AN ANALYSIS FOR GROUPED SURVIVAL DATA WITH SPLIT PLOT VARIANCE COMPONENT MODELS

By

DALAL ABDULRAZZAK ABDULLATIF

Bachelor of Science
University of Baghdad
Baghdad, Iraq
1975

Master of Science
Mississippi State University
Mississippi State, Mississippi
1980

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1984

AN ANALYSIS FOR GROUPED SURVIVAL DATA WITH SPLIT PLOT VARIANCE COMPONENT MODELS

Thesis Approved:

William H. Stewart
Thesis Adviser

Donald Holbert

P. Larry Claypool

H. O. Burchard

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my thesis adviser, Dr. William H. Stewart, for his valuable advice, guidance and encouragement that enabled me to complete my thesis. Appreciation is also expressed to my committee chairman, Dr. Larry Claypool, and my committee members, Dr. Donald Holbert and Dr. Hermann Burchard, for their helpful suggestions and assistance.

Special gratitude is given to my late mother, my father, my brothers and sisters, and my husband, Jaffar, and my lovely daughter, Reem, for their love and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Several experimental situations give rise to analyzing time to response on observational units (survival data) using split plot in time models. The general structure of such experiments is that the observation of the time of occurrence of an event (called a death, failure, or response) is of interest. The observational units are grouped into whole units and the treatments are randomized to whole units. If time to the occurrence of an event T is a continuous random variable then whole units would be considered as subsamples. If time response was grouped into intervals in the above setting, then the sufficient statistics in this case would be the counts of observed occurrences of an event (number of deaths, failure) within intervals. The experiment can then be viewed as a split plot over time where time intervals (periods) are subunits and whole units would be the same as in continuous time setting, and the response variable is some function of the counts. For the split plot over time model we are interested in estimating survival curves rather than means for the usual structure of split plot model.

In this chapter we outline the type of data and the process in data collection that defines such experimental situations using a fish experiment, where studying the effect of treatment combinations on the survival times of fish in aquarium water was desired. In an experiment presented by Pierce, Steware and Kopecky (1978), fish were subjected to three levels of zinc concentration in aquarium water, and approximate times-to-death were observed. It was desired to study the effect of either one or two week's acclimation in the test aquaria before introduction of the zinc. There were initially two tanks for each of the treatment combinations. The experiment was a 2×3 factorial for the treatment combinations structure. The 2×3 treatment combinations were assigned to tanks in a completely randomized design. From this point onwards we use CRD to designate this design. The experiment was carried on for 10 days and mortality was observed on a daily basis. Three hundred fish were randomized to 12 tanks, 25 fish to each tank. The 2×3 treatment combinations were assigned so that 2 tanks received each treatment. Table I gives the daily mortalities, where the numbers are the observed numbers of deaths in each day, for 10 days.

Days 1 and 2, and days 8, 9, and 10, were each combined giving $K = 7$ class intervals. Now let us denote the points defining the time intervals by

$$0 = t_0 < t_1 < t_2 < t_3 < t_4 < t_5 < t_6 < t_7.$$

The number of failures or deaths in each day or combined days would be the number of failures or deaths in time intervals $(t_{k-1}, t_k]$ for $k = 1,2,\dots,7$. Also, define

n_{ijk}: number assigned (at risk) to trt i, time interval k and tank j,

s_{ijk}: number of survivors during interval k on trt i and tank j,

r_{ijk}: number of failures on trt i during interval k for tank j,

p_{ijk}: conditional probability that a unit on trt i fails in time interval k given that it survived k-1 time intervals for a given tank j, and

q_{ijk}: conditional probability that a unit on trt i survives time interval k given that it survived k-1 time intervals for a given tank j, where $q_{ijk} = 1 - p_{ijk}$.

Therefore, Table I of the observed number of deaths or failures will represent the table of the values of r_{ijk}, after combining days.

TABLE I

OBSERVED NUMBER OF DEATHS

Acclimation Time:	One Week						Two Weeks					
Zinc Concentration:	Lo		Med		Hi		Lo		Med		Hi	
Tank:	1	2	1	2	1	2	1	2	1	2	1	2
Day/Mortality												
1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	2	3	0	1	1	0	0	1	0	3	0
3	5	7	7	10	12	10	9	4	12	9	12	12
4	7	4	9	7	7	8	4	4	5	3	3	7
5	1	2	0	5	4	3	0	0	3	2	2	2
6	0	0	0	1	0	1	0	0	0	0	1	0
7	0	0	0	0	1	1	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0

Since the 50 fish were randomly assigned to each treatment combination with 2 tanks for each treatment then 25 fish were assigned to each tank. Thus the number at risk for the first time interval is 25 fish and the size of this risk set, n_{ijk}, decreases as time advances. For the no censoring case, the number at risk for time interval k would be the number at risk for time interval k-1 minus the number of deaths for time interval k-1. Therefore, Table II represents risk sets (values of n_{ijk}).

Now let us define $\hat{q}_{ijk} = s_{ijk}/n_{ijk}$, where $s_{ijk} = n_{ijk} - r_{ijk}$ assuming no censoring. Table III represents the values of $\hat{q}_{ijk}$. If $s_{ijk} = n_{ijk}$ then use $s_{ijk} - .5$, if $s_{ijk} = 0$ then use .5.

TABLE II

RISK SET TABLE

Acclimation Time:	One Week						Two Weeks					
Zinc Concentration:	Lo		Med		Hi		Lo		Med		Hi	
Tank:	1	2	1	2	1	2	1	2	1	2	1	2
Interval/n_{ijk}												
1	25	25	25	25	25	25	25	25	25	25	25	25
2	24	23	22	22	24	24	25	25	24	25	22	25
3	19	16	15	15	12	14	16	21	12	16	10	13
4	12	12	6	8	5	6	12	17	7	13	7	6
5	11	10	6	3	1	3	12	17	4	11	5	4
6	11	10	6	2	1	2	12	17	4	11	4	4
7	11	10	6	2	0	1	12	17	4	11	4	4

TABLE III

OBSERVED VALUES OF q_{ijk}

Acclimation Time:	One Week						Two Weeks					
Zinc Concentration:	Lo		Med		Hi		Lo		Med		Hi	
Tank:	1	2	1	2	1	2	1	2	1	2	1	2
Interval/$\hat{q}_{ijk}$												
1	0.960	0.920	0.880	0.980	0.960	0.960	0.980	0.980	0.960	0.980	0.880	0.980
2	0.792	0.696	0.682	0.600	0.500	0.583	0.640	0.840	0.500	0.640	0.455	0.520
3	0.632	0.950	0.400	0.533	0.417	0.420	0.750	0.810	0.583	0.813	0.700	0.962
4	0.917	0.833	0.917	0.375	0.200	0.500	0.958	0.971	0.571	0.816	0.714	0.667
5	0.955	0.950	0.917	0.667	0.500	0.667	0.958	0.971	0.875	0.955	0.800	0.875
6	0.955	0.950	0.917	0.500	0.500	0.500	0.958	0.971	0.875	0.955	0.875	0.875
7	0.955	0.950	0.917	0.500	0.500	0.500	0.958	0.971	0.875	0.955	0.875	0.875

Assume that tank effects increase or decrease the survivals, i.e. assume that there is tank variability involved, since treatment combinations were applied to main units (tanks). Also assume that failure time T is a discrete random variable since time responses were grouped into intervals 1,2,...,K, where K = 7 for the experiment presented. The response for the discrete setting would be some function of the number of deaths or the number of survivors. This will give us a split plot in time where subplot units are time intervals. Failure time variability will arise from the fact that 25 fish were randomly assigned to each tank. Assuming that conditional on being in the same tank survival times of different fish are independent, then the model to be considered is:

$$\text{Response} = \mu + \alpha_i + \varepsilon_{ij} + \beta_k + (\alpha\beta)_{ik} + \delta_{ijk},$$

$$i = 1,\ldots,I, \quad j = 1,\ldots,J, \quad \text{and} \quad k = 1,\ldots,K \qquad \ldots \quad (1.1)$$

where, μ is an overall mean,

α_i is treatment combination i effect,

ε_{ij} is main unit variability (tank variability) with

$$E(\varepsilon_{ij}) = 0, \quad E(\varepsilon_{ij}\varepsilon_{ij'}) = \begin{cases} \sigma_\varepsilon^2 & \text{for } j = j' \\ 0 & \text{for } j \neq j', \end{cases}$$

β_k is the subplot treatment or time interval effect,

$(\alpha\beta)_{ik}$ is the interaction between treatment and time interval,

δ_{ijk} is the variability due to different fish in each tank with

$$E(\delta_{ijk}) = 0, \quad E(\delta_{ijk}\delta_{ijk'}) = \begin{cases} \sigma_{\delta_{ijk}}^2 & \text{for } k = k' \\ 0 & \text{for } k \neq k', \end{cases}$$

and $E(\varepsilon_{ij}\delta_{ijk}) = 0$.

The response of the above model will depend on the model assumed for the hazard function for time interval k and trt i. The hazard function $\lambda_i(t_k)$ is the conditional probability of failing in an interval given surviving until that interval. The choice for the response is Response = $f(\hat{q}_{ijk})$. Two possible choices for this function that will be considered are:

$f(\hat{q}_{ijk}) = \log(-\log \hat{q}_{ijk})$, and

$f(\hat{q}_{ijk}) = \log(\hat{q}_{ijk})$.

These responses are derived from continuous random variable models as will be seen later. From this point onwards we use log(x) to denote $\log_e(x)$.

Individuals at risk during time interval k may fail, be censored, or survive to the start of the following time period. Assuming that there is no censoring, the observed number at risk for time interval k on a given trt i and a given tank j is n_{ijk}, and the number of individuals failing is r_{ijk}. Define $n_{ij(k+1)} = n_{ijk} - r_{ijk}$, which is denoted by s_{ijk} (the number of individuals surviving interval k). Thus, individuals surviving interval k will be individuals at risk for the next time interval, i.e., $s_{ijk} = n_{ij(k+1)}$. For a given trt i and a given tank j, and for K time intervals, number of deaths or failures $r_{ij1}, r_{ij2}, \ldots, r_{ijK}$ in time intervals $(t_0,t_1], (t_1,t_2], \ldots, (t_{K-1},t_K]$ with $t_0 = 0$ among n_{ij1} starters, follow a multinomial distribution with probability function:

$$\Pr(r_{ij1}, r_{ij2}, \ldots, r_{ijK} | \varepsilon_{ij}) = \frac{n_{ij1}!}{r_{ij1}! r_{ij2}! \ldots r_{ij(K+1)}!} \prod_{k=1}^{K+1} \pi_{ijk}^{r_{ijk}}, \quad \ldots \quad (1.2)$$

where

$$r_{ij1} + r_{ij2} + \dots + r_{ij(K+1)} = n_{ij1} \text{ starters, and}$$

$$\pi_{ij1} + \pi_{ij2} + \dots + \pi_{ij(K+1)} = 1.$$

Now define

$P_{ijk} = \prod_{\ell=1}^{k} q_{ij\ell}$ is the probability an individual on trt i and in tank j survives beyond interval k,

$\pi_{ijk} = P_{ij(k-1)} - P_{ijk}$ is the probability an individual fails in interval k for a given tank j on trt i,

q_{ijk} is the conditional probability an individual on trt i and in in tank j survives beyond interval k given that it survives beyond interval k-1, where

$q_{ijk} = P_{ijk}/P_{ij(k-1)}$,

$p_{ijk} = 1 - q_{ijk}$ is the conditional probability an individual on trt i and in tank j fails in interval k given that it survives beyond interval k-1, and

$r_{ij(K+1)} = s_{ijK}$ is the number of individuals surviving at end of study.

Therefore, we have

$$\pi_{ijk} = P_{ij(k-1)} - P_{ijk} = q_{ij1}q_{ij2}\cdots q_{ij(k-1)}p_{ijk}, \qquad \dots \quad (1.3)$$

for k = 1,2,...,K. The likelihood function for the multinomial distribution is given by

$$\Pr(r_{ij1}, r_{ij2}, \dots, r_{ijK} | \varepsilon_{ij}) \propto \prod_{k=1}^{K+1} \pi_{ijk}^{r_{ijk}}$$

$$\propto \{ \prod_{k=1}^{K} (q_{ij1}\cdots q_{ij(k-1)}p_{ijk})^{r_{ijk}} \} \times$$

$$\{ (q_{ij1}\cdots q_{ij(K-1)}q_{ijK})^{r_{ij(K+1)}} \}$$

$$\propto p_{ij1}^{r_{ij1}} p_{ij2}^{r_{ij2}} p_{ij3}^{r_{ij3}} \cdots p_{ijK}^{r_{ijK}}$$

$$q_{ij1}^{r_{ij2}+r_{ij3}+\cdots+r_{ij(K+1)}}$$

$$q_{ij2}^{r_{ij3}+r_{ij4}+\cdots+r_{ij(K+1)}} \cdots q_{ijK}^{r_{ij(K+1)}} .$$

Recall that $r_{ij1} + r_{ij2} + \cdots + r_{ij(K+1)} = n_{ij1}$, and $n_{ijk} = n_{ij1} - r_{ij1} - r_{ij2} - \cdots - r_{ij(k-1)}$ for $k = 1,2,\ldots,K+1$. Therefore the likelihood is proportional to

$$\Pr(r_{ij1}, r_{ij2}, \ldots, r_{ijK} | \varepsilon_{ij}) \propto p_{ij1}^{r_{ij1}} q_{ij1}^{n_{ij1}-r_{ij1}} \; p_{ij2}^{r_{ij2}} q_{ij2}^{n_{ij1}-r_{ij1}-r_{ij2}} \cdots$$

$$p_{ijK}^{r_{ijK}} q_{ijK}^{n_{ij1}-r_{ij1}-r_{ij2}-\cdots-r_{ijK}}$$

$$\propto p_{ij1}^{r_{ij1}} q_{ij1}^{n_{ij1}-r_{ij1}} \; p_{ij2}^{r_{ij2}} q_{ij2}^{n_{ij2}-r_{ij2}} \cdots$$

$$p_{ijK}^{r_{ijK}} q_{ijK}^{n_{ijK}-r_{ijK}}$$

$$\propto \prod_{k=1}^{K} p_{ijk}^{r_{ijk}} q_{ijk}^{n_{ijk}-r_{ijk}}$$

$$\propto \prod_{k=1}^{K} q_{ijk}^{s_{ijk}} (1 - q_{ijk})^{n_{ijk}-s_{ijk}} \quad \ldots \quad (1.4)$$

Therefore, conditioning on n_{ijk}, the number of survivors s_{ijk} in time interval k on trt i and a given tank j is distributed as a binomial random variable with parameters n_{ijk} and q_{ijk}. Furthermore, the covariance between s_{ijk} and $s_{ijk'}$ is zero. Also, the mean and variance of $\hat{q}_{ijk}$ given that n_{ijk} is fixed by its observed number and for a given tank are given by

$E(\hat{q}_{ijk} | n_{ijk}, \varepsilon_{ij}) = q_{ijk}$, and

$\mathrm{Var}(\hat{q}_{ijk} | n_{ijk}, \varepsilon_{ij}) = p_{ijk} q_{ijk} / n_{ijk}$, respectively.

Now, for $k < k'$, assuming that $n_{ijk} > 0$ we have

$$
\begin{aligned}
E((\hat{q}_{ijk} - q_{ijk}) | \hat{q}_{ijk'}, \varepsilon_{ij}, n_{ijk}) &= E_{s_{ijk}}((\frac{s_{ijk}}{n_{ijk}} - q_{ijk}) | \hat{q}_{ijk'}, \varepsilon_{ij}, n_{ijk}) \\
&= E_{n_{ijk}}(E_{s_{ijk} | n_{ijk}}((\frac{s_{ijk}}{n_{ijk}} - q_{ijk}) | \hat{q}_{ijk'}, \\
&\quad \varepsilon_{ij}, n_{ijk})) \\
&= E_{n_{ijk}}(0 | \hat{q}_{ijk'}, \varepsilon_{ij}, n_{ijk}) = 0.
\end{aligned}
$$

Hence, for $k < k'$ -

$$
\begin{aligned}
\mathrm{cov}(\hat{q}_{ijk}, \hat{q}_{ijk'} | \varepsilon_{ij}, n_{ijk}) &= E_{\hat{q}_{ijk'}}((\hat{q}_{ijk'} - q_{ijk'}) E(\hat{q}_{ijk} - q_{ijk} | \hat{q}_{ijk'}, \\
&\quad \varepsilon_{ij}, n_{ijk})) \\
&= 0.
\end{aligned}
$$

Using (1.4), for a given tank j and a fixed risk set (n_{ijk}), we have

$s_{ijk} | \varepsilon_{ij}, n_{ijk} \sim \text{Binomial}\ (n_{ijk}, q_{ijk})$, and

$\mathrm{cov}(s_{ijk}, s_{ijk'} | \varepsilon_{ij}, n_{ijk}, n_{ijk'}) = 0.$

Thus, for a large sample size the asymptotic distribution is given by

$$s_{ijk} | \varepsilon_{ij}, n_{ijk} \overset{\bullet}{\underset{\bullet}{\sim}} \text{Normal } (n_{ijk}q_{ijk}, n_{ijk}q_{ijk}(1-q_{ijk})), \quad \text{and}$$

$$\hat{q}_{ijk} = \frac{s_{ijk}}{n_{ijk}} | \varepsilon_{ij}, n_{ijk} \overset{\bullet}{\underset{\bullet}{\sim}} \text{Normal } (q_{ijk}, q_{ijk}(1-q_{ijk})/n_{ijk}).$$

Therefore, equal variance structure of $\hat{q}_{ijk}$'s would be inappropriate since these variances depend on q_{ijk}'s which may vary over time, and the fact that the risk sets decrease over time (we begin with 25 fish at risk for the first time interval and we might end up with, say, only 4 fish at risk for the last time interval). Hence, we are going to look at a way to estimate the survivor functions for different treatment combinations using split plot model with unequal subplot variances.

Another experimental situation occurs in studying the effect of some treatment combinations on patients in several hospitals. The patients per hospital will be selected randomly. Time-to-the introduction of a result will be of interest. The treatment combinations will be randomly applied to each hospital. Thus all selected patients in the same hospital will receive the same treatment. If time T is a discrete random variable, then time response will be grouped into intervals and the response variable will be the number of patients on which a result occurs (number of deaths). This will give a split plot in time where subplot units are time intervals. Main unit variability arises from the fact that we randomly apply treatments to hospitals. Subunit variability arises from the fact that there is more than one patient to be selected from each hospital. In such a case estimating survival curves is of interest.

Another experimental situation that is related to our type of study is a seed germination trial, where time-to-germinate on observational units (carrot seeds) is considered. Seeds are randomly assigned to dishes (whole units) in a CRD such that seeds that are applied to a dish are of the same kind. Treatments (stored seeds against control seeds) are randomly applied to dishes. If time to germinate T is a discrete random variable, then time response is grouped into intervals. The response variable for this setting is the number of seeds germinated for each time interval. This will give us a split plot in time where subplot units are time intervals. The variability due to applying the same treatment to more than one dish is the main unit variability and failure time variability or subunit variability arises from randomly assigning seeds to dishes. Estimates of the probabilities that seeds will germinate after a specified time are of interest. Some of the seeds will not germinate ever therefore the probability functions will not be exact survival curves as is the situation with the other two examples. However, estimating these probability functions is similar to the idea of estimating survival curves.

CHAPTER II

LITERATURE REVIEW

2.1 Survival Analysis Literature

In life testing and medical follow up, the observation of the time of occurrence of the event (called death, failure, or response) is of interest. Sometimes these occurrences may be prevented for some of the items of the sample by the occurrence of some other event (called loss or censoring). Kaplan and Meier (1958) assumed that the life time is independent of the potential loss time, and they provided, for random samples of size N, the product-limit (PL) estimate that can be defined as follows. List and label the N observed lifetimes (whether to death or loss) in order so that one has $0 \leq t'_1 \leq t'_2 \leq \ldots \leq t'_N$. Then $\hat{P}(t) = \prod_r [(N-r)/(N-r+1)]$, where r assumes those values for which $t'_r \leq t$, and for which t'_r measures the time to death. This is the distribution-free estimator which maximizes the likelihood function.

Cox (1972) considered the analysis of censored failure times. He suggested a regression model for the failure time T of an individual when values of one or more explanatory variables were available. For T continuous, the hazard function is given by

$$\lambda(t,z) = \lambda_0(t) \exp(\beta' z),$$

which is known as the proportional hazard function. It is also known as the multiplicative form of the hazard function with β being the vector of the unknown parameters, and $\lambda_0(t)$ is the underlying hazard function when $z \equiv 0$. For T discrete, the logistic model was suggested. A conditional likelihood and maximum likelihood estimates were obtained. However, Cox (1972) proportional hazard regression model does not handle grouped survival data or large data sets with many ties (many individuals failed at the same time).

Kalbfleisch and Prentice (1973) obtained a marginal likelihood for the regression parameters by restricting the class of models presented by Cox (1972) to those that possessed a strictly monotone survivor function or, equivalently, to those for which the hazard function $\lambda_0(t)$ was not identically zero over an open interval. The invariance of this restricted class under the group of monotone increasing transformations on T was exploited to derive a marginal likelihood function for β. If no ties occur their results and the results of Cox (1972) are the same with a simple justification. But if ties occur in the data the results obtained by Kalbfleisch and Prentice (1973) are different from those suggested by Cox (1972).

Prentice and Glocker (1978) considered the grouped data version of the proportional hazards model (Kalbfleisch and Prentice (1973)) in an attempt to develop computationally feasible estimators of the relative risk function and the corresponding survivor function in the presence of many tied failure times. Asymptotic likelihood results were given for both the estimation of the regression coefficients and the survivor functions.

Regression models of the proportional hazard were used for analyzing some data arising from a clinical trial in medicine. Kay (1977) considered applying the regression models of the proportional hazards to the analysis of censored survival data. Many forms of the proportional hazard model and a search for a model fitting were carried out. When the number of independent variables was large, selecting those independent variables to be included in the model was achieved by a forward stepwise procedure.

Usual regression techniques were widely used to analyze survival data. Such work was done by Krane (1963) and Pierce, Stewart and Kopecky (1978). Krane (1963) introduced a type of statistical analysis of survival data applicable under the conditions that usually the available data were grouped, most commonly in yearly intervals, and more serious was the fact that the data was often "censored". Assuming that there exists a survivor-function, S(t), such function is given by exp[-y(t)], where y(t) is the time integral of the failure rate which was approximated by a polynomial. For large samples it was found that the covariance structure for y(t) may be obtained from the multinomial distribution when the data was grouped. Thus the method of weighted least squares may be employed to fit y(t). "Censored" data in no way vitiate the method.

Pierce, Stewart and Kopecky (1978) provided a method based on regression model for the proportional hazards to obtain, by making an approximation, a maximum likelihood function involving only the regression parameters. The authors presented an example for analyzing toxicology data.

