+PS2(q)-1)^-tho)-((PS1(j+1)+PS2(q+1)-1)^-tho);
    elseif stat(i,1)==0 & stat(i,2)==0
      l=((PS1(j+1)+PS2(q+1)-1)^-tho);
    end
    L=[L;l];
  end
  like=sum(log(L))
  u=1; a=0;
  while u>a
    nb1=randn(1,5)*0.05+B1;
    nb0=randn(1,2)*0.05+B0;
    tel1=tel1+1;
    nL=[];
    for i=1:n
      nPS1=[1 cumprod(exp(-exp((LInt-Z0(i)*nb0(1,1)-Z(i,:)*nb1')/sig)))].^-Itho;
      nPS2=[1 cumprod(exp(-exp((LInt-Z0(i)*nb0(1,2)-Z(i,:)*nb1')/sig)))].^-Itho;
      j=alp(i,1); q=alp(i,2);
      if stat(i,1)==1 & stat(i,2)==1
        l=((nPS1(j+1)+nPS2(q+1)-1)^-tho)+((nPS1(j)+nPS2(q)-1)^-tho)-((nPS1(j)+nPS2(q+1)-1)^-tho)-
((nPS1(j+1)+nPS2(q)-1)^-tho);
      elseif stat(i,1)==1 & stat(i,2)==0
        l=((nPS1(j)+nPS2(q+1)-1)^-tho)-((nPS1(j+1)+nPS2(q+1)-1)^-tho);
      elseif stat(i,1)==0 & stat(i,2)==1
        l=((nPS1(j+1)+nPS2(q)-1)^-tho)-((nPS1(j+1)+nPS2(q+1)-1)^-tho);
      elseif stat(i,1)==0 & stat(i,2)==0
        l=((nPS1(j+1)+nPS2(q+1)-1)^-tho);
      end
      nL=[nL;l];
    end
    npos=sum(log(nL));
    a=min(exp(npos-pos),1);
    u=rand(1,1);
  end
  B1=nb;
  B0=nb0;
  BB1=[BB1;B1];
  BB0=[BB0;B0];
  for i=1:n
    PS1=[1 cumprod(exp(-exp((LInt-Z(i,:)*B1')/sigma)))].^-Itho;
    PS2=[1 cumprod(exp(-exp((LInt-Z(i,:)*B2')/sigma)))].^-Itho;
    j=alp(i,1); q=alp(i,2);
    if stat(i,1)==1 & stat(i,2)==1
      l=((PS1(j+1)+PS2(q+1)-1)^-tho)+((PS1(j)+PS2(q)-1)^-tho)-((PS1(j)+PS2(q+1)-1)^-tho)-
((PS1(j+1)+PS2(q)-1)^-tho);
    elseif stat(i,1)==1 & stat(i,2)==0
      l=((PS1(j)+PS2(q+1)-1)^-tho)-((PS1(j+1)+PS2(q+1)-1)^-tho);
    elseif stat(i,1)==0 & stat(i,2)==1
      l=((PS1(j+1)+PS2(q)-1)^-tho)-((PS1(j+1)+PS2(q+1)-1)^-tho);
    elseif stat(i,1)==0 & stat(i,2)==0
```

```
    l=((PS1(j+1)+PS2(q+1)-1)^-tho);
   end
   L=[L;l];
  end
  like=sum(log(L))
  u=1; a=0;
  while u>a
   lsig=0.01*randn(1,1)+log(sigma1);
   nsig=exp(lsig);
   Ltho=0.01*randn(1,1)+log(tho);
   ntho=exp(Ltho);
   nItho=1/ntho;
   tel1=tel1+1;
   nL=[];
   for i=1:n
    nPS1=[1 cumprod(exp(-exp((LInt-Z0(i)*B0(1,1)-Z(i,:)*B1')/nsig)))].^-nItho;
    nPS2=[1 cumprod(exp(-exp((LInt-Z0(i)*B0(1,2)-Z(i,:)*B1')/nsig)))].^-nItho;
    j=alp(i,1); q=alp(i,2);
    if stat(i,1)==1 & stat(i,2)==1
     l=((nPS1(j+1)+nPS2(q+1)-1)^-ntho)+((nPS1(j)+nPS2(q)-1)^-ntho)-((nPS1(j)+nPS2(q+1)-1)^-
ntho)-((nPS1(j+1)+nPS2(q)-1)^-ntho);
    elseif stat(i,1)==1 & stat(i,2)==0
     l=((nPS1(j)+nPS2(q+1)-1)^-ntho)-((nPS1(j+1)+nPS2(q+1)-1)^-ntho);
    elseif stat(i,1)==0 & stat(i,2)==1
     l=((nPS1(j+1)+nPS2(q)-1)^-ntho)-((nPS1(j+1)+nPS2(q+1)-1)^-ntho);
    elseif stat(i,1)==0 & stat(i,2)==0
     l=((nPS1(j+1)+nPS2(q+1)-1)^-ntho);
    end
    nL=[nL;l];
   end
   npos=sum(log(nL));
   a=min(exp(npos-pos),1);
   u=rand(1,1);
  end
  sigma1=nsig1;
  tho=ntho;
  SIGMA=[SIGMA;sigma];
  THO=[THO;tho];
  end
```

Indpos: A Program for simulating posterior distributions for bivariate data using Metropolis-Hastings algorithm on IW model .

```
**********************************************************************************************************
clear all
% Data file
% Starting values for all parameters
%Specify Intervals, # of intervals

M=2;
BB=[]; PPr=[]; PPc=[]; LI=[]; K=20000; tel1=0; Like=[];
for It=1:K
  It
  RLI=[]; CLI=[];
  for i=1:n
    PcZ=1./(1+exp(cdelta+Z(i,:)*B'));
    PrZ=1./(1+exp(rdelta+Z(i,:)*B'));
    if cdalp(i,1)==0
      Hc=0;  sGc=0;
      for j=1:cdalp(i,2)
        Gc=(1-PcZ(j))*prod(PcZ(1:j-1));
        sGc=sGc+Gc;
        hc=Ac(i,j)*Gc;
        Hc=Hc+hc;
      end
      cli=Hc+(1-sGc);
    else
      cli=0;
      for j=1:cdr
        cl=Ac(i,j)*(1-PcZ(j))*prod(PcZ(1:j-1));
        cli=cli+cl;
      end
    end
    if rnalp(i,1)==0
      Hr=0;  sGr=0;
      for q=1:rnalp(i,2)
        Gr=(1-PrZ(q))*prod(PrZ(1:q-1));
        sGr=sGr+Gr;
        hr=Ar(i,q)*Gr;
        Hr=Hr+hr;
      end
      rli=Hr+(1-sGr);
    else
      rli=0;
      for q=1:rnr
        rl=Ar(i,q)*(1-PrZ(q))*prod(PrZ(1:q-1));
        rli=rli+rl;
      end
    end
    RLI=[RLI;rli];CLI=[CLI;cli];
  end
  Li=sum(log(RLI))+sum(log(CLI))
  lpost=Li;
```

```
ZPc=min(Pc,0.9999); Pc=max(ZPc,0.0001);
ZPr=min(Pr,0.9999); Pr=max(ZPr,0.0001);
u=1; a=0;
while u>a
  nb=randn(1,14)*diag(varb)*0.01+B;
  nCLI=[]; nRLI=[]; tel1=tel1+1;
  for i=1:n
    PoZc=1./(1+exp(cdelta+Z(i,:)*nb'));
    PoZr=1./(1+exp(rdelta+Z(i,:)*nb'));
    if rnalp(i,1)==0
      Hr=0;     sGr=0;
      for q=1:rnalp(i,2)
        Gr=(1-PoZr(q))*prod(PoZr(1:q-1));
        sGr=sGr+Gr;
        hr=Ar(i,q)*Gr;
        Hr=Hr+hr;
      end
      rli=Hr+(1-sGr);
    else
      rli=0;
      for q=1:rnr
        rl=Ar(i,q)*(1-PoZr(q))*prod(PoZr(1:q-1));
        rli=rli+rl;
      end
    end
    if cdalp(i,1)==0
      Hc=0;     sGc=0;
      for j=1:cdalp(i,2)
        Gc=(1-PoZc(j))*prod(PoZc(1:j-1));
        sGc=sGc+Gc;
        hc=Ac(i,j)*Gc;
        Hc=Hc+hc;
      end
      cli=Hc+(1-sGc);
    else
      cli=0;
      for j=1:cdr
        cl=Ac(i,j)*(1-PoZc(j))*prod(PoZc(1:j-1));
        cli=cli+cl;
      end
    end
    nCLI=[nCLI;cli];
    nRLI=[nRLI;rli];
  end
  sLi=sum(log(nRLI))+sum(log(nCLI));
  lposto=sLi;
  a=min(exp(lposto-lpost),1);
  u=rand(1,1);
end
B=nb;
BB=[BB;B];
```

```
% simulating baseline parameters
lpr=0;
for q=1:rnr
  for j=1:cdr
    csal=sum(cAl(j+1:cdr+1));
    sal1=sum(cAl(j:cdr+1));
    %cpr=(1/beta(csal,cAl(1,j)))*exp(cAl*cdelta(1,j))/(1+exp(cdelta(1,j)))^sal2;
     rsal=sum(rAl(q+1:rnr+1));
     sal2=sum(rAl(q:rnr+1));
    %rpr=(1/beta(rsal,rAl(1,q)))*exp(rAl*rdelta(1,q))/(1-exp(rdelta(1,q)))^sal2;
    cpr=exp(cdelta(j))/(1+exp(cdelta(j)))^2;
    rpr=exp(rdelta(q))/(1+exp(rdelta(q)))^2;
    lpr=lpr+log(cpr)+log(rpr);
  end
end
Cli=[]; Rli=[]; PC=[];
for i=1:n
  PcZ=1./(1+exp(cdelta+Z(i,:)*B'));
  PrZ=1./(1+exp(rdelta+Z(i,:)*B'));
  if cdalp(i,1)==0
    Hc=0; sGc=0;
    for j=1:cdalp(i,2)
      Gc=(1-PcZ(j))*prod(PcZ(1:j-1));
      sGc=sGc+Gc;
      hc=Ac(i,j)*Gc;
      Hc=Hc+hc;
    end
    cli=Hc+(1-sGc);
  else
    cli=0;
    for j=1:cdr
      cl=Ac(i,j)*(1-PcZ(j))*prod(PcZ(1:j-1));
      cli=cli+cl;
    end
  end
  if rnalp(i,1)==0
    Hr=0; sGr=0;
    for q=1:rnalp(i,2)
      Gr=(1-PrZ(q))*prod(PrZ(1:q-1));
      sGr=sGr+Gr;
      hr=Ar(i,q)*Gr;
      Hr=Hr+hr;
    end
    rli=Hr+(1-sGr);
  else
    rli=0;
    for q=1:rnr
      rl=Ar(i,q)*(1-PrZ(q))*prod(PrZ(1:q-1));
      rli=rli+rl;
    end
  end
  Rli=[Rli;rli];Cli=[Cli;cli];
```

```
end
Li=sum(log(Rli))+sum(log(Cli));
lpost=Li+lpr;
ZPc=min(Pc,0.9999); Pc=max(ZPc,0.0001);
ZPr=min(Pr,0.9999); Pr=max(ZPr,0.0001);
u=1; a=0;
while u>a
  Poc=0.01*randn(1,cdr)*diag(varc)+cdelta;
  Por=0.01*randn(1,rnr)*diag(varr)+rdelta;
  lpro=0;
  for q=1:rnr
    for j=1:cdr
      csal=sum(cAl(j+1:cdr+1));
      sal1=sum(cAl(j:cdr+1));
      %ncpr=(1/beta(csal,cAl(1,j)))*exp(cAl(1,j)*Poc(j))/(1+exp(Poc(j)))^sal1;
      rsal=sum(rAl(q+1:rnr+1));
      sal2=sum(rAl(q:rnr+1));
      %nrpr=(1/beta(rsal,rAl(1,q)))*exp(rAl(1,q)*Por(q))/(1+exp(Por(q)))^rsal2;
      ncpr=exp(Poc(j))/(1+exp(Poc(j)))^2;
      nrpr=exp(Por(q))/(1+exp(Por(q)))^2;
      lpro=lpro+log(ncpr)+log(nrpr);
    end
  end
  nCli=[ ]; nRli=[ ]; tel1=tel1+1;
  for i=1:n
    PoZc=1./(1+exp(Poc+Z(i,:)*B'));
    PoZr=1./(1+exp(Por+Z(i,:)*B'));
    if rnalp(i,1)==0
      Hr=0;     sGr=0;
      for q=1:rnalp(i,2)
        Gr=(1-PoZr(q))*prod(PoZr(1:q-1));
        sGr=sGr+Gr;
        hr=Ar(i,q)*Gr;
        Hr=Hr+hr;
      end
      rli=Hr+(1-sGr);
    else
      rli=0;
      for q=1:rnr
        rl=Ar(i,q)*(1-PoZr(q))*prod(PoZr(1:q-1));
        rli=rli+rl;
      end
    end
    if cdalp(i,1)==0
      Hc=0;     sGc=0;
      for j=1:cdalp(i,2)
        Gc=(1-PoZc(j))*prod(PoZc(1:j-1));
        sGc=sGc+Gc;
        hc=Ac(i,j)*Gc;
        Hc=Hc+hc;
      end
      cli=Hc+(1-sGc);
```

```
      else
        cli=0;
        for j=1:cdr
          cl=Ac(i,j)*(1-PoZc(j))*prod(PoZc(1:j-1));
          cli=cli+cl;
        end
      end
      nCli=[nCli;cli];
      nRli=[nRli;rli];
    end
    sLi=sum(log(nRli))+sum(log(nCli));
    lposto=sLi+lpro;
    a=min(exp(lposto-lpost),1);
    u=rand(1,1);
  end
  cdelta=Poc;
  rdelta=Por;
  PPc=[PPc;cdelta];
  PPr=[PPr;rdelta];
end
HSB=sort(BB);
HPB=[HSB((K*0.025),:);HSB((K*0.975),:)]
```

Conpos: **A Program for simulating posterior distributions for bivariate data using Metropolis-Hastings algorithm on Conditional bivariate (CB) model.**

```
**********************************************************************************************************
clear all
% Data file
% Starting values for all parameters
%Specify Intervals, # of intervals

M=2;
BB=[]; CDdelta=[]; RNdelta=[]; LI=[]; K=5000;
for It=1:K
  It
  P1=[];
  for i=1:n
    p1=exp(-exp(fdelta+Z(i,:)*B'));
    P1=[P1;p1];
  end
  for q=1:rnr
    for i=1:n
      for j=1:cdr
        P2(i,j)=exp(-exp(cdelta(q,j)+Z(i,:)*B'));
      end
      cl=cdalp(i,1);cu=cdalp(i,2);
      fl=rnalp(i,1);fu=rnalp(i,2);
      if cl==0
        li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*prod(P2(i,1:cu));
      elseif cl==cu
        li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1));
      elseif cl~=0 & cl==cu-1
        li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2));
      elseif cl~=0 & cl==cu-2
        li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3));
      elseif cl~=0 & cl==cu-4
        li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3))+(1-
P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-3))*prod(P2(i,1:cu-4))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-
4))*prod(P2(i,1:cu-5));
      end
    end
  end
  for i=1:n
    cl=cdalp(i,1);cu=cdalp(i,2);
    fl=rnalp(i,1);fu=rnalp(i,2);
    if cl==0 & fl==0
      lli(i)=log(sum(li(i,fu+1:rnr),2));
    elseif fl>0 & cl==0
      lli(i)=log(sum(li(i,fl:fu)));
    elseif fl==0 & cl>0
      lli(i)=log(sum(li(i,fu+1:rnr),2));
    elseif fl>0 & cl>0
```

```
    lli(i)=log(sum(li(i,fl:fu)));
  end
end
lb=sum(lli)
u=1; a=0;
while u>a
  nb=randn(1,14)*0.01*diag(varb)+B;
  nP1=[];
  for i=1:n
    np1=exp(-exp(fdelta+Z(i,:)*nb'));
    nP1=[nP1;np1];
  end
  for q=1:rnr
    for i=1:n
      for j=1:cdr
        nP2(i,j)=exp(-exp(cdelta(q,j)+Z(i,:)*nb'));
      end
      cl=cdalp(i,1);cu=cdalp(i,2);
      fl=rnalp(i,1);fu=rnalp(i,2);
      if  cl==0
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*prod(nP2(i,1:cu));
      elseif cl==cu
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1));
      elseif cl~=0 & cl==cu-1
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2));
      elseif cl~=0 & cl==cu-2
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3));
      elseif cl~=0 & cl==cu-4
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-3))*prod(nP2(i,1:cu-
4))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-4))*prod(nP2(i,1:cu-5));
      end
    end
  end
  for i=1:n
    cl=cdalp(i,1);cu=cdalp(i,2);
    fl=rnalp(i,1);fu=rnalp(i,2);
    if cl==0 & fl==0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl==0
      nlli(i)=log(sum(nli(i,fl:fu)));
    elseif fl==0 & cl>0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl>0
      nlli(i)=log(sum(nli(i,fl:fu)));
    end
  end
  nlb=sum(nlli);
```

```
  a=min(exp(nlb-lb),1);
  u=rand(1,1);
end
B=nb;
BB=[BB;B];

%SIMULATING PARAMETERS FOR RNA PARAMETERS
priof=0;
sfpr=0;
for q=1:rnr
  fsal=sum(fAl(q+1:rnr+1));
  %fpr=(1/beta(fsal,fAl(1,q)))*exp(fdelta(1,q)-fsal*
                exp(fdelta(1,q)))*(1-exp(-exp(fdelta(1,q))))^(fAl(1,q)-1);
  fpr=exp(-exp(fdelta(1,q)))*exp(fdelta(1,q));
  sfpr=sfpr+log(fpr);
end
priof=sfpr;
P1=[];
for i=1:n
  p1=exp(-exp(fdelta+Z(i,:)*B'));
  P1=[P1;p1];
end
for q=1:rnr
  for i=1:n
    P2=[];
    for j=1:cdr
      P2(i,j)=exp(-exp(cdelta(q,j)+Z(i,:)*B'));
    end
    cl=cdalp(i,1);cu=cdalp(i,2);
    fl=rnalp(i,1);fu=rnalp(i,2);
    if cl==0
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*prod(P2(i,1:cu));
    elseif cl==cu
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1));
    elseif cl~=0 & cl==cu-1
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2));
    elseif cl~=0 & cl==cu-2
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3));
    elseif cl~=0 & cl==cu-4
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3))+(1-
P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-3))*prod(P2(i,1:cu-4))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-
4))*prod(P2(i,1:cu-5));
    end
  end
end
for i=1:n
  cl=cdalp(i,1);cu=cdalp(i,2);
  fl=rnalp(i,1);fu=rnalp(i,2);
  if cl==0 & fl==0
```

```
    lli(i)=log(sum(li(i,fu+1:rnr),2));
  elseif fl>0 & cl==0
    lli(i)=log(sum(li(i,fl:fu)));
  elseif fl==0 & cl>0
    lli(i)=log(sum(li(i,fu+1:rnr),2));
  elseif fl>0 & cl>0
    lli(i)=log(sum(li(i,fl:fu)));
  end
end
lf=sum(lli)+priof
u=1; a=0;
while u>a
  nfdelta=randn(1,rnr)*0.01*diag(fvar)+fdelta;
  npriof=0;
  nfpr=0;
  for q=1:rnr
    fsal=sum(fAl(q+1:rnr+1));
    %nfr=(1/beta(fsal,fAl(1,q)))*exp(nfdelta(1,q)-fsal*
                    exp(nfdelta(q)))*(1-exp(-exp(nfdelta(1,q))))^(fAl(1,q)-1);
    nfr=exp(-exp(nfdelta(1,q)))*exp(nfdelta(1,q));
    nfpr=nfpr+log(nfr);
  end
  npriof=nfpr;
  nP1=[];
  for i=1:n
    np1=exp(-exp(nfdelta+Z(i,:)*B'));
    nP1=[nP1;np1];
  end
  nP2=[];
  for q=1:rnr
    for i=1:n
      for j=1:cdr
        nP2(i,j)=exp(-exp(cdelta(q,j)+Z(i,:)*B'));
      end
      cl=cdalp(i,1);cu=cdalp(i,2);
      fl=rnalp(i,1);fu=rnalp(i,2);
      if  cl==0
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*prod(nP2(i,1:cu));
      elseif cl==cu
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1));
      elseif cl~=0 & cl==cu-1
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2));
      elseif cl~=0 & cl==cu-2
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3));
      elseif cl~=0 & cl==cu-4
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-3))*prod(nP2(i,1:cu-
4))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-4))*prod(nP2(i,1:cu-5));
```

```
      end
    end
  end
  for i=1:n
    cl=cdalp(i,1);cu=cdalp(i,2);
    fl=rnalp(i,1);fu=rnalp(i,2);
    if cl==0 & fl==0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl==0
      nlli(i)=log(sum(nli(i,fl:fu)));
    elseif fl==0 & cl>0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl>0
      nlli(i)=log(sum(nli(i,fl:fu)));
    end
  end
  nlf=sum(nlli)+npriof;
  a=min(exp(nlf-lf),1);
  u=rand(1,1);
 end
 fdelta=nfdelta;
 RNdelta=[RNdelta;fdelta];

%Simulating for CD4 failure parameters
 prioc=0;
 scpr=0;
 for q=1:rnr
   for j=1:cdr
     csal=sum(cAl(j+1:cdr+1));
     %cpr=(1/beta(csal,cAl(q,j)))*exp(cdelta(q,j)-csal*
                exp(cdelta(q,j)))*(1-exp(-exp(cdelta(q,j))))^(cAl(q,j)-1);
     cpr=exp(-exp(cdelta(q,j)))*exp(cdelta(q,j));
     scpr=scpr+log(cpr);
   end
 end
 prioc=scpr;
 P1=[];
 for i=1:n
   p1=exp(-exp(fdelta+Z(i,:)*B'));
   P1=[P1;p1];
 end
 P2=[];
 for q=1:rnr
   for i=1:n
     for j=1:cdr
       P2(i,j)=exp(-exp(cdelta(q,j)+Z(i,:)*B'));
     end
     cl=cdalp(i,1);cu=cdalp(i,2);
     fl=rnalp(i,1);fu=rnalp(i,2);
     if cl==0
       li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*prod(P2(i,1:cu));
     elseif cl==cu
```

```
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))* prod(P2(i,1:cu-1));
    elseif cl~=0 & cl==cu-1
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))* prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2));
    elseif cl~=0 & cl==cu-2
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))* prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3));
    elseif cl~=0 & cl==cu-4
      li(i,q)=(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu))*prod(P2(i,1:cu-1))+(1-P1(i,q))*prod(P1(i,1:q-
1))*(1-P2(i,cu-1))*prod(P2(i,1:cu-2))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-2))*prod(P2(i,1:cu-3))+(1-
P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-3))*prod(P2(i,1:cu-4))+(1-P1(i,q))*prod(P1(i,1:q-1))*(1-P2(i,cu-
4))*prod(P2(i,1:cu-5));
    end
  end
end
for i=1:n
  cl=cdalp(i,1);cu=cdalp(i,2);
  fl=rnalp(i,1);fu=rnalp(i,2);
  if cl==0 & fl==0
    lli(i)=log(sum(li(i,fu+1:rnr),2));
  elseif fl>0 & cl==0
    lli(i)=log(sum(li(i,fl:fu)));
  elseif fl==0 & cl>0
    lli(i)=log(sum(li(i,fu+1:rnr),2));
  elseif fl>0 & cl>0
    lli(i)=log(sum(li(i,fl:fu)));
  end
end
lc=sum(lli)+prioc;
u=1; a=0;
while u>a
  ndel=randn(1,cdr*rnc r)*0.01+[cdelta(1,:) cdelta(2,:) cdelta(3,:)  delta(4,:) cdelta(5,:) ];
  ncdelta=[ndel(1,1:14);ndel(1,15:28);ndel(1,29:42);ndel(1,43:56); ndel(1,57:70) ];
  nprioc=0;
  ncpr=0;
  for q=1:rnr
    for j=1:cdr
      csal=sum(cAl(j+1:cdr+1));
      %ncr=(1/beta(csal,cAl(q,j)))*exp(ncdelta(q,j)-csal*
            exp(ncdelta(q,j)))*(1-exp(-exp(ncdelta(q,j))))^(cAl(q,j)-1);
      ncr=exp(-exp(ncdelta(q,j)))*exp(ncdelta(q,j));
      ncpr=ncpr+log(ncr);
    end
  end
  nprioc=ncpr;
  nP1=[];
  for i=1:n
    np1=exp(-exp(fdelta+Z(i,:)*B'));
    nP1=[nP1;np1];
  end
  nP2=[];
  for q=1:rnr
```

```
    for i=1:n
      for j=1:cdr
        nP2(i,j)=exp(-exp(ncdelta(q,j)+Z(i,:)*B'));
      end
      cl=cdalp(i,1);cu=cdalp(i,2);
      fl=rnalp(i,1);fu=rnalp(i,2);
      if cl==0
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*prod(nP2(i,1:cu));
      elseif cl==cu
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1));
      elseif cl~=0 & cl==cu-1
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2));
      elseif cl~=0 & cl==cu-2
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3));
      elseif cl~=0 & cl==cu-4
        nli(i,q)=(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu))*prod(nP2(i,1:cu-1))+(1-
nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-1))*prod(nP2(i,1:cu-2))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-
nP2(i,cu-2))*prod(nP2(i,1:cu-3))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-3))*prod(nP2(i,1:cu-
4))+(1-nP1(i,q))*prod(nP1(i,1:q-1))*(1-nP2(i,cu-4))*prod(nP2(i,1:cu-5));
      end
    end
  end
  for i=1:n
    cl=cdalp(i,1);cu=cdalp(i,2);
    fl=rnalp(i,1);fu=rnalp(i,2);
    if cl==0 & fl==0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl==0
      nlli(i)=log(sum(nli(i,fl:fu)));
    elseif fl==0 & cl>0
      nlli(i)=log(sum(nli(i,fu+1:rnr),2));
    elseif fl>0 & cl>0
      nlli(i)=log(sum(nli(i,fl:fu)));
    end
  end
  nlc=sum(nlli)+nprioc;
  a=min(exp(nlc-lc),1);
  u=rand(1,1);
 end
 cdelta=ncdelta;
 CDdelta=[CDdelta;cdelta];
end
codel1=CDdelta(1:5:5*K-4,:);
codel2=CDdelta(2:5:5*K-3,:);
codel3=CDdelta(3:5:5*K-2,:);
codel4=CDdelta(4:5:5*K-1,:);
codel5=CDdelta(5:5:5*K,:);
```

Clapos: **A Program for simulating posterior distributions for bivariate data using Metropolis-Hastings algorithm on Clayton copula (CC) model. (Nonparametric)**

```
**********************************************************************************************************
clear all
% Data file
% Starting values for all parameters
%Specify Intervals, # of intervals.
M=2;
cAl=0.5*ones(1,cdr+1);
rAl=0.5*ones(1,rnr+1);
BB=[]; CDdelta=[]; RNdelta=[]; THO=[]; POS=[]; K=10000; tel1=0;
for It=1:K
  It
  for i=1:n
    PSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
    PSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-Itho;
     jl=cdalp(i,1);ju=cdalp(i,2);ql=rnalp(i,1);qu=rnalp(i,2);
    if STATC(i)==0 & STATR(i)==0
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==0 & STATR(i)==1
      li=((PSC(ju+1)+PSR(ql)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==0
      li=((PSC(jl)+PSR(qu+1)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==1
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho)+((PSC(jl)+PSR(ql)-1)^-tho)-((PSC(jl)+PSR(qu+1)-1)^-tho)-
((PSC(ju+1)+PSR(ql)-1)^-tho);
    end
    L=[L;li];
  end
  ll=sum(log(L))
  fb=ll;
  POS=[POS;fb];
  u=1; a=0;
  while u>a
    nb=randn(1,14)*0.01*diag(varb)+B;
    tel1=tel1+1;
    NL=[];
    for i=1:n
      NPSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*nb')))].^-Itho;
      NPSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*nb')))].^-Itho;
      jl=cdalp(i,1); ju=cdalp(i,2); ql=rnalp(i,1); qu=rnalp(i,2);
      if STATC(i)==0 & STATR(i)==0
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==0 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(ql)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==0
        l=((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho)+((NPSC(jl)+NPSR(ql)-1)^-tho)-
((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(ql)-1)^-tho);
      end
      NL=[NL;l];
```

```
    end
    nll=sum(log(NL));
    nfb=nll+nprior;
    a=min(exp(nfb-fb),1);
    u=rand(1,1);
  end
  B=nb;
  BB=[BB;B];

  %SIMULATING PARAMETERS FOR CD4 VARIABLE;
  prioc=0;
  for j=1:cdr
     csal=sum(cAl(j+1:cdr+1));
     cpr=beta(csal,cAl(1,j))^-1*exp(cdelta(1,j)-csal*exp(cdelta(1,j)))*(1-exp(-
exp(cdelta(1,j))))^(cAl(1,j)-1);
      prioc=prioc+log(cpr);
    end
  end
  L=[];
  for i=1:n
    PSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
    PSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-Itho;
     jl=cdalp(i,1);ju=cdalp(i,2);ql=rnalp(i,1);qu=rnalp(i,2);
    if STATC(i)==0 & STATR(i)==0
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==0 & STATR(i)==1
      li=((PSC(ju+1)+PSR(ql)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==0
      li=((PSC(jl)+PSR(qu+1)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==1
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho)+((PSC(jl)+PSR(ql)-1)^-tho)-((PSC(jl)+PSR(qu+1)-1)^-tho)-
((PSC(ju+1)+PSR(ql)-1)^-tho);
    end
    L=[L;li];
  end
  ll=sum(log(L))
  fc=ll+prioc;
  POSC=[POSC;fc];
  u=1; a=0;
  while u>a
    nprioc=0;
        ncdelta=randn(1,14)*0.01*diag(cocd,0)+cdelta;
        tel1=tel1+1;
    for j=1:cdr
       csal=sum(cAl(j+1:cdr+1));
       ncpr=(beta(csal,cAl(1,j))^-1)*exp(ncdelta(j)-csal*exp(ncdelta(j)))*(1-exp(-
exp(ncdelta(j))))^(cAl(j)-1);
       nprior=nprior+log(ncpr);
     end
    end
    NL=[];
    for i=1:n
```

```
      NPSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
      NPSC=[1 cumprod(exp(-exp(ncdelta+Z(i,:)*B')))].^-Itho;
      jl=cdalp(i,1); ju=cdalp(i,2); ql=rnalp(i,1); qu=rnalp(i,2);
      if STATC(i)==0 & STATR(i)==0
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==0 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(ql)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==0
        l=((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho)+((NPSC(jl)+NPSR(ql)-1)^-tho)-
((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(ql)-1)^-tho);
      end
      NL=[NL;l];
    end
    nll=sum(log(NL));
    nfc=nll+nprioc;
    a=min(exp(nfc-fc),1);
    u=rand(1,1);
  end
  cdelta=ncdelta;
  CDdelta=[CDdelta;cdelta];

  %SIMULATING PARAMETER ESTIMATES FOR RNA
  prior=0;
  for q=1:rnr
    rsal=sum(rAl(q+1:rnr+1));
    rpr=beta(rsal,rAl(1,q))^-1*exp(rdelta(1,q)-rsal*exp(rdelta(1,q)))*(1-exp(-
exp(rdelta(1,q))))^(rAl(1,q)-1);;
    prior=prior+log(rpr);
  end
  L=[];
  for i=1:n
    PSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
    PSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-Itho;
     jl=cdalp(i,1);ju=cdalp(i,2);ql=rnalp(i,1);qu=rnalp(i,2);
    if STATC(i)==0 & STATR(i)==0
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==0 & STATR(i)==1
      li=((PSC(ju+1)+PSR(ql)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==0
      li=((PSC(jl)+PSR(qu+1)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==1
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho)+((PSC(jl)+PSR(ql)-1)^-tho)-((PSC(jl)+PSR(qu+1)-1)^-tho)-
((PSC(ju+1)+PSR(ql)-1)^-tho);
    end
    L=[L;li];
  end
  ll=sum(log(L))
  fr=ll+prior;
  POSR=[POSR;fr];
  u=1; a=0;
```

```
  while u>a
    nprior=0;
    nrdelta=randn(1,5)*0.01*diag(corn,0)+rdelta;
    tel1=tel1+1;
    for q=1:rnr
      rsal=sum(rAl(q+1:rnr+1));
      nrpr=(beta(rsal,rAl(1,q))^-1)*exp(nrdelta(q)-rsal*exp(nrdelta(q)))*(1-exp(-
exp(nrdelta(q))))^(rAl(q)-1); %assume Al(m,j)=1 for all m,j
      nprior=nprior+log(nrpr);
    end
    NL=[];
    for i=1:n
      NPSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
      NPSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-Itho;
      jl=cdalp(i,1); ju=cdalp(i,2); ql=rnalp(i,1); qu=rnalp(i,2);
      if STATC(i)==0 & STATR(i)==0
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==0 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(ql)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==0
        l=((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-tho);
      elseif STATC(i)==1 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-tho)+((NPSC(jl)+NPSR(ql)-1)^-tho)-
((NPSC(jl)+NPSR(qu+1)-1)^-tho)-((NPSC(ju+1)+NPSR(ql)-1)^-tho);
      end
      NL=[NL;l];
    end
    nll=sum(log(NL));
    nfr=nll+nprior;
    a=min(exp(nfr-fr),1);
    u=rand(1,1);
  end
  rdelta=nrdelta;
  RNdelta=[RNdelta;rdelta];
  %SIMULATING THE DEPEDENCE PARAMETER

  Itho=1/tho;
  L=[];
  for i=1:n
    PSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-Itho;
    PSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-Itho;
     jl=cdalp(i,1);ju=cdalp(i,2);ql=rnalp(i,1);qu=rnalp(i,2);
    if STATC(i)==0 & STATR(i)==0
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==0 & STATR(i)==1
      li=((PSC(ju+1)+PSR(ql)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==0
      li=((PSC(jl)+PSR(qu+1)-1)^-tho)-((PSC(ju+1)+PSR(qu+1)-1)^-tho);
    elseif STATC(i)==1 & STATR(i)==1
      li=((PSC(ju+1)+PSR(qu+1)-1)^-tho)+((PSC(jl)+PSR(ql)-1)^-tho)-((PSC(jl)+PSR(qu+1)-1)^-tho)-
((PSC(ju+1)+PSR(ql)-1)^-tho);
    end
```

```
    L=[L;li];
  end
  ll=sum(log(L))
  fd=ll;
  u=1; a=0;
  while u>a
    nho=0.01*randn(1,1)+log(tho);
    ntho=exp(nho);
    nItho=1/ntho; tel1=tel1+1;
    nprior=nprior+log(ncpr)+log(nrpr);
    NL=[];
    for i=1:n
      NPSR=[1 cumprod(exp(-exp(rdelta+Z(i,:)*B')))].^-nItho;
      NPSC=[1 cumprod(exp(-exp(cdelta+Z(i,:)*B')))].^-nItho;
      jl=cdalp(i,1); ju=cdalp(i,2); ql=rnalp(i,1); qu=rnalp(i,2);
      if STATC(i)==0 & STATR(i)==0
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-ntho);
      elseif STATC(i)==0 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(ql)-1)^-ntho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-ntho);
      elseif STATC(i)==1 & STATR(i)==0
        l=((NPSC(jl)+NPSR(qu+1)-1)^-ntho)-((NPSC(ju+1)+NPSR(qu+1)-1)^-ntho);
      elseif STATC(i)==1 & STATR(i)==1
        l=((NPSC(ju+1)+NPSR(qu+1)-1)^-ntho)+((NPSC(jl)+NPSR(ql)-1)^-ntho)-
((NPSC(jl)+NPSR(qu+1)-1)^-ntho)-((NPSC(ju+1)+NPSR(ql)-1)^-ntho);
      end
      NL=[NL;l];
    end
    nll=sum(log(NL));
    nfd=nll;
    a=min(exp(nfd-fd),1);
    u=rand(1,1);
  end
  tho=ntho;
  THO=[THO;tho];
end
```

Visposaid.m A Matlab program for computing interval visting visiting probabilities for patients being monitored for CD4 cell count in ACTG75 study.

```
**********************************************************************************************************
clear
X=Visit;
D1=usaids;
n=length(D1);
D2=treat;
V=X(:,6:19);
cdlass=D1(:,3:4);
rnlass=D1(:,5:6);
for i=1:n
  for j=1:3
    if D1(i,9)==j
      Z(i,j)=1;
    else
      Z(i,j)=0;
    end
  end
end
A=[];
for i=1:n
  if X(i,4)==0
    a=[zeros(1,X(i,5)) ones(1,14-X(i,5))];
  else
    a=[zeros(1,X(i,4)-1) ones(1,X(i,5)-X(i,4)+1) zeros(1,14-X(i,5))];
  end
  A=[A;a];
end
Z1=D1(:,7);      %AGE
Z2=D1(:,8);      %SEX M=1;F=0
Z3=D1(:,10);     %KARN
Z4=D1(:,11);     %Previous ARV(ZDV) use Y=1;N=0
Z5=D1(:,19);     %Weight
Z6=D1(:,15);     %Homosexuality Y=1; N=0
Z7=D1(:,21);     %Symptomatic Y=1; N=0
Z8=D1(:,17);     %Intravenous Drug use Y=1; N=0
Z9=Z(:,1);       %Race1
Z10=Z(:,2);      %Race2
Z11=Z(:,3);      %Race3
Z12=D2(:,1);     %Didanosine ARV
Z13=D2(:,2);     %Zidovudine+Didanosine ARV
Z14=D2(:,3);     %Zidovudine+Zalcitabine ARV
Z=[Z1 Z2 Z3 Z4 Z5 Z6 Z7 Z8 Z9 Z10 Z11 Z12 Z13 Z14];
M=2;
Int=56:84:1148;
CDin=[0 Int];
cdcat=length(CDin);
r=cdcat-1;
Ri=X(:,2);
for m=1:M
  for j=1:r
```

```
    for i=1:n
      if cdlass(i,m)>=CDin(j) & cdlass(i,m)<CDin(j+1)
        cdalp(i,m)=j;
      elseif cdlass(i,1)==0
        cdalp(i,1)=0;
      end
    end
  end
end
%STARTING VALUES
%---------------
B=[0.03 -0.2 -0.029 0.55 -0.003 0.13 1.05  -0.2 1.2 1.2 0.5 -0.65 -1.1 -0.7];
cdelta=[-9.3 -1.8 -1.7 -1.7 -2 -2 -2 -1.5 -1.4 -1.1 -1.5 -0.9 -2.2 -0.9];
cAl=[0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5];
pa=[0.3 0.99 0.5 0.5 0.4 0.4 0.4 0.3 0.3 0.22 0.28 0.3 0.33 0.7];
pb=[0.91 0.99 0.92 0.91 0.91 0.88 0.89 0.92 0.88 0.91 0.89 0.94 0.95 0.7];
varb=[0.2 0.50 0.091 0.3 0.2 0.61 0.3  1.5 1.5  1.5  1.5  0.246  0.297  0.251];
cvpa=[0.0387   0.0331   0.0499   0.0380   0.0258   0.0734   0.0827   0.0588   0.0564   0.0367   0.0838
0.0528   0.0762   0.0698];
pa=0.1*ones(1,r); pb=0.93*ones(1,r);
Like=[];
P=exp(-exp(cdelta));
PP=[]; Pb=[]; Pa=[]; BB=[]; tel1=0; tel2=0; tel3=0;
for k=1:2000
  k
  %SIMULATING P
  %------------
  L=[];
  for i=1:n
    PZ=P.^exp(Z(i,:)*B');
    FB=(pb.^V(i,:)).*((1-pb).^(1-V(i,:)));
    FA=(pa.^V(i,:)).*((1-pa).^(1-V(i,:)));
    if X(i,4)==0
      H=0; sG=0;
      for j=1:Ri(i)-1
        E=prod(FB(1:j));
        F=prod(FA(j+1:Ri(i)));
        G=(1-PZ(j))*prod(PZ(1:j-1));
        sG=sG+G;
        h=A(i,j)*G*E*F;
        H=H+h;
      end
      ll=H+(1-sG)*prod(FB(1:Ri(i)));
    else
      ll=0;
      for j=1:r
        E=prod(FB(1:min(j,Ri(i))));
        F=prod(FA(j+1:Ri(i)));
        l=A(i,j)*(1-PZ(j))*E*F;
        ll=ll+l;
      end
    end
```

```
  L=[L;ll];
end
LL=sum(log(L))
Like=[Like;LL];
lnpr=0;
for j=1:r
  pr=(P(j)^(0.5*(r-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pr);
end
lnpost=LL+lnpr;
ZP=min(P,0.9999); P=max(ZP,0.0001);
%P=[Z2 0];
r1=1; a1=0;
while r1>=a1;
  gam=0.05*randn(1,r)+log(P(1:r)./(1-P(1:r)));
  Po=[1./(1+exp(-gam))];
  L=[]; tel1=tel1+1;
  for i=1:n
    PoZ=Po.^exp(Z(i,:)*B');
    FB=(pb.^V(i,:)).*((1-pb).^(1-V(i,:)));
    FA=(pa.^V(i,:)).*((1-pa).^(1-V(i,:)));
    if X(i,4)==0
      H=0; sG=0;
      for j=1:Ri(i)-1
        E=prod(FB(1:j));
        F=prod(FA(j+1:Ri(i)));
        G=(1-PoZ(j))*prod(PoZ(1:j-1));
        sG=sG+G;
        h=A(i,j)*G*E*F;
        H=H+h;
      end
      ll=H+(1-sG)*prod(FB(1:Ri(i)));
    else
      ll=0;
      for j=1:r
        l=A(i,j)*(1-PoZ(j))*prod(FB(1:min(j,Ri(i))))*prod(FA(j+1:Ri(i)));
        ll=ll+l;
      end
    end
    L=[L;ll];
  end
  LL=sum(log(L));
  lnpr=0;
  for j=1:r
    pro=(Po(j)^(0.5*(r-j-1)))*((1-Po(j))^(-0.5));
    lnpr=lnpr+log(pro);
  end
  lnposto=LL+lnpr;
  a1=exp(lnposto-lnpost)
  r1=rand(1,1);
end
P=Po;
```

```
PP=[PP;P];

%SIMULATING PB AND PA
%---------------------
L=[];
for i=1:n
  PZ=P.^exp(Z(i,:)*B');
  FB=(pb.^V(i,:)).*((1-pb).^(1-V(i,:)));
  FA=(pa.^V(i,:)).*((1-pa).^(1-V(i,:)));
  if X(i,4)==0
    H=0; sG=0;
    for j=1:Ri(i)-1
      E=prod(FB(1:j));
      F=prod(FA(j+1:Ri(i)));
      G=(1-PZ(j))*prod(PZ(1:j-1));
      sG=sG+G;
      h=A(i,j)*G*E*F;
      H=H+h;
    end
    ll=H+(1-sG)*prod(FB(1:Ri(i)));
  else
    ll=0;
    for j=1:r
      l=A(i,j)*(1-PZ(j))*prod(FB(1:min(j,Ri(i))))*prod(FA(j+1:Ri(i)));
      ll=ll+l;
    end
  end
  L=[L;ll];
end
LL=sum(log(L));
lnpr=0;
for j=1:r
  pr=(P(j)^(0.5*(r-j-1)))*((1-P(j))^(-0.5));
  lnpr=lnpr+log(pr);
end
lnpost=LL+lnpr;
r2=1; a2=0;
while r2>=a2;
  bb1=.5*randn(1,r)+log(pb./(1-pb));
  aa1=.5*randn(1,r)+log(pa./(1-pa));
  pbb=1./(1+exp(-bb1));
  paa=1./(1+exp(-aa1));
  L=[]; tel2=tel2+1;
  for i=1:n
    FB=(pbb.^V(i,:)).*((1-pbb).^(1-V(i,:)));
    FA=(paa.^V(i,:)).*((1-paa).^(1-V(i,:)));
    if X(i,4)==0
      H=0; sG=0;
      for j=1:Ri(i)-1
        E=prod(FB(1:j));
        F=prod(FA(j+1:Ri(i)));
        G=(1-PZ(j))*prod(PZ(1:j-1));
```

```
        sG=sG+G;
        h=A(i,j)*G*E*F;
        H=H+h;
      end
      ll=H+(1-sG)*prod(FB(1:Ri(i)));
    else
      ll=0;
      for j=1:r
        l=A(i,j)*(1-PZ(j))*prod(FB(1:min(j,Ri(i))))*prod(FA(j+1:Ri(i)));
        ll=ll+l;
      end
    end
    L=[L;ll];
  end
  LL=sum(log(L));
  lnpr=0;
  for j=1:r
    pro=(P(j)^(0.5*(r-j-1)))*((1-P(j))^(-0.5));
    lnpr=lnpr+log(pro);
  end
  lnposto=LL+lnpr;
  a2=exp(lnposto-lnpost)
  r2=rand(1,1);
end
pb=pbb;
pa=paa;
Pb=[Pb;pb];Pa=[Pa;pa];

%SIMULATING BETA
%---------------
L=[];
for i=1:n
  FB=(pb.^V(i,:)).*((1-pb).^(1-V(i,:)));
  FA=(pa.^V(i,:)).*((1-pa).^(1-V(i,:)));
  if X(i,4)==0
    H=0; sG=0;
    for j=1:Ri(i)-1
      E=prod(FB(1:j));
      F=prod(FA(j+1:Ri(i)));
      G=(1-PZ(j))*prod(PZ(1:j-1));
      sG=sG+G;
      h=A(i,j)*G*E*F;
      H=H+h;
    end
    ll=H+(1-sG)*prod(FB(1:Ri(i)));
  else
    ll=0;
    for j=1:r
      l=A(i,j)*(1-PZ(j))*prod(FB(1:min(j,Ri(i))))*prod(FA(j+1:Ri(i)));
      ll=ll+l;
    end
  end
```

```
    L=[L;ll];
  end
  LL=sum(log(L));
  lnpr=0;
  r3=1; a3=0;
  while r3>=a3;
    nb=B+randn(1,14)*0.01*diag(varb);
    L=[]; tel3=tel3+1;
    for i=1:n
      PZ=P.^exp(Z(i,:)*nb');
      FB=(pb.^V(i,:)).*((1-pb).^(1-V(i,:)));
      FA=(pa.^V(i,:)).*((1-pa).^(1-V(i,:)));
      if X(i,4)==0
        H=0; sG=0;
        for j=1:Ri(i)-1
          E=prod(FB(1:j));
          F=prod(FA(j+1:Ri(i)));
          G=(1-PZ(j))*prod(PZ(1:j-1));
          sG=sG+G;
          h=A(i,j)*G*E*F;
          H=H+h;
        end
        ll=H+(1-sG)*prod(FB(1:Ri(i)));
      else
        ll=0;
        for j=1:r
          l=A(i,j)*(1-PZ(j))*prod(FB(1:min(j,Ri(i))))*prod(FA(j+1:Ri(i)));
          ll=ll+l;
        end
      end
      L=[L;ll];
    end
    LLo=sum(log(L));
    a3=exp(LLo-LL)
    r3=rand(1,1);
  end
  B=nb;
  BB=[BB;B];
  %pause
end
%save c:\matlabr11\work\LES\cd4.mat PP Pa Pb
```

REFERENCES

Agresti, A. (1990): *Categorical Data Analysis*. John Wiley and Sons: New York, NY.

Albert, J.H. and Chibb, S. (1993): Bayesian Analysis of Binary and Polychotomous Response data, *Journal of American Statistical Association*, **88**, 669-679.

Albert, J.H. and Chibb, S. (2001): Sequential Ordinal modelling with Applications to Survival Data, *Biometrics*, **57**, 829-836.

Amemiya, T. (1981): Qualitative Response Models: a survey *Journal of Economic Literature*, **19**, 1483-1536

Besag, J., Green, P., Higdon, D., and Mengersen, K. (1995): Bayesian computation and stochastic systems, *Statistical Science*, **10**, 3-41.

Betensky, R.A. and Finkelstein D.M. (2002): Testing for Dependence between Failure Time and visit Compliance with Interval Censored data, *Biometrics* **58**, 58-63.

Cai, J. and Prentice, R.L. (1992): Covariance and Survivor function estimation usingcensored multivariate failure time data, *Biometrika*, **79**, 495-512.

Casella, G. and George E.I. (1996): An Introduction to Gibbs Sampling, *Technical Report '96'* University of Chicago, Illinois, USA.

Casella, G. and Berger, R. L. (1990): *Statistical Inference*. Duxbury Press: Belmont, California.

Cheng, S. C., Wei, L. J., and Ying, Z. (1995): Analysis of transformation models with censored data. *Biometrika*, **82**, 835-845

Chib S. and Greenberg E. (1995): Understanding the Metropolis Hastings algorithm, *Journal of American Statistical Association*, **49**, 327-335.

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 M.A. and Cuzick, J. (1985): Multivariate generalizations of Proportional hazards model (with discussion), *Journal of the Royal Statistical Society, Series*, **A 148**, 82-117.

Collett, D. (1994): *Modelling Survival data in Medical Research*. Chapman and Hall: London.

Colosimo, E.A., Chalita, L.V.A.S., and Deme'trio, C.G.B. (2000): Tests of Proportional Hazards and Proportional Odds Models for Grouped Survival Data, *Biometrics*, **56**, 1233-1240.

Congdon, P. (2001): *Bayesian Statistical Modelling*. Jonn Wiley and Sons: Baffins Lane, Chichester.

Cox, D.R. (1970): *Analysis of Binary Data*. Metheun: London.

Cox, D.R. (1972): Regresion Models and Life-tables, *J. R. Statistical Soc., Series* B, **47**, 447-452.

Cox, D.R. (1975): Partial likelihood, *Biometrika*, **62,** 269-276.

Cox, D.R., and Oakes, D. (1984): *Analysis of Survival Data*. Chapman and Hall: New York.

Crowder, M.J. (2001): *Classical Competing Risks*. Chapman and Hall: Boca Raton, Florida.

David and Moeschberger M.L. (1978): *The Theory of Competing Risks*. Charles Griffin: London.

Devroye L. (1986): *Non-Uniform Random Variate Generation*. Springer-Verlag: New York, NY.

Efromovich, S. (1999): *Nonparametric Curve Estimation, Methods, Theory and Applications*. Springer-Verlag: New York, NY.

Fahrmeir L. and Kaufmann H. (1985): Consistency and asymptotic Normality of the Maximum likelihood estimator in Generalized Linear Models, *The Annals of Statistics*, **13**, 342-368.

Fairmeir, L. and Tutz, G. (1994): *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer- Verlag: New York.

Faucett, C.L., Schenker, N. and J.M.G. Taylor (2002): Survival Analysis Using Auxillary Variables Via Multiple Imputation, with Application to AIDS Clinical Trial Data, *Biometrics*, **58**, 37-47.

Finkelstein, D.M. (1986): A Proportional Hazards Model for Interval Failure Time data, *Biometrics*, **42**, 845-854.

Finkelstein, D.M., Goggins, W.B. and Schoenfeld, D.A. (2002): Analysis of Failure Time with Dependent Interval Censoring, *Biometrics*, **58**, 298-304.

Fleming, T.R. and Lin, D.Y. (2000): Survival Analysis in Clinical Trials: Past Developments and Future Directions, *Biometrics*, **56**, 971-983.

Gelfand, A.E., Hills, S.E., Racine-Poon, A. and Smith, A.F.M. (1990): Illustration on Bayesian Inference in Normal Data models using Gibbs Sampling, *Journal of American Statistical Association*, **85**, No. 412, 972-985.

Gelfand, A.E., and Smith, A.F.M. (1990): Sampling-based approaches to calculating marginal densities, *Journal of American Statistical Association*, **85**, 398-409.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (1995): *Bayesian Data Analysis*. Chapman and Hall: London.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996): *Markov Chain Monte Carlo in Practice*. Chapman and Hall: London.

Goggins, W.B., and Finkelstein, D.M. (2000): A Proportional Hazards Model for Multivariate Interval-censored Failure Time Data, *Biometrics*, **56**, 940-943.

Guo, S.W., and Lin, D.Y. (1994): Regression Analysis of Multivariate Grouped Survival Data, *Biometrics*, **50**, 632-639.

Genest, C., and Rivest, L. (1993): Statistical Inference procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association,* **88,** 1034-1043.

Groenewald, P.C.N. and Mokgatlhe, L.L. (2002): Bayesian Analysis of Bivariate Interval Censored Survival Data, *Technical Report*, **313**, University of Free State, Bloemfontein.

Gumbel, E.J. (1960): Bivariate Exponential Distribution, *J. Am. Statist. Assoc.* **55**, 698-707.

Hammer, S. M., Katzenstein, D.A., Hughes, M.D., Gundacker H., Schooley, R.T., Haubrich, R. H., Henry, W.K., Lederman M.M., Phair J.P., Niu, M., Hirsch, M.S., and Merigan, T.C. (1996): A Trial Comparing Nucleoside Monotherapy With Combination Therapy in HIV-Infected Adults with CD4 Cell counts from 200 to 500 per cubic millimetre, *The New England Journal of Medicine,* **335**, 1081-1090.

Heckerman, D. (1996): A Tutorail on learning Bayesian Networks, *Technical Report*: Microsoft Research, Redmond, WA.

Hosmer, D.W. and Lemeshow, S. (1989): *Applied Logistic Regression.* John Wiley and Sons: New York, NY.

Hougaard, P. (2000): *Analysis of Multivariate Survival Data.* Springer-Verlag: New York, NY.

Hougaard, P. (1987b): A class of Multivariate failure time distributions, *Biometrika,* **73,** 671-678.

Izenman, A.J. (1991): Recent Developments in Nonparametric Density Estimation, *Journal of American Statistical Association*, **86**, 205-222.

Kalbfleisch, J.D. and Prentice, R.L. (2002): *The Statistical Analysis of Failure Time Data*. John Wiley and sons: Hoboken, New Jersey.

Katzenstein D.A., Scott M.H., Hughes, M.D., Gundacker, H., Jackson, J. B., Fiscus, S., Rasheed, S., Elbeik, T., Reichman, R., Japour, A., Merigan, T.C., Hirsch, M.S. (1996): The Relation of Virologic and Immunologic Markers to Clinical Outcomes after Nucleoside Therapy in HIV-Infected Adults with 200 to 500 CD4 Cells per cubic millimeter, *The New England Journal of Medicine*, **335**, 1091-1098.

Klein, J.P. and Moeschberger, M.L. (1997): *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag: New York, N.Y.

Lawless, J.F. (1982): *Statistical Models & Methods for Lifetime data*. John Wiley and Sons: Toronto, Canada.

Lee, P. (1997): *Bayesian Statistics: An Introduction*. John Wiley and Sons: New York, N.Y.

Lehmann, E.L. (1999): *Elements of Large –sample Theory*. Springer-Verlag: New York, NY.

Mallick B.K. and Gelfand A.E. (1994): Generalized Linear Models with Unknown Link Functions, *Biometrics*, **81**, 237-245.

Matlab (1992): *Reference Guide*, The Math-works: Natick, Massachusetts.

McCullagh, P. and Nelder J.A. (1983): *Generalized Linear Models.* Chapman and Hall: London.

McGilchrist, C.A, and Aisbett, C. W. (1991): Regression with Frailty in Survival Analysis, *Biometrics*, **47**, 461-466.

Nelder, J.A. and Welderburn, R.M.W. (1972): Generalized Linear Models, *Journal of the Royal Statistical Society*, **A135**, 370-384.

Oakes, D. (1982): Model for association in Bivariate Survival data, *J.R. statistical Soc.*, **B44**, 414-422.

Pan, W. and Chapell, R. (2002): Estimation in the Cox Proportional Hazards Model with left-Truncated and Interval Censored data, *Biometrics*, **58**, 64-70.

Prentice, R.L. and Cai, J. (1992): Covariance and Survivor function estimation using Censored Multivariate failure time data, *Biometrika*, **79**, 495-512.

Prentice, R.L. and Gloeckler, L.A. (1978): Regression analysis of grouped survival data with application to breast cancer data, *Biometrics*, **34**, 57-67.

Powers, D.A and Xie, Y. (2000): *Statistical Methods for Categorical Data Analysis.* Academic Press: San Diego, CA.

SAS (1990): *User's Guide.* SAS Institute Inc.: Cary, NC.

Satten, G.A. (1996): Rank-based inference in the proportional hazards model for interval censored data, *Biometrika*, **83**, 355-370.

Sinha, D., Chen, M., and Ghosh, S.K. (1999): Bayesian Analysis and Model Selection for Interval-Censored Survival Data, *Biometrics*, **55**, 585-590.

Sinha, D., Chen, M., and Ibrahim, J.G. (2001): *Bayesian Survival Analysis*. Springer-Verlag: New York, NY.

Tanner, M.A. and Wong (1987): The calculation of posterior distributions by data augmentation, *Journal of American Statistical Association*, **82**, 528-550.

Turnbull, B.W. (1976): The empirical distribution Function with arbitrarily Grouped, Censored and Truncated Data, *Journal of The Royal Statistical Society, Series B*, **38**, 290-295.

White, H. (1982): Maximum Likelihood Estimation of Mis-specified Models, *Econometrika*, **50**, 1-25.

Worku, Z.B. (1997): *A Bayesian Approach to child Survival Analysis*, Ph.D Dissertation, University of Orange Free State, Bloemfontein.

Zellner, A. and Rossi, P.E. (1984): Bayesian Analysis of Dichotomous Quantal Response Models, *Journal of Econometrics*, **25**, 365-393.

OPSOMMING

Tyd tot 'n gebeurtenis (faling) van eenhede op 'n regiment van longitudinale kliniese besoeke kan nie altyd presies bepaal word nie. Gewoonlik kan net 'n interval waarin 'n gebeurtenis plaasgevind het bepaal word. In so 'n geval word elke eenheid se faling beskryf deur 'n enkele interval wat lei tot gegroepeerde data oor die hele steekproef. Verder, as gevolg van nie-nakomings van besoeke deur sommige eenhede, kan die falings slegs beskryf word deur die eindpunte van die tydperk waarin die gebeurtenis plaasgevind het. Hierdie eindpunte mag verskeie intervalle insluit, en kry ons dus oorvleuende tydperke oor eenhede. Verder mag die gebeurtenis van belang by sekere eenhede nie plaasvind binne die voorafbepaalde tydperk van studie nie, en is dus gesensoreerd. Laastens, verskeie gebeurtenisse van belang kan ondersoek word op 'n enkele eenheid. Die resultaat is dan meervoudige falingstye wat uiteraard dan afhanklik is. Bogenoemde situasie word dan beskryf as interval-gesensoreerde oorlewingsdata met meervoudige falingstye.

Drie modelle vir die analise van interval-gesensoreerde oorlewingsdata met twee falingstye is toegepas op vier data stele. Vir verdelingsvrye metodes is Cox se gevaarfunksie met of 'n log-log transformasie of 'n logit transformasie op die basislyn voorwaardelike oorlewingswaarskynlikhede gebruik om die aanneemlikheidsfunksie af te lei. Die Onafhanklikheidsaanname model (IW) neem aan dat die leeftye inherent onafhanklik is en dat afhanklikheid slegs ingebring

word deur gemeenskaplike koveranderlikes. Die tweede model aanvaar nie onafhanklikheid nie, maar bereken die gesamentlike falingswaarskynlikhede deur die voorwaardelike waarskynlikheid vir die interval van een leeftyd gegee die ander leeftyd se interval, te bereken. Dit is die Voorwaardelike tweeveranderlike model (CB). Die Clayton en Farley-Morgenstern tweeveranderlike Copulas (CC) met ingeboude afhanklikheidsparameters is die derde model. Vir parametriese modelle is die IW en CC metodes toegepas op die data onder die aanname dat die randverdelings van die leeftye Weibull is.

Die traditionele klassieke beramingsmetode van Newton-Raphson is gebruik om die optimale beramers of modus van die afgeleide aanneemlikheidsfunksie te vind waar moontlik. Bayes metodes kombineer die data met a priori informasie. Vir elk van die twee transformasies is twee nie-inligtende prior verdelings algelei, wat se kombinasie met die aanneemlikheidsfunksie lei tot 'n posterior funksie. Om die volledige verdeling van 'n parameter te beraam uit nie-standaard posterior funksies is twee Markov Ketting Monte Carlo (MCMC) metodes gebruik. Die Gibbs steekproefnemingsmetode neem waarnemings uit die voorwaardelike verdeling van 'n parameter, gegee die ander parameters. Vir nie-standaard komplekse posterior funksies is die Metropolis-Hastings metode gebruik deur 'n vector van moontlike parameter waardes in 'n blok uit 'n surrogaat verdeling te trek.

Die analise van ACTG175 dui aan dat toename in vlakke van MIV RNS die afname van CD4 sell tellings voorafgaan. Daar is 'n sterk afhanklikheid tussen die twee falingstye, wat dus die gebruik van die onafhanklikheidsaanname model beperk. Die meer aanvaarbare modelle gebruik copulas en ook die voorwaardelike tweeveranderlike model. Dit is aangetoon dat die gebruik van ARV 'n pasiënt se leeftyd kan verleng, met kombinasie behandelings wat die beste resultate gee. 'n Sorgwekkende resultaat is dat die MIV virus 'n weerstand teen die middels ontwikkel. Dit blyk uit die nadelige effek wat vorige gebruik van ARV op 'n pasiënt het, deurdat 'n nuwe middel dan minder effek het. Laastens is dit belangrik dat 'n pasiënt op 'n vroë stadium behandeling begin aangesien pasiënte wat al tekens van VIGS wys negatief kan reageer op behandeling.

ABSTRACT

KEYWORDS: *AIDS data, Bayes, Copula, dependence, hazard rate, Interval-censoring, Multiple-failure, Metropolis-Hastings Algorithm, Visiting-compliance, Weibull.*

The measure of time to event (failure) for units on longitudinal clinical visits cannot always be ascertained exactly. Instead only time intervals within which the event occurred may be recorded. That being the case, each unit's failure will be described by a single interval resulting in grouped interval data over the sample. Yet, due to non-compliance to visits by some units, failure will be described by endpoints within which the event has occurred. These endpoints may encompass several intervals, hence overlapping intervals across units. Furthermore, some units may not realize the event of interest within the preset duration of study, hence are censored. Finally, several events of interest can be investigated on a single unit resulting in several failure times that inevitably are dependent. All these prescribe an interval-censored survival data with multiple-failure times.

Three models of analysing interval-censored survival data with two failure times were applied to four sets of data. For the distribution free methods, Cox's hazard with either a log-log transform or logit transform on the baseline conditional survival probabilities was used to derive the likelihood. The Independence assumption model (IW) work under the assumption that the lifetimes are independent and any dependence exists through the use of common covariates.

The second model that do not necessarily assume independence, computes the joint failure probabilities for two lifetimes by Bayes' rule of conditioning on the interval of failure for one lifetime, hence Conditional Bivariate model (CB). The use of Clayton and Farley-Morgenstern bivariate Copulas (CC) with inbuilt dependence parameter was the other model. For parametric models the IW and CC methods were applied to the data sets on the assumption that the marginal distribution of the lifetimes is Weibull.

The traditional classical estimation method of Newton-Raphson was used to find optimum parameter estimates and their variances stabilized using a sandwich estimator, where possible. Bayesian methods combine the data with prior information. Thus for either transforms, two proper priors were derived, of which their combination with the likelihood resulted in a posterior function. To estimate the entire distribution of a parameter from non-standard posterior functions, two Markov Chain Monte Carlo (MCMC) methods were used. The Gibbs Sampler method samples in turn observations from the conditional distribution of a parameter in question, while holding other parameters constant. For intractably complex posterior functions, the Metropolis-Hastings method of sampling vectors of parameter values in blocks from a Multivariate Normal proposal density was used.

The analysis of ACTG175 data revealed that increase in levels of HIV RNA precede decline in CD4 cell counts. There is a strong dependence between the two failure times, hence restricting the use of the independence model. The most preferred models are using copulas and the conditional bivariate model. It was shown that ARV's actually improves a patient's lifetime at varying rates, with combination treatment performing better. The worrying issue is the resistance that HIV virus develops against the drugs. This is evidenced by the adverse effect the previous use of ARV's has on patients, in that a new drug used on them has less effect. Finally it is important that patients start therapy at early stages since patients displaying signs of AIDS at entry respond negatively to drugs.