RESTRICTED ALTERNATIVES TESTS IN COMPETING RISK ANALYSES

by

Cicilia Yuko Wada

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1847T

March 1988
RESTRICTED ALTERNATIVES TESTS IN COMPETING RISK ANALYSES

by

Cicilia Yuko Wada

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics.

Chapel Hill, 1988

Approved by:

Advisor

Reader

Reader
ABSTRACT

CICILIA YUKO WADA. Restricted Alternatives Tests in Competing Risk Analyses.
(Under the direction of Dr. P.K. SEN)

The objective of the present study is to derive a union-intersection test statistic for subhypotheses against restricted alternatives in a competing risk situation with k causes of failure and p concomitant variables. Asymptotic properties of the proposed tests are investigated and compared to those of the usual score test.

For the Cox regression model involving p concomitant variables and two treatments, testing of the subhypotheses against orthant restricted alternatives based on partial likelihood is considered. Roy's union-intersection principle is applied and the Kuhn-Tucker-Lagrange minimization technique is utilized for the derivation of the proposed test. The distribution of the proposed test under the null and under "local" alternatives is studied. The asymptotic relative efficiency of the proposed test with respect to the classical score test is investigated.

The above derivation is extended to the comparison of r+1 subsamples by considering tests of subhypotheses against restricted alternatives, such as the orthant alternative, the ordered orthant alternative and finally the ordered alternative. Asymptotic properties of these tests are also studied.

The derivation of the tests using the same approach is also considered when the logistic model and a dependent parametric model are utilized to assess hypotheses of no difference between the two treatments against the restricted alternative that one treatment is superior to the another.
A numerical illustration of the Cox regression model in a competing risk situation with two causes of failure, two treatments and with two covariates is performed. Monte-Carlo simulation is utilized to show the power superiority of the proposed test compared to the score test.
ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to Dr. Pranab Kumar Sen, advisor of this dissertation. I am grateful for his firm guidance, for his patience, and for his valuable support and encouragement in difficult times.

I express my deep appreciation to Dr. Timothy Morgan, for his help, comments and suggestions. His interest and solicitude will be always be remembered with gratitude. I also thank Dr. Dana Quade, for his kindness and patience in helping to edit this dissertation. I extend my thanks to other members of my committee, Drs. Shrikant Bangdiwala, C. Edward Davis and Sherman James.

I am indebted to the University of São Paulo State (UNESP) and to Conselho Nacional de Desenvolvimento Científico e Tecnológico in Brazil (CNPq) for financial support during the first years of my doctoral program.

I thank my family: my parents, Hisae and Wasaji Tsuji, and especially my sister Mitie Tsuji, for their love, care and support. I appreciate the friendship and care given to me by my friends from Brazil and those from here during the years I spent here.

Finally, I wish to thank Ms. Ans Janssens for helping me to format this dissertation.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS**

1. **INTRODUCTION AND LITERATURE REVIEW**

1.1. Introduction .......................................................... 1

1.2. Basic Formulation and Notation in Competing Risk Problems .............................................. 3

1.3 Review of Competing Risk Models .................................. 7

1.3.1. Parametric Models .................................................. 8

1.3.1.1. The Likelihood Function ..................................... 11

1.3.1.2. The Maximum Likelihood Estimators ......................... 13

1.3.2. Semiparametric Approach ........................................ 15

1.4. Review of Restricted Alternatives and Union-Intersection Tests .......................................... 18

1.5. Proposed Research .................................................... 23

2. **RESTRICTED ALTERNATIVES TESTS IN A PROPORTIONAL HAZARD MODEL WITH COMPETING RISK DATA**

2.1. Introduction .......................................................... 24

2.2. The Union-Intersection Test Based on Partial Maximum Likelihood Estimators ...................... 26

2.2.1. Partial Maximum Likelihood Estimators .................................................. 27

2.2.2. The Union-Intersection Test ........................................ 41

2.3 Asymptotic Distribution of the Union-Intersection Test Statistic $I_1^*$ .................................. 51

2.4. Local Asymptotic Power and Efficiency of $I_1^*$ .......................................................... 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. RESTRICTED ALTERNATIVES TEST IN A PARAMETRIC MODEL AND IN A LOGISTIC MODEL WITH COMPETING RISK DATA</td>
<td></td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>60</td>
</tr>
<tr>
<td>3.2. Union-Intersection Test for a Parametric Model in Competing Risk Situation</td>
<td>61</td>
</tr>
<tr>
<td>3.2.1. The Problem and Assumptions</td>
<td>64</td>
</tr>
<tr>
<td>3.2.2. The Union-Intersection Test</td>
<td>75</td>
</tr>
<tr>
<td>3.3. Union-Intersection Test for a Logistic Model in Competing Risk Situation</td>
<td>84</td>
</tr>
<tr>
<td>3.3.1. Notation and Assumptions</td>
<td>86</td>
</tr>
<tr>
<td>3.3.2. The Union-Intersection Test</td>
<td>92</td>
</tr>
<tr>
<td>4. ORDERED (ORTHANT) ALTERNATIVE TESTS IN A PROPORTIONAL HAZARD MODEL WITH COMPETING RISK DATA</td>
<td></td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>98</td>
</tr>
<tr>
<td>4.2. The Problem and Notation</td>
<td>98</td>
</tr>
<tr>
<td>4.3. The Union-Intersection Test for Orthant (Ordered) Alternative $l^*_3$</td>
<td>102</td>
</tr>
<tr>
<td>4.4. The Union-Intersection Test for Ordered Alternative $l^*_3$</td>
<td>104</td>
</tr>
<tr>
<td>4.4.1. Union-Intersection Test Statistic $l^*_3$ Derivation</td>
<td>104</td>
</tr>
<tr>
<td>4.4.2. Asymptotic Distribution of the Statistic $l^*_3$</td>
<td>113</td>
</tr>
<tr>
<td>4.4.3. Local Asymptotic Power and Efficiency</td>
<td>118</td>
</tr>
<tr>
<td>5. NUMERICAL ILLUSTRATION AND FURTHER RESEARCH</td>
<td></td>
</tr>
<tr>
<td>5.1. Introduction</td>
<td>123</td>
</tr>
<tr>
<td>5.2. Numerical Illustration</td>
<td>123</td>
</tr>
<tr>
<td>5.3. Results</td>
<td>125</td>
</tr>
<tr>
<td>5.4. Suggestions for Further Research</td>
<td>137</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>141</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

A considerable amount of statistical research has focused on problems concerning the failure or deaths of individuals from any one of the several potential causes. Such problems occur in many areas, such as Epidemiology, Demography, Clinical Trials and Industry and are called Competing risk problems. Data for such problems usually consist of the cause of failure or death and the time of the failure for the study unit. Other information may also be available, as in the case of additional variables describing different conditions in which the experiment was conducted (e.g. exposure levels, treatment, etc.). These additional variables are called concomitant variables or covariates. One example for this situation was given in Hoel (1972), in a study of mortality an experiment in which two groups of mice, one kept in a germfree environment and another in a conventional laboratory environment, receiving a certain dose of radiation. The cause of death was determined by autopsy and could be: tymic lymphoma \((C_1)\), reticulum cell sarcoma \((C_2)\) and all other causes \((C_3)\). The available data for each mice are the failure and a concomitant variable which is the indicator of group. The objective in such situation may be to estimate the survival rates for each cause as well as for all causes and to perform comparisons among groups. A second example which was presented by Holt (1978), illustrates the extension of the competing risk situation to paired data (multivariate failure data): in a study of smoking discor-dance in pairs of twins two causes of death were recorded, vascular disease \((C_1)\) and other \((C_2)\). The available data for each twin are the failure times, the causes of death and the vector of concomitant variables indicating the smoking habit and whether or not the pair is
monozygotic. The objective was to assess whether or not the smoking effects on the two risks of deaths differ for monozygotic and dizygotic twins. Lagakos (1977) studied data from a lung cancer clinical trial being conducted at that time by the Eastern Cooperative Oncology Group. The data contain failure times of 194 patients with squamous cell carcinoma as well as the cause of failure $C_1$, local spread of disease, and cause $C_2$, metastatic spread of disease. Three covariates were considered: $Z_1 =$ performance status (ambulatory, not ambulatory), $Z_2 =$ treatment (A and B) and $Z_3 =$ age in years. The objective was to assess the differences between treatments A and B.

In all the examples presented above, the comparison of the two groups could be denoted by the hypothesis $H_0: \theta = 0$ and the alternative hypothesis $H_{a1}: \theta \neq 0$, where $\theta$ represents the difference between the first and the second groups. Now, the problem is to test the null hypothesis of no difference between the groups against the alternative that the first treatment is better than the second, that is, against the restricted alternative $H_{a2}: \theta > 0$.

If we consider the $r+1$ sample problem involving a control and $r$ treatment groups, then, $\theta = (\theta_1, \theta_2, \ldots, \theta_r)$ is the vector of the unknown parameters, where $\theta_j$ represents the j-th treatment effect over the control. Against the null hypothesis of no treatment effect, that is, $H_0: \theta = 0$, we may be interested in the alternative none of the treatments is inferior to the control, that is,

$$H_0^{(1)}: \theta_j \geq 0, \ j = 1, 2, \ldots, r$$

with at least one strict inequality.

A second hypothesis of interest may be the ordering of the treatment effects when $H_0$ may not hold, that is,

$$H_0^{(2)}: \theta_1 \leq \ldots \leq \theta_r, \text{ with at least one strict inequality.}$$
A third hypothesis of interest may be the ordered orthant alternative, that is,

\[ H_a^{(3)}: 0 \leq \theta_1 \leq \ldots \leq \theta_r, \text{ with at least one strict inequality.} \]

We are interested in testing the hypothesis of no treatment effect against restricted alternatives stated above and comparing the efficiency of such tests against the test with unrestricted alternatives, in a competing risk situation with concomitant variables.

1.2. Basic Formulation and Notation in Competing Risk Problems

Initially we introduce a notation which is primarily due to Prentice et al (1978).