Most of the preceeding literature seems to pay most attention to the multiplicative form of the hazard function, and less work has been done with the additive form. Elandt-Johnson (1980) used the additive model for the hazard function to demonstrate techniques in deriving posterior distributions by assuming a normal prior distribution for the variables influencing the hazard function. The hazard rate function in the additive form is given by

$$\lambda(t,z_0) = \lambda(t) + \sum_{i=1}^{k} h_i(t) z_{0i} \quad (>0),$$

where $\lambda(t)$ (>0) is the, so-called, underlying hazard rate, $h_i(t)$'s are functions of t alone, and z_{0i}'s are the covariates influencing the survival.

An extension of the proportional hazard models was suggested by Aranda-Ordaz (1983) where a family of transformations for probabilities was considered for the analysis of grouped survival data. Additive and multiplicative models for the hazard function were compared.

Similar work-was done by Tibshirani and Ciampi (1983) where a family of proportional and additive hazards models for the analysis of grouped survival data was developed. They generalized the work of Aranda-Ordaz (1983) by allowing time trends to enter the hazards. This generalization proved to be useful in the case of crossing hazards.

From the preceeding literature, it seems that most work has been done for continuous time random variables. Our grouped time models that are used for inference are chosen to relate to these well known continuous time models. We have generalized the Cox (1972) model to include main unit variability to be able to get the split plot in time model as we will see in Chapter III.

A general approach to the analysis of categorical data was provided by Grizzle, Starmer and Koch (1969) by assuming that there were $n_{i\cdot}$, $i=1,2,\ldots,s$, samples from a multinomial distributions each having r categories of response. They defined any r-1 functions of the unknown true cell probabilities $\{\pi_{ij}: i=1,2,\ldots,s;\ j=1,2,\ldots,r,$ where $\sum_{j=1}^{r} \pi_{ij} = 1\}$ that have up to the second order derivatives with respect to π_{ij}. A noniterative weighted least squares procedure was described to fit these functions to a linear model, along with testing hypotheses about the parameters and testing the goodness-of-fit of the model.

This general procedure for analyzing categorical data can be applied to survival data in the case that the variance-covariance matrix is a diagonal matrix with the binomial variances on the diagonal.

For our grouped time model, we use a similar approach to the approach used by Grizzle, Starmer and Koch (1969) since we have the same general structure. The only difference is that we have unequal binomial variances; further we add an extra term in the variance-covariance matrix which is the main-unit variability.

2.2 Split Plot and Variance Component Literature

Our model for survival analysis is based on using a split plot in time model, and therefore we need to consider the related literature. What we need in the variance component analysis is a method for split-plot models with uneuqal sub-plot variances. We must mention here that we could not find any work in the literature that has been done for this particular study. However, a list and a presentation of the literature that has been done in both split plot model and variance component areas separately and combined will be considered. Some of the listed literature might not be of direct relation to our study, and some are related in the sense that they gave us an idea on the approach that we have used for variance component estimation.

Early work on variance component estimation has been done by many authors. Crump (1946) discussed and pointed out the hypotheses appropriate to the two uses of the analysis of variance as to obtain tests of significance of treatment effects and provided estimates of variance components. The estimation of variance components was accomplished by equating the mean squares in the standard analysis of variance to their expectations and then solving for the unknown variances. This method of estimating variance components dealt with the one-way classification, nested classification, and factorial classifications having equal subclass numbers. But often the subclasses are of unequal size. Therefore a need for some other methods of estimation was raised.

Henderson (1953) developed three methods for estimating variance components in the non-orthogonal case. The three methods can be described as follows. Method (1): Compute sums of squares as in the standard analysis of variance of corresponding orthogonal data. Equate these sums to their expectations and solve for the unknown variances. This method leads to biased estimates if certain elements of the model are fixed or if some are correlated. Method (2): Obtain least squares estimates of fixed effects, "correct" the data according to these estimates, then use the corrected data and proceed as in Method (1). This method gives estimates which are free of the first of these biases, but not of the second. Method (3): Compute mean squares by least squares analysis of non-orthogonal data. Equate these mean squares to their expectations and solve for the unknown variances. This method yields unbiased estimates, but the computations involved may be prohibitive.

Henderson's (1953) methods were discussed and reformulated in matrix theory by Searle (1968). Also a fourth method for variance component estimation was introduced.

A common assumption in split plot experiments is that the error variances for subplot treatments are the same. Curnow (1957) provided tests of significance for the departure from equality of the variances for different subplot treatments. Also, an estimate of the ratio of a pair of such variances was provided in this paper.

Rao (1970) considered the problem of estimating the different variances for the linear model $y = X\beta + e$, where

$$D(e) = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_n^2 \end{bmatrix},$$

by introducing a new principle called Minimum Norm Quadratic Unbiased Estimation (MINQUE). This principle of estimation can be summarized as follows. Let $\Sigma p_i \sigma_i^2$ be a linear function of the variances to be estimated, where all σ_i^2 may not be distinct. The quadratic form $Y'AY$ is said to be a MINQUE of $\Sigma p_i \sigma_i^2$ if the matrix $A = (a_{ij})$ is chosen such that $||A||$, the Euclidean norm of A, which is the same as the square root of trace A^2, is minimum subject to the conditions

$$AX = 0, \quad \text{and} \quad \sum_{i=1}^{n} a_{ii}\sigma_i^2 = \sum_{i=1}^{n} p_i\sigma_i^2.$$

Hartley and Jayatillake (1973) pointed out that Rao's (1970) MINQUE estimators suffer from three defects, namely (1) The MINQUE estimator $\hat{\sigma}_i^2$, although unbiased, may be negative, (2) The residuals $y - X\hat{\beta}$ employed for

σ_i^2 estimation were based on $\hat{\beta}$ of β which are known not to be BLUE in case the true σ_i^2 differ, and (3) Unless the matrix X has a standard structure MINQUE estimation requires the inversion of an $n \times n$ matrix to be specially computed from the elements of the observed X in each problem. With normal assumption of the residuals added to the model presented by Rao (1970), Hartley and Jayatillake (1973) examined the method of maximum likelihood under the normality assumption for the estimation problem of β and σ_i^2 which are free from the three disadvantages that MINQUE estimators have. Therefore the elements of β and σ_i^2 are estimated by maximum likelihood under the assumption of a lower bound for the σ_i^2 of the form $0 < \delta_i^2 \leq \sigma_i^2$ so that the likelihood is finite in the restricted parameter space. The authors also considered a second problem in which the Y vector splits into subvectors Y_j's such that all elements of Y_j have equal variances.

For the balanced two-way layout split plot design Li and Klotz (1978) compared maximum likelihood estimators and restricted maximum likelihood estimators with minimum variance unbiased estimators of variance components. Performance was compared in terms of mean squared error for the three estimators.

For a general mixed-effects model Brown (1978) viewed the problem of estimating variance components in the context of linear model theory. The approach was to estimate the unknown vector of parameters β by some vector b and thus obtain a vector of residuals $e = Y - Xb$. A vector of the squares and cross products of the residuals was then obtained, the expectation of which was a known linear transformation of the variance components.

For categorical data, Manton, Woodbury and Stallard (1981) presented maximum likelihood procedures for the estimation of the model parameters based on the assumption that the distribution function for each cell death count is the negative binomial probability function. This assumption is equivalent to assuming a mixture of poisson processes with the differential risk levels among individuals within cells being a two parameter gamma distribution.

CHAPTER III

MODELS FOR ANALYSIS

3.1 Grouped Time, Multiplicative and Additive Hazard Conditional on Main Unit with Normal Main Unit Error

As presented in Chapter I, the structure for the design that will be considered is that we have J main units per treatment combination according to a CRD, n_{ij} observational units in each main unit and time to response on each observational unit is measured. Time to response is grouped into intervals where the points defining the time intervals are denoted by $0 = t_0 < t_1 < t_2 < \ldots < t_K$. The number of failures or deaths in time interval k, $k = 1,\ldots,K$ is the number of failures or deaths in time interval $(t_{k-1}, t_k]$.

Define

n_{ij}: number of individuals assigned to main unit j of trt i,

r_{ijk}: number of individuals failed on trt i, main unit j during time interval k,

s_{ijk}: number of individuals survived interval k for trt i and main unit j, and

n_{ijk}: number of individuals at risk for trt i, main unit j and time interval k.

For the no censoring case we have

$n_{ij} = n_{ij1}$, and $n_{ijk} = s_{ij(k-1)}$ for $k > 1$.

For the censoring case we have to define c_{ijk} as the number censored during the <u>kth</u> interval, then

$n_{ij1} = n_{ij} - c_{ij1}$, and

$n_{ijk} = s_{ij(k-1)} - c_{ijk} = n_{ij(k-1)} - r_{ij(k-1)} - c_{ijk}$ for $k > 1$.

Also, define

p_{ijk}: conditional probability that an individual on trt i and main unit j fails in interval k given that it survived k-1 time intervals, and

$q_{ijk} = 1 - p_{ijk}$: conditional probability of surviving interval k given survival of k-1 time intervals for an individual on trt i and main unit j.

Now, let $F_{ij}(t)$ be the cumulative distribution function for the continuous response time random variable T for a given main unit j. Define $S_{ij}(t) = 1 - F_{ij}(t)$ to be the survival function for trt i and main unit j.

By definition, p_{ijk} can be written as

$$p_{ijk} = \frac{\Pr\left(\begin{array}{l}\text{failing in time interval k for an individual on trt i}\\ \text{and main unit j}\end{array}\right)}{\Pr\left(\begin{array}{l}\text{surviving (k-1) time intervals for an individual on}\\ \text{trt i and main unit j}\end{array}\right)}$$

$$= \frac{F_{ij}(t_k) - F_{ij}(t_{k-1})}{1 - F_{ij}(t_{k-1})}$$

$$= \frac{[1 - F_{ij}(t_{k-1})] - [1 - F_{ij}(t_k)]}{1 - F_{ij}(t_{k-1})}$$

$$= 1 - \frac{1 - F_{ij}(t_k)}{1 - F_{ij}(t_{k-1})}, \text{ and}$$

$$q_{ijk} = 1 - p_{ijk} = \frac{1 - F_{ij}(t_k)}{1 - F_{ij}(t_{k-1})} = \frac{S_{ij}(t_k)}{S_{ij}(t_{k-1})} . \qquad \ldots \quad (3.1)$$

Define the hazard function $\lambda(t)$ as the limiting conditional probability of failing in an interval given surviving until that interval as the interval shrinks, to be

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)},$$

where $f(t)$ and $S(t)$ are the density function and the survival function, respectively, for the continuous response time random variable T. Cox (1972) suggested a regression model for the failure time T of an individual when values of one or more explanatory variables are available. For T continuous the hazard function is of the form

$$\lambda(t,z) = \lambda_0(t) \exp(\beta' z),$$

which is known as the proportional hazard function, also known as a multiplicative form of the hazard function, where $\lambda_0(t)$ is the underlying hazard function when $z \equiv 0$. β is the vector of unknown parameters.

For our problem we generalize Cox's (1972) model to include the extra variability involved. In other words we will try to model the continuous time variable in a way related to Cox's (1972) model to include the random component ε_{ij}.

The multiplicative hazard function for trt i and main unit j that will be considered is as follows

$$\lambda_{ij}(t) = \lambda_0(t) \exp(\beta' x_i(t) + \varepsilon_{ij}),$$

where $\lambda_0(t)$ is the underlying hazard when $x_i(t) \equiv 0$ and $\varepsilon_{ij} \equiv 0$, β is the vector of unknown parameters and $x_i(t)$ are the variables influencing failure times. Also, the survival function for trt i given main unit j is given by

$$S_{ij}(t) = 1 - F_{ij}(t) = \exp(-\int_0^t \lambda_{ij}(u)\ du),$$

and therefore

$$S_{ij}(t) = \exp(-\int_0^t \lambda_0(u) \exp(\beta' x_i(u) + \varepsilon_{ij}) du). \qquad \ldots \quad (3.2)$$

Substituting (3.2) in (3.1) we get

$$q_{ijk} = \exp(-\int_{t_{k-1}}^{t_k} \lambda_0(u) \exp(\beta' x_i(u) + \varepsilon_{ij}) du).$$

Now, let us assume that $x_i(t)$ is constant on interval k, i.e., let x_{ik} = value of $x_i(t)$ on interval k. Then we have

$$q_{ijk} = \exp(-\exp(\beta' x_{ik} + \varepsilon_{ij}) \int_{t_{k-1}}^{t_k} \lambda_0(u)\ du),$$

and this leads to

$$\log(-\log q_{ijk}) = \beta' x_{ik} + \varepsilon_{ij} + \log \int_{t_{k-1}}^{t_k} \lambda_0(u)\ du. \qquad \ldots \quad (3.3)$$

Let

$$\tau_k = \log \int_{t_{k-1}}^{t_k} \lambda_0(u)\ du, \quad \text{then}$$

$\log(-\log q_{ijk}) = \beta' x_{ik} + \varepsilon_{ij} + \tau_k$, where $\beta \in \mathbb{R}^p$, $\tau_k \in \mathbb{R}$, and
$\log(-\log \hat{q}_{ijk}) = \log(-\log q_{ijk}) + \delta_{ijk}$, where $\hat{q}_{ijk} = \frac{s_{ijk}}{n_{ijk}}$, and δ_{ijk} is a random error defined by $\delta_{ijk} = \log(-\log \hat{q}_{ijk}) - \log(-\log q_{ijk})$.

For the additive form of the hazard function we generalize the model presented by Elandt-Johnson (1980) to include the random component ε_{ij} in an additive fashion. Now we derive the model that will be used later in analysis using the additive hazard model which is given by

$$\lambda_{ij}(t) = \lambda_0(t) + \beta' x_i(t) + \varepsilon_{ij}.$$

The survivor function is then given by

$$\begin{aligned} S_{ij}(t) &= 1 - F_{ij}(t) \\ &= \exp(-\textstyle\int_0^t \lambda_{ij}(u)\, du) \\ &= \exp(-\textstyle\int_0^t (\lambda_0(u) + \beta' x_i(u) + \varepsilon_{ij}) du). \qquad \ldots \quad (3.4) \end{aligned}$$

Substituting (3.4) in (3.1) we get

$$q_{ijk} = \exp(-\int_{t_{k-1}}^{t_k} (\lambda_0(u) + \beta' x_i(u) + \varepsilon_{ij}) du).$$

Again assume that $x_i(t)$ is constant on interval k. In this case we have

$$q_{ijk} = \exp((\beta' x_{ik} + \varepsilon_{ij})(t_{k-1} - t_k)) \cdot \exp(-\int_{t_{k-1}}^{t_k} \lambda_0(u)\, du),$$

and this leads to

$$\log(q_{ijk}) = \beta' x_{ik}(t_{k-1} - t_k) + \varepsilon_{ij}(t_{k-1} - t_k) + [-\int_{t_{k-1}}^{t_k} \lambda_0(u)\, du]. \qquad \ldots \quad (3.5)$$

Define $z_{ik} = x_{ik}(t_{k-1} - t_k)$, $\varepsilon'_{ij} = \varepsilon_{ij}(t_{k-1} - t_k)$, and $\tau_k = -\int_{t_{k-1}}^{t_k} \lambda_0(u)\, du$. Then

$$\log(q_{ijk}) = \beta' z_{ik} + \varepsilon'_{ij} + \tau_k, \quad \text{where } \beta \in \mathbb{R}^p,\ \tau_k \in \mathbb{R}, \quad \text{and}$$

$$\log(\hat{q}_{ijk}) = \log(q_{ijk}) + \delta_{ijk}, \quad \text{where } \hat{q}_{ijk} = \frac{s_{ijk}}{n_{ijk}}, \text{ and } \delta_{ijk} \text{ is a}$$

random error defined by $\delta_{ijk} = \log(\hat{q}_{ijk}) - \log(q_{ijk})$.

Our grouped time model given by (1.1) is similar to these continuous models in the sense of having similar set of parameters. Therefore we can start with continuous setting for response time T and still end up with grouped time model that we considered for analysis although in our case response time T is discrete random variable.

It is appropriate here to mention that the proportional hazards model is convenient, e.g., the log(-log) model is to be preferred over the log model for the following two reasons:

1) Using the proportional hazards model leads to work with log(-log) model specified by the equation

$$\log(-\log q_{ijk}) = \beta' x_{ik} + \varepsilon_{ij} + \log \int_{t_{k-1}}^{t_k} \lambda_0(u)\, du.$$

However, using the additive form for the hazard leads to work with log model specified by the equation

$$\log(q_{ijk}) = \beta' x_{ik}(t_{k-1} - t_k) + \varepsilon_{ij}(t_{k-1} - t_k) + [-\int_{t_{k-1}}^{t_k} \lambda_0(u)\, du].$$

Therefore, inference with log(-log) transform is directly related to the parameters of the continuous time interpretation. The log(-log) model is to be preferred since β is invariant to time grouping,

2) The log model has a restricted range. $\hat{q}_{ijk}$'s are observed proportions and thus $0 \leq \hat{q}_{ijk} \leq 1$, which implies that $\log(\hat{q}_{ijk}) < 0$.

3.2 The Conditional Likelihood Function-
The No Censoring Case

As we have seen in Chapter I and section 3.1, n_{ij} is the number of individuals assigned to main unit j on trt i (n_{ij1} starters). For the no censoring case we have

$$n_{ij} = n_{ij1} \text{ and } n_{ijk} = s_{ij(k-1)} \text{ for } k > 1.$$

Using equation (1.4) and assuming that individuals in a main unit survive independently of other individuals we then have independent multinomial distributions over i and j. Therefore the conditional likelihood function for the observed data can be written as:

$$L(\underset{\sim}{q}|\underset{\sim}{\varepsilon}) \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} q_{ijk}^{s_{ijk}}(1 - q_{ijk})^{n_{ijk}-s_{ijk}}, \text{ where } \underset{\sim}{q} \text{ is } \ldots \quad (3.6)$$

a vector of q_{ijk}'s. For (1.1) q_{ijk} is a function of $\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\varepsilon}$ depending on the form assumed for the two hazard functions. Now let $q_{ijk}=g(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\varepsilon})$ then,

$$L(\underset{\sim}{\alpha},\underset{\sim}{\beta}|\underset{\sim}{\varepsilon}) \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} \{g(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\varepsilon})\}^{s_{ijk}} \{1 - g(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\varepsilon})\}^{r_{ijk}}.$$

The form of the conditional likelihood is the same as the likelihood function for product binomial random variables for fixed n_{ijk}. Therefore the asymptotic results for both cases are the same. For simplicity we will act as if we had a product of binomial random variables with fixed n_{ijk}, even though the n_{ijk} are random. In other words the asymptotic results are the same for the fixed n_{ijk} or the random n_{ijk} problem. This is one motivation for treating the n_{ijk} as fixed. Also it is difficult to see how there will be any information in the n_{ijk} about the q_{ijk} that is not already obtained in the s_{ijk}. This could be given as another motivation for treating n_{ijk} fixed.

3.3 Handling Censored Data

As discussed in section 3.1, n_{ijk} for the censoring case can be defined by the following relations

$$n_{ij1} = n_{ij} - c_{ij1}, \text{ and}$$

$$n_{ijk} = s_{ij(k-1)} - c_{ijk} = n_{ij(k-1)} - r_{ij(k-1)} - c_{ijk} \quad \text{for } k > 1,$$

while for the no censoring case we had

$$n_{ij} = n_{ij1}, \text{ and } n_{ijk} = s_{ij(k-1)} \quad \text{for } k > 1.$$

Therefore, the only difference between the censored and uncensored data is that risk sets at each time interval can be obtained somwhat differently.

Computationally, the case of censoring will not effect our parameters of interest nor the structure of our layout since our methods are based on the knowledge that

$$s_{ijk} | \varepsilon_{ij}, n_{ij\underline{k}} \sim \text{Binomial } (n_{ijk}, q_{ijk}).$$

Hence handling censored data will be straightforward.

In general, the idea behind handling censored time can be formulated as follows. It is often assumed that each individual has a life time T and a censoring time C, where T and C are independent continuous random variables with survivor functions $S(t,\theta)$ and $G(c,\phi)$ and probability density functions $f(t,\theta)$ and $g(c,\phi)$, respectively. θ is the vector of parameters of interest and ϕ is the vector of parameters on censoring time C. Let us assume T_i and C_i are independent for all i and define

$$Y_i = \min(T_i, C_i), \text{ and } \delta_i = \begin{cases} 1 & \text{if } Y_i = T_i \\ 0 & \text{if } Y_i = C_i \end{cases}.$$

The data from observations on individuals consist of the pair (Y_i, δ_i). Further, assume that the Y_i's are independent then if an individual failed then the "likelihood contribution" of observing a failure given $Y_i < C_i$ is given by the product $f(y_i,\theta)G(y_i,\phi)$. If, on the other hand, an individual is censored then the "likelihood contribution" of observing a censored given $Y_i < T_i$ is given by the product $g(y_i,\phi)S(y_i,\theta)$. Therefore, the full likelihood for all individuals in the study can be written as

$$L(\theta,\phi) = L(y_i,\theta,\phi) = \prod_{\substack{\text{Set of all}\\ \text{failures}}} f(y_i,\theta)G(y_i,\phi) \prod_{\substack{\text{Set of all}\\ \text{censored}}} g(y_i,\phi)S(y_i,\theta).$$

Since the parameter of interest is θ then we might consider working with the following marginal likelihood function

$$L(\theta) = \prod_{\substack{\text{Set of all}\\ \text{failures}}} f(y_i,\theta) \prod_{\substack{\text{Set of all}\\ \text{censored}}} S(y_i,\theta).$$

Note that $L(\theta,\phi) = L(\theta)\cdot K(\phi)$. Thus, for inference on θ alone $K(\phi)$ acts as a constant. $K(\phi)$ will not be used in solving for MLE's of θ or likelihood inference on θ. Therefore we consider working with the marginal likelihood function $L(\theta)$ rather than the full likelihood $L(\theta,\phi)$.

In what follows, we derive the form of the full and the marginal likelihoods given ε_{ij} for our grouped time model. From the way the data has been collected we have

$$(r_{ij1},c_{ij1},\ldots,r_{ijK},c_{ijK}|\varepsilon_{ij}) \sim \text{Multinomial } (n_{ij},\alpha_{ij1},\phi_{ij1},\ldots,\alpha_{ijK},\phi_{ijK}), \quad \ldots \quad (3.7)$$

where

$$\sum_{k=1}^{K+1} (r_{ijk} + c_{ijk}) = n_{ij} \text{ starters,}$$

$$\sum_{k=1}^{K+1} (\alpha_{ijk} + \phi_{ijk}) = 1,$$

α_{ijk} is the probability of an individual fails in interval k for a given main unit j on trt i with censoring, and

ϕ_{ijk} is the probability of an individual censored in interval k for a given main unit j on trt i.

Recall that π_{ijk} is the probability of an individual fails in interval k for a given main unit j on trt i for the no censoring case. Assuming that all censors take place at the start of an interval then α_{ijk} and π_{ijk} are related in the following form

$$\pi_{ijk} = \alpha_{ijk} / \sum_{k=1}^{K+1} \alpha_{ijk}. \qquad \ldots \quad (3.8)$$

We know that the conditional multinomial likelihood function is given by

$$\Pr(r_{ij1}, c_{ij1}, \ldots, r_{ijK}, c_{jjK} | \varepsilon_{ij}) = \frac{n_{ij1}!}{\prod_{k=1}^{K+1} r_{ijk}! \prod_{k=1}^{K+1} c_{ijk}!} \prod_{k=1}^{K+1} \alpha_{ijk}^{r_{ijk}} \phi_{ijk}^{c_{ijk}}. \qquad \ldots \quad (3.9)$$

Now, combining (1.3), (3.7), and (3.8) we get

$$\alpha_{ijk} = q_{ij1} \cdots q_{ij(k-1)} p_{ijk} (1 - \sum_{k=1}^{K+1} \phi_{ijk}), \qquad \ldots \quad (3.10)$$

and

$$\alpha_{ij(K+1)} = q_{ij1} \cdots q_{ij(K-1)} q_{ijK} (1 - \sum_{k=1}^{K+1} \phi_{ijk}). \qquad \ldots \quad (3.11)$$

Substituting (3.10), and (3.11) in (3.9) we have

$$P(r_{ij1}, c_{ij1}, \ldots, r_{ijK}, c_{ijK} | \varepsilon_{ij}) \propto \{p_{ij1}^{r_{ij1}} p_{ij2}^{r_{ij2}} \cdots p_{ijK}^{r_{ijK}} q_{ij1}^{r_{ij2}+\cdots+r_{ij(K+1)}}$$

$$q_{ijK}^{r_{ij(K+1)}} (1 - \sum_{k=1}^{K+1} \phi_{ijk})^{r_{ij1}+\cdots+r_{ij(K+1)}}\} \times$$

$$\{\prod_{k=1}^{K} \phi_{ijk}^{c_{ijk}}\}$$

$$\propto \{\prod_{k=1}^{K} p_{ijk}^{r_{ijk}} q_{ijk}^{n_{ijk}-r_{ijk}}\} \times$$

$$\{\prod_{k=1}^{K+1} \phi_{ijk}^{c_{ijk}} (1 - \sum_{k=1}^{K+1} \phi_{ijk})^{n_{ij1} - \sum_{k=1}^{K+1} c_{ijk}}\}.$$

Assuming that individuals in a main unit survive independently of other individuals we then have independent multinomial distributions over i and j. Therefore the conditional likelihood function can be written as

$$L(\underset{\sim}{q},\underset{\sim}{\phi}|\underset{\sim}{\varepsilon}) \propto \{\prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} q_{ijk}^{s_{ijk}} (1 - q_{ijk})^{r_{ijk}}\} \times$$

$$\{\prod_{i=1}^{I} \prod_{j-1}^{J} \prod_{k=1}^{K+1} \phi_{ijk}^{c_{ijk}} (1 - \sum_{k=1}^{K+1} \phi_{ijk})^{n_{ij1} - \sum_{k=1}^{K+1} c_{ijk}}\}.$$

Since our parameter of interest is $\underset{\sim}{q}$ then we might consider working with the following conditional marginal likelihood function.

$$L(\underset{\sim}{q}|\underset{\sim}{\varepsilon}) \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} q_{ijk}^{s_{ijk}} (1 - q_{ijk})^{r_{ijk}}$$

Therefore, for inference on $\underset{\sim}{q}|\underset{\sim}{\varepsilon}$ alone we consider working with the conditional marginal likelihood function $L(\underset{\sim}{q}|\underset{\sim}{\varepsilon})$ rather than the full conditional likelihood function $L(\underset{\sim}{q},\underset{\sim}{\phi}|\underset{\sim}{\varepsilon})$. This conditional marginal likelihood function for censored data has the same form as the conditional likelihood function for uncensored data as given by (3.6). Again note that the form of this conditional marginal likelihood function is the same as the likelihood function of independent binomial random variables with fixed n_{ijk} as is the case with no censored observations. Therefore and for simplicity we act as if we had a product of independent binomial random variables with fixed n_{ijk}, even though the n_{ijk} are actually random. Again the asymptotic results will be the same for both fixed n_{ijk} or random n_{ijk}.

3.4 Unconditional Survival Functions and Likelihood Functions

As we have seen in section 3.2, the conditional likelihood function of our parameters of interest $\underset{\sim}{q}=g(\underset{\sim}{\alpha},\underset{\sim}{\beta})$ is also a function of the unknown random vector $\underset{\sim}{\varepsilon}$. Thus our aim in this section is to find a likelihood function that is free from these unknown values. The purpose is that if we can get an unconditional likelihood function, we can then find a maximum likelihood estimator for our parameters of interest. But there are some difficulties with this approach, and hence least squares estimates will be discussed in Chapter IV.

The approach to find the unconditional likelihood function is outlined below.

1. Use the conditional hazard function to get the unconditional one by assuming a normal distribution for ε_{ij}.
2. Get q_{ijk} as a function of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ only.
3. Write the likelihood function which is free of $\underset{\sim}{\varepsilon}$.

This approach will be carried out for both multiplicative and additive hazard functions.

The conditional hazard function in the multiplicative form is given by

$$\lambda_{i|\varepsilon_{ij}}(t) = \lambda_0(t) \exp(\beta' x_i(t) + \varepsilon_{ij}).$$

Assume that $x_i(t) = x_{ik}$ then we have

$$S_{i|\varepsilon_{ij}}(t) = 1 - F_{i|\varepsilon_{ij}}(t) = \exp(-\int_0^t \lambda_0(u) \exp(\beta' x_{ik} + \varepsilon_{ij})du)$$

$$= \exp(-\exp(\beta' x_{ik} + \varepsilon_{ij})\int_0^t \lambda_0(u)du)$$

$$= \exp(-\Lambda_0(t) \exp(\beta' x_{ik} + \varepsilon_{ij})),$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$. Now if we assume that $\varepsilon_{ij} \sim N(0,\sigma_\varepsilon^2)$ then we have

$$S_i(t) = \int_{-\infty}^{\infty} S_{i|\varepsilon_{ij}}(t) f(\varepsilon_{ij}) d\varepsilon_{ij}$$

$$= \int_{-\infty}^{\infty} \exp(-\Lambda_0(t) \exp(\beta' x_{ik} + \varepsilon_{ij})) f(\varepsilon_{ij}) d\varepsilon_{ij}$$

$$= M_Y(-\Lambda_0(t)\exp(\beta' x_{ik})),$$

where $M(\cdot)$ is the moment generating function of $(\cdot)$, and $Y = \exp(\varepsilon_{ij}) \sim \text{lognormal}(0,\sigma_\varepsilon^2)$. We should mention that there is no closed formula for the moment generating function of a lognormal random variable. At this stage we can use an approximation by using Taylor expansion of second degree for

$$g(\varepsilon_{ij}) = \exp(-\Lambda_0(t) \exp(\beta' x_{ik} + \varepsilon_{ij})),$$

and expand it around $\varepsilon_{ij} = 0$. Hence we have

$$g(0) = \exp(-\Lambda_0(t)\exp(\beta' x_{ik})),$$

$$g'(\varepsilon_{ij}) = -\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})\ \exp(-\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})),$$

$$g'(0) = -\Lambda_0(t)\exp(\beta' x_{ik})\exp(-\Lambda_0(t)\exp(\beta' x_{ik})),$$

$$g''(\varepsilon_{ij}) = -\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})\ \exp(-\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})) + [\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})]^2\ \exp(-\Lambda_0(t)\ \exp(\beta' x_{ik} + \varepsilon_{ij})),$$

$$g''(0) = -\Lambda_0(t)\ \exp(\beta' x_{ik})[1 - \Lambda_0(t)\exp(\beta' x_{ik})]\ \exp(-\Lambda_0(t)\ \exp(\beta' x_{ik})).$$

Therefore,

$$S_i(t) \doteq \{1 - \Lambda_0(t)\ \exp(\beta' x_{ik})[1 - \Lambda_0(t)\ \exp(\beta' x_{ik})]\frac{\sigma_\varepsilon^2}{2}\} \times \exp(-\Lambda_0(t)\ \exp(\beta' x_{ik})).$$

$S_i(t)$ satisfies the following properties $S_i(0) = 1$, $\lim_{t\to\infty} S_i(t) = 0$, and $S_i(t)$ is nonincreasing, and left continuous. Therefore, $S_i(t)$ is a survivor function even though it is only an approximation of the true survivor function. The only restriction on this survivor function is that it always intersects the underlying function $\Lambda_0(t)$ at height $= e^{-1}$.

The conditional hazard function in an additive form is given by

$$\lambda_{i|\varepsilon_{ij}}(t) = \lambda_0(t) + \beta' x_i(t) + \varepsilon_{ij}.$$

Assume that $x_i(t) = x_{ik}$ then we have

$$S_{i|\varepsilon_{ij}}(t) = 1 - F_{i|\varepsilon_{ij}}(t) = \exp[-\textstyle\int_0^t (\lambda_0(u) + \beta' x_{ik} + \varepsilon_{ij})du]$$

$$= \exp(-(\beta' x_{ik} + \varepsilon_{ij})t - \Lambda_0(t)),$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$. Again if we assume that $\varepsilon_{ij} \sim N(0,\sigma_\varepsilon^2)$ then we

$$
\begin{aligned}
S_i(t) &= \int_{-\infty}^{\infty} S_{i|\varepsilon_{ij}}(t)f(\varepsilon_{ij})d\varepsilon_{ij} \\
&= \int_{-\infty}^{\infty} \exp(-(\beta' x_{ik}t + \Lambda_0(t)) - \varepsilon_{ij}t)f(\varepsilon_{ij})d\varepsilon_{ij} \\
&= \exp(-(\beta' x_{ik}t + \Lambda_0(t)))\int_{-\infty}^{\infty} e^{-\varepsilon_{ij}t} f(\varepsilon_{ij})d\varepsilon_{ij} \\
&= \exp(-\beta' x_{ik}t - \Lambda_0(t))\ E(e^{-\varepsilon_{ij}t}) \\
&= \exp(-\beta' x_{ik}t - \Lambda_0(t) + \frac{t^2}{2}\sigma_\varepsilon^2).
\end{aligned}
$$

However, $S_i(t)$ is not a survivor function since

$$S_i(t) \longrightarrow \infty \text{ as } t \longrightarrow \infty.$$

Thus, from now on we are going to emphasize on working with log(-log) model rather than with log model, because of the restriction on the parameters that log model has, and that treatment effects are free from time grouping as we have seen in deriving the model to be used in discrete setting from the continuous setting. Also, we should mention here that least squares estimates are going to be considered over maximum likelihood estimates because of the following reasons:

1. There is no closed form for the likelihood function.
2. The quality of the approximation is in doubt.

CHAPTER IV

AN APPROACH TO INFERENCE

4.1 Estimation of Variance Components

4.1.1 Estimation of the Binomial Variabilities

As we have seen in Chapter I,

$$s_{ijk} | \varepsilon_{ij}, n_{ijk} \sim \text{Binomial}(n_{ijk}, q_{ijk}), \text{ with}$$

$$\text{cov}(s_{ijk}, s_{ijk'} | \varepsilon_{ij}, n_{ijk}, n_{ijk'}) = 0.$$

For a large sample size, the asymptotic distribution is given by

$$s_{ijk} | \varepsilon_{ij}, n_{ijk} \mathrel{\dot{\sim}} \text{Normal}(n_{ijk}q_{ijk}, n_{ijk}q_{ijk}(1 - q_{ijk})), \text{ and}$$

$$\hat{q}_{ijk} = \frac{s_{ijk}}{n_{ijk}} | \varepsilon_{ij}, n_{ijk} \mathrel{\dot{\sim}} \text{Normal}(q_{ijk}, q_{ijk}(1 - q_{ijk})/n_{ijk}).$$

Let $g(q_{ijk}) = \log(-\log q_{ijk})$,

$$\dot{g}(q_{ijk}) = 1/q_{ijk}(\log q_{ijk}),$$

where $\dot{g}$ is defined to be the derivative of the function g. Then

$$g(\hat{q}_{ijk}) | \varepsilon_{ij}, n_{ijk} \mathrel{\dot{\sim}} \text{Normal}(g(q_{ijk}), [\dot{g}(q_{ijk})]^2 q_{ijk}(1 - q_{ijk})/n_{ijk}),$$

that is,

$$\log(-\log \hat{q}_{ijk}) | \varepsilon_{ij}, n_{ijk} \mathrel{\dot{\sim}} \text{Normal}(\log(-\log q_{ijk}), \frac{1 - q_{ijk}}{n_{ijk}q_{ijk}(\log q_{ijk})^2}).$$

Going back to the general model,

$$\log(-\log \hat{q}_{ijk}) = \log(-\log q_{ijk}) + \delta_{ijk},$$

where δ_{ijk} is a random error then we have

$$\delta_{ijk} | \varepsilon_{ij}, n_{ijk} \overset{\cdot}{\underset{\cdot}{\sim}} \text{Normal}(0, \frac{1 - q_{ijk}}{n_{ijk} q_{ijk} (\log q_{ijk})^2}).$$

In constructing the variance-covariance matrix for the model we need $\text{Var}(\delta_{ijk})$ rather than $\text{Var}(\delta_{ijk} | \varepsilon_{ij}, n_{ijk})$. Therefore, we need to consider ε_{ij} and δ_{ijk} to be uncorrelated but not necessarily independent since $\text{Var}(\delta_{ijk})$ and $\text{Var}(\delta_{ijk} | \varepsilon_{ij}, n_{ijk})$ may be different. Thus we have

$$\text{Var}(\delta_{ijk}) = E(\text{Var}(\delta_{ijk} | \varepsilon_{ij}, n_{jik})) + \text{Var}(E(\delta_{ijk} | \varepsilon_{ij}, n_{ijk})).$$

Since $E(\varepsilon_{ij}) = 0$ and $E(\delta_{ijk}) = 0$ then $E(\delta_{ijk} | \varepsilon_{ij}, n_{ijk}) = 0$ and hence

$$\text{Var}(\delta_{ijk}) = E(\text{Var}(\delta_{ijk} | \varepsilon_{ij}, n_{ijk})).$$

This suggests averaging over all main units on the same trt and the same interval to get an estimate of $\text{Var}(\delta_{ijk})$ given as

$$\hat{\text{Var}}(\delta_{ijk}) = \sum_{j=1}^{J} \hat{\text{Var}}(\delta_{ijk} | \varepsilon_{ij}, n_{ijk}) / \# \text{ main units on trt } i. \quad \ldots \quad (4.1)$$

For the rest of all arguments we will designate $\text{Var}(\delta_{ijk})$ by $\sigma^2_{\delta_{ijk}}$ and its estimate by $\hat{\sigma}^2_{\delta_{ijk}}$.

4.1.2 Equating SS to ESS to estimate terms in V

Using a method similar to the fitting constant method provided by Henderson (1953) and Searle (1968), an estimate of σ^2_ε can be obtained as follows

Step (1) Fit the full model

$$\begin{aligned}\underset{\sim}{y} &= M\underset{\sim}{B} + \underset{\sim}{\delta} \\ &= [X \,\vdots\, Z]\begin{bmatrix} \underset{\sim}{b} \\ \underset{\sim}{\varepsilon} \end{bmatrix} + \underset{\sim}{\delta} \\ &= X\underset{\sim}{b} + Z\underset{\sim}{\varepsilon} + \underset{\sim}{\delta},\end{aligned}$$

where $\underset{\sim}{b}$ is the vector of fixed effects, $\underset{\sim}{\varepsilon}$ is the single vector of random effects, and $\underset{\sim}{\delta} \overset{\cdot}{\underset{\cdot}{\sim}} N(\underset{\sim}{0},\hat{W})$, where $\hat{W}$ is a diagonal matrix of $\hat{\sigma}^2_{\delta_{ijk}}$'s being the diagonal elements. Now, the reduction sum of squares for fitting the full model is given by

$$\begin{aligned}R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) &= \underset{\sim}{y}'\hat{W}^{-1}[X \,\vdots\, Z]\left[\begin{bmatrix} X' \\ --- \\ Z' \end{bmatrix}\hat{W}^{-1}[X \,\vdots\, Z]\right]^{-}\begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix}\hat{W}^{-1}\underset{\sim}{y} \\ &= \underset{\sim}{y}'\hat{W}^{-1}[X \,\vdots\, Z]\begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}Z \end{bmatrix}^{-}\begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix}\hat{W}^{-1}\underset{\sim}{y}.\end{aligned}$$

Step (2) Fit the reduced model

$$\underset{\sim}{y} = X\underset{\sim}{b} + \underset{\sim}{\delta}.$$

Then, the reduction sums of squares for fitting the reduced model is given by

$$R(\underset{\sim}{b}) = \underset{\sim}{y}'\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\,X'\hat{W}^{-1}\underset{\sim}{y}.$$

Step (3) Get the expectation of the quadratic form under the full model. The expectation of $\underset{\sim}{y}'Q\underset{\sim}{y}$ under the model $\underset{\sim}{y} = MB + \delta$ is given by

$$E(\underset{\sim}{y}'Q\underset{\sim}{y}) = \text{tr}\left[\begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix} Q[X \,\vdots\, Z]E(BB')\right] + \text{tr}[Q\hat{W}].$$

Then, the expectation of $R(\underset{\sim}{b},\underset{\sim}{\varepsilon})$ under the full model is given by

$$E(R(\underset{\sim}{b},\underset{\sim}{\varepsilon})) = \text{tr}\left[\begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix} \hat{W}^{-1}[X \mid Z] \begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}Z \end{bmatrix}^{-} \begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix} \hat{W}^{-1}[X \mid Z]E(BB')\right]$$

$$+ \text{ tr}\left[\hat{W}^{-1}[X \mid Z] \begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}Z \end{bmatrix}^{-} \begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix}\right] .$$

If we let

$$\hat{W}^{-1}[X \mid Z] \begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}Z \end{bmatrix}^{-} \begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix} = C_1,$$

then we have

$$E(R(\underset{\sim}{b},\underset{\sim}{\varepsilon})) = \text{tr}\left[\begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}Z \end{bmatrix}^{-} E(BB')\right] + \text{tr}[C_1].$$

Similarly, the expectation of $R(\underset{\sim}{b})$ under the full model is given by

$$E(R(\underset{\sim}{b})) = \text{tr}\left[\begin{bmatrix} X' \\ -- \\ Z' \end{bmatrix} \hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'\hat{W}^{-1}[X \mid Z]E(BB')\right] +$$

$$\text{tr}[\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'].$$

If we let $\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X' = C_2$, then we have

$$E(R(\underset{\sim}{b})) = \text{tr}\left[\begin{bmatrix} X'\hat{W}^{-1}X & X'\hat{W}^{-1}Z \\ Z'\hat{W}^{-1}X & Z'\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'\hat{W}^{-1}Z \end{bmatrix} E(BB')\right] + \text{tr}[C_2].$$

Hence

$$E(R(\underset{\sim}{\varepsilon} \mid \underset{\sim}{b})) = E(R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) - R(\underset{\sim}{b}))$$

$$= \text{tr}\left[\begin{bmatrix} 0 & 0 \\ 0 & Z'\hat{W}^{-1}Z - Z'\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}X'\hat{W}^{-1}Z \end{bmatrix} E(BB')\right] + \text{tr}[C_1]$$

$$- \text{ tr}[C_2]$$

$$= \mathrm{tr}[Z'(\hat{W}^{-1}-\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'\hat{W}^{-1})ZE(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}')] + \mathrm{tr}[C_1] - \mathrm{tr}[C_2]$$

$$= \sigma_{\varepsilon}^2 \,\mathrm{tr}\ [Z'(\hat{W}^{-1}-\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'\hat{W}^{-1})Z] + \mathrm{tr}[C_1] - \mathrm{tr}[C_2].$$

If we let $Z'(\hat{W}^{-1}-\hat{W}^{-1}X(X'\hat{W}^{-1}X)^{-}\ X'\hat{W}^{-1})Z = C$, then we have

$$E(R(\underset{\sim}{\varepsilon}|\underset{\sim}{b})) = \sigma_{\varepsilon}^2\,\mathrm{tr}\ [C] + \mathrm{tr}[C_1] - \mathrm{tr}[C_2].$$

Step (4) Equate $ER(\underset{\sim}{\varepsilon}|\underset{\sim}{b})$ to $R(\underset{\sim}{b},\varepsilon) - R(b)$ and solve for the unknown parameter σ_{ε}^2 gives us

$$\hat{\sigma}_{\varepsilon}^2 = [R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) - R(\underset{\sim}{b}) - \mathrm{tr}[C_1] + \mathrm{tr}[C_2]]/\mathrm{tr}[C]. \qquad \ldots \quad (4.2)$$

4.2 Weighted Least Squares and Survival Functions Estimation

The split plot model with unequal subplot variances will be analyzed first for the purpose of finding a good fitting model. This model has main treatment effects, time effects, and treatment by time interaction effects. The model obtained using the proportional hazard model did not have treatment by time effects interaction. So, our modification here can be formulated as follows.

1. If the best fitted model has no treatment by time interaction, the proportional hazard model, which we started with, is then the appropriate one.

2. If the best fitted model has treatment by time interaction, then we have to modify the proportional hazard model. In this case we can use the general Cox (1972) model with time dependent variables, i.e., the covariables are functions of time. In other words we can start with the following model for the hazard function. Let

$$\lambda_{ij}(t) = \lambda_0(t) \exp[\beta' x_i(t) + \varepsilon_{ij}],$$

where $x_i(t)$ are the variables influencing failure times. In general, these variables can be functions of time. However, in deriving the model to be used it can be assumed that $x_i(t)$ are constant on interval k, as we have seen before. Therefore, if the best fitted model has a treatment by time interaction, we can use the general time dependent variables for Cox (1972) model. Otherwise, the assumption that variables are free from time effects would be appropriate.

The split plot model with $\log(-\log \hat{q}_{ijk})$ response can be written in a matrix form as given below.

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u}$$

where

$\underset{\sim}{y}$ is an IJK × 1 vector of known values and $y_{ijk} = \log(-\log \hat{q}_{ijk})$,

X is an IJK × P design matrix,

$\underset{\sim}{\beta}$ is a P × 1 vector of unknown parameters,

$\underset{\sim}{u}$ is an IJK × 1 vector of random components, and

$\underset{\sim}{u} \overset{\cdot}{\underset{\cdot}{\sim}} N_{IJK}(\underset{\sim}{0}, V)$,

where V is an IJK × IJK variance-covariance matrix having both $v(\varepsilon_{ij}) = \sigma^2_\varepsilon$ and $v(\delta_{ijk}) = \sigma^2_{\delta_{ijk}}$ in the following form

$$V = \begin{bmatrix} \begin{matrix} \sigma_\varepsilon^2+\sigma_{\delta_{111}}^2 & \sigma_\varepsilon^2 \cdots & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2+\sigma_{\delta_{112}}^2 \cdots & \sigma_\varepsilon^2 \\ \vdots & \vdots & \vdots \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \sigma_\varepsilon^2+\sigma_{\delta_{11K}}^2 \end{matrix} & 0 & \\ 0 & \begin{matrix} \sigma_\varepsilon^2+\sigma_{\delta_{121}}^2 & \sigma_\varepsilon^2 \cdots & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2+\sigma_{\delta_{122}}^2 \cdots & \sigma_\varepsilon^2 \\ \vdots & \vdots & \vdots \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \sigma_\varepsilon^2+\sigma_{\delta_{12K}}^2 \end{matrix} & \\ & & \ddots \\ 0 & 0 & \text{IJ times} \end{bmatrix}$$

Similar to the procedure provided by Grizzle, Starmer and Koch (1969), generalized least squares estimates after estimating the variance-covariance matrix V will be considered and are given by

$$\hat{\underset{\sim}{B}} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}\underset{\sim}{y} \quad \text{with} \quad \widehat{\text{Cov}}(\hat{\underset{\sim}{B}}) = (X'\hat{V}^{-1}X)^{-1}.$$

For T discrete, the conditional survivor function for treatment i and time interval t_k is given by

$$S_{i|\varepsilon_{ij}}(t_k) = \prod_{\ell=1}^{k} q_{ij\ell}. \qquad \ldots \quad (4.3)$$

Estimates of the above conditional survivor functions can be easily found from the estimates of q_{ijk} for all $i = 1,\ldots,I$, $j = 1,\ldots,J$, and $k = 1,\ldots,K$. It is important to mention here that estimates of the survivor functions given $\varepsilon_{ij} = 0$ will be considered. The reasons for that can be summarized as follows:

1. We assumed that tank variability (ε_{ij}) increase or decrease the survivor function,
2. Usually treatment means comparisons are done holding all other conditions as constants. Similarly treatment comparisons for survival data can be done using survival curves for different treatments holding all other conditions constant, e.g., $\varepsilon_{ij} = 0$.

CHAPTER V

ASYMPTOTIC PROPERTIES

5.1 Introduction

The asymptotic properties for the estimators presented in Chapter IV are based on the behavior of the estimators when both the number of main units, J, and the number of subunits in each main unit, $n_{ij} = n$ for all i and j, approach ∞. As we have seen in Chapter IV, the variance-covariance matrix V for grouped time model is a block diagonal matrix that involves both $\sigma^2_{\delta_{ijk}}$ and σ^2_{ε}. $\sigma^2_{\delta_{ijk}}$ is a function of n_{ijk} and J, and V is of order $IJK \times IJK$, and thus V depends on J. From this point onwards we write $V_{(J)}$ to denote this dependency. A reasonable approach to prove the asymptotic properties for our estimators is as follows.

Step (1) Prove that

$$(n_{ijk}\hat{\sigma}^2_{\delta_{ijk}} | \varepsilon_{ij}) - (n_{ijk}\sigma^2_{\delta_{ijk}} | \varepsilon_{ij}) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Step (2) Assume as a working approximation that

$$\hat{\sigma}^2_{\delta_{ijk}} | \varepsilon_{ij} \doteq \sigma^2_{\delta_{ijk}} | \varepsilon_{ij} \quad \text{and} \quad E(\underset{\sim}{y}) \doteq X\underset{\sim}{\beta}.$$

Step (3) Prove that

$$\hat{\sigma}^2_{\delta_{ijk}} - \sigma^2_{\delta_{ijk}} \xrightarrow{P} 0 \text{ as } J \to \infty \text{ and that } \hat{W}_{(J)} - W_{(J)} \xrightarrow{P} 0 \text{ as } J \to \infty.$$

Step (4) $\hat{\sigma}_\varepsilon^2$ using $W_{(J)}$ as the variance-covariance matrix, $\hat{\sigma}_\varepsilon^2(W)$ say, has the following property

$$\hat{\sigma}_\varepsilon^2(W) - \sigma_\varepsilon^2 \xrightarrow{P} 0 \quad \text{as} \quad J \longrightarrow \infty.$$

Step (5) From steps (3) and (4), prove, using $\hat{W}_{(J)}$ that

$$\hat{\sigma}_\varepsilon^2(\hat{W}) - \sigma_\varepsilon^2 \xrightarrow{P} 0 \quad \text{as} \quad J \longrightarrow \infty.$$

Step (6) With known $V_{(J)}$ prove that $\hat{\underset{\sim}{\beta}}$ has an asymptotic normal distribution as $J \to \infty$.

Step (7) Prove that $\hat{V}_{(J)} - V_{(J)} \xrightarrow{P} 0$ as $J \to \infty$.

Step (8) With unknown $V_{(J)}$ prove that $\hat{\underset{\sim}{\beta}}$ has an asymptotic normal distribution as $J \to \infty$.

We discuss the above steps in detail in section 5.2. All tests of hypotheses and confidence limits are discussed in section 5.3. In section 5.4 we discuss confidence limits for survivor functions and test for the appropriateness of assuming the binomial variances.

5.2 Consistency of Variance Components Estimators and Asymptotic Distribution of $\hat{\underset{\sim}{\beta}}$

As discussed in Chapter I we have

$$(r_{ij1},\ldots,r_{ijK},s_{ijK}|\varepsilon_{ij}) \sim \text{Multinomial}(n,\pi_{ij1},\pi_{ij2},\ldots,\pi_{ij(K+1)}).$$

Therefore, by the Weak Law of Large Numbers we have

$$\begin{bmatrix} r_{ij1}/n \\ r_{ij2}/n \\ \\ r_{ijk}/n \\ \\ s_{ijK}/n \end{bmatrix} \mid \varepsilon_{ij} \xrightarrow{P} \begin{bmatrix} \pi_{ij1} \\ \pi_{ij2} \\ \\ \pi_{ijk} \\ \\ \pi_{ij(K+1)} \end{bmatrix} \quad \text{as } n\to\infty,$$

where

$$\pi_{ij\ell} = \left(\prod_{\ell=1}^{K-1} q_{ij\ell}\right)(1-q_{ij\ell}) \text{ for } \ell = 1,2,\ldots,K, \text{ and}$$

$$\pi_{ij(K+1)} = \prod_{\ell=1}^{K} q_{ij\ell}.$$

Now, let

$$h(\underset{\sim}{\pi}_{ijk}) = q_{ijk} = \sum_{\ell=k+1}^{K+1} \pi_{ij\ell} \Big/ \sum_{\ell=k}^{K+1} \pi_{ij\ell}$$

be a continuous function. Then

$$\hat{q}_{ijk}\mid\varepsilon_{ij} = h\left(\frac{r_{ijk}}{\underset{\sim}{n}}\right) \xrightarrow{P} h(\underset{\sim}{\pi}_{ijk}) = q_{ijk} \quad \text{as } n\to\infty.$$

We now let

$$g(q_{ijk}) = \frac{1-q_{ijk}}{q_{ijk}(\log q_{ijk})^2}$$

be another continuous function. Then

$$\frac{1-\hat{q}_{ijk}}{\hat{q}_{ijk}(\log \hat{q}_{ijk})^2}\mid\varepsilon_{ij} \xrightarrow{P} \frac{1-q_{ijk}}{q_{ijk}(\log q_{ijk})^2} \quad \text{as } n\to\infty.$$

Recall that

$$\hat{\sigma}^2_{\delta_{ijk}}\mid\varepsilon_{ij} = \frac{1-\hat{q}_{ijk}}{n_{ijk}\hat{q}_{ijk}(\log \hat{q}_{ijk})^2}\mid\varepsilon_{ij}.$$

Therefore, we have

$$n_{ijk}\hat{\sigma}^2_{\delta_{ijk}}|\varepsilon_{ij} = \frac{1-\hat{q}_{ijk}}{\hat{q}_{ijk}(\log \hat{q}_{ijk})^2}|\,\varepsilon_{ij}.$$

Hence, we get

$$n_{ijk}\hat{\sigma}^2_{\delta_{ijk}}|\varepsilon_{ij} \xrightarrow{P} g(q_{ijk}) \quad \text{as } n\to\infty.$$

Now, if we let

$$\frac{1-q_{ijk}}{n_{ijk}q_{ijk}(\log q_{ijk})^2}$$

be the true variance of $\delta_{ijk}|\varepsilon_{ij}$ then we have

$$\sigma^2_{\delta_{ijk}}|\varepsilon_{ij} = \frac{1-q_{ijk}}{n_{ijk}q_{ijk}(\log q_{ijk})^2}\,.$$

Therefore, we have

$$(n_{ijk}\hat{\sigma}^2_{\delta_{ijk}}|\varepsilon_{ij}) - (n_{ijk}\sigma^2_{\delta_{ijk}}|\varepsilon_{ij}) \xrightarrow{P} 0 \quad \text{as } n\to\infty. \qquad \ldots \quad (5.1)$$

This completes the proof of step (1).

Assuming step (2) is true we can act as if

$$\hat{\sigma}^2_{\delta_{ijk}}|\varepsilon_{ij} \doteq \sigma^2_{\delta_{ijk}}|\varepsilon_{ij} \quad \text{and} \quad E(\underset{\sim}{y}) \doteq X\underset{\sim}{\beta}, \text{ when the } n_{ijk} \text{ are large.}$$

Now we prove step (3). Recall that q_{ijk} is a random variable and it depends on ε_{ij}. Then we can view the sequence $\{(1-q_{ijk})/q_{ijk}(\log q_{ijk})^2\}$ for all $j = 1,2,\ldots,J$ as $J\to\infty$ as a sequence of independently and identically distributed random variables such that $E((1-q_{ijk})/q_{ijk}(\log q_{ijk})^2) < \infty$. Therefore, by the Strong Law of Large Numbers we have

$$\frac{1}{J}\sum_{j=1}^{J}((1-q_{ijk})/q_{ijk}(\log q_{ijk})^2) - E((1-q_{ijk})/q_{ijk}(\log q_{ijk})^2)$$

$\xrightarrow{a.s.} 0$ as $J \to \infty$.

Hence, the convergence in probability is also true. Therefore,

$$\hat{\sigma}^2_{\delta_{ijk}} - \sigma^2_{\delta_{ijk}} \xrightarrow{P} 0, \qquad \ldots \quad (5.2)$$

as $J \to \infty$ and for all $i = 1,\ldots,I$, $j = 1,\ldots,J$ and $k = 1,\ldots,K$. By (5.2) and Arnold's (1981) result on page 341 we have

$$\hat{W}_{(J)} - W_{(J)} \xrightarrow{P} 0, \qquad \ldots \quad (5.3)$$

as $J \to \infty$ where

$$W_{(J)} = (\sigma^2_{\delta_{ijk}})_{IJK}, \qquad \ldots \quad (5.4)$$

for $i = 1,\ldots,I$, $j = 1,\ldots,J$, and $k = 1,\ldots,K$. This completes the proof of step (3).

For the proof of steps (4), (5), (6), (7), and (8) we proceed as follows. For our grouped time model, $\underset{\sim}{y}_{(J)} = X_{(J)}\underset{\sim}{\beta} + \underset{\sim}{u}$, where $\text{cov}(\underset{\sim}{u}) = V_{(J)}$ and $V_{(J)}$ is a block diagonal matrix defined as before, we partition $\underset{\sim}{y}_{(J)}$, $X_{(J)}$ and $V_{(J)}$ as follows

$$\underset{\sim}{y}_{(J)} = \underset{\sim}{y}_{IJK\times 1} = [\underset{\sim}{y}'_1 \cdots \underset{\sim}{y}'_j \cdots \underset{\sim}{y}'_J]',$$

$$X_{(J)} = X_{IJK\times P} = [X'_1 \cdots X'_j \cdots X'_J]',$$

where X_j's are all IK×P identical matrices. Also

$$V_{(J)} = V_{IJK\times IJK} = \begin{bmatrix} V_1 & & & & 0 \\ & \ddots & & & \\ & & V_j & & \\ & & & \ddots & \\ 0 & & & & V_J \end{bmatrix},$$

where V_j's are all IK×IK identical matrices. Therefore, $\underset{\sim}{y}_j$'s are independently and identically distributed IK-dimensional random vectors with $E(\underset{\sim}{y}_j) = X_j\underset{\sim}{\beta}$ and $\text{Cov}(\underset{\sim}{y}_j) = V_j$. Using Arnold's (1981) theorem 18.16 (b) we have

$$\sqrt{J}\,(\bar{\underset{\sim}{y}}_{(J)} - X_j\underset{\sim}{\beta}) \xrightarrow{L} N_{IK}(\underset{\sim}{0},V_j) \qquad \ldots \quad (5.5)$$

as J → ∞, where $\bar{\underset{\sim}{y}}_{(J)} = \frac{1}{J}\sum_{j=1}^{J}\underset{\sim}{y}_j$.

To prove step (4), we recall the estimation procedure used to get the estimate of σ_ε^2. Full and a reduced model were both fitted. The full model is of the form

$$\underset{\sim}{y}_{(J)} = [X\,\vdots\,Z]_{(J)}\begin{bmatrix}\underset{\sim}{b}\\ \underset{\sim}{\varepsilon}\end{bmatrix} + \underset{\sim}{\delta},$$

where $\underset{\sim}{b}$ is the vector of fixed effects, $\underset{\sim}{\varepsilon}$ is the vector of random effects, and $\text{Cov}(\underset{\sim}{\delta}) = W_{(J)}$, where W is given by (5.4). For this model we write

$$\underset{\sim}{y}_{(J)} = M_{(J)}\underset{\sim}{B} + \underset{\sim}{\delta}, \text{ where } M_{(J)} = [X\,\vdots\,Z]_{(J)}.$$

The reduced model is of the form

$$\underset{\sim}{y}_{(J)} = X_{(J)}\underset{\sim}{b} + \underset{\sim}{\delta}.$$

$\hat{\sigma}_\varepsilon^2$ is unbiased for σ_ε^2 since

$$E(\hat{\sigma}_\varepsilon^2) = \{E(R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) - R(\underset{\sim}{b})) - tr[c_1] + tr[c_2]\}/tr[c].$$

Recall that $E(R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) - R(\underset{\sim}{b})) = \sigma_\varepsilon^2 tr[c] + tr[c_1] - tr[c_2]$ then we have $E(\hat{\sigma}_\varepsilon^2) = \sigma_\varepsilon^2$. Now, to prove that $\hat{\sigma}_\varepsilon^2$ is a consistent estimator for σ_ε^2 we need to show that $\text{Var}(\hat{\sigma}_\varepsilon^2) \to 0$ as J → ∞. From (4.2) we have

$$\text{Var}(\hat{\sigma}_\varepsilon^2) = \text{Var}((R(\underset{\sim}{b},\underset{\sim}{\varepsilon}) - R(\underset{\sim}{b}))/tr[c]).$$

But $\hat{\sigma}^2_\varepsilon$ is a quadratic form with $E(\underset{\sim}{y}) = X\underset{\sim}{\beta}$ and $V(\underset{\sim}{y}) = V$ then we have

$$\begin{aligned}\mathrm{Var}(\hat{\sigma}^2_\varepsilon) = {} & \frac{2}{(\mathrm{tr}[c])^2}\,\mathrm{tr}\,[W^{-1}_{(J)}\{M_{(J)}(M'_{(J)}W^{-1}_{(J)}M_{(J)})^{-}M'_{(J)}-X_{(J)}(X'_{(J)}W^{-1}_{(J)}X_{(J)})^{-} \\ & X'_{(J)}\}W^{-1}_{(J)}V_{(J)}]^2 + \frac{4}{(\mathrm{tr}[c])^2}\,\underset{\sim}{\beta}'X'_{(J)}W^{-1}_{(J)}\{M_{(J)}(M'_{(J)}W^{-1}_{(J)}M_{(J)})^{-} \\ & M'_{(J)}-X_{(J)}(X'_{(J)}W^{-1}_{(J)}X_{(J)})^{-}X'_{(J)}\}W^{-1}_{(J)}V_{(J)}W^{-1}_{(J)}\{M_{(J)}(M'_{(J)}W^{-1}_{(J)}M_{(J)})^{-} \\ & M'_{(J)}-X_{(J)}(X'_{(J)}W^{-1}_{(J)}X_{(J)})^{-}X'_{(J)}\}W^{-1}_{(J)}X_{(J)}\underset{\sim}{\beta}. \qquad \ldots \quad (5.6)\end{aligned}$$

Using the partition described for the grouped time model and then applying it on both fitted models and partitioning Z, M, and W accordingly we have

$$\begin{aligned}X'_{(J)}W^{-1}_{(J)}X_{(J)} &= [X'_1\ldots X'_j\ldots X'_J]\begin{bmatrix} W_1^{-1} & & & & 0 \\ & \ddots & & & \\ & & W_j^{-1} & & \\ & & & \ddots & \\ 0 & & & & W_J^{-1}\end{bmatrix}\begin{bmatrix} X'_1 \\ \vdots \\ X'_j \\ \vdots \\ X'_J\end{bmatrix} \\ &= \sum_{j=1}^{J} X'_jW_j^{-1}X_j = JX'_jW_j^{-1}X_j.\end{aligned}$$

Similarly, we have

$$X'_{(J)}W^{-1}_{(J)}Z_{(J)} = JX'_jW_j^{-1}Z_j,$$

$$Z'_{(J)}W^{-1}_{(J)}Z_{(J)} = JZ'_jW_j^{-1}Z_j,$$

$$M'_{(J)}W^{-1}_{(J)}M_{(J)} = JM'_jW_j^{-1}M_j,$$

$$X'_{(J)}W^{-1}_{(J)}V_{(J)} = [X'_1W_1^{-1}V_1 \ \ldots\ X'_jW_j^{-1}V_j \ \ldots\ X'_JW_J^{-1}V_J],$$

$$M'_{(J)}W^{-1}_{(J)}V_{(J)} = [M'_1W^{-1}_1V_1 \; \ldots \; M'_jW^{-1}_jV_j \; \ldots \; M'_JW^{-1}_JV_J],$$

$$X'_{(J)}W^{-1}_{(J)} = [X'_1W^{-1}_1 \; \ldots \; X'_jW^{-1}_j \; \ldots \; X'_JW^{-1}_J], \text{ and}$$

$$M'_{(J)}W^{-1}_{(J)} = [M'_1W^{-1}_1 \; \ldots \; M'_jW^{-1}_j \; \ldots \; M'_JW^{-1}_J].$$

Also we have

$$W^{-1}_{(J)}X_{(J)}(X'_{(J)}W^{-1}_{(J)}X_{(J)})^{-}X'_{(J)}W^{-1}_{(J)}V_{(J)} =$$

$$\frac{1}{J}W^{-1}_jX_j(X'_jW^{-1}_jX_j)^{-}X'_jW^{-1}_jV_j\begin{bmatrix} I & I & \ldots & I \\ I & I & \ldots & I \\ \vdots & & & \\ I & I & \ldots & I \end{bmatrix}_{IJK} ,$$

$$W^{-1}_{(J)}M_{(J)}(M'_{(J)}W^{-1}_{(J)}M_{(J)})^{-}M'_{(J)}W^{-1}_{(J)}V_{(J)} =$$

$$\frac{1}{J}W^{-1}_jM_j(M'_jW^{-1}_jM_j)^{-}M'_jW^{-1}_jV_j\begin{bmatrix} I & I & \ldots & I \\ I & I & \ldots & I \\ \vdots & & & \\ I & I & \ldots & I \end{bmatrix}_{IJK} ,$$

$$W^{-1}_{(J)}M_{(J)}(M'_{(J)}W^{-1}_{(J)}M_{(J)})^{-}M'_{(J)}W^{-1}_{(J)} =$$

$$\frac{1}{J}W^{-1}_jM_j(M'_jW^{-1}_jM_j)^{-}M'_jW^{-1}_j\begin{bmatrix} I & I & \ldots & I \\ I & I & \ldots & I \\ \vdots & & & \\ I & I & \ldots & I \end{bmatrix}_{IJK} , \text{ and}$$

$$W^{-1}_{(J)}X_{(J)}(X'_{(J)}W^{-1}_{(J)}X_{(J)})^{-}X'_{(J)}W^{-1}_{(J)} =$$

$$\frac{1}{J}W_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}X_j'W_j^{-1}\begin{bmatrix} I & I & \cdots & I \\ I & I & \cdots & I \\ \vdots & & & \\ I & I & \cdots & I \end{bmatrix}_{IJK}.$$

Therefore, the first term in (5.6) can be written as

$$2\,\mathrm{tr}\,[W_j^{-1}M_j(M_j'W_j^{-1}M_j)^{-}M_j'W_j^{-1}V_jW_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}X_j'W_j^{-1}V_j]^2 \div$$

$$J^2(\mathrm{tr}\,[Z_j'W_j^{-1}Z_j - Z_j'W_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}X_j'W_j^{-1}Z_j])^2$$

$\longrightarrow 0$ as $J \longrightarrow \infty$. Also, the second term in (5.6) can be written as

$$4\underset{\sim}{\beta}'X_j'[W_j^{-1}M_j(M_j'W_j^{-1}M_j)^{-}M_j'W_j^{-1}V_j + W_j'M_j(M_j'W_j^{-1}M_j)^{-}W_j^{-1} -$$

$$W_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}X_j'W_j^{-1}V_j - W_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}W_j^{-1}]X_j\underset{\sim}{\beta} \div$$

$$J(\mathrm{tr}\,[Z_j'W_j^{-1}Z_j - Z_j'W_j^{-1}X_j(X_j'W_j^{-1}X_j)^{-}X_j'W_j^{-1}Z_j])^2$$

$\longrightarrow 0$ as $J \longrightarrow \infty$. Therefore, we have $\mathrm{Var}(\hat{\sigma}_\varepsilon^2) \longrightarrow 0$ as $J \longrightarrow \infty$, which implies that $\hat{\sigma}_\varepsilon^2$ with a known $W_{(J)}$, $\hat{\sigma}^2(W)$ say, is a consistent estimator for σ_ε^2 as $J \longrightarrow \infty$. This completes the proof of step (4).

To prove step (5) we use steps (3) and (4) as follows. Let $\hat{\sigma}_\varepsilon^2(\hat{W})$ be a continuous function of $W_{(J)}$, and since $\hat{W}_{(J)} - W_{(J)} \xrightarrow{P} 0$ as $J \longrightarrow \infty$, then we have

$$\hat{\sigma}_\varepsilon^2(\hat{W}) - \sigma_\varepsilon^2 \xrightarrow{P} 0 \text{ as } J \longrightarrow \infty.$$

For the proof of step (6), we proceed as follows.

$$X_{(J)}'V_{(J)}^{-1}X_{(J)} = J(X_j'V_j^{-1}X_j), \quad \text{and} \quad X_{(J)}'V_{(J)}^{-1}\underset{\sim}{y}_{(J)} = \sum_{j=1}^{J} X_j'V_j^{-1}\underset{\sim}{y}_j.$$

Therefore $\hat{\beta}$ based on known $V_{(J)}$, $\hat{\beta}(V)$ say, is given by

$$\hat{\beta}(V) = (X'_{(J)}V^{-1}_{(J)}X_{(J)})^{-1}X'_{(J)}V^{-1}_{(J)}y_{(J)}$$

$$= [(X'_jV_j^{-1}X_j)^{-1}\sum_{j=1}^{J} X'_jV_j^{-1}y_j]/J.$$

Now,

$\sum_{j=1}^{J} X'_jV_j^{-1}y_j/J \xrightarrow{P} E(X'_jV_j^{-1}y_j)$ as $J \longrightarrow \infty$ by the weak law of large numbers. Therefore, we have

$$\hat{\beta}(V) \xrightarrow{P} (X'_jV_j^{-1}X_j)^{-1}E(X'_jV_j^{-1}y_j) =$$

$$(X'_jV_j^{-1}X_j)^{-1}X'_jV_j^{-1}E(y_j) = (X'_jV_j^{-1}X_j)^{-1}X'_jV_j^{-1}X_j\beta = \beta$$

as $J \longrightarrow \infty$. Thus $\hat{\beta}(V)$ is a consistent estimator for β as $J \longrightarrow \infty$. Now we define a function of $\bar{y}_{(J)}$ as

$$\hat{\beta}(V) = (X'_jV_j^{-1}X_j)^{-1}X'_jV_j^{-1}\bar{y}_{(J)}$$

and use (5.5). Applying the delta method then we obtain the following result.

$$\sqrt{J}(\hat{\beta}(V) - \beta) \xrightarrow{L} N_p(0,(X'_jV_j^{-1}X_j)^{-1}), \qquad \ldots \quad (5.7)$$

as $J \longrightarrow \infty$. This completes the proof of step (6).

Using steps (3) and (5) we have

$\hat{\sigma}^2_{\delta_{ijk}} - \sigma^2_{\delta_{ijk}} \xrightarrow{P} 0$ as $J \longrightarrow \infty$ for all $i = 1,\ldots,I$, $j = 1,\ldots,J$, and $k = 1,\ldots,K$, and

$\hat{\sigma}^2_\varepsilon - \sigma^2_\varepsilon \xrightarrow{P} 0$ as $J \longrightarrow \infty$. Then by Arnold's (1981) result on page 341

$\hat{V}_j \xrightarrow{P} V_j$ as $J \longrightarrow \infty$ and also we get

$\hat{V}_{(J)} - V_{(J)} \xrightarrow{P} 0$ as $J \longrightarrow \infty$. This completes the proof of step (7).

From the last two results we have

$$X_j'\hat{V}_j^{-1} \xrightarrow{P} X_j'V_j^{-1}, \qquad \ldots \quad (5.8)$$

as $J \longrightarrow \infty$.

$$X_j'\hat{V}_j^{-1}X_j \xrightarrow{P} X_j'V_j^{-1}X_j, \qquad \ldots \quad (5.9)$$

as $J \longrightarrow \infty$. Also

$$(X_j'\hat{V}_j^{-1}X_j)^{-1} \xrightarrow{P} (X_j'V_j^{-1}X_j)^{-1}, \qquad \ldots \quad (5.10)$$

as $J \longrightarrow \infty$.

Now, $\hat{\underset{\sim}{\beta}}$ based on unknown $V_{(J)}$, $\hat{\underset{\sim}{\beta}}(\hat{V})$ say, is given by

$$\hat{\underset{\sim}{\beta}}(\hat{V}) = [(X_j'\hat{V}_j^{-1}X_j)^{-1} \sum_{j=1}^{J} X_j'\hat{V}_j^{-1}\underset{\sim}{y}_j]/J$$

$$= (X_j'\hat{V}_j^{-1}X_j)^{-1}X_j'\hat{V}_j^{-1}(\sum_{j=1}^{J} \underset{\sim}{y}_j/J).$$

Thus, using (5.8), (5.10) and since $\sum_{j=1}^{J} \underset{\sim}{y}_j/J \xrightarrow{P} E(\underset{\sim}{y}_j)$ as $J \longrightarrow \infty$ we have

$$\hat{\underset{\sim}{\beta}}(\hat{V}) \xrightarrow{P} (X_j'V_j^{-1}X_j)^{-1}X_j'V_j^{-1}X_j\underset{\sim}{\beta} = \underset{\sim}{\beta} \quad \text{as } J \longrightarrow \infty.$$

Therefore $\hat{\underset{\sim}{\beta}}(\hat{V})$ is consistent estimator for $\underset{\sim}{\beta}$ as $J \longrightarrow \infty$. Now, since $\hat{\underset{\sim}{\beta}}(V)$ and $\hat{\underset{\sim}{\beta}}(\hat{V})$ are both consistent estimators for $\underset{\sim}{\beta}$ then we have

$$\hat{\underset{\sim}{\beta}}(\hat{V}) - \hat{\underset{\sim}{\beta}}(V) \xrightarrow{P} 0 \quad \text{as} \quad J \longrightarrow \infty.$$

Therefore

$$\sqrt{J}(\hat{\underset{\sim}{\beta}}(\hat{V}) - \hat{\underset{\sim}{\beta}}(V)) =$$

$$[(X_j'\hat{V}_j^{-1}X_j)^{-1}X_j'\hat{V}_j^{-1}\sum_{j=1}^{J} \underset{\sim}{y}_j - (X_j'V_j^{-1}X_j)^{-1}X_j'V_j^{-1}\sum_{j=1}^{J} \underset{\sim}{y}_j]/\sqrt{J} =$$

$$= [(X_j'\hat{V}_j^{-1}X_j)^{-1}X_j'\hat{V}_j^{-1} - (X_j'V_j^{-1}X_j)^{-1}X_j'V_j^{-1}]\sqrt{J}\bar{\underset{\sim}{y}}_{(J)}.$$

Combining (5.5), (5.8) and (5.10) leads to

$$\sqrt{J}(\hat{\underset{\sim}{\beta}}(\hat{V}) - \hat{\underset{\sim}{\beta}}(V)) \xrightarrow{P} 0, \qquad \ldots \quad (5.11)$$

as $J \longrightarrow \infty$. From (5.7) we have

$$\sqrt{J}(\hat{\underset{\sim}{\beta}}(V) - \underset{\sim}{\beta}) \overset{\bullet}{\underset{\bullet}{\sim}} N_p(0,(X_j'V_j^{-1}X_j)^{-1}),$$

which implies that for all $\underset{\sim}{a} \in \mathbb{R}^p$ we have

$$\sqrt{J}\underset{\sim}{a}'(\hat{\underset{\sim}{\beta}}(V) - \underset{\sim}{\beta})[\underset{\sim}{a}'(X_j'V_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{-\frac{1}{2}} \xrightarrow{L} N(0,1), \qquad \ldots \quad (5.12)$$

as $J \longrightarrow \infty$.

Now, using (5.10) and for all $\underset{\sim}{a} \in \mathbb{R}^p$ we have

$$\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a} \xrightarrow{P} \underset{\sim}{a}'(X_j'V_j^{-1}X_j)^{-1}\underset{\sim}{a} \quad \text{as} \quad J \longrightarrow \infty.$$

This can be written equivalently as

$$\frac{[\underset{\sim}{a}'(X_j'V_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{1/2}}{[\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{1/2}} \xrightarrow{P} 1, \qquad \ldots \quad (5.13)$$

as $J \longrightarrow \infty$. Multiplying (5.12) and (5.13) and using Slutsky's theorem imply that

$$\sqrt{J}\ \underset{\sim}{a}'(\hat{\underset{\sim}{\beta}}(V) - \underset{\sim}{\beta})]\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{-\frac{1}{2}} \xrightarrow{L} N(0,1)$$

as J ⟶ ∞. By adding and subtracting $\hat{\underset{\sim}{\beta}}(\hat{V})$, the last result can be rewritten as

$$\frac{\sqrt{J}\ \underset{\sim}{a}'(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta})}{[\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{1/2}} + \frac{\sqrt{J}\ \underset{\sim}{a}'(\hat{\underset{\sim}{\beta}}(V) - \hat{\beta}(\hat{V}))}{[\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{1/2}} \xrightarrow{L} N(0,1), \qquad \ldots \quad (5.14)$$

as J ⟶ ∞.

Combining (5.11) and (5.14) we have

$$\sqrt{J}\ \underset{\sim}{a}'(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta})[\underset{\sim}{a}'(X_j'\hat{V}_j^{-1}X_j)^{-1}\underset{\sim}{a}]^{-\frac{1}{2}} \xrightarrow{L} N(0,1)$$

as J ⟶ ∞ for all $\underset{\sim}{a} \in \mathbb{R}^p$. Thus we have

$$\sqrt{J}(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta}) \overset{\bullet}{\underset{\bullet}{\sim}} N_p(\underset{\sim}{0},(X_j'\hat{V}_j^{-1}X_j)^{-1}), \qquad \ldots \quad (5.15)$$

as J ⟶ ∞. This completes the proof of step (8).

5.3 Testing Hypotheses and Confidence Intervals

5.3.1 Testing the Hypothesis H_0: $H\underset{\sim}{\beta} = \underset{\sim}{0}$ vs. H_1: $H\underset{\sim}{\beta} \neq \underset{\sim}{0}$

The test statistic for testing the above hypothesis is given by

$$SS(H\underset{\sim}{\beta} = \underset{\sim}{0}) = \hat{\underset{\sim}{\beta}}'(\hat{V})H'[H(X'\hat{V}^{-1}X)^{-1}H']^{-1}H\hat{\underset{\sim}{\beta}}(\hat{V}).$$

Using (5.15) we have

$$\sqrt{J}\ H(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta}) \overset{\bullet}{\underset{\bullet}{\sim}} N_p(0,H(X_j'\hat{V}_j^{-1}X_j)^{-1}H')$$

as $J \longrightarrow \infty$. Then we have

$$J(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta})'H'[H(X_j'\hat{V}_j^{-1}X_j)^{-1}H']^{-1}H(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta}) \overset{\cdot}{\underset{\cdot}{\sim}} \chi^2(\text{rank}(H))$$

as $J \longrightarrow \infty$. This then implies that

$$(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta})'H'[H(X'\hat{V}^{-1}X)^{-1}H']^{-1}H(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta}) \overset{\cdot}{\underset{\cdot}{\sim}} \chi^2(\text{rank}(H))$$

as $J \longrightarrow \infty$. Hence, $SS(H\underset{\sim}{\beta} = \underset{\sim}{0})$ has an asymptotic χ^2 distribution with rank(H) d.f. if H_0 is true.

5.3.2 Goodness of Fit Test

The test statistic is given by

$$GOF = (\underset{\sim}{y} - X\hat{\beta}(\hat{V}))'\hat{V}^{-1}(\underset{\sim}{y} - X\hat{\beta}(\hat{V})) = \underset{\sim}{y}'[\hat{V}^{-1} - \hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}]\underset{\sim}{y}.$$

GOF is the sum of squares to test the hypothesis that $\log(-\log \underset{\sim}{q}) = X\underset{\sim}{\beta}$ and that δ_{ijk} are with distribution derived from conditional binomial assumption. If $X\underset{\sim}{\beta}$ is a saturated model, then this is a test for binomial assumption or independence of survival times conditional on main unit. Using Arnold's (1981) theorem 10.3 we have

$$GOF \overset{\cdot}{\underset{\cdot}{\sim}} \chi^2(\text{d.f.}) \text{ as } J \longrightarrow \infty, \text{ where}$$

$$\text{d.f.} = \text{rank } (\hat{V}^{-1} - \hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}).$$

Therefore, GOF has an asymptotic χ^2 distribution if the model fits.

5.3.3 Confidence Limits for β_i and Confidence Region for β

Using (5.15) and defining an appropriate function of $\hat{\underset{\sim}{\beta}}(\hat{V})$ we have

$$(c_{ii})^{-\frac{1}{2}}(\hat{\beta}_i - \beta_i) \xrightarrow{L} N(0,1) \quad \text{as} \quad J \longrightarrow \infty,$$

where c_{ii} is the appropriate diagonal element of $(X'\hat{V}^{-1}X)^{-1}$ corresponding to the element $\hat{\beta}_i$. Thus an approximate 100(1-α)% confidence interval for β_i is given by

$\hat{\beta}_i \pm z_{\alpha/2}(c_{ii})^{1/2}$, where $z_{\alpha/2}$ is the value such that a standard normal variate falls within $-z_{\alpha/2}$ and $z_{\alpha/2}$ with probability 1-α.

Also, using (5.15) and defining an appropriate function of $\hat{\underset{\sim}{\beta}}(\hat{V})$ we can conclude that any subvector $\hat{\underset{\sim}{\beta}}(\hat{V})$ is asymptotically normally distributed and then an approximate 100(1-α)% confidence region is given by

$$\{\underset{\sim}{\beta} \in \mathbb{R}^P : (\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta})'X'\hat{V}^{-1}X(\hat{\underset{\sim}{\beta}}(\hat{V}) - \underset{\sim}{\beta}) \leq \chi^2_\alpha(P)\},$$

where $\chi^2_\alpha(P)$ is a tabulated chi-square value with α-level and P degrees of freedom.

5.4 Other Asymptotic Properties

5.4.1 Confidence Limits for the Conditional Survivor Functions

To construct confidence limits for the conditional survivor function given by

$$S_{i|\varepsilon_{ij}=0}(t_k) = \prod_{\ell=1}^{k} q_{ij\ell},$$

Prentice and Gloeckler (1978) suggested using the following function

$$\gamma_i(\hat{\underset{\sim}{\beta}}) = \log(-\log S_{i|\varepsilon_{ij}=0}(t_k,\hat{\underset{\sim}{\beta}})).$$

Then the distribution of $\gamma_i(\hat{\underset{\sim}{\beta}})$ may be approximated by a normal distribution with mean $\gamma_i(\underset{\sim}{\beta})$ and variance given by

$\sigma_i^2 = [g_i(\hat{\underset{\sim}{\beta}})]'\mathrm{cov}(\hat{\underset{\sim}{\beta}})[g_i(\hat{\underset{\sim}{\beta}})]$, where

$$g_i(\underset{\sim}{\beta}) = \frac{\partial \gamma_i(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}}.$$

Therefore, approximate $100(1-\alpha)\%$ confidence interval for $\gamma_i(\underset{\sim}{\beta})$ is given by

$[\gamma_i(\hat{\underset{\sim}{\beta}}) - z_{\alpha/2}\sigma_i\, , \gamma_i(\hat{\underset{\sim}{\beta}}) + z_{\alpha/2}\sigma_i]$, where $z_{\alpha/2}$ is the value such that a standard normal variate falls within $-z_{\alpha/2}$ and $z_{\alpha/2}$ with probability $1-\alpha$. Thus, an approximate $100(1-\alpha)\%$ confidence interval for $S_{i|\varepsilon_{ij}=0}(t_k,\underset{\sim}{\beta})$ is given by

$$[\{S_{i|\varepsilon_{ij}=0}(t_k,\hat{\underset{\sim}{\beta}})\}^{\exp(z_\alpha\sigma_i)}\, , \{S_{i|\varepsilon_{ij}=0}(t_k,\hat{\underset{\sim}{\beta}})\}^{\exp(-z_\alpha\sigma_i)}]. \quad \ldots \quad (5.20)$$

5.4.2 Test for the Binomial Variabilities

An ad-hoc test to see whether the binomial variances, that were used to estimate main unit variance, are appropriate or not is provided. This test can be summarized as follows. Add, as covariate, the values of n_{ijk} to the fitted model $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u}$, say. Then obtain a new model $\underset{\sim}{y} = X\underset{\sim}{\beta} + \gamma\underset{\sim}{n} + \underset{\sim}{u}$, where $\underset{\sim}{n}$ is the vector of n_{ijk} values for $i = 1,\ldots,I$, $j = 1,\ldots,J$, and $k = 1,\ldots,K$. If $\gamma=0$ then the estimates of q_{ijk}'s as functions of the parameters for the above model are unaffected by risk sets, and hence risk sets do not affect the estimated survivor functions. However, if $\gamma > 0$ then the estimates of q_{ijk}'s are larger with larger number of subunits at risk, and hence risk sets will affect the estimated survivor functions. On the other hand if $\gamma < 0$ then the estimates of q_{ijk}'s are smaller with larger number of subunits at risk, and hence risk sets will affect the estimated survivor functions. Therefore a test on γ is a test for the validity of the assumption that binomial variances are appropriate. A test for H_0: $\gamma = 0$ against H_1: $\gamma \neq 0$ is now given by $z = \hat{\gamma}/s_{\hat{\gamma}}$, where $\hat{\gamma}$ and $s_{\hat{\gamma}}$ are the corresponding estimate and standard error of the estimate of γ, respectively. If units fail independently, then the binomial distribution applies. If q_{ijk} depends on the number at risk, then the units will not fail independently. Small p-values indicate lack of independence of subunits within main unit.

CHAPTER VI

EXAMPLE OF MODEL AND ANALYSIS APPLIED TO A REAL DATA SET USING SAS

In this chapter analysis of the fish experiment presented and fully explained in Chapter I will be considered. SAS was used for all computations. The full model is given by

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u}, \quad \text{where } \operatorname{cov}(y) = V \text{ and}$$

V is a block diagonal variance-covariance matrix involving both $\sigma^2_{\delta_{ijk}}$ for $i = 1,\ldots,I$, $j = 1,\ldots,J$, and $k = 1,\ldots,K$, and σ^2_{ε}. $\underset{\sim}{\beta}$ is the vector of of unknown parameters to be estimated, X is a design matrix of known constants, and $\underset{\sim}{y}$ is a vector of transformed values of the observed $\hat{q}_{ijk}$'s. The function of these $\hat{q}_{ijk}$'s that was considered in the analysis is given by $y_{ijk} = \log(-\log \hat{q}_{ijk})$ for all $i = 1,\ldots,I$, $j = 1,\ldots,J$ and $k = 1,\ldots,K$.

Estimates of $\sigma^2_{\delta_{ijk}}$ can be found using (4.1). Table IV represents the calculations involved to get $\hat{\sigma}^2_{\delta_{ijk}}$ and $y_{ijk} = \log(-\log \hat{q}_{ijk})$. It should be mentioned here that in order to avoid having values for $\hat{q}_{ijk}$ as one or zero adjusted survival, $s_{ijk}(AD)$ say, can be used instead of the s_{ijk} values as follows. If $s_{ijk} = n_{ijk}$ then take $s_{ijk}(AD) = s_{ijk} - .5$. If $s_{ijk} = 0$ then take $s_{ijk}(AD) = .5$. On the other hand for occasional $n_{ijk} = 0$, we follow Grizzle, Starmer, and Koch (1969), and suggest that s_{ijk} can be replaced by $1/K$, where K is the number of time intervals.

TABLE IV

ESTIMATES OF BINOMIAL VARIANCES AND VALUES OF THE RESPONSE VARIABLE

Accl.	Conc.	Tank	Time	r_{ijk}	n_{ijk}	s_{ijk}	$\hat{\sigma}^2_{\delta_{ijk}}$	y_{ijk}
1	1	1	1	1	25	24	0.74715	-3.1985
1	1	1	2	5	24	19	0.1728	-1.4559
1	1	1	3	7	19	12	0.19855	-0.779
1	1	1	4	1	12	11	0.7531	-2.4459
1	1	1	5	0	11	11	2.08315	-3.0782
1	1	1	6	0	11	11	2.08315	-3.0782
1	1	1	7	0	11	11	2.08315	-3.0782
1	1	2	1	2	25	23	0.74745	-2.4843
1	1	2	2	7	23	16	0.1728	-1.015
1	1	2	3	4	16	12	0.19855	-1.2459
1	1	2	4	2	12	10	0.7531	-1.6998
1	1	2	5	0	10	10	2.08315	-2.9702
1	1	2	6	0	10	10	2.08315	-2.9702
1	1	2	7	0	10	10	2.08315	-2.9702
1	2	1	1	3	25	22	1.1878	-2.057
1	2	1	2	7	22	15	0.12345	-0.9604
1	2	1	3	9	15	6	0.1333	-0.0874
1	2	1	4	0	6	6	1.1138	-2.4459
1	2	1	5	0	6	6	1.5129	-2.4459
1	2	1	6	0	6	6	2.012	-2.4459
1	2	1	7	0	6	6	1.5258	-2.4459
1	2	2	1	0	25	25	1.1878	-3.9019
1	2	2	2	10	25	15	0.12345	-0.6717
1	2	2	3	7	15	8	0.1333	-0.4633
1	2	2	4	5	8	3	1.1138	-0.0194
1	2	2	5	1	3	2	1.5129	-0.904
1	2	2	6	0	2	2	2.012	-1.2459
1	2	2	7	1	2	1	1.5258	-0.3665
1	3	1	1	1	25	24	0.9804	-3.1985
1	3	1	2	12	24	12	0.0946	-0.3665
1	3	1	3	7	12	5	0.1425	0.1339
1	3	1	4	4	5	1	0.3279	0.4759
1	3	1	5	0	1	1	1.5479	-0.3665
1	3	1	6	1	1	0	1.5609	-0.3665
1	3	1	7	0	0	0	2.0812	-0.3665
1	3	2	1	1	25	24	0.9804	-3.1985
1	3	2	2	10	24	14	0.0946	-0.617
1	3	2	3	8	14	6	0.1425	-0.1669
1	3	2	4	3	6	3	0.3279	-0.3665
1	3	2	5	1	3	2	1.5479	-0.904
1	3	2	6	1	2	1	1.5609	-0.3665
1	3	2	7	0	1	1	2.0812	-0.3665
-1	1	1	1	0	25	25	2.0408	-3.9019
-1	1	1	2	9	25	16	0.1818	-0.8068
-1	1	1	3	4	16	12	0.2516	-1.2459
-1	1	1	4	0	12	12	1.9909	-3.1487
-1	1	1	5	0	12	12	1.9909	-3.1487
-1	1	1	6	0	12	12	1.9909	-3.1487
-1	1	1	7	0	12	12	1.9909	-3.1487
-1	1	2	1	0	25	25	2.0408	-3.9019
-1	1	2	2	4	25	21	0.1818	-1.7467
-1	1	2	3	4	21	17	0.2516	-1.5572
-1	1	2	4	0	17	17	1.9909	-3.5258

TABLE IV (Continued)

Accl.	Conc.	Tank	Time	r_{ijk}	n_{ijk}	s_{ijk}	$\hat{\sigma}^2_{\delta_{ijk}}$	y_{ijk}
-1	1	2	5	0	17	17	1.9909	-3.5258
-1	1	2	6	0	17	17	1.9909	-3.5258
-1	1	2	7	0	17	17	1.9909	-3.5258
-1	2	1	1	1	25	24	1.5106	-3.1985
-1	2	1	2	12	24	12	0.0999	-0.3665
-1	2	1	3	5	12	7	0.27	-0.617
-1	2	1	4	3	7	4	0.421	-0.5792
-1	2	1	5	0	4	4	2.0231	-2.0134
-1	2	1	6	0	4	4	2.0231	-2.0134
-1	2	1	7	0	4	4	2.0231	-2.0134
-1	2	2	1	0	25	25	1.5106	-3.9019
-1	2	2	2	9	25	16	0.0999	-0.8068
-1	2	2	3	3	16	13	0.27	-1.5749
-1	2	2	4	2	13	11	0.421	-1.7883
-1	2	2	5	0	11	11	2.0231	-3.0782
-1	2	2	6	0	11	11	2.0231	-3.0782
-1	2	2	7	0	11	11	2.0231	-3.0782
-1	3	1	1	3	25	22	1.1877	-2.057
-1	3	1	2	12	22	10	0.0871	-0.2389
-1	3	1	3	3	10	7	0.2436	-1.0309
-1	3	1	4	2	7	5	0.5058	-1.0881
-1	3	1	5	1	5	4	1.5052	-1.4999
-1	3	1	6	0	4	4	2.0064	-2.0134
-1	3	1	7	0	4	4	2.0064	-2.0134
-1	3	2	1	0	25	25	1.1877	-3.9019
-1	3	2	2	12	25	13	0.0871	-0.4248
-1	3	2	3	7	13	6	0.2436	-0.2585
-1	3	2	4	2	6	4	0.5058	-0.904
-1	3	2	5	0	4	4	1.5052	-2.0134
-1	3	2	6	0	4	4	2.0064	-2.0134
-1	3	2	7	0	4	4	2.0064	-2.0134

From Table IV and (4.2), we have $\hat{\sigma}_{\varepsilon}^{2} = 0.0095$. Although this estimate is small it will be considered throughout. After constructing $\hat{V}$, a weighted least squares procedure is used to fit the full model.

Test for the appropriateness of using the binomial variabilities is considered as proposed in Chapter V. The full model with n_{ijk} as a covariable has been fitted. Estimate and standard error for the appropriate coefficient on this covariable are given by $\hat{\beta}_{n_{ijk}} = -0.0683$ and s.e. $= 0.0553$. Then a test for H_0: $\beta_{n_{ijk}} = 0$ against H_0: $\beta_{n_{ijk}} \neq 0$ is given by $z = -1.23541$ with a p-value of 0.11. This relatively large p-value indicates that there is no evidence that the response variable $\underset{\sim}{y}$ depends on risk sets through anything other than the binomial variances. Table V represents an analysis of deviation table with corresponding χ^2 and p-values obtained from fitting the full model without n_{ijk} as covariable. From Table V, it is obvious that the model fits very well. However, a need for another model that might fit just as well but with less factors was considered. Thus, a search for a better fitted model was carried out. The test of significance for each factor is obtained by considering its performance with all other factors included in the model. The model that has been chosen to be the appropriate one for this experiment is the one that has only main treatment effects along with time effects. Since this best fitting model has no time by treatments interactions the proportional hazard for continuous time setting is appropriate.

For the chosen model, Table VI represents the analysis of deviation with the resulting χ^2 and p-values. Table VII represents the estimate of $\underset{\sim}{\beta}$ and the standard errors of the estimates, where

$$\underset{\sim}{\hat{\beta}} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}\underset{\sim}{y} \quad \text{and} \quad \widehat{\mathrm{Cov}}(\underset{\sim}{\hat{\beta}}) = (X'\hat{V}^{-1}X)^{-1}.$$

TABLE V

ANALYSIS OF DEVIATION FOR THE FULL MODEL

Source	d.f.	χ^2	p-value
Accl.	1	11.8354	<0.005
Conc.	2	19.7773	<0.005
Accl. × Conc.	2	3.5608	0.290
Time	6	18.3412	0.005
Accl. × Time	6	7.6430	0.380
Conc. × Time	12	8.1059	0.780
Goodness of Fit	48	20.9127	0.995

All $\hat{\sigma}_{\delta_{ijk}}$ and $\hat{\sigma}^2_{\varepsilon}$ are consistent estimators as the number of tanks $J \longrightarrow \infty$. Furthermore, $\hat{\underset{\sim}{\beta}}(\hat{\underset{\sim}{V}})$ has all the asymptotic properties discussed in Chapter V. We should mention here that the asymptotic properties discussed in Chapter V are all valid approximations if we have more than one repetition per treatment. As the number of those repetitions goes to ∞, the estimators are consistent. These results are valid for the split plot design with main treatments being completely randomized to main units. Also the asymptotic properties hold for the split plot design with main units having a completely randomized block structure for the case of more than one repetition per treatment within blocks.

Estimates and plots of the survivor functions along with their confidence limits for each treatment combination are presented in Table VIII and Figures 1, 2, 3, 4, 5, and 6, respectively. Also estimated survivor functions of the two acclimation times for low, medium, and high levels of zinc concentration are presented in Figures 7, 8, and 9, respectively. We should mention here that there are two acclimation times namely one week and two weeks. Also there are three levels of zinc concentration namely low, medium, and high. Treatments 1, 3, and 5 represent one week of acclimation time with low, medium, and high levels of zinc concentration. On the other hand treatments 2, 4, and 6 represent two weeks of acclimation time with low, medium, and high levels of zinc concentration, respectively.

From the above analysis we conclude that for $\varepsilon_{ij} = 0$, the effect of the acclimation time was important in explaining the data. For the first two time intervals there was practically no difference in survival rates between acclimation times of one week and two weeks. Fish under two week acclimations survived better than those with one week acclimation time in the sense that the effect became greater with time. This suggests it is better to collect the data (count the number of deaths) after a period of at least three days. There was also an effect due to zinc concentration which indicates that fish survive better with low levels of zinc concentration than for higher levels.

TABLE VI

ANALYSIS OF DEVIATION FOR THE CHOSEN MODEL

Source	d.f.	χ^2	p-value
Accl.	1	3.203	0.0782
Conc.	2	29.515	< 0.0050
Time	6	81.841	< 0.0001
Goodness of fit	68	46.972	0.9000

TABLE VII

ESTIMATE OF $\underset{\sim}{\beta}$

Parameter	Estimate	s.e.
μ	-1.741860	0.11802
a_1	-0.561034	0.11291
c_1	0.103323	0.10440
c_2	0.133520	0.07461
τ_1	-1.432720	0.28642
τ_2	0.945690	0.14188
τ_3	1.016890	0.15675
τ_4	0.574602	0.22061
τ_5	-0.362336	0.34149
τ_6	-0.388477	0.35749

TABLE VIII

ESTIMATES OF SURVIVOR FUNCTIONS

	$S_{i\|\varepsilon_{ij}=0}$	L.L.	U.L.	Time
Trt 1	0.973095	0.9487	0.986	1
	0.725115	0.6342	0.797	2
	0.528728	0.4157	0.6295	3
	0.431604	0.312	0.5455	4
	0.398606	0.28	0.5149	5
	0.368886	0.2515	0.4865	6
	0.340447	0.2255	0.4588	7
Trt 2	0.979334	0.9599	0.9894	1
	0.781842	0.7676	0.8767	2
	0.613888	0.5055	0.7056	3
	0.525531	0.405	0.6326	4
	0.494482	0.3721	0.6054	5
	0.466	0.3428	0.5799	6
	0.438235	0.3151	0.5545	7
Trt 3	0.948384	0.9026	0.973	1
	0.535495	0.4242	0.6346	2
	0.289875	0.1907	0.3964	3
	0.195403	0.1115	0.2966	4
	0.167422	0.0905	0.2645	5
	0.144021	0.0738	0.2368	6
	0.123231	0.06	0.2104	7
Trt 4	0.960235	0.9237	0.9794	1
	0.619895	0.5212	0.7041	2
	0.387467	0.2782	0.4953	3
	0.286478	0.1844	0.3966	4
	0.254508	0.1574	0.3632	5
	0.226797	0.1345	0.3337	6
	0.201279	0.1147	0.3053	7
Trt 5	0.927249	0.8647	0.9612	1
	0.410589	0.3014	0.5173	2
	0.171204	0.0971	0.263	3
	0.0975879	0.0464	0.1715	4
	0.0782952	0.0346	0.1454	5
	0.0631745	0.026	0.1238	6
	0.0505878	0.0193	0.1049	7
Trt 6	0.943808	0.87	0.9708	1
	0.505825	0.3333	0.6031	2
	0.258899	0.2544	0.3657	3
	0.168352	0.1157	0.2642	4
	0.142224	0.0998	0.233	5
	0.120676	0.0676	0.2063	6
	0.101797	0.0232	0.1818	7

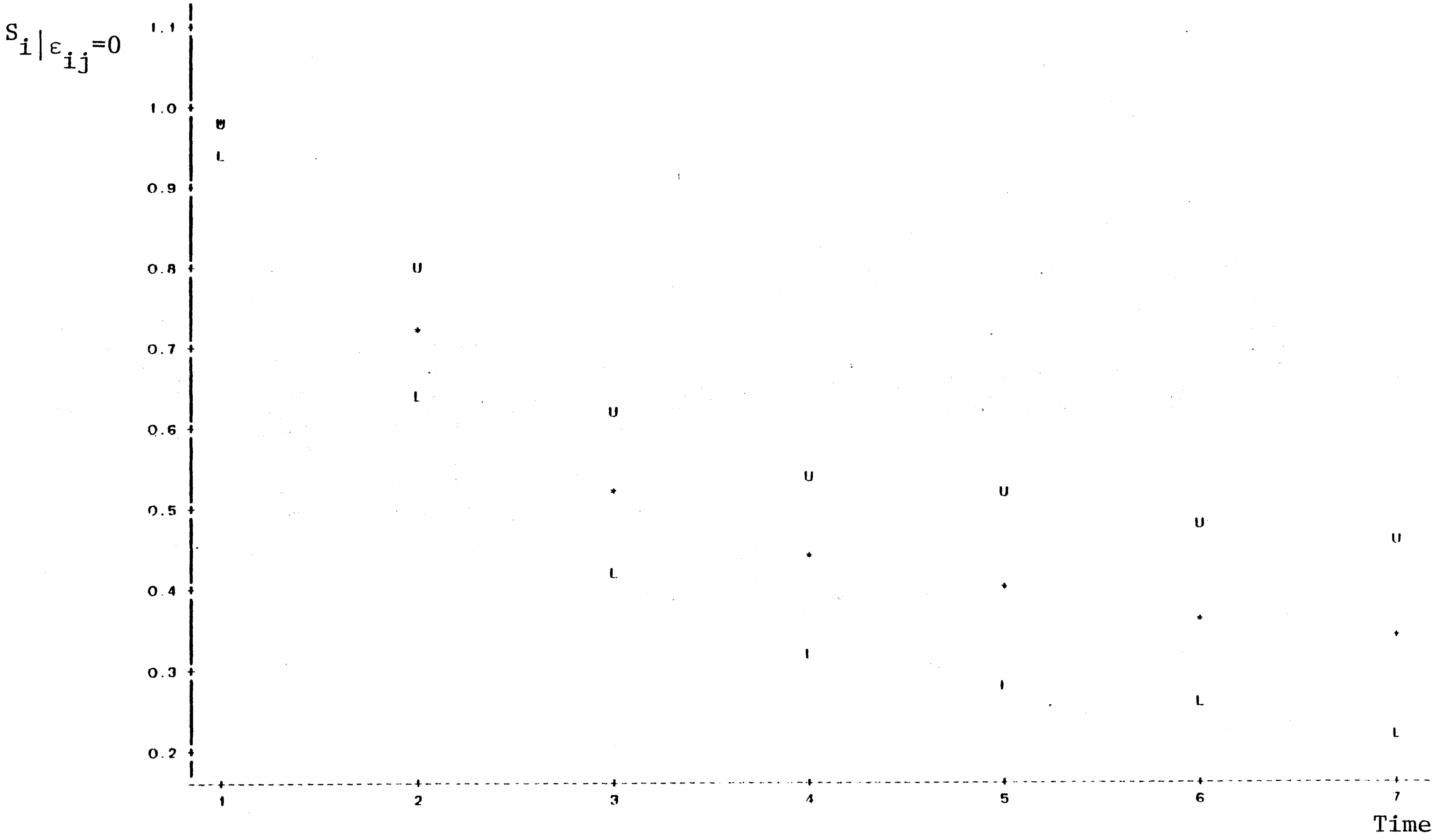

Figure 1. Estimated Survivor Function with Its Confidence Limits for Trt 1

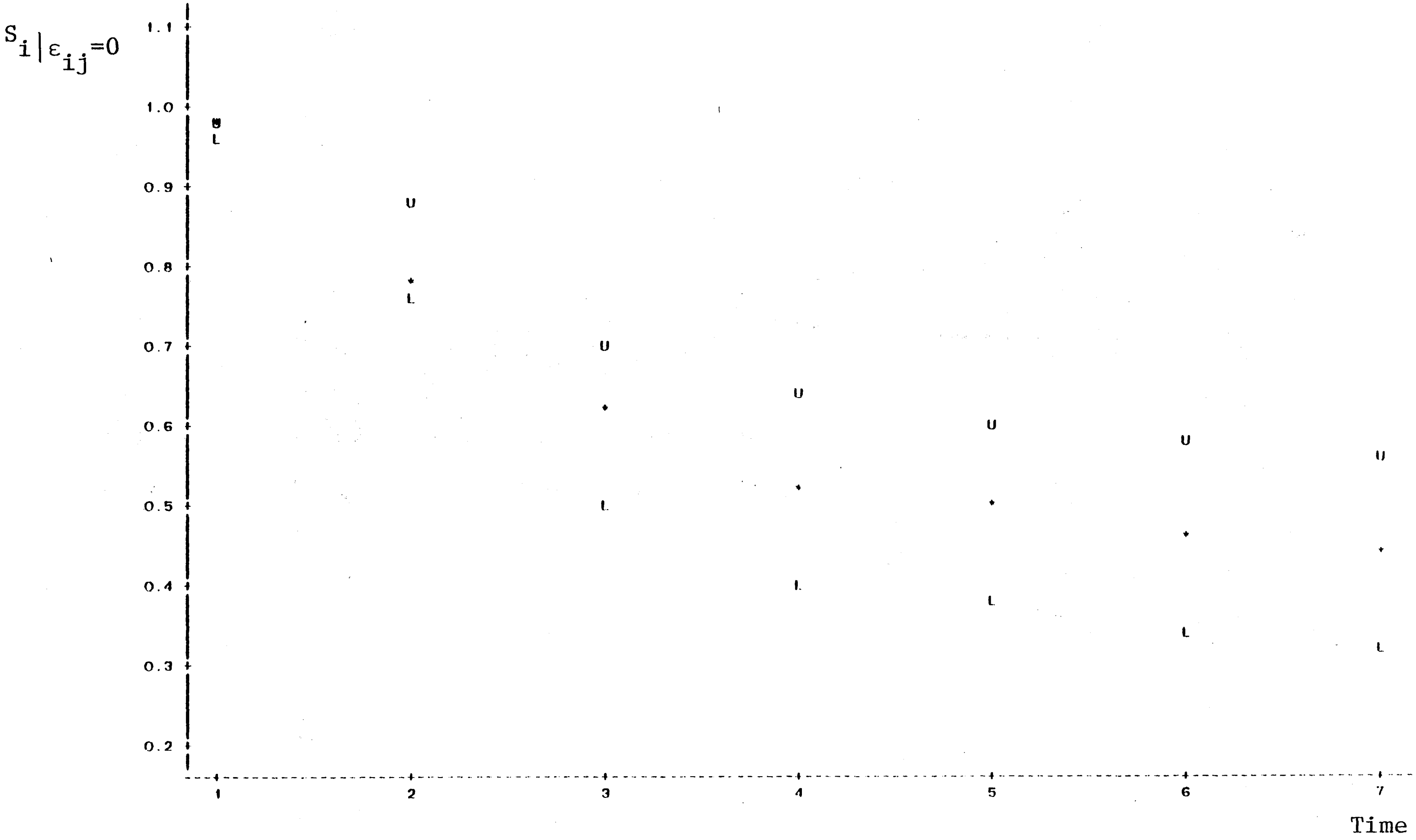

Figure 2. Estimated Survivor Function with Its Confidence Limits for Trt 2

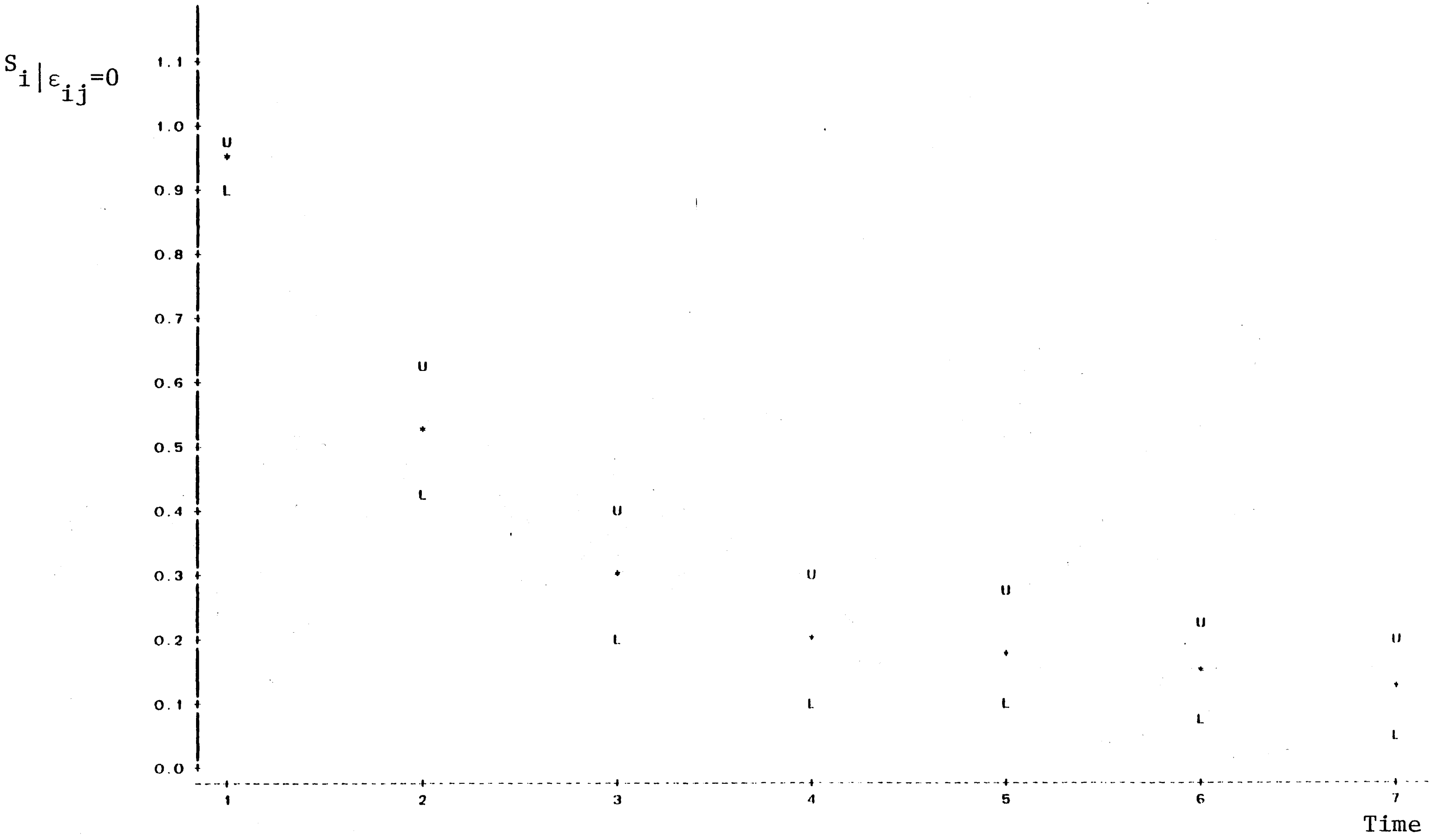

Figure 3. Estimated Survivor Function with Its Confidence Limits for Trt 3

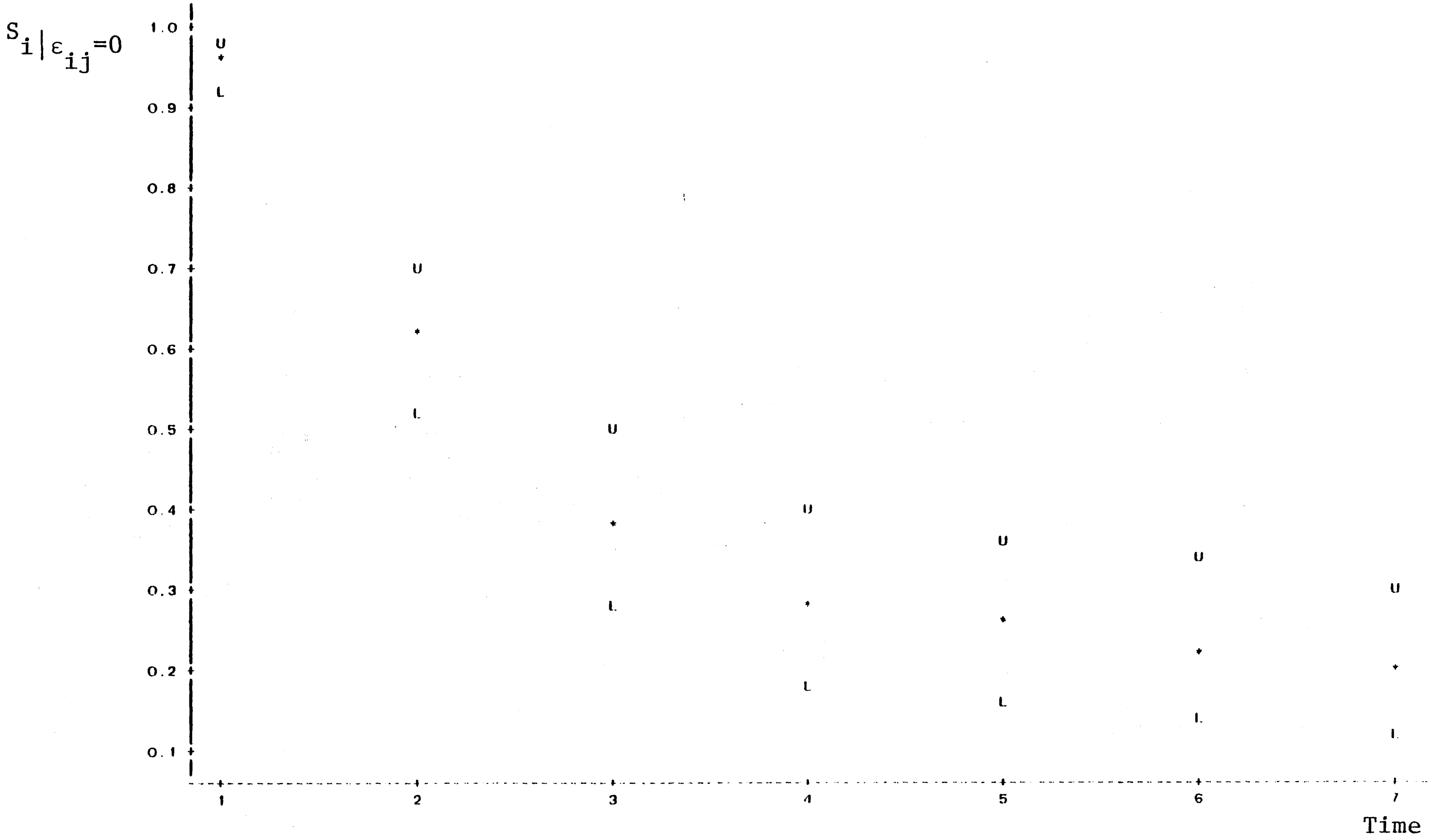

Figure 4. Estimated Survivor Function with Its Confidence Limits for Trt 4

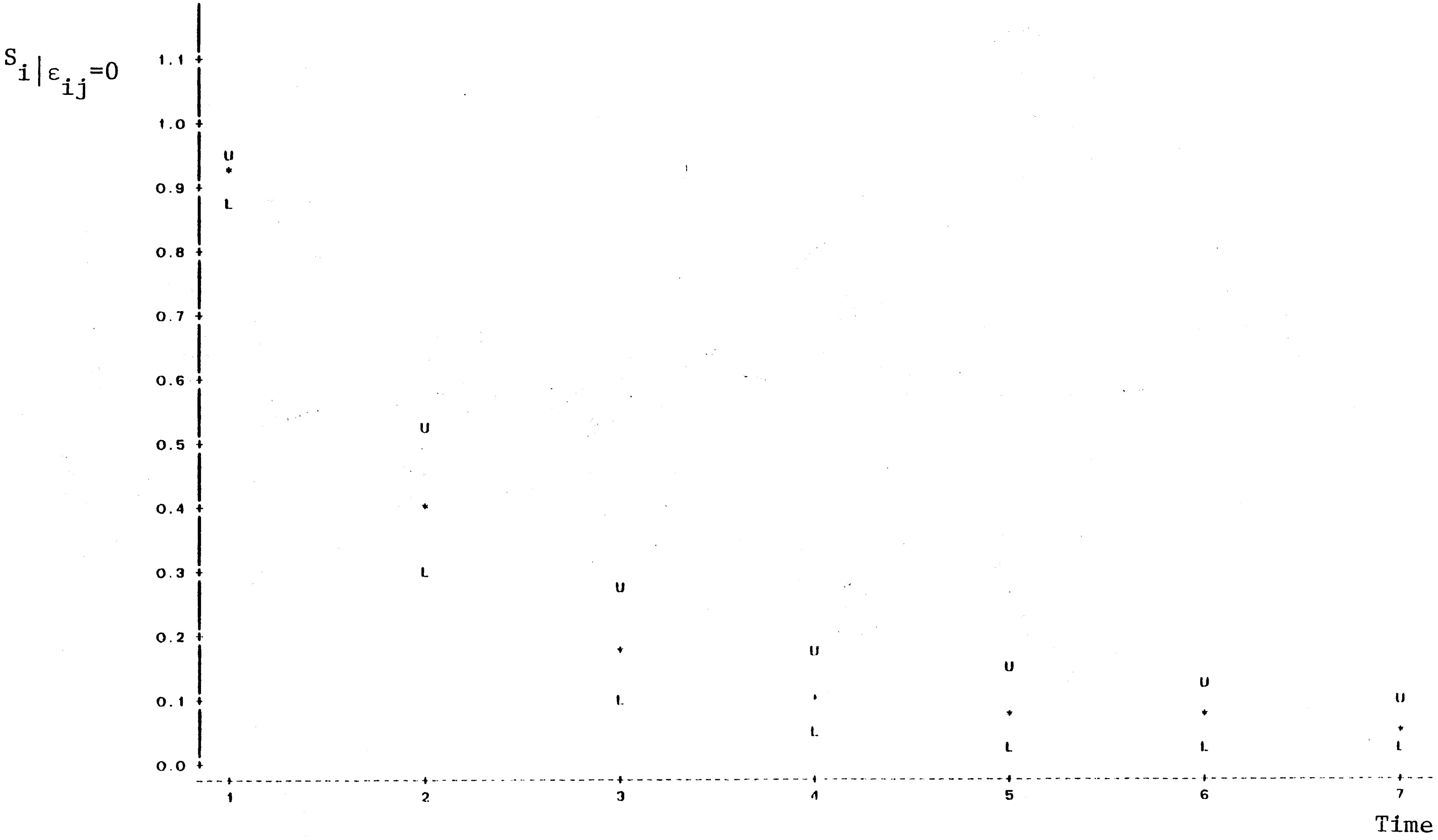

Figure 5. Estimated Survivor Function with Its Confidence Limits for Trt 5

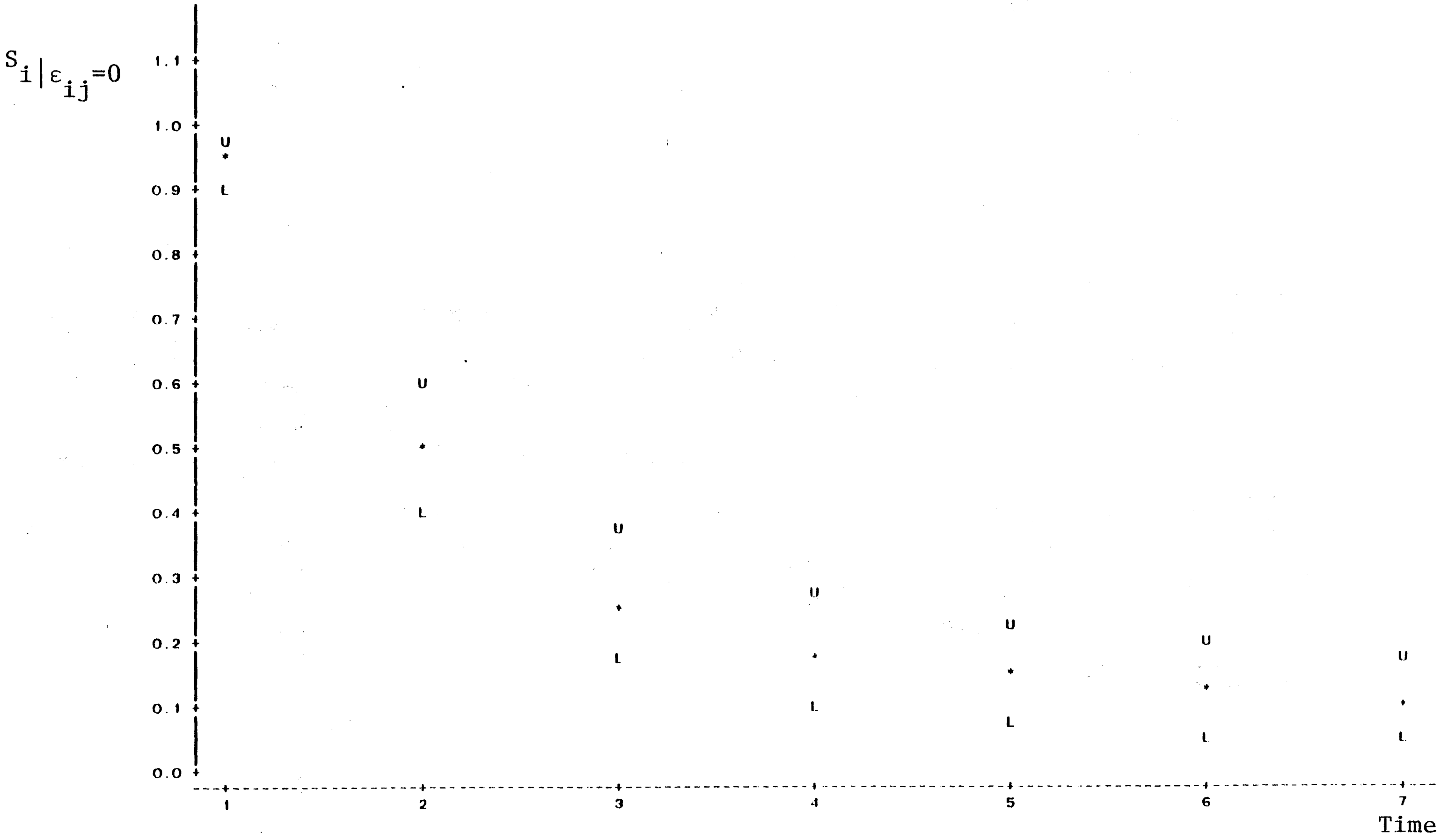

Figure 6. Estimated Survivor Function with Its Confidence Limits for Trt 6

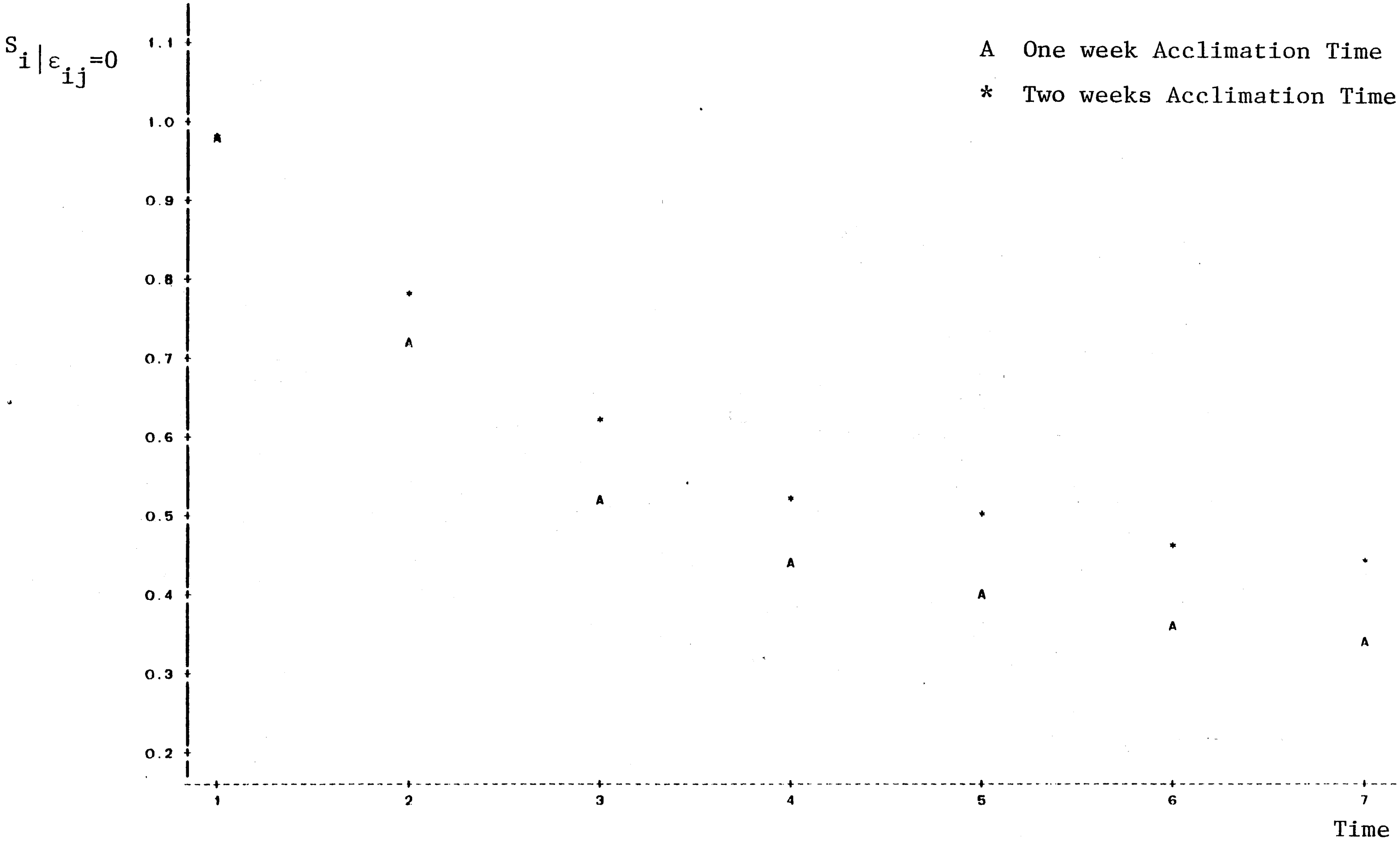

Figure 7. Estimated Survivor Functions for Trt 1 and Trt 2

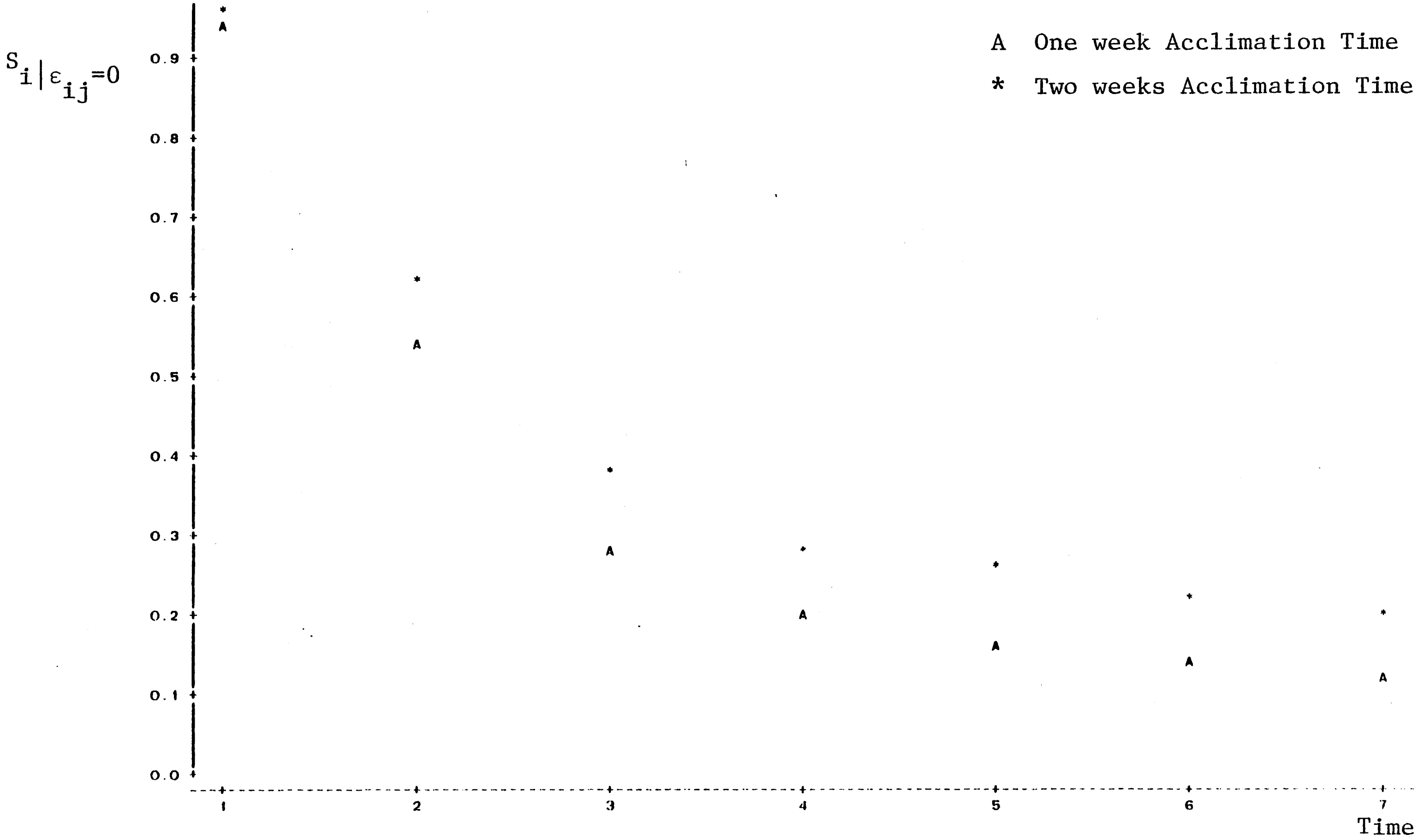

Figure 8. Estimated Survivor Functions for Trt 3 and Trt 4

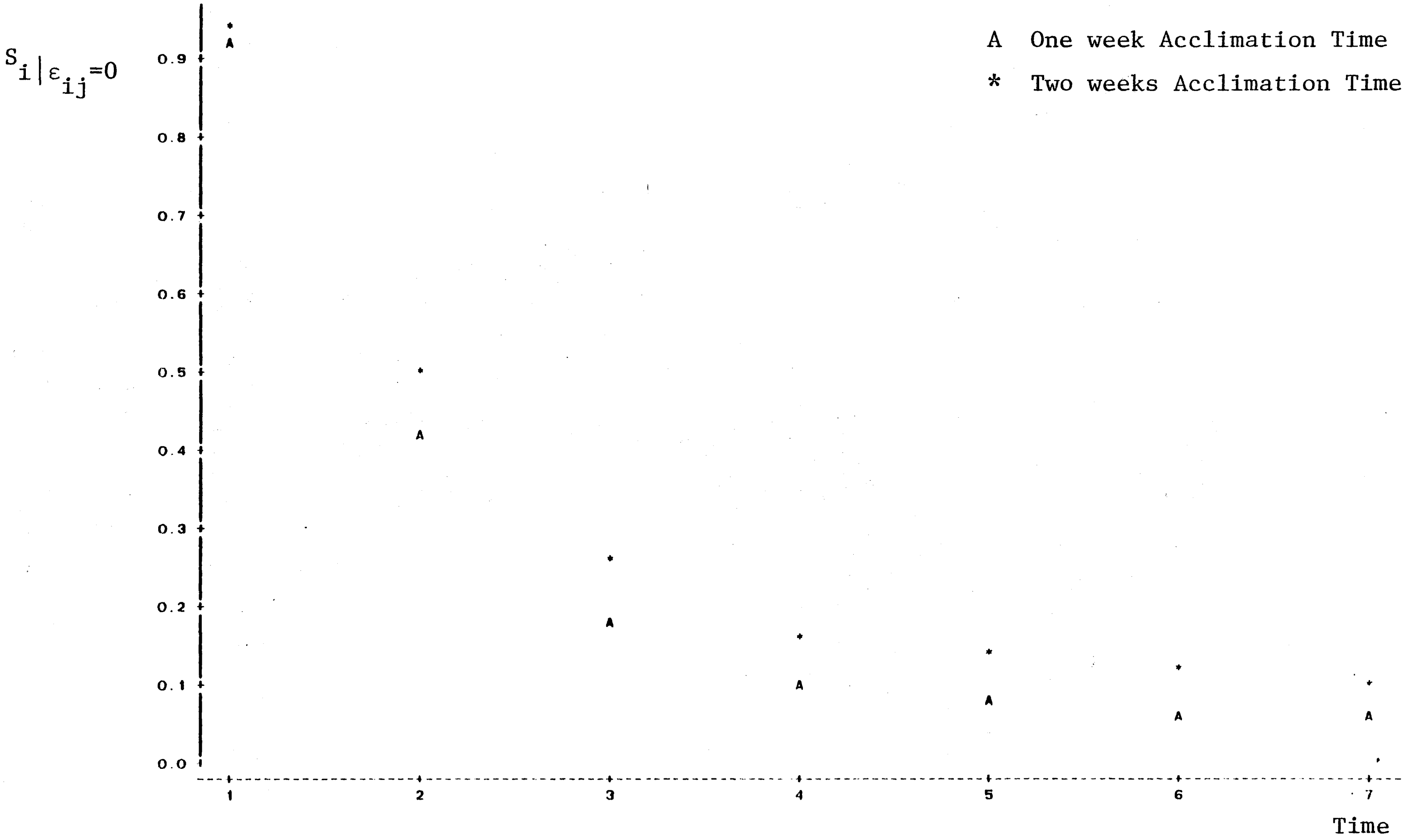

Figure 9. Estimated Survivor Functions for Trt 5 and Trt 6

CHAPTER VII

COMPARISON TO OTHER APPROACHES

7.1 Introduction

The problem that we consider in this study is to make inferences such as point estimation, interval estimations, and hypotheses tests where the split plot model with unequal subunit variances is used. The choice for such a model is based on the process that has been used in data collection, the way that the experiment was conducted, and the assumptions that main unit variability is present and that subunits in a main unit survive independently of other subunits. In other words our method is based on the experimental procedure which indicates that the variances of $\hat{q}_{ijk}$ are not the same for different intervals. The method provides estimates for the variances and least squares estimates for $\underset{\sim}{\beta}$ which have the asymptotic properties discussed in Chapter V. It also provides estimates and confidence limits for the survivor functions for each treatment combination.

Another approach to this problem is to integrate out the random component ε_{ij}, find the unconditional likelihood function, and then find the maximum likelihood estimates. This approach was considered by Marton, Woodbury and Stallard (1981) for the Poisson case. However, it is not applicable to our problem and the reader is referred to section 4 of Chapter III.

Three different simplified approaches that were also considered to solve the problem are listed below

i. No main unit effect

ii. Unweighted least squares split plot.

iii. Unequal subplot error depending on risk sets for different time intervals.

A comparison of our approach with the above three approaches is carried out. The comparison with the first approach is discussed in section 7.2 where we compare the variance of $\hat{\underset{\sim}{\beta}}$ for some cases of interest. The comparison with the other two approaches is done using a generated example in section 7.3.

7.2 Comparison with Inference that Ignores Main Unit Variability

In this section we assume that split plot model with unequal subunit variances is true. We also assume that the model with no main unit variability is the fitted one. The true model is of the form

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u}, \quad \text{where} \quad \text{Cov}(\underset{\sim}{u}) = V$$

and V is a block diagonal matrix of the form

$$V = \begin{bmatrix} \sigma^2_{\delta_{111}} & & & & & 0 \\ & \sigma^2_{\delta_{112}} & & & & \\ & & \ddots & & & \\ & & & \sigma^2_{\delta_{ijk}} & & \\ & & & & \ddots & \\ 0 & & & & & \sigma^2_{\delta_{IJK}} \end{bmatrix} + \sigma^2_{\varepsilon} \begin{bmatrix} J_K & & & & & \\ & J_K & & & & 0 \\ & & \ddots & & & \\ & & & J_K & & \\ 0 & & & & \ddots & \\ & & & & & J_K \end{bmatrix}$$

where $J_K = \underset{\sim}{j}_K \underset{\sim}{j}_K'$, $\underset{\sim}{j}_K' = (1\ 1\ \ldots\ 1)$. Now if we let

$$N = \sigma_\varepsilon^2 \begin{bmatrix} J_K & & & 0 \\ & J_K & & \\ & & \ddots & \\ 0 & & & J_K \end{bmatrix}$$

and use (5.4) then V can be written as

$$V = W + N. \qquad \ldots \quad (7.1)$$

The fitted model has the form

$$\underset{\sim}{y} = X\underset{\sim}{\beta}' + \underset{\sim}{u}', \quad \text{where} \quad \mathrm{Cov}(\underset{\sim}{u}') = W$$

and W is as given in (7.1). Now define

$$\hat{\underset{\sim}{\beta}} = (X'V^{-1}X)^{-1}X'V^{-1}\underset{\sim}{y} \quad \text{and} \quad \tilde{\underset{\sim}{\beta}} = (X'W^{-1}X)^{-1}X'W^{-1}\underset{\sim}{y}.$$

Therefore, the variance-covariance matrices of these estimates under either the full or the fitted model are given by

$$\mathrm{Cov}_V(\hat{\underset{\sim}{\beta}}) = (X'V^{-1}X)^{-1},$$

$$\mathrm{Cov}_W(\tilde{\underset{\sim}{\beta}}) = (X'W^{-1}X)^{-1}, \quad \text{and}$$

$$\mathrm{Cov}_V(\tilde{\underset{\sim}{\beta}}) = (X'W^{-1}X)^{-1}X'W^{-1}VW^{-1}X(X'W^{-1}X)^{-1}.$$

To compare the variance-covariance matrices of the two estimates for both models, the case of two treatments, two main units, and two time intervals is considered below. For convenience in calculations we assume that the over all average of log(-log) model is zero. Then we have

$$X' = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}.$$

$W = (\sigma^2_{\delta_{ijk}})_8$. These variances are affected by treatment and time only, therefore, W can be written as

$$W = \begin{bmatrix} C_1 & & & 0 \\ & C_1 & & \\ & & C_2 & \\ 0 & & & C_2 \end{bmatrix}, \text{ where } C_i = \begin{bmatrix} \sigma^2_{\delta_{i\cdot 1}} & 0 \\ 0 & \sigma_{\delta_{i\cdot 2}} \end{bmatrix}$$

for $i = 1,2$. And V is of the form

$$V = W + \sigma^2_\varepsilon \begin{bmatrix} J_2 & & & 0 \\ & J_2 & & \\ & & J_2 & \\ 0 & & & J_2 \end{bmatrix}, \text{ where } J_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Now, V^{-1} can be written as

$$V^{-1} = W^{-1} - R, \text{ where } R = \begin{bmatrix} R_1 & & & 0 \\ & R_1 & & \\ & & R_2 & \\ 0 & & & R_2 \end{bmatrix} \text{ with}$$

$$R_j = \frac{\sigma^2_\varepsilon}{1 + \sigma^2_\varepsilon(\sigma^{-2}_{\delta_{j\cdot 1}} + \sigma^{-2}_{\delta_{j\cdot 2}})} \begin{bmatrix} (\sigma^2_{\delta_{j\cdot 1}})^{-2} & (\sigma^2_{\delta_{j\cdot 1}} \sigma^2_{\delta_{j\cdot 2}})^{-1} \\ (\sigma^2_{\delta_{j\cdot 1}} \sigma^2_{\delta_{j\cdot 2}})^{-1} & (\sigma^2_{\delta_{j\cdot 2}})^{-2} \end{bmatrix}$$

for $j = 1,2$.

For the purpose of comparisons, we now consider the following cases.

Case 1: Assume that $\sigma^2_{\delta_{ijk}} = \sigma^2_\delta$ and $\sigma^2_\varepsilon = \sigma^2_\delta$. Thus we have

$$(X'W^{-1}X)^{-1} = \frac{1}{8}\sigma_{\delta}^{2}I_2,$$

$$(X'V^{-1}X)^{-1} = \frac{3}{8}\sigma_{\delta}^{2}I_2, \text{ and}$$

$$(X'W^{-1}X)^{-1}X'W^{-1}VW^{-1}X(X'W^{-1}X)^{-1} = \sigma_{\delta}^{2}\begin{bmatrix} \frac{3}{8} & 0 \\ 0 & \frac{5}{8} \end{bmatrix} .$$

Let us now use the notation $A \leq B$ to denote that each element of the matrix A is less than or equal to the corresponding element of the matrix B. Therefore, we conclude the following relation

$$\text{Cov}_W(\tilde{\beta}) < \text{Cov}_V(\hat{\beta}) \leq \text{Cov}_V(\tilde{\beta}).$$

Case 2:

a) Assume all $\sigma_{\delta_{ijk}}^2 = \sigma_{\delta}^2$ and $\sigma_{\varepsilon}^2 > \sigma_{\delta}^2$.

Then

$$(X'W^{-1}X)^{-1} = \frac{1}{8}\sigma_{\delta}^{2}I_2,$$

$$(X'V^{-1}X)^{-1} = \frac{1}{8}\sigma_{\delta}^{2}(1+\sigma_{\varepsilon}^{2})I_2, \text{ and}$$

$$(X'W^{-1}X)^{-1}X'W^{-1}VW^{-1}X(X'W^{-1}X)^{-1} = \begin{bmatrix} \frac{1}{8}\sigma_{\delta}^{2}+\frac{1}{4}\sigma_{\varepsilon}^{2} & 0 \\ 0 & \frac{1}{8}\sigma_{\delta}^{2}+\frac{1}{8}\sigma_{\varepsilon}^{2} \end{bmatrix} .$$

Therefore, we conclude the following relation

$$\text{Cov}_W(\tilde{\beta}) < \text{Cov}_V(\hat{\beta}) \leq \text{Cov}_V(\tilde{\beta}).$$

b) Now assume all $\sigma_{\delta_{ijk}}^2 = \sigma_{\delta}^2$ and $0 \leq \sigma_{\varepsilon}^2 < \sigma_{\delta}^2$. Then we conclude

$$\text{Cov}_W(\tilde{\beta}) \leq \text{Cov}_V(\hat{\beta}) \leq \text{Cov}_V(\tilde{\beta}).$$

Case 3: Assume all $\sigma_{\delta_{ijk}}^2$ are the same over trt but increase over time and $\sigma_{\varepsilon}^2 = \max\{\sigma_{\delta_{ijk}}^2\}$ for all i, j, and k. For this case we are assuming that $\sigma^2_{\delta_{1\cdot 1}} = \sigma^2_{\delta_{2\cdot 1}}$, $\sigma^2_{\delta_{1\cdot 2}} = \sigma^2_{\delta_{2\cdot 2}}$, and $\sigma^2_{\delta_{2\cdot 2}} > \sigma^2_{\delta_{2\cdot 1}}$. Then we have

$$(X'W^{-1}X)^{-1} = \frac{1}{4}\cdot\frac{\sigma^2_{\delta_{2\cdot 1}}\sigma^2_{\delta_{2\cdot 2}}}{\sigma^2_{\delta_{2\cdot 1}} + \sigma^2_{\delta_{2\cdot 2}}} I_2 ,$$

$$(X'V^{-1}X)^{-1} = \frac{1}{2}\sigma^2_{\delta_{2\cdot 2}} I_2, \text{ and}$$

$$(X'W^{-1}X)^{-1}X'W^{-1}VW^{-1}X(X'W^{-1}X)^{-1} =$$

$$\{\tfrac{1}{2}\sigma^2_{\delta_{2\cdot 2}}\}\{\tfrac{1}{2}(\sigma^2_{\delta_{2\cdot 1}} + \sigma^2_{\delta_{2\cdot 2}})^{-1}\} \begin{bmatrix} 2\sigma^2_{\delta_{2\cdot 1}} + \sigma^2_{\delta_{2\cdot 2}} & \sigma^2_{\delta_{2\cdot 2}} - \sigma^2_{\delta_{2\cdot 1}} \\ 0 & \dfrac{2\sigma^4_{\delta_{2\cdot 1}} - 2\sigma^2_{\delta_{2\cdot 1}}\sigma^2_{\delta_{2\cdot 2}} + \sigma^4_{\delta_{2\cdot 2}}}{\sigma^2_{\delta_{2\cdot 1}}\sigma^2_{\delta_{2\cdot 2}}} \end{bmatrix}$$

Therefore, we conclude the relation

$$\mathrm{Cov}_W(\underset{\sim}{\tilde{\beta}}) \le \mathrm{Cov}_V(\underset{\sim}{\hat{\beta}}) \le \mathrm{Cov}_V(\underset{\sim}{\tilde{\beta}}).$$

Case 4: Assume all $\sigma^2_{\delta_{ijk}}$ are the same over trt but increase over time and $\sigma^2_\varepsilon > \max\{\sigma^2_{\delta_{ijk}}\}$ for all i, j, and k. Then

$$(X'W^{-1}X)^{-1} = \frac{1}{4}\,\frac{\sigma^2_{\delta_{2\cdot 1}}\sigma^2_{\delta_{2\cdot 2}}}{\sigma^2_{\delta_{2\cdot 1}} + \sigma^2_{\delta_{2\cdot 2}}} I_2,$$

$$(X'V^{-1}X)^{-1} = \frac{1}{4}\left(\frac{\sigma^2_{\delta_{2\cdot 1}}\sigma^2_{\delta_{2\cdot 2}}}{\sigma^2_{\delta_{2\cdot 1}} + \sigma^2_{\delta_{2\cdot 2}}} + \sigma^2_\varepsilon\right) I_2, \text{ and}$$

$$(X'W^{-1}X)^{-1}X'W^{-1}VW^{-1}X(X'W^{-1}X)^{-1} =$$

$$\frac{1}{4}\frac{\sigma^2_{\delta_{2\cdot1}}\sigma^2_{\delta_{2\cdot2}}}{\sigma^2_{\delta_{2\cdot1}}+\sigma^2_{\delta_{2\cdot2}}} I_2 + \begin{bmatrix} \frac{1}{4}\sigma^2_\varepsilon & \frac{1}{4}\sigma^2_\varepsilon \dfrac{\sigma^2_{\delta_{2\cdot2}}-\sigma^2_{\delta_{2\cdot1}}}{\sigma^2_{\delta_{2\cdot1}}+\sigma^2_{\delta_{2\cdot2}}} \\ 0 & \dfrac{\sigma^2_\varepsilon(\sigma^2_{\delta_{2\cdot2}}-\sigma^2_{\delta_{2\cdot1}})^2}{\sigma^2_{\delta_{2\cdot1}}\sigma^2_{\delta_{2\cdot2}}(\sigma^2_{\delta_{2\cdot1}}+\sigma^2_{\delta_{2\cdot2}})^2} \end{bmatrix}.$$

Therefore, we conclude the following relation

$$\mathrm{Cov}_W(\tilde{\underset{\sim}{\beta}}) \leq \mathrm{Cov}_V(\hat{\underset{\sim}{\beta}}) \leq \mathrm{Cov}_V(\tilde{\underset{\sim}{\beta}}).$$

From the above four cases we conclude that if W is used when in fact V is the true variance-covariance matrix then this will under estimate the variances of the estimators.

7.3 A Generated Example for Comparison with Inference Using Unweighted Least Squares or Inference Weighting by Risk Sets

In this section we use a generated example which has the properties presented in Chapter I for the problem under study. These properties are listed below.

1. Main unit variability has a symmetric distribution for the discrete case.
2. Risk sets decrease over time.
3. There is a trend in the $\sigma^2_{\delta_{ijk}}$ in that these variances increase over time.

Two treatments, two main units, and three time intervals are used. Let ε_{ij} have the following frequency distribution

ε_{ij}:	-1.5	-1	0	1	1.5
$f(\varepsilon_{ij})$:	1/9	2/9	3/9	2/9	1/9

Therefore, $E(\varepsilon_{ij}) = 0$ and $\mathrm{Var}(\varepsilon_{ij}) = 0.9444$. Also let the parameters for this experiment be known and are given by $\mu = -1$, $\alpha_1 = .5$, $\tau_1 = .4$, and $\tau_2 = .1$. Also we let the number of subunits in each main unit be fixed by $n_{ij} = 100$. For $I = 2$, $J = 2$, and $K = 3$ the design matrix has the following form

$$X' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{bmatrix}.$$

With this generated example we would like to compare our model with unweighted least squares split plot model and also with a split plot model having unequal subplot errors depending on risk sets for different time intervals. Before proceeding with the calculations, we mention the structure of the models to be compared. Our model, which is assumed to be the true model, has the following form

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u}, \quad \text{where} \quad \mathrm{Cov}(\underset{\sim}{u}) = V$$

and V is as given by (7.1). The unweighted least squares split plot model has the following form

$$\underset{\sim}{y} = X\underset{\sim}{\beta}'' + \underset{\sim}{u}'', \quad \text{where } \mathrm{Cov}(\underset{\sim}{u}'') = \Sigma$$

and Σ is the variance-covariance matrix and is of the form

$$\Sigma = \sigma_\varepsilon^2 \begin{bmatrix} J_K & & & 0 \\ & J_K & & \\ & & \ddots & \\ 0 & & & J_K \end{bmatrix} + \sigma_\delta^2 I.$$

The split plot model with unequal subplot errors depending on risk sets has the following form

$$\underset{\sim}{y} = X\underset{\sim}{\beta}''' + \underset{\sim}{u}''', \quad \text{where} \quad \mathrm{Cov}(\underset{\sim}{u}''') = D$$

and D is the variance-covariance matrix and is of the form

$$D = \sigma_\varepsilon^2 \begin{bmatrix} J_K & & & 0 \\ & J_K & & \\ & & \ddots & \\ 0 & & & J_K \end{bmatrix} + \sigma^2 \begin{bmatrix} n_{111}^{-1} & & & & 0 \\ & \ddots & & & \\ & & n_{ijk}^{-1} & & \\ & & & \ddots & \\ 0 & & & & n_{IJK}^{-1} \end{bmatrix}.$$

Now, we proceed with the calculations involved to get the variance-covariance matrices of $\underset{\sim}{\beta}$ under either the true model or one of the two fitted models. For each value of ε_{ij} we calculate the corresponding values of

$$\log(-\log q_{ijk}) = \mu + \alpha_i + \tau_k + \varepsilon_{ij}.$$

Therefore, values of q_{ijk}, and $(1-q_{ijk})/q_{ijk}(\log q_{ijk})^2$ are also calculated. We start with $n_{ij} = 100$ then we use the equation $E(n_{ijk}) = n_{ij(k-1)} q_{ij(k-1)}$ for $k = 1,2,3$, to obtain values of n_{ijk} for each value of ε_{ij}. Hence values of $(1-q_{ijk})/n_{ijk}q_{ijk}(\log q_{ijk})^2$ can also be calculated. Now, using the frequency distribution of ε_{ij} we get

$$\sigma^2_{\delta_{ijk}} = \sum_{\ell=1}^{5} [(1-q_{ijk})/n_{ijk}q_{ijk}(\log q_{ijk})^2]_\ell f(\ell),$$

$$\sigma^2_\delta = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{3} \sigma^2_{\delta_{ijk}} /12,$$

$$\sigma^2 = \sum_{\ell=1}^{5} [(1-q_{ijk})/q_{ijk}(\log q_{ijk})^2]_\ell f(\ell), \text{ and}$$

$$\overline{n_{ijk}} = \sum_{\ell=1}^{5} (n_{ijk})_\ell f(\ell).$$

Therefore, we have the following results.

$$(X'V^{-1}X)^{-1} = \begin{bmatrix} .2500 & .00289 & -.01146 & -.00228 \\ .00289 & .24433 & -.00468 & .00040 \\ -.01146 & -.00468 & .01935 & -.00124 \\ -.00228 & .00040 & -.00124 & .02411 \end{bmatrix},$$

$$(X'\Sigma^{-1}X)^{-1} = \begin{bmatrix} .2778 & 0 & 0 & 0 \\ 0 & .2778 & 0 & 0 \\ 0 & 0 & .08317 & -.04158 \\ 0 & 0 & -.04158 & .08317 \end{bmatrix}, \text{ and}$$

$$(X'D^{-1}X)^{-1} = \begin{bmatrix} 0.24592 & -0.00070 & -0.00658 & -0.00196 \\ -0.00070 & 0.24242 & -0.00140 & 0.00057 \\ -0.00658 & -0.00140 & 0.01301 & -0.00125 \\ -0.00196 & 0.00057 & -0.00125 & 0.01766 \end{bmatrix}.$$

Then, a summary for comparison is given in Table IX. From Table IX it seems that all methods are close for Var($\hat{\alpha}_1$). Σ does not get Var($\hat{\tau}$) well which may cause problems with Var($\hat{S}(t)$). Also, D looks like a possible simplifying approach unless $\hat{q}_{ijk}$'s are widely different.

TABLE IX

SUMMARY OF COMPARISONS RESULTS

Parameter	$Var_V(\beta_i)$	$Var_\Sigma(\beta_i)$	$Var_D(\beta_i)$
μ	0.25000	0.27780	0.24592
α_1	0.24433	0.27780	0.24242
τ_1	0.01935	0.08317	0.01301
τ_2	0.02411	0.08317	0.01766
τ_3	0.04098	0.08317	0.02816

BIBLIOGRAPHY

Aranda-Ordaz, F. J. (1983). An extension of the proportional-hazards model for grouped data. *Biometrics*, *39*, 109-117.

Arnold, S. F. (1981). *The Theory of Linear Models and Multivariate Analysis*. John Wiley & Sons, Inc., New York.

Brown, K. G. (1978). Estimation of variance components using residuals. *J. Amer. Statist. Assoc.*, *73*, 141-146.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. R. Statist. Soc.*, *B*, *34*, 187-220.

Crump, S. L. (1946). The estimation of variance components in the analysis of variance. *Biometrics Bul.*, *2*, *1*, 7-11.

Curnow, R. N. (1957). Hetrogeneous error variances in split plot experiments. *Biometrika*, *44*, 378-383.

Elandt-Johnson, R. C. (1980). Some prior distributions in survival analysis: a critical insight on relationship derived from cross-sectional data. *J. R. Statist. Soc.*, *B*, *42*, 96-106.

Grizzle, J. E., Starmer, C. F., and Koch, G. G. (1969). Analysis of categorical data by linear models. *Biometrics*, *25*, 489-504.

Hartley, H. O. and Jayatillake, K. S. E. (1973). Estimation of linear models with unequal variances. *J. Amer. Statist. Assoc.*, *68*, 189-192.

Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics*, *9*, 226-252.

Kalbfleisch, J. D. and Prentice, R. L. (1973). Marginal likelihoods based on Cox's regression and life model. *Biometrika*, *60*, 267-278.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.*, *53*, 457-481.

Kay, R. (1977). Proportional hazard regression models and the analysis of censored survival data. *Appl. Statist.*, *26*, 227-237.

Krane, S. A. (1963). Analysis of survival data by regression techniques. *Technometrics*, *5*, 161-174.

Li, S. H. and Klotz, J. H. (1978). Components of variance estimation for the split plot design. J. Amer. Statist. Assoc., 73, 147-152.

Manton, K. G., Woodbury, M. A. and Stallard, E. (1981). A variance components approach to categorical data models with heterogeneous cell populations: Analysis of spatial gradients in lung-cancer mortality rates in North Carolina counties. Biometrics, 37, 259-269.

Pierce, D. A., Stewart, W. H. and Kopecky, K. J. (1979). Distribution-free regression analysis of grouped survival data. Biometrics, 35, 785-793.

Prentice, R. L. and Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. Biometrics, 34, 57-67.

Rao, C. R. (1965). Linear Statistical Inference and its Application. John Wiley and Sons, Inc., New York.

Rao, C. R. (1970). Estimation of heteroscedastic variances in linear models. J. Amer. Statist. Assoc., 65, 161-172.

Searle, S. R. (1968). Another look at Henderson's methods of estimating variance components. Biometrics, 24, 749-778.

Searle, S. R. (1971). Linear Models. John Wiley and Sons, Inc., New York.

Tibshirani, R. J. and Ciampi, A. (1983). A family of proportional- and additive-hazards models for survival data. Biometrics, 39, 141-147.

APPENDIX

A COMPUTER PROGRAM FOR ESTIMATING MAIN UNIT VARIANCE

The following computer program can be used to estimate σ_{ε}^{2} as given by 4.2 using SAS procedure MATRIX as follows. Values for treatments, main units, time, transformed values of $\hat{q}_{ijk}$'s, and estimates of $\sigma_{\delta_{ijk}}^{2}$'s using 4.1 were entered as TRT, REP, TIME, Y, and W, respectively.

```
RY=NROW(Y);
M=J(RY,1,1);
DTRT=DESIGN(TRT);
DTM=DESIGN(TIME);
DTRTTM=DTRT@|DTM;
DR=DESIGN(REP);
X=M||DTRT||DTM||DTRTTM;
Z=DR;
N=X||Z;
D=DIAG(W);
DI=INV(D);
NUM1=DI*N*GINV(N'*DI*N)*N';
TRNUM1=TRACE(NUM1);
NUM2=DI*X*GINV(X'*DI*X)*X';
TRNUM2=TRACE(NUM2);
TRNUM=TRNUM1-TRNUM2;
DEN1=Z'*DI*Z;
TRDEN1=TRACE(DEN1);
DEN2=Z'*DI*X*GINV(X'*DI*X)*X'*DI*Z:
TRDEN2=TRACE(DEN2);
TRDEN=TRDEN1-TRDEN2;
RSSFUL=Y'*DI*N*GINV(N'*DI*N)*N'*DI*Y;
RSSRED=Y'*DI*X*GINV(X'*DI*X)*X'*DI*Y;
ESVARCOM=(RSSFUL-RSSRED-TRNUM)#/TRDEN;
```

VITA

Dalal Abdulrazzak Abdullatif

Candidate for the Degree of

Doctor of Philosophy

Thesis: AN ANALYSIS FOR GROUPED SURVIVOR DATA WITH SPLIT PLOT VARIANCE COMPONENT MODELS

Major Field: Statistics

Biographical:

Personal Data: Born in Baghdad, Iraq, March 23, 1953.

Education: Graduated from Alkansa'a High School, Baghdad, Iraq, in July, 1970; received Bachelor of Science from the University of Baghdad, Baghdad, Iraq, in June, 1975; received Master of Science degree in Statistics from Mississippi State University in the Spring of 1980; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1984.

Professional Experience: Instructor, Department of Statistics, University of Baghdad, Baghdad, Iraq, 1975 - 1978.

Professional Organization: Member of Mu Sigma Rho.