Suppose a subject from a population is susceptible to \( k \) causes of failure, denoted by \( C_1, C_2, \ldots, C_k \). A sample of size \( n \) from this population is subjected to an experiment. For each subject, survival experience can be represented by \((T,I,z)\), where \( T \) is a random variable corresponding to the observed failure time, \( I \) is a random index defined by

\[
I = \begin{cases} 
  i & \text{if failure is due to } C_i, i=1, \ldots, k \\
  0 & \text{otherwise}
\end{cases}
\]

and \( z'=(z_1, z_2, \ldots, z_m) \) is the vector of concomitant variables. The overall survival function is represented by

\[
S_T(t;z) = P(T \geq t; z)
\] (2.1)

and the overall hazard function by

\[
\lambda_T(t;z) = \lim_{\Delta t \to 0} \frac{P\{t \leq T \leq t+\Delta t \mid T \geq t; z\}}{\Delta t} = \frac{f_T(t; z)}{S_T(t; z)}
\] (2.2)

where \( f_T(t; z) \) is the probability density function of \( T \). Cause specific hazard functions
are defined by
\[ \lambda_i(t; z) = \lim_{\Delta t \to 0} \frac{P\{t \leq T \leq t + \Delta t, \ I = i| T \geq t; z\}}{\Delta t} \] (2.3)

for \( i = 1, 2, \ldots, k \). The function \( \lambda_i(t; z) \) is the instantaneous failure rate from cause \( C_i \) at the time \( t \), given the vector of concomitant variables \( z \), in the presence of all other types of failure. The overall hazard function (2.1) can be expressed in terms of cause-specific hazard functions as
\[ \lambda_T(t) = \sum_{i=1}^{k} \lambda_i(t; z) \] (2.4)

and the overall survival function (2.1) can be written as
\[ S_T(t; z) = \exp \left\{ - \int_0^t \lambda_T(u; z) \, du \right\} = \prod_{i=1}^{k} \exp \left\{ - \int_0^t \lambda_i(u; z) \, du \right\} = \prod_{i=1}^{k} G_i(t; z) \] (2.5)

where
\[ G_i(t; z) = \exp \left\{ - \int_0^t \lambda_i(u; z) \, du \right\} \] (2.6)

represents a distribution associated with cause \( C_i \) alone (Elandt-Johnson, 1980).

Alternatively, competing risk problems have also been formulated using the concept of latent failure times (Cox, 1959, Moeschberger and David, 1971 and Gail, 1975). Following this concept, presumably at birth subject is endowed with a set of \( k \) latent failure times, denoted by \( X_1, X_2, \ldots, X_k \) each one corresponding to \( k \) acting causes \( C_1, C_2, \ldots, C_k \). In this formulation, the joint distribution of the latent survival times, called multiple decrement function, is given by
\[ S(t_1, t_2, \ldots, t_k) = P\{X_1 > t, X_2 > t, \ldots, X_k > t\} = P\left\{ \bigcap_{i=1}^{k} (X_i > t) \right\} \] (2.7)
where \(0 < t_i < \infty, i = 1, 2, \ldots, k\), \(S(0, 0, \ldots, 0) = 1\) and \(S(\infty, \ldots, \infty) = 0\). Since \((X_1, X_2, \ldots, X_k)\) can not be observed jointly, the same holds for the multiple decrement function \(S(t_1, t_2, \ldots, t_k)\). The observable quantity is

\[
T = \min (X_1, X_2, \ldots, X_k),
\]

(2.8)

the overall survival distribution, already given in (2.1), can be expressed as

\[
S_T(t; z) = P\{T \geq t\} = P \{\bigcap_{i=1}^{k} (X_i > t)\} = S(t, t, \ldots, t)
\]

(2.9)

and the cause-specific hazard functions given in (2.3) can be written as

\[
\lambda_i(t; z) = \lim_{\Delta t \to 0} \frac{P\{t \leq T \leq t + \Delta t, I = i | T \geq t; z\}}{\Delta t}
\]

\[
= \left[ \frac{\partial \log S(t_1, t_2, \ldots, t_k)}{\partial t_i} \right]_{t_1 = t_2 = \ldots = t_k = t}
\]

(2.10)

for \(i = 1, 2, \ldots, k\). Elandt-Johnson (1980) called this hazard function the "crude hazard rate". The overall hazard function given in (2.2) can also be written as

\[
\lambda_T(t) = \left[ - \frac{\partial \log S_T(t)}{\partial t} \right] = \left[ - \frac{\partial S_T(t)}{\partial t} \right]
\]

(2.11)

The marginal distribution of the latent failure times \(X_i\) is given by

\[
S_i(t) = P (X_i > t) = S(0, 0, \ldots, 0, t, 0, \ldots, 0)
\]

(2.12)
and its corresponding hazard function (net hazard rate) is defined by

\[ h_i(t) = \left[ - \frac{\partial \log S_i(t)}{\partial t} \right] = \left[ - \frac{\partial S_i(t)}{\partial t} / S_i(t) \right] \quad (2.13) \]

for \( i = 1, 2, \ldots, k \). Gail (1975) has interpreted the marginal survival distribution given in (2.12) as the survival distribution for cause \( C_i \) alone.

When the \( X_i \) are independent, the joint survival distribution of \( X_1, X_2, \ldots, X_k \) can be expressed by

\[ S(t_1, t_2, \ldots, t_k) = \prod_{i=1}^{k} S_i(t_i) \quad (2.14) \]

and the overall survival distribution by

\[ S_T(t) = \prod_{i=1}^{k} S_i(t). \quad (2.15) \]

Furthermore, when the \( X_i \) are independent,

\[ \lambda_i(t) = h_i(t) \quad (2.16) \]

and

\[ G_i(t) = S_i(t). \quad (2.17) \]

Elandt-Johnson (1975) called (2.16) the "identity of forces of mortality". Note that when (2.16) holds, the \( X_i \) are not necessarily independent (Gail, 1975); however, the events \( \{X_i > t\}, i = 1, 2, \ldots, k \) are independent (Elandt-Johnson, 1980).

Suppose now that we have \( r+1 \) subpopulations subject to \( k \) competing causes of failures. The survival experience for a subject from the \( j \)-th subpopulation, \( j = 1, 2, \ldots, r+1 \), can be represented by \( (T, I, z, \delta) \), where \( T \in (0, \infty) \) is the observed survival time, \( \delta \) is a censoring indicator which is one if the observation is not censored and zero if censored, \( I \) indicate the cause of failure, \( i = 1, 2, \ldots, k \), and \( z \) is a \((p \times 1)\) vector of covariates. The cause-specific hazard function can be defined as in (2.3) or as in (2.10) adding the
subscript $j$, that is, can be represented as $\lambda_{ij}(t)$, for $i=1, 2, \ldots, k$ and $j=1, 2, \ldots, r+1$. We may formulate the model as in the latent time formulation, postulating the existence of hypothetical survival times $X_i$ for each $C_i$ under the actual study conditions. Then \{${X_{ij(l)}}, 1 \leq i \leq k, 1 \leq j \leq r+1, l=1, 2, \ldots, n_i$\} denote the survival times from one of $r+1$ independent random samples subject to $k$ competing causes of failures. The observed failure time is defined by $T_{j(l)} = \min (X_{1j(l)}, X_{2j(l)}, \ldots, X_{kj(l)})$. As shown by Prentice et al (1978), observations on $(T, I, z, \delta)$ permit only the estimation of those cause-specific functions without further assumptions. These in turn permit the estimation of related quantities, such as the crude survival probability, the probability that an individual will survive up to age $t$ and die from cause $C_i$,

$$Q_{ij}(t) = P\{T > t, I = i | J = j\} = \int_t^\infty - \lambda_{ij}(u) \exp \{- \Lambda(u; z) du\}$$  \hspace{1em} (2.18)

where

$$\Lambda(t; z) = \int_0^\infty \lambda_T(u; z) du, \quad \lambda_T(t) = \sum_{j=1}^{r+1} \sum_{i=1}^k \lambda_{ij}(t; z)$$  \hspace{1em} (2.19)

and the crude survival function $G_{ij}(t; z)$ defined in (2.6). The functions $Q_{ij}(t)$ and $G_{ij}(t)$ are the basic distributions for cause $C_i$ (Elandt-Johnson, 1980). Certainly the overall survival distribution $S_T(t)$ defined in (2.5) is also estimable.

1.3. Review of Competing Risk Models

In this section we present a review of some parametric, semiparametric and nonparametric models for a specific class of the competing risk problems. The principal issues that competing risk model has been aimed at studying are (following Prentice et al., 1980):

1. Inference on the effects of treatments, exposure or other variables on specific types of failures;

2. Study of the interrelation among failure types; and,
3. Estimation of failure rates for certain types of failure given the "removal" of some or all other failure types.

We intend to present a review of the pertinent material focusing on the first class of problems, in parametric and semiparametric approaches, in subsections (1.3.1) to (1.3.2). Before the review of competing risk models, we need to review some concepts useful in the study of likelihood. Some specific problems can arise when mortality with human populations is dealt in an experimental type setting, where the patients could be alive at the end of the study or could be lost to observation for some reason (e.g. change of address, disinterest, etc.). In these cases, mortality data are incomplete since the time of failure of those individuals is unknown. In the terminology of reliability studies, those subjects which are alive at the end of the study are said to have "censored" or "truncated" lifetimes. This censoring scheme is called Type I censoring or truncation. In animal studies or clinical trials, patients may enter the study at different times. Censoring may occur in the following forms: loss to follow-up, dropout or termination of the study. The censoring times, that is, the lengths of time until one of the censoring forms above, are random. This censoring scheme is called random censoring and Type I is a particular case of the random censoring scheme. In this type of censoring it is usual to assume that time of failure and censoring are independent (independent random censorship model), which in many situations is not a realistic assumption. Another form of incomplete data can arise when a preassigned number of failures determine the end of the study. The subjects who remain alive are said to have Type II censoring. This led to increased concern with models that allow for the incorporation of censoring and concomitant variables in the analysis of single risk survival data as well as of competing risk data.

1.3.1. Parametric Models

Parametric models for competing risk problems assume the underlying distribution
of failure times $S(t_1, t_2, \ldots, t_k)$ known.

One of the problems here in competing risk situation consists then in the estimation and inference about the parameters of the underlying distribution from the distribution of the identified minimum $(T, I)$. Given the joint probability distribution of $(T,I)$,

$$P_i(t) = P(T < t) = P(X_i < t, X_i < X_j, i = j) = \int_0^t \left[ \int_{0}^{\infty} \int_{t_i}^{\infty} f(t_1, \ldots, t_k) \prod_{r=1}^{k} dt_r \right] dt_i \quad (3.1.1)$$

estimable from the data, the joint survival distribution $S(t_1, t_2, \ldots, t_k)$ can not be estimated nor its marginals, without the assumption of the independence of failure times (Berman, 1963, Tsiatis, 1975, Elandt-Johnson, 1979 among others). With the knowledge of the form of parametric model, it is possible to estimate all the parameters given the distribution of $(T,I)$ defined above (3.1.1). Some particular results concerning the study of identifiability of the parameters of $S(t_1, \ldots, t_k)$ given the distribution of $(T,I)$ are known, for the case of dependent failure times. Nadas (1971), Basu and Ghosh (1978, 1980) have proved that the distribution of $(T,I)$ identifies the parameters of normal distribution and bivariate exponential distribution of Marshall and Olkin (1967), bivariate distribution of Gumbel (1960) and bivariate Weibull distribution. Moeschberger and David (1971), Moeschberger (1974), David and Moeschberger (1978) and more recently Moeschberger and Klein (1985) have presented some studies dealing with parametric independent and dependent models in competing risk situations.

Another type of problem arises when concomitant variables are present and the underlying survival distribution of failure times is assumed to be known, which is equivalent to specifying the form of $\lambda_{ei}(t)$, the hazard function of the underlying survival when covariates are ignored. In general, the hazard function in competing risk problems can be written as

$$\lambda_i(t; z) = \lambda_{ei}(t) c(z; \beta) \quad i = 1, 2, \ldots, k \quad (3.1.2)$$
where $\beta$ is the vector of parameters corresponding to $z'=(z_1, z_2, \ldots, z_p)$ vector of covariates, $c(z; \beta)$ is any function of $z$ and $\beta$ so that $c(z; 0)=1$. The covariates and their corresponding parameters can be introduced in the likelihood and estimation and inference can proceed as usual. However, parametric models with covariates have not been considered extensively. Lagakos (1977) has extended Glasser’s exponential model (1967) for competing risk problems, allowing for type I censoring. The cause-specific hazard function was defined by

$$\lambda_i(t) = \exp \{\alpha_i + \beta_i z\} \quad i = 1, 2, \ldots, k$$

(3.1.3)

where $\alpha_i$ and $\beta_i=(\beta_{i1}, \beta_{i2}, \ldots, \beta_{ip})$ are the parameters to be estimated; a full likelihood was derived for estimating and testing the parameters. Kalbfleisch and Prentice (1980) have suggested an "accelerated failure times model" in which the covariate $z$ accelerates or decelerates the time of failure with hazard function given by

$$\lambda_i(t | z) = \lambda_0(t e^{-\beta'z}) e^{-\beta'z} \quad i = 1, 2, \ldots, k$$

(3.1.4)

In particular, when $\beta_i \cap \beta_j = 0, i \neq j$, the above model can be written as

$$\lambda_i(t | z) = \lambda_0(t e^{-\beta'_i z}) e^{-\beta'_i z} \quad i = 1, 2, \ldots, k$$

(3.1.5)

In this case, due to the factorization of the likelihood function, survival analysis for a single risk can be used to estimate the $\beta_i$, $i=1, 2, \ldots, k$. Farewell and Prentice (1977) consider the exponential, Weibull, and log-normal distributions as special cases of the model.
1.3.1.1 The Likelihood Function

Suppose that the underlying distribution $S(t_1, t_2, ..., t_k)$ with $p$ parameters $\gamma' = (\gamma_1, \gamma_2, ..., \gamma_p)$ is assumed to be known. The hazard function $\lambda_i(t; \gamma)$ in (2.3) and the survival function for cause $G_i(t)$ in (2.6) are estimable from the data. Likelihood methods are employed for the estimation of the $r$ parameters of the underlying distribution.

David and Moeschberger (1971, 1978) have provided a formal justification of the likelihood function, in the case of independent failure times and when the data are complete, that is, without censoring. When $n$ individuals are observed and $d_i$ of them fail due to $C_i$, the likelihood function can be written as

$$L_1 = \frac{n}{\prod_{i=1}^{n}} \prod_{i=1}^{k} \prod_{r=1}^{d_i} \lambda_i(t_{i(r)}; \gamma_i) S_{T}(t_{i(r)})$$

(3.1.1.1)

where $t_{i(r)}$ denote the time to failure of the $r$-th subject who fails from $C_i$. For the case of incomplete data, they formulated the likelihood incorporating the data from Type I censoring if his lifetime $T_i = \min (X_{i1}, X_{i2}, ..., X_{ik})$, $l=1, 2, ..., n$ is greater than the censoring time $\tau_i$. For the case of type II censoring, only a preassigned number of failures are observed ($m < n$). The remaining $(n-m)$ subjects are considered under type II censoring. They have assumed that all subjects were under observation for the same length of time. Let $d_i$ denote the number of subjects who fail due to $C_i$ and $m = \sum d_i$, the total number of failures; $s$ denotes the total number of survivals such that $n=m+s$, where $n$ is the total number of observed subjects. Let $t_{i(r)}$ denote the time to failure of the subject with the $r$-th lifetime among those who fail due to $C_i$, $i=1, 2, ..., k$ and $r=1, 2, ..., n_i$. The likelihood function allowing for type I and type II censoring can be written as

$$L_2 \propto \prod_{i=1}^{k} \prod_{r=1}^{d_i} \lambda_i(t_{i(r)}; \gamma_i) S_{T}(t_{i(r)}) \prod_{u=1}^{s} S_{T}(\tau(u))$$

(3.1.1.2)
where
\[ t_{i(1)} < t_{i(2)} < \ldots < t_{i(d_i)}, \quad i = 1, 2, \ldots, k, \quad t_{i(d_i)} < \tau_{i(r)}, \quad r = 1, 2, \ldots, d_i, \quad \tau_{i(u)}, \quad u = 1, 2, \ldots, s \]
denotes the censoring times of the \( s \) survivors and \( \tau_{i(r)} \) denotes the fixed censoring times of the subjects whose failure time is \( t_{i(r)} \). Using (2.5) the likelihood function (3.1.1.1) and (3.1.1.2) can be written, respectively, as

\[
L_1 \propto \prod_{i=1}^{k} \left\{ \left[ \frac{d_i}{\prod_{r=1}^{d_i}} \lambda_i(t_{i(r)}; \gamma_i) \right]^2 \left[ \frac{1}{\prod_{l=1}^{d_i}} \prod_{r=1}^{d_i} S_T(t_{l(r)}; \gamma_i) \right] \right\} \tag{3.1.1.3}
\]

\[
L_2 \propto \prod_{i=1}^{k} \left\{ \left[ \frac{1}{\prod_{r=1}^{d_i}} \lambda_i(t_{i(r)}; \gamma_i) \right]^2 \left[ \frac{1}{\prod_{l=1}^{d_i}} \prod_{r=1}^{d_i} S_T(t_{l(r)}; \gamma_i) \right] \left[ \frac{1}{\prod_{r=1}^{d_i}} G_i(\tau_{i(u)}; \gamma_i) \right] \right\} \tag{3.1.1.4}
\]

Alternatively, Prentice et al (1978), assuming the cause-specific hazard functions \( \lambda_i(t_j, z_j), \quad j = 1, 2, \ldots, n \) for each individual are known, and consequently the overall survival function, arrived at a similar likelihood function, under an independent censoring mechanism (which includes type I and type II censoring):

\[
L_3 \propto \prod_{j=1}^{n} \left\{ \lambda_i(t_{i(r)}; \gamma_i) \right\}^{\delta_j} S(t_j; z_j) \tag{3.1.1.5}
\]

\[
= \prod_{j=1}^{n} \left[ \lambda_i(t_{i(r)}; \gamma_i) \right]^{\delta_j} \prod_{i=1}^{k} \exp \left\{ - \int_{0}^{t_j} \lambda_i(u; z(u)) \, du \right\}
\]

where

\[
\delta_j = \begin{cases} 
1 & \text{if the j-th individual fails} \\
0 & \text{if censored.}
\end{cases}
\]

David and Moeschberger (1978) as well as Prentice et al (1978) have pointed out
the likelihood factorization which allows one to perform a separate maximization of each of \( L_i, \ i=1, 2, \ldots, k \) when each \( \lambda_i(t; \gamma_i) \) has a different set of parameters, that is, \( \gamma_i \cap \gamma_j = \emptyset, \ i \neq j \). Also, factorization as in (3.1.1.3) and (3.1.1.4) justifies the common procedure of treating failure from other causes than \( C_i \) as well as live withdrawals as censored.

1.3.1.2. The Maximum Likelihood Estimators

Suppose that we have \( n \) independent observations vectors on \( x, x_1, x_2, \ldots, x_n \) from a distribution with probability density function \( f(x; \theta) \), where \( \theta \) is the vector of unknown parameters lying in a \( k \)-dimensional parameter space \( \Omega \) and \( x \) is a \( p \)-dimensional random vector with values over a region \( R \) independent of \( \theta \). Suppose further that \( \theta \) is partitioned as \( \theta = (\theta_1, \theta_2)' \), where \( \theta_1 \) is \( p \times 1 \) and \( \theta_2 \) is \((k-p)\times 1\). The likelihood function for \( \theta \) is defined as

\[
L(\theta) = \prod_{i=1}^{k} f(x_i; \theta)
\]  

(3.1.2.1)

Let \( \hat{\theta} \) be a point of \( \Omega \) at which \( L(\theta) \) is maximized; \( \hat{\theta} \) is called a maximum likelihood estimate (m.l.e) of \( \theta \). It is common to work with \( \log L(\theta) \), which is also maximized at \( \hat{\theta} \), and in many cases \( \hat{\theta} \) can be readily found by solving the likelihood equations

\[
U_i(\theta) = 0 \quad (i = 1, 2, \ldots, k),
\]

where

\[
U_i = \frac{\partial \log L(\theta)}{\partial \theta_i} \quad i = 1, 2, \ldots, k
\]

The \( U_i(\theta) \)'s are called scores and the \( k \times 1 \) vector \( U(\theta) = (U_1 \ U_2) \) is called the score vector. The matrix

\[
I(\theta) = \begin{bmatrix}
I_{11}(\theta) & I_{12}(\theta) \\
I_{21}(\theta) & I_{22}(\theta)
\end{bmatrix}
\]
with entries
\[ I_{ij}(\theta) = E \left( -\frac{\partial^2 \log L(\theta)}{\partial \theta_i \partial \theta_j} \right) \quad i, j = 1, 2, \ldots, k \]
is called the Fisher information matrix.

If the likelihood \( L(\theta) \) satisfies the regularity conditions (assumptions A1-A4), some large sample properties of \( \theta \) hold. The regularity conditions are:

A1. For almost all \( x \in \mathbb{R} \) and for all \( \theta \in \Omega \),
\[
\frac{\partial \log f(x; \theta)}{\partial \theta_r}, \frac{\partial^2 \log f(x; \theta)}{\partial \theta_r \partial \theta_s} \text{ and } \frac{\partial^3 \log f(x; \theta)}{\partial \theta_r \partial \theta_s \partial \theta_t}
\]
exist for \( i, j, t = 1, 2, \ldots, k \).

A2. For almost all \( x \in \mathbb{R} \) and for every \( \theta \in \Omega \),
\[
|\partial f/\partial \theta_r| < F_r(x) \quad \text{and} \quad |\partial^2 f/\partial \theta_r \partial \theta_s| < F_{rs}(x),
\]
where \( F_r(x) \) and \( F_{rs}(x) \) are integrable over \( \mathbb{R}, r, s = 1, \ldots, k \).

These assumptions permit certain interchanges of order of differentiation and integration.

A3. For every \( \theta \in \Omega \), the matrix \( I(\theta) = [I_{rs}(\theta); r, s = 1, 2, \ldots, k] \) with
\[
I_{rs}(\theta) = E_\theta \left[ \frac{\partial \log f}{\partial \theta_r} \frac{\partial \log f}{\partial \theta_s} \right]
\]
is positive definite with finite determinant.

A4. For almost all \( x \in \mathbb{R} \) and for all \( \theta \in \Omega \),
\[
\left| \frac{\partial^3 \log f}{\partial \theta_r \partial \theta_s \partial \theta_t} \right| < H_{rst}(x),
\]
where there exists a positive real number \( M \) such that
\[ \mathbb{E}_\theta \left[ H_{rst}(x) \right] < M < \infty, \]
for all \( \theta \in \Omega \) and \( r,s = 1, 2, \ldots, k. \)

Condition A4 can be replaced by the following condition:

\[ \lim_{\delta \to 0} \mathbb{E} \left\{ \sup_{\theta : ||\theta - \theta_0|| < \delta} \left| \frac{\partial^2 \log f}{\partial \theta_r \partial \theta_s} \theta_0 - \frac{\partial^2 \log f}{\partial \theta_r \partial \theta_s} \theta_0 \right| \right\} = 0. \]

Under the regularity conditions above, \( \hat{\theta} \) is a consistent estimator of \( \theta \). In addition, several other asymptotic results hold that lead to useful inference procedures. First, \( U(\theta) \) is asymptotically \( \mathbb{N}_k[0, I(\theta)]. \) Second, \( I(\hat{\theta}) \) is a consistent estimator of \( I(\theta) \). As a result, under the hypothesis \( H_0: \theta = \theta_0 \) the score test statistic \( U'(\theta_0) I(\theta_0)^{-1} U(\theta_0) \) is \( \chi^2_k \). Note that this test does not require the calculation of \( \hat{\theta} \). Tests and estimates for a subset of the \( \theta_i \)'s can also be obtained. Suppose that we want to test the following hypothesis: \( H_0: \theta_1 = \theta_{10}, \theta_2 \) unspecified. Denote \( \hat{\theta} = [\theta_{10}, \hat{\theta}_2(\theta_{10})] \), where \( \hat{\theta}_2(\theta_{10}) \) is the m.l.e. of \( \theta_2 \), obtained by maximizing \( L(\theta_{10}, \theta_2) \). Then, under \( H_0: \theta_1 = \theta_{10}, \theta_2 \) unspecified, \( U_1(\hat{\theta})' [I_1(\hat{\theta})]^{-1} U_1(\hat{\theta}) \) is asymptotically \( \chi^2_{(p)}, I_1(\hat{\theta}) = I_{11}(\theta) - I_{12}(\theta) I_{22}(\theta) I_{21}(\theta). \)

Since in only a few situations can m.l.e.'s be found analytically, iterative methods are required for obtaining \( \hat{\theta} \). Newton-Raphson iteration and the method of scoring are the most commonly used.

1.3.2. Semiparametric Approach

The most attractive model which incorporates concomitant variables in hazard function was proposed by Cox (1972) for a single risk survival problem. It was called the proportional hazard model and it assumes that

\[ \lambda(t; z) = \lambda_0(t) \exp(\beta'z) \] (3.2.1)
where the underlying hazard function $\lambda_0(t)$ is assumed to be the same for all subjects in the study but not assumed to be known, $z$ is the $(p \times 1)$ vector of covariates, and $\beta$ is the vector of regression parameters. In Cox's formulation, the vector $z$ can be time dependent in which case should be denoted by $z(t)$. He proposed a "partial likelihood" approach in order to estimate the coefficients of the regression variables and studied the asymptotic properties of $\beta$. He proposed to use the score statistic test for assessing the hypothesis $H_0: \beta = 0$. An important feature of Cox's formulation is the possibility of estimation of the underlying survival distribution $S_0(t)$ by the estimation of $\lambda_0(t)$ and $\beta$. Cox (1972), Kalbfleisch and Prentice (1973), Breslow (1975) and more recently Tsiatis (1981) and Link (1984) have suggested methods of estimation of the underlying survival function. Asymptotic properties of the estimates have been studied by Tsiatis (1978, 1981), Bailey (1983) and Link (1984). Tsiatis gives a proof of asymptotic normality of $\beta$ and $A_0(t) = \int \lambda_0(u) du$ using stochastic processes in a random sampling and random censoring setting. The estimates of the asymptotic variances calculated from the limiting processes are the same as obtained by himself in 1978 using a heuristic model. Bailey utilized a projection technique due to Hajek leading to an asymptotic representation of the score function as a sum of independent (but identically distributed) random variables. For proving asymptotic normality of the cumulative hazard function $A_0(t)$, he utilized an asymptotic representation of the estimator as the sum of an independent increment process and another term which is the product of $(\hat{\beta} - \beta)$ and a fixed function of time. Finally, Link has proposed a smooth estimator of Breslow estimator of the cumulative hazard function and used a heuristic argument to derive its asymptotic variance. No formal study of asymptotic normality was provided for this estimator.

Extension to competing risk problems has been proposed by Holt (1978) and by Kalbfleisch and Prentice (1978). The cause-specific hazard functions at time $t$ for a subject with covariates $z' = (z_1, z_2, \ldots, z_p)$ was formulated as
\[ \lambda_i(t; z) = \lambda_{0i}(t) \exp(\beta_i'z) \quad i = 1, 2, \ldots, k \] (3.2.2)

Here, \( \lambda_{0i}(t) \geq 0 \) are the hazards for cause \( C_i \) if covariates are ignored and may vary for different causes \( C_i \). Using the conditional argument, that is, computing the product of conditional probabilities for each cause \( C_i \) at time \( t_{i(j)} \) conditional on the risk set \( R\{t_{i(j)}\} \), the partial likelihood can be written as

\[ L_1(\beta) \propto \prod_{i=1}^{n} \frac{\exp\{\beta_{i(l)}'z_l\}}{\sum_{l \in R\{t_{i(j)}\}} \exp\{\beta_{i(l)}'z_l\}} = \prod_{i=1}^{k} \prod_{j=1}^{d_i} \frac{\exp\{\beta_{i}z_j\}}{\sum_{l \in R\{t_{i(j)}\}} \exp\{\beta_{i}z_l\}} \] (3.2.3)

where \( t_{i(1)} \), \ldots, \( t_{i(d_i)} \) are the \( d_i \) ordered observed survival times for cause \( C_i \). This likelihood allow \( z \) to be time dependent. Due to the factorization into \( k \) components, estimation of \( \beta_i \)'s can be obtained by applying standard asymptotic likelihood techniques in \( k \) factors, assuming no common parameters among the \( \beta_i \)'s exist.

When \( \lambda_{0i}(t) \geq 0 \) do not vary for different causes, the cause-specific hazard function can be written as

\[ \lambda_i(t; z) = \lambda_{0}(t) \exp(\beta_i'z) \quad i = 1, 2, \ldots, k \] (3.2.4)

and the Kalbfleisch and Prentice (1973) invariant argument of marginal distribution of ranks of observations leads to the marginal likelihood

\[ L_2(\beta) \propto \prod_{i=1}^{n} \frac{\exp\{\beta_{i(l)}'z\}}{\sum_{h \in R\{t_i\}} \sum_{i=1}^{k} \exp\{\beta_i'z_h\}}. \] (3.2.5)

The likelihoods \( L_1(\beta) \) and \( L_2(\beta) \) are identical when the risks are assumed to be independent.
1.4. **Restricted Alternatives and the Union-Intersection Test.**

Suppose we want to test a hypothesis against a restricted alternative, that is,

\[
\{H_0: \omega \in \Omega_0\} \text{ against } \{H_a^{(1)}: \omega \in \Omega_1\},
\]

where \( \Omega_0 \) and \( \Omega_1 \) are subsets of the parameter space \( \Omega \), such that \( \Omega_0 \cup \Omega_1 \) is a proper subset of \( \Omega \). The test statistic derived for testing the hypothesis above is usually more powerful than the statistic derived for the test with unrestricted alternative

\[
\{H_0: \omega \in \Omega_0\} \text{ against } \{H_a^{(2)}: \omega \not\in \Omega_1\}
\]

when actually \( \omega \in \Omega_1 \). Barlow et al (1972) present a thorough discussion of various parametric and nonparametric tests in univariate setups.

In the multivariate setup, Kudo (1963) derived the likelihood ratio statistic for testing the hypothesis that the mean vector of a \( p \)-variate normal distribution is null against an orthant restriction when the covariance matrix is known. If \( X_1, X_2, \ldots, X_n \) are independent, identically distributed \( N_p(\mu, \Sigma) \), then the hypothesis above can be formulated as \( H_0: \mu = 0 \) against \( H_a: \mu \geq 0 \). Kudo has shown that the LR test rejects the hypothesis \( H_0 \) in favor of \( H_a \) above for large values of

\[
Q^2 = N \{ \bar{X}' \Sigma^{-1} \bar{X} - (\bar{X} - \mu^*)' \Sigma^{-1}(\bar{X} - \mu^*) \},
\]

where \( \mu^* \) is the maximum likelihood estimator of \( \mu \) in the region \( \Omega_1 = \{ \mu: \mu \geq 0 \} \). Kudo outlined an algorithm for calculating \( Q^2 \). Let \( J \) be any subset of \( \mathcal{P} = \{1, 2, \ldots, p\} \) and \( \bar{J} \) its complement. For each \( J, \emptyset \subseteq J \subseteq \mathcal{P} \), let
\[ \mathbf{X} = \begin{bmatrix} \mathbf{X}_{(j)} \\ \mathbf{X}_{(j')} \end{bmatrix} \]

and

\[ \Sigma = \begin{bmatrix} \Sigma_{(jj)} & \Sigma_{(jj')} \\ \Sigma_{(j'j)} & \Sigma_{(j'j')} \end{bmatrix} \]

denote the corresponding partitions of \( \mathbf{X} \) and \( \Sigma \). In addition, we need to define for each set \( J, \emptyset \subseteq J \subseteq P \),

\[ \Sigma_{(jj; j')} = \Sigma_{(jj)} - \Sigma_{(jj')} (\Sigma_{(j'j')}^{-1}) \Sigma_{(j'j)} \tag{4.3} \]

\[ \mathbf{X}_{(j; j')} = \mathbf{X}_j - \Sigma_{(jj')} (\Sigma_{(j'j')}^{-1}) \mathbf{X}_{(j')} \tag{4.4} \]

Kudo has shown that

\[ Q^2 = \sum_{\emptyset \subseteq J \subseteq P} N \mathbf{X}_{(j; j')} (\Sigma_{(jj; j')})^{-1} \mathbf{X}_{(j; j')} I(\mathbf{X}_{(j; j')} > 0, (\Sigma_{(j'j')}^{-1}) \mathbf{X}_{(j')} < 0) \tag{4.5} \]

where \( I(A) \) represents the indicator function for the set \( A \). Under \( H_0: \mu = 0 \), the distribution of \( Q^2 \) in (4.5) reduces to

\[ P(Q^2 \leq q) = \sum_{j=0}^{p} w_j P(\chi_j^2 \leq q) \tag{4.6} \]

where for each \( j, 0 \leq j \leq p \), \( \chi_j^2 \) represents a chi-squared random variable with \( j \) degree of freedom and \( w_j \) is the sum of \( \binom{p}{j} \) orthant probabilities of the normal random vectors \( [\mathbf{X}_{(j; j')}, - (\Sigma_{(j'j')}^{-1}) \mathbf{X}_{(j')} ] \) for all subsets \( J \) with \( j \) elements.

Nuesch (1966) simplified the derivation of \( Q^2 \) and its distribution under the null hypothesis, by reducing it from a maximum likelihood problem to a quadratic programming problem. He also developed the LR test statistic and its distribution for \( H_0: \mu = 0 \) against \( H_a: \mu \geq 0 \) when the covariance matrix \( \Sigma = \sigma^2 \Sigma_0 \), where \( \Sigma_0 \) is completely specified and \( \sigma^2 \) is unknown. For this particular case, the LR test rejects \( H_0: \mu = 0 \) in
favor of $H_a: \mu \geq 0$ for large values of

$$Q_0^2 = Q^2 / \sum_{i=1}^{N} X_i' \Sigma_0^{-1} X_i$$  \hspace{1cm} (4.7)$$

where $Q^2$ is the statistic in (4.5) with $\Sigma_0$ replacing $\Sigma$. Under $H_0: \mu = 0$, the distribution of $Q_0^2$ reduces to

$$P(Q_0^2 \leq q) = \sum_{j=0}^{p} w_j P(\beta(\frac{3}{2}, \frac{(Np-j)}{2}) \leq q)$$  \hspace{1cm} (4.8)$$

where $\beta(s,t)$ represents a Beta random variable with parameters $s$ and $t$ ($\beta(0,t) = 0$).

Chatterjee and De (1972) first utilized the union-intersection (UI) principle to derive a bivariate linear rank test and also Chinchilli and Sen (1981a), using the UI principle, derive asymptotically distribution-free tests for a class of restricted alternative problems in the general linear multivariate model.

The union-intersection principle of test construction suggested by Roy (1953) is intended for multivariate situations. Suppose that the composite null hypothesis $H_0$ and the composite alternative $H_a$ can be decomposed as

$$H_0 = \bigcap_{\nu \in \Gamma} H_{0\nu}, \quad H_a = \bigcup_{\nu \in \Gamma} H_{a\nu}$$  \hspace{1cm} (4.9)$$

where $\Gamma$ is an index set such that reasonable tests for $H_{0\nu}$ vs $H_{a\nu}$ are available. Specifically, suppose that the acceptance region for $H_{0\nu}$ against $H_{a\nu}$ is $A_{\nu}$, $\nu \in \Gamma$. The union-intersection principle dictates that $H_0$ be accepted against $H_a$ over $A = \bigcap_{\nu \in \Gamma} A_{\nu}$ and rejected on its complement. The UI test for $H_0$ vs $H_a$ depends both upon the decomposition of the null and alternative hypothesis and upon the family of tests for component hypotheses. For example, if $X \sim N(\mu, \Sigma)$, then $a'X \sim N_1(a', a\Sigma a)$ for every $a \in \mathbb{R}^p$ and the hypothesis $H_0: \mu = \mu_0$ has natural decomposition.
\[ H_0 = \bigcap_{a \in \mathbb{R}^p} \{ H_0(a): a'\mu = a'\mu_0 \} \]

The original multivariate hypothesis of \( \mu = 0 \) is true if and only if \( a'\mu = a'\mu_0 \) for all non-null \( a \). Acceptance is equivalent to accepting all univariate hypotheses for varying \( a \).

Since the univariate acceptance region is \( t^2(a) \leq t^2_{\alpha/2, N-1} \), the multivariate acceptance region is the intersection \( \bigcap_a [t^2(a) \leq t^2_{\alpha/2, N-1}] \), which is the same as specifying that \( \max_a [t^2(a) \leq t^2_{\alpha/2, N-1}] \).

Therefore, using the UI principle it is necessary to maximize some function. For tests against ordered alternatives the function is nonlinear and must be maximized subject to inequality constraints. The nonlinear programming maximization methods may be applied here in order to obtain the maximum of that function. In nonlinear programming terminology the problem is as follows.

Determine the vector \( \lambda \in \Lambda \) such that \( h(\lambda) = \inf_{\lambda \in \Lambda} h(\lambda) \), where

1) \( h(\lambda) \) is a scalar valued function of \( \lambda \), and

2) \( \Lambda = \{ \lambda \in \mathbb{R}^p : h_1(\lambda) \leq 0, h_2(\lambda) = 0 \} \), where \( h_1 \) and \( h_2 \) are \( r_1 \) and \( r_2 \) vector valued functions of \( \lambda \).

This is known as minimization problem with constraints. \( h \) is called the objective function and \( h_1 \) and \( h_2 \) are called the constraint functions. Another important function which assists in solving the minimization problem is the Lagrangean function

\[ L(\lambda, t_1, t_2) = h(\lambda) + t_1' h_1(\lambda) + t_2' h_2(\lambda) \quad (4.10) \]

where \( t_1 \) and \( t_2 \) are \( r_1 \) and \( r_2 \) vectors.

A Kuhn-Tucker-Lagrange (KTL) point is any point \( (\lambda^*, t_1, t_2) \) which satisfies the system of conditions
\begin{align*}
\text{i) } t_1 & \geq 0 \\
\text{ii) } h_1(\lambda) & \leq 0 \\
\text{iii) } h_2(\lambda) & = 0 \\
\text{iv) } t_1' h_1(\lambda) & = 0 \\
\text{v) } \partial L(\lambda, t_1, t_2)/\partial \lambda & = 0 \\
\end{align*} 
\tag{4.11}

Sen (1984) has proposed UI principle test statistics for a class of subhypotheses testing against restricted alternatives for the Cox regression model. Sen considered the general regression model due to Cox (1972), involving both design variables (controlled and non-stochastic) and concomitant variables (stochastic). The cause-specific hazard function was defined as

\begin{equation}
\begin{split}
h_i(t; c_i; z_i) = h_0(t) \exp(\beta' c_i + \gamma' z_i), \quad i = 1, 2, \ldots, N, \\
\end{split}
\end{equation} 
\tag{4.12}

where \( h_0(t) \), the hazard rate for \( c_i = 0, z_i = 0 \), is an unknown, arbitrary nonnegative function, \( \beta \) is the vector of unknown parameters relating to the design-effects and \( \gamma \) is the vector of unknown parameters relating to the effect of concomitant variables. For the null hypothesis \( H_0: \beta = 0 \), the following alternatives have been considered: \( H_{a1}: \beta_j \geq 0, \, j = 1, \ldots, r \), \( H_{a2}: \beta_1 \leq \ldots \leq \beta_r \) and \( H_{a3}: 0 \leq \beta_1 \leq \ldots \leq \beta_r \), all the alternative with at least one strict inequality. The partial likelihood was considered and utilizing the UI principle he proposed the following test

\begin{equation}
\begin{split}
L_N^1 = \sup \{ a' \hat{U}_N / (a' \hat{V}_{11} a)^{1/2} : a \in A \} \\
\end{split}
\end{equation} 
\tag{4.13}

with result similar to the statistic given in (4.5).
1.5. Proposed Research

The focus of this dissertation is on the development of statistical tests with restricted alternatives by applying the UI principle in the context of the competing risk situation. Comparisons to tests with unrestricted alternatives are also a focus of this study.

In Chapter II Cox's semiparametric model is formulated in a two-sample competing risk setting with two causes of failures, in the presence of concomitant variables. The UI test is derived for the null hypothesis of no difference between the two groups against the alternative hypothesis that the survival experience in one group is greater than in the second group. The distribution of the test statistic under the null and alternative hypothesis is studied. Efficiency is investigated by the comparison with a test statistic with a unrestricted alternative.

In Chapter III the parametric bivariate exponential model of Sarkar (1987) is considered as well as a logistic model in a two-sample competing risk situation with \( k \) causes of failure. UI tests are derived for the hypothesis of no difference between the two groups.

In Chapter IV ordered alternatives are considered in a similar context as in Chapter II, that is, with the Cox regression model but \( r+1 \) subpopulations subject to \( k \) competing risks. Again, tests statistics are derived using the UI principle and distributions under the null and local alternatives are investigated. Also, the power of the test is analyzed and compared with the power for the unrestricted test.

Finally, in Chapter V, results of a numerical application concerning the performance of the UI test with respect to the score test are presented. Here we present some final comments and a few suggestions for further research.
CHAPTER II
RESTRICTED ALTERNATIVES TESTS IN A PROPORTIONAL HAZARD
MODEL WITH COMPETING RISK DATA

2.1. Introduction

In this Chapter, we consider the partial likelihood approach for the comparison of
the survival distributions derived from two independent subpopulations in a competing
risk situation with two causes of failure. A Union-Intersection test based on partial
maximum likelihood estimators is derived for a restricted alternative test and its
asymptotic distributions under the null and alternative hypothesis are studied. Also the
efficiency of the proposed test with respect to the one for unrestricted alternative is
studied.

The survival experience for an individual to k competing causes of failure is repre-
sented by the bivariate random variable \((T, I)\), where \(T \in (0, \infty)\) is the observed survival
time and \(I \in \{1, 2, \ldots, k\}\) indicates which of the k potential causes actually caused the
failure. The cause-specific hazard function for a subject with covariate
\(z'=(z_1, z_2, \ldots, z_p)\) is taken to be

\[
\lambda_i(t | z) = \lambda_{0i}(t) \exp (\beta_i' z) \tag{1.1}
\]

where the underlying hazard function \(\lambda_{0i}(t)\) may be different for different causes and the
regression parameters are different for each lifetime, that is, if \(\beta_i'=(\beta_{i1}, \ldots, \beta_{ip})\) denotes
the set of the parameters associated with cause \(C_i\), then \(\beta_i \cap \beta_j=\emptyset\) for \(i \neq j\) and
\(i,j=1,2, \ldots, k\).

Let \(t_{i1} < t_{i1} < \ldots < t_{id_i}\) denote the \(d_i\) observed failures of type \(i\) and \(z_{i(t)}\) be the
regression function for the subject that fail at \( t_{i(l)} \). The partial likelihood can be written as in Holt (1978) as
\[
L(\beta) = \prod_{l=1}^{n} \prod_{i=1}^{k} \left\{ \frac{\exp(\beta_i^2 z_{i(l)})}{\sum_{u \in R(t_{i(l)})} \exp(\beta_i^2 z_u)} \right\}^\delta_{i(l)}
\]
(1.2)

where \( n = \sum_{i=1}^{k} \delta_{i(l)} \) is an indicator function which assumes the value 1 if the \( l \)-th subject fails from \( C_i \) and zero otherwise.

Suppose now we have two subpopulations and we are interested in testing the equality of survival distributions from the two subpopulations. We have a set \( \beta_{ij} = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ip}) \) of vectors of parameters for each subject who fails from \( C_i \) and comes from the \( j \)-th population. Let \( \beta_j = (\beta_{1j}, \beta_{2j}) \), \( j = 1, 2 \) represent the \( 2p \)-vector of the parameters corresponding to the \( j \)-th population. In order to compare two subpopulations, the test may be on the parameters \( \beta_{ij} \), specifically of
\[
H_0: \beta_1 = \beta_2
\]
(1.3)

against the alternative
\[
H_a: \beta_1 \neq \beta_2
\]
(1.4)

We may also use the following reparameterization
\[
\beta_1 = \beta_2 + \Delta
\]
(1.5)

where
\[
\Delta = (\Delta_1 \Delta_2) \quad \text{with} \quad \Delta'_j = (\Delta_{i1}, \Delta_{i2}, \ldots, \Delta_{ip}).
\]
(1.6)

Then the hypothesis (1.4) and the alternative (1.5) can be stated as
\[
H_0: \Delta = 0
\]
(1.7)

and
\[
H_a: \Delta \neq 0
\]
(1.8)
considering $\beta_2 = \gamma$ as a nuisance parameter vector.

In order to derive the union-intersection test for the hypothesis and the alternative above, we need to check whether or not the regularity conditions of the maximum likelihood estimator are satisfied when partial maximum likelihood is utilized. The regularity conditions are discussed in Section 2.2.1, whereas the derivation of the statistic is presented in Section 2.2.2. In Section 2.3 we study the asymptotic distribution of the union-intersection test for the null and alternative hypothesis above. Finally, in Section 2.4 we present the study of local asymptotic power and efficiency.

2.2. The Union-Intersection Test Based on Partial Maximum Likelihood Estimators

The construction of the likelihood is one of the central issues in statistical inference. The partial likelihood has been discussed in many studies, especially after Cox (1972).

Consider a random variable $Y$ having probability density function $f_Y(y; \theta)$, where $\theta = (\phi, \eta)$. Suppose that $Y$ is transformed to a new random variable $(v, w)$ by a transformation not depending on the unknown parameters. Also, suppose that we are interested in inference about $\phi$ treating $\eta$ as a nuisance parameter vector. The factorization of the full likelihood leads us to a marginal and conditional likelihood

$$f_Y(y; \theta) = f_Y(v; \phi) \cdot f_{w|v}(w; \eta)$$

The existence of such a factorization is of key importance to the nuisance parameter problem: besides the simplicity of the analysis achieved by the reduction of dimensionality with many nuisance parameters there is a gain in robustness. In this particular case, the likelihood suggested by Cox (1972) is the first term of the full likelihood below, in a situation with hazard function $\lambda_i(t|z)$ as given in (1.1).

$$L(\beta, \eta) = \prod_{i=1}^n \left[ \lambda_i(t_i|z_i) \right] S_{T_i}(t_i, z_i)$$
\[
\prod_{l=1}^{n} \prod_{i=1}^{k} \left\{ \frac{\exp(\beta_i^l z_{i(l)})}{\sum_{u \in \mathcal{R}(t_{i(l)})} \exp(\beta_i^l z_u)} \right\}^{\delta_{i(l)}} \times \n\prod_{l=1}^{n} \prod_{i=1}^{k} \left\{ G_{0i}(t) \times \left[ \sum_{u \in \mathcal{R}(t_{i(l)})} \exp(\beta_i^l z_u) \right]^{-\delta_{i(l)}} \right\}
\]

where the vector of nuisance parameters \( \eta = \{ \lambda_{0i}(t), i = 1, 2, \ldots, k \} \). Therefore, in Cox’s partial likelihood there is no need to estimate the underlying hazard function \( \lambda_i(t) \) and the corresponding survival functions \( G_{0i}(t) \). With the reparameterization in (1.5), the cause specific hazard function (1.1) may be written as

\[
\lambda_i(t|z) = \lambda_{0i}(t). \exp \{ \gamma_i^l + w_{i(l)} \Delta_i^l z_{i(l)} \}
\]

where \( w_{i(l)} \) is equal to 1 if the \( l \)-th subject comes from the first population and is equal to zero if comes from the second population, \( \beta_i^l = (\gamma_i^l, \Delta_i^l) \) \( i = 1, 2, \ldots, k \), \( \beta_i \cap \beta_j = \emptyset \), for \( i \neq j \), where \( \gamma_i(i) = (\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{ip}) \), \( \gamma_{ij} = (\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2}) \) with \( \gamma_{ij0} = 0 \) and \( \Delta_i^l = (\Delta_{i1}, \Delta_{i2}, \ldots, \Delta_{ip}) \), \( \Delta_{ij} = (\Delta_{ij0}, \Delta_{ij1}, \Delta_{ij2}) \). The partial likelihood (1.3) can be written as

\[
L(\gamma, \Delta) = \prod_{l=1}^{n} \prod_{i=1}^{k} \left\{ \frac{\exp(\gamma_i^l + \Delta_i^l w_{i(l)} z_{i(l)})}{\sum_{u} \exp(\gamma_i^l + \Delta_i^l w_{i(u)} z_{i(u)})} \right\}^{\delta_{i(l)}}
\]

\[
= \prod_{l=1}^{n} \left\{ \prod_{i=1}^{k} \left[ p_{i(l)}(\gamma, \Delta) \right]^{\delta_{i(l)}} \right\}
\]

where

\[
p_{i(l)}(\gamma, \Delta; z, t) = \left\{ \frac{\exp(\gamma_i^l + \Delta_i^l w_{i(l)} z_{i(l)})}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \exp(\beta_i^l z_{i(l)})} \right\} \quad i = 1, 2 \text{ and } l = 1, 2, \ldots, n.
\]

where \( I(T_{i(l)} \geq t_{i(l)}) \) is a indicator function which assumes the value 1 if the \( l \)-th subject survived until time \( t \) (belong to the risk set \( \mathcal{R}(t_{i(l)}) \)) and 0 otherwise.
For $l=1$, 2, ..., $n$ we assume the covariates $z$ are independent and identically distributed with $\mu = \mathbb{E}(z)$ and $\sigma = \mathbb{E}(x-\mu) (x-\mu)'$, and $\sigma$ is positive definite with $\text{det}(\sigma) < \infty$.

### 2.2.1 Partial Maximum Likelihood Estimators

Consider the partial likelihood defined in (2.3). If $L(\gamma, \Delta)$ has a unique maximum as a function of $\beta$, $\hat{\beta}$ is the partial maximum likelihood estimator (p.m.l.e). The p.m.l.e of $\gamma$ and $\Delta$ can be found by solving the system of simultaneous equations

\[
\frac{\partial \log L(\gamma, \Delta)}{\partial \gamma_i} = \sum_{l=1}^{n} \delta_{i(l)} \left\{ z_{i(l)} - \frac{\sum_{u} z_{u(l)} \exp (\gamma_u + \Delta_u w_{u(l)}) z_{u(l)}}{\sum_{u} (\gamma_{u(l)} + \Delta_u w_{u(l)}) z_{u(l)}} \right\}
\]  

(2.1.1)

and

\[
\frac{\partial \log L(\gamma, \Delta)}{\partial \Delta_i} = \sum_{l=1}^{n} \delta_{i(l)} \left\{ w_{i(l)} z_{i(l)} - \frac{\sum_{u} z_{u(l)} \exp (\gamma_u + \Delta_u w_{u(l)}) z_{u(l)}}{\sum_{u} (\gamma_{u(l)} + \Delta_u w_{u(l)}) z_{u(l)}} \right\}
\]  

(2.1.2)

Cox treats the partial likelihood as an ordinary likelihood function for the purposes of inference on $\beta$. Then, the usual assumptions of the M.L.E might be satisfied also for the p.m.l.e. They are studied below.

1. For the $l$-th subject and $i$-th cause, the log of the conditional probability of failure given the risk set is

\[
\log p_{i(l)}(\gamma, \Delta; z, t) = \left\{ (\gamma_i + \Delta_i w_{i(l)}) z_{i(l)} - \log \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \exp (\beta_i z_{i(l)}) \right\}
\]

(2.1.3)

which is a continuous function in $\beta \in \beta$ and differentiable. Thus, first and second partial derivatives exist. The existence of the third derivatives can be replaced by the following
condition:

$$
\lim_{\epsilon \to \infty} E \left\{ \sup_{|\beta^* - \beta| < \epsilon} \left| \frac{\partial^2 \log p_1(\beta^*)}{\partial \beta_i^* \partial \beta_i} - \frac{\partial^2 \log p_1(\beta)}{\partial \beta_i \partial \beta_i} \right| \right\} = 0
$$

Let us consider a $p \times p$ matrix whose elements are

$$
\frac{\partial^2 \log p_1(\beta)}{\partial \gamma_i \partial \gamma_i'} =
$$

$$
= \left\{ \sum_{l=1}^{n} \frac{I(T_{i(l)} \geq t_{i(l)}) \ z_{i(l)} \ z_{i(l)}'}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \ exp(\gamma_i + \Delta_i' \ w_{i(l)})} \ z_{i(l)} \right\}
$$

$$
- \left\{ \sum_{l=1}^{n} \frac{I(T_{i(l)} \geq t_{i(l)}) \ z_{i(l)} \ exp(\gamma_i + \Delta_i' \ w_{i(l)})}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \ exp(\gamma_i + \Delta_i' \ w_{i(l)})} \ z_{i(l)} \right\}
$$

$$
\times \left\{ \sum_{l=1}^{n} \frac{I(T_{i(l)} \geq t_{i(l)}) \ z_{i(l)} \ exp(\gamma_i + \Delta_i' \ w_{i(l)})}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \ exp(\gamma_i + \Delta_i' \ w_{i(l)})} \ z_{i(l)} \right\}
$$

(2.1.4)

We can write the last product term as

$$
\left( \sum_{l=1}^{n} a_i b_l \ z_{i(l)} \right) \times \left( \sum_{l=1}^{n} a_i b_l \ z_{i(l)}' \right)
$$

$$
= \left[ \left( \sum_{l=1}^{n} a_i b_l \ (z_{i1(l)} \ z_{i2(l)} \ldots \ z_{ip(l)})' \right) \right] \left[ \left( \sum_{l=1}^{n} a_i b_l \ (z_{i1(l)} \ z_{i2(l)} \ldots \ z_{ip(l)}) \right) \right]
$$

$$
= \left[ \left( \sum_{l=1}^{n} a_i b_l \ z_{i1(l)} \right) \sum_{l=1}^{n} a_i b_l \ z_{i2(l)} \ldots \sum_{l=1}^{n} a_i b_l \ z_{ip(l)} \right]'
$$

$$
\times \left[ \left( \sum_{l=1}^{n} a_i b_l \ z_{i1(l)} \right) \sum_{l=1}^{n} a_i b_l \ z_{i2(l)} \ldots \sum_{l=1}^{n} a_i b_l \ z_{ip(l)} \right]
$$

(2.1.5)
The pq-th element of this matrix can be written as

\[
\left( \sum_{l=1}^{n} a_l b_l z_{i,p(l)} \right) \left( \sum_{l=1}^{n} a_l b_l z_{i,q(l)} \right) \geq \sum_{l=1}^{n} a_l^2 b_l^2 z_{i,p} z_{i,q}
\]

Then

\[
\left( \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) z_{i,l} \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)} \right) \times \left( \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) z_{i,l}^* \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)} \right) \geq \left( \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) z_{i,l} z_{i,l}^* \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)} \right).
\]

(2.1.6)

Then

\[
\frac{\partial^2 \log p_i(\beta^*)}{\partial \gamma_i^* \partial \gamma_i^*} - \frac{\partial^2 \log p_i(\beta)}{\partial \gamma_i \partial \gamma_i} \leq \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) |z_{i,l} z_{i,l}^*|
\]

\[
x \left\{ \frac{\exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}} \right. \\
- \left. \frac{\exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}} \right\} \\
+ \sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) |z_{i,l} z_{i,l}^*| \\
x \left\{ \frac{\exp 2(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}}{\sum_{l=1}^{n} I(T_{i(l)} \geq t_{i(l)}) \exp(\gamma_i^* + \Delta_i^* w_{i(l)}) z_{i(l)}} \right\}
\]
\[ - \frac{\exp \left( \gamma_i' + \Delta_i' \epsilon_{i(i)} \right) \cdot z_{i(i)}}{\sum_{l=1}^{n} I(T_{i(i)} \geq t_{i(i)}) \exp(\gamma_i' + \Delta_i' \epsilon_{i(i)}) z_{i(i)}} \right) \} ,
\]

and

\[ \sup_{|\gamma^* - \gamma| < \epsilon} \left\{ \frac{\partial^2 \log p_i(\beta^*)}{\partial \gamma_i^* \partial \gamma_i^*} - \frac{\partial^2 \log p_i(\beta)}{\partial \gamma_i \partial \gamma_i'} \right\} \]
\[ \leq \sup_{|\gamma^* - \gamma| < \epsilon} \left\{ \frac{\sum_{l=1}^{n} I(T_{i(i)} \geq t_{i(i)}) \cdot z_{i(i)} \cdot z_{i(i)}'}{x} \right\} \times \]
\[ \left\{ \frac{\exp(\gamma_i' \cdot z_{i(i)})}{\exp(\Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}} \right\} \]
\[ \left\{ \frac{\sum_{l=1}^{n} \exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}' \cdot z_{i(i)}'}{\exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}'} \right\} \]
\[ \exp(\gamma_i' \cdot z_{i(i)}) \left( \frac{\sum_{l=1}^{n} \exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}' \cdot z_{i(i)}'}{\exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}'} \right) \]
\[ \left\{ \frac{\sum_{l=1}^{n} \exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}' \cdot z_{i(i)}'}{\exp(\epsilon_i' + \Delta_i' \epsilon_{i(i)}) \cdot z_{i(i)}'} \right\} \} \} .
\]

Then

\[ \lim_{\epsilon \to \infty} \left\{ \sup_{|\gamma^* - \gamma| < \epsilon} \left\{ \frac{\partial^2 \log p_i(\beta^*)}{\partial \gamma_i^* \partial \gamma_i^*} - \frac{\partial^2 \log p_i(\beta)}{\partial \gamma_i \partial \gamma_i'} \right\} \right\} = 0
\]

and

\[ E\left\{ \lim_{\epsilon \to \infty} \sup_{|\gamma^* - \gamma| < \epsilon} \left\{ \frac{\partial^2 \log p_i(\beta^*)}{\partial \gamma_i^* \partial \gamma_i^*} - \frac{\partial^2 \log p_i(\beta)}{\partial \gamma_i \partial \gamma_i'} \right\} \right\} = 0 \]
The same can be verified for all $\beta \in \beta$. Note that for

$$
\mathbb{E}\left\{ \lim_{\epsilon \to 0} \sup_{|\beta^* - \beta| < \epsilon} |x^*_n - x_n| \right\} = \lim_{\epsilon \to 0} \mathbb{E}\left\{ \sup_{|\beta^* - \beta| < \epsilon} |x^*_n - x_n| \right\}
$$

the condition that $\mathbb{E}\left\{ \sup_n |x^*_n| \right\} < \infty$ suffices. Therefore, if

$$
\mathbb{E}\left\{ \sup_{\beta} I(T_{i(i)} \geq t_{i(i)}) \mid z_{i(i)} \mid z'_{i(i)} \mid \exp(2\beta' z_{i(i)}) \right\} < \infty \quad (2.1.7)
$$

with $k \in \mathbb{I}^+$, then

$$
\lim_{\epsilon \to \infty} \mathbb{E}\left\{ \sup_{|\gamma^* - \gamma| < \epsilon} \left| \frac{\partial^2 \log p_{i(i)}(\beta^*)}{\partial \gamma_{i}^* \partial \gamma_{i}^*} - \frac{\partial^2 \log p_{i(i)}(\beta)}{\partial \gamma_{i} \partial \gamma_{i}} \right| \right\} = 0 \quad (2.1.8)
$$

2. In order to exchange integral and derivatives signs we need the conditions

$$
\left| \frac{\partial p_{i(i)}}{\partial \beta_R} \right| \leq T(z) \quad \text{and} \quad \left| \frac{\partial^2 p_{i(i)}}{\partial \beta^R \partial \beta_R} \right| \leq H(z)
$$

where $\int T(z) \, dz < \infty$ and $\int h(z) \, dz < \infty$.

Since $\frac{\partial p_{i(i)}}{\partial \beta_R} = \frac{\partial \log p_{i(i)}}{\beta_R} x_{i(i)}$, with $p_{i(i)}$ given in (2.1.3), we have for $\gamma_{iR} \in \beta$

$$
\frac{\partial p_{i(i)}}{\partial \gamma_{iR}} = p_{i(i)} \left\{ z_{i(i)} - \frac{\sum z_{i(r)} \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu}}{\sum \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu}} \right\}
$$

$$
= \frac{\exp(\gamma_{i} + \Delta_i w_{i(i)} z_{i(i)} - \sum z_{i(r)} \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu})}{\sum \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu}} \left\{ z_{i(i)} - \frac{\sum z_{i(r)} \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu}}{\sum \exp(\gamma_{i} + \Delta_i w_{iu}) z_{iu}} \right\}
$$

$$
(2.1.9)
$$
Then, we have (2.1.9)

\[
\left| \frac{\partial p_{i(t)}}{\partial \gamma_{ir}} \right| \leq \frac{\exp(\gamma_i' + \Delta_i' w_{i(t)}) z_{i(t)}}{\sum_u \exp(\gamma_{i'} + \Delta_i' w_{i(t)}) z_{i(t)}} \times \left\{ \left| z_{i(t)} \right| - \sum_{l=1}^{\frac{\|z_{i(t)}\|}{\sum_u \exp(\gamma_{i'} + \Delta_i' w_{i(t)}) z_{i(t)}}} \right\} = T(z)
\]

In order that $T(z)$ is integrable, we need to have

\[
E \left\{ I(t_{i(t)} \geq t_{i(t)}) \mid z_{i(t)} \right\} \exp(\gamma_i' + \Delta_i' w_{i(t)}) z_{i(t)} \right\} < \infty. \tag{2.1.10}
\]

Now, since

\[
\frac{\partial^2 p_{i(t)}}{\partial \gamma_i \partial \gamma_j} = \frac{\partial p_{i(t)}}{\partial \gamma_i} \times \frac{\partial p_{i(t)}}{\partial \gamma_j} \times \frac{1}{p_{i(t)}} + \frac{\partial^2 \log p_{i(t)}}{\partial \gamma_i \partial \gamma_j} \times p_{i(t)}, \quad \tag{2.1.11}
\]

we have

\[
\left| \frac{\partial^2 p_{i(t)}}{\partial \gamma_i \partial \gamma_j} \right| \leq \left| \frac{\partial p_{i(t)}}{\partial \gamma_i} \right| \times \left| \frac{\partial p_{i(t)}}{\partial \gamma_j} \right| \times \frac{1}{p_{i(t)}} + \left| \frac{\partial^2 \log p_{i(t)}}{\partial \gamma_i \partial \gamma_j} \right| \times p_{i(t)}.
\]

From (2.1.6) we know that

\[
\left| \frac{\partial^2 \log p_{i(t)}}{\partial \gamma_i \partial \gamma_j} \right| \leq \sum_{l=1}^{\frac{\|z_{i(t)}\|}{\sum_u \exp(\gamma_{i'} + \Delta_i' w_{i(t)}) z_{i(t)}}} I(T_{i(t)} \geq t_{i(t)}) \exp(\gamma_i' + \Delta_i' w_{i(t)}) z_{i(t)} \right| z_{i(t)} z_{i(t)}' \right| \\
+ \sum_{l=1}^{\frac{\|z_{i(t)}\|}{\sum_u \exp(\gamma_{i'} + \Delta_i' w_{i(t)}) z_{i(t)}}} I(T_{i(t)} \geq t_{i(t)}) \exp^2(\gamma_i' + \Delta_i' w_{i(t)}) z_{i(t)} \right| z_{i(t)} z_{i(t)}' \right| \\
= H_1(z) + H_2(z). \tag{2.1.12}
\]
Then, from (2.1.9) and (2.1.12)

\[ \left| \frac{\partial^2 P_{i(l)}}{\partial \gamma_i \partial \gamma'_j} \right| \leq T(z) + H_1(z) + H_2(z) = H(z) \]

and \( H(z) \) is integrable if

\[ \mathbb{E} \left\{ I(T_{i(l)} \geq t_{i(l)}) \ | \ z_{i(l)} \ z'_{i(l)} \ | \ | \exp (2 \beta'_i z_{i(l)}') \right\} < \infty. \]  \hspace{1cm} (2.1.13)

3. The regularity condition \( (A_3) \) concerns the variance-covariance matrix \( I(\beta) \): it is required that

\[ \mathbb{E} \left| \left( \frac{\partial \log P_{i(l)}}{\partial \beta_r} \right) \left( \frac{\partial \log P_{i(l)}}{\partial \beta_s} \right) \right| < \infty \]

is satisfied. From (2.1.11) and

\[ \frac{\partial P_{i(l)}}{\partial \beta_r} = \frac{\partial \log P_{i(l)}}{\beta_r} \times P_{i(l)} \]

we have for

\[ \frac{\partial^2 P_{i(l)}}{\partial \gamma_i \partial \gamma'_j} = \frac{\partial \log P_{i(l)}}{\partial \gamma_i} \times \frac{\partial \log P_{i(l)}}{\partial \gamma'_j} \times P_{i(l)} + \frac{\partial^2 \log P_{i(l)}}{\partial \gamma_i \partial \gamma'_j} \times P_{i(l)}. \]

Taking the expected value of both sides, we have that

\[ \mathbb{E} \left( \frac{\partial^2 \log P_{i(l)}}{\partial \gamma_i \partial \gamma'_j} \right) = - \mathbb{E} \left( \frac{\partial \log P_{i(l)}}{\partial \gamma_i} \right) \left( \frac{\partial \log P_{i(l)}}{\partial \gamma'_j} \right) \]

since \( \mathbb{E} \left( \frac{\partial^2 P_{i(l)}}{\partial \gamma_i \partial \gamma'_j} \right) = \mathbb{E} \left( \frac{\partial P_{i(l)}}{\partial \gamma_i} \right) = 0 \). Then, by (2.1.12)
\[ E \left[ \left( \frac{\partial \log P_{i(t)}}{\partial \gamma_i} \right) \left( \frac{\partial \log P_{i(t)}}{\partial \gamma_j} \right) \right] < \infty. \]

From the study of the regularity conditions on the MLE from the partial likelihood function as given in (2.4), we verify that there is a need for additional assumptions which can be stated as

\[ E \left[ \sup_{\beta} I(T_{i(t)} \geq \tau_{i(t)} \mid z_{i(t)} \exp(k \beta' z_{i(t)}) \right] < \infty, \text{ for } k=1, 2. \]

Consider the efficient scores vector

\[ \hat{U}_i = \left[ \frac{1}{\sqrt{n}} \frac{\partial \log L(\gamma, \Delta)}{\partial \Delta} \right]_{\Delta=0, \gamma=\hat{\gamma}} \]

\[ = \sum_{l=1}^{n} \delta_{i(l)} \left\{ w_{i(l)} z_{i(l)} - \frac{\sum w_{i(l)} z_{i(l)} \exp(\hat{\gamma}' z_{i(l)})}{\sum \exp(\hat{\gamma}' z_{i(l)})} \right\} i = 1, 2. \]  

(2.1.14)

and the matrix

\[ I(\hat{\beta}) = \begin{bmatrix} I_{\gamma \gamma} & I_{\gamma \Delta} \\ I_{\Delta \gamma} & I_{\Delta \Delta} \end{bmatrix} \]  

(2.1.15)

where

\[ -\left[ \frac{1}{\sqrt{n}} \frac{\partial^2 \log L(\gamma, \Delta)}{\partial \gamma_n \partial \gamma_{i(m)}} \right]_{\Delta=0, \gamma=\hat{\gamma}} = \]

\[ = \sum_{l=1}^{n} \delta_{i(l)} \left\{ \frac{\sum z_{i(m)} z_{i(m)} \exp(\hat{\gamma}' z_{i(l)})}{\sum \exp(\hat{\gamma}' z_{i(l)})} \right\} \]
\[
\sum_u z_{in(u)} \exp(\gamma_i z_{i(l)}) \sum_u z_{im(u)} \exp(\gamma_i z_{i(l)}) \bigg/ \left[ \sum_u \exp(\gamma_i z_{i(l)}) \right]^2
\]

is the \((n,m)\)-th element of \(I_{\gamma \gamma}\), \(n,m=1, 2, \ldots, p, i=1, 2\).

\[
- \left[ \frac{1}{\sqrt{n}} \frac{\partial^2 \log L(\gamma, \Delta)}{\partial \Delta_{in} \partial \Delta_{im}} \right]_{\Delta=0, \gamma=\tilde{\gamma}}
\]

\[
= \sum_{l=1}^n \delta_{i(l)} \left\{ \frac{\sum_u z_{in(u)} z_{im(u)} w_{i(l)} \exp(\gamma_i z_{i(l)})}{\sum_u \exp(\gamma_i z_{i(l)})} \right\}
\]

\[
- \sum_u z_{in(u)} \exp(\gamma_i z_{i(l)}) \sum_u z_{im(u)} w_{i(l)} \exp(\gamma_i z_{i(l)}) \bigg/ \left[ \sum_u \exp(\gamma_i z_{i(l)}) \right]^2
\]

is the \((n,m)\)-th element of \(I_{\Delta \Delta}\) and finally

\[
- \left[ \frac{1}{\sqrt{n}} \frac{\partial^2 \log L(\gamma, \Delta)}{\partial \gamma_{in} \partial \Delta_{im}} \right]_{\Delta=0, \gamma=\tilde{\gamma}}
\]

\[
= \sum_{l=1}^n \delta_{i(l)} \left\{ \frac{\sum_u z_{in(u)} z_{im(u)} w_{i(l)} \exp(\gamma_i z_{i(l)})}{\sum_u \exp(\gamma_i z_{i(l)})} \right\}
\]

\[
- \sum_u z_{in(u)} \exp(\gamma_i z_{i(l)}) \sum_u z_{im(u)} w_{i(l)} \exp(\gamma_i z_{i(l)}) \bigg/ \left[ \sum_u \exp(\gamma_i z_{i(l)}) \right]^2
\]

is the \((n,m)\)-th element of \(I_{\gamma \Delta}\) or \(I_{\Delta \gamma}\), \(n,m=1, 2, \ldots, p, i=1, 2\). Since \(\Delta_i \cap \gamma_i = \emptyset\) for
i\neq i' and \( \gamma_i \cap \gamma_i' = \emptyset \) for \( i \neq i' \). The submatrices \( I_{\gamma\gamma}, I_{\gamma\Delta}, I_{\Delta\gamma} \) and \( I_{\Delta\Delta} \) are block diagonal matrices with \( I_{\gamma\gamma}, I_{\gamma\Delta}, I_{\Delta\gamma}, \) and \( I_{\Delta\Delta} \), \( i=1, 2 \) on the diagonal. Under regularity conditions and \( H_0: \Delta = 0 \), the following properties of \( \hat{\gamma}_0 \) hold.

1. There exists a vector \( \hat{\gamma} \) of solutions of the likelihood equations (2.1.1) and (2.1.2) which converges in probability to \( \gamma \). Also, we have that for any estimator \( \hat{\beta}_0 \) of \( \beta \) that \( I(\hat{\beta}) \) converges in probability to \( I(\beta) \), that is

\[
P\left\{ \lim_{\beta \to \beta_0} \begin{bmatrix} \partial^2 \log L(\beta) \\ \partial \beta \partial \beta_0 \end{bmatrix}_{\beta = \beta_0} \right\} = I(\beta).
\]

2. \( \sqrt{n} (\hat{\gamma} - \gamma) \) has a limiting multivariate normal distribution with mean 0 and variance-covariance matrix \( I_{\gamma\gamma} \).

Consider now a sequence of local alternative \( K_n: \Delta_n = \Delta = \frac{a}{\sqrt{n}} \) for some \( a \in \mathbb{R}^{2p} \). In order to derive the asymptotic properties of \( \hat{\gamma} \), the p.m.l.e of \( \gamma \) under \( K_n \), we need to verify contiguity of a sequence of local alternative \( \{K_n\} \) to the null hypothesis. The concept of contiguity is due to LeCam (1960). A very detailed discussion of the contiguity is contained in Chapter VI of Hajek and Sidak (1967). Basically, if \( \{P_n\} \) and \( \{Q_n\} \) are two sequences of (absolutely continuous) probability measure on measure space \((X_n, A_n, \mu_n)\) such that for any sequence of events \( A_n \in A_n, \{P_n(A_n) - 0\} - \{Q_n(A_n) - 0\}, \) then the sequence of measures \( \{Q_n\} \) is said to be contiguous to \( \{P_n\} \). Consider the likelihood functions \( L(\gamma, \Delta_0) \) and \( L(\gamma, \Delta_0 + \frac{a}{\sqrt{n}}) \). The log of the likelihood ratio is

\[
\log L = \log L(\beta_a) - \log L(\beta_0)
\]

where

\[
\beta_a = \left( \gamma', \Delta_0' + \frac{a'}{\sqrt{n}} \right) \quad \text{and} \quad \beta_0 = \left( \gamma', \Delta_0' \right).
\]

\[
\log L(\beta_a) = \frac{1}{n} \sum_{i=1}^{n} \log f_i(x_i; \beta_a) 
= \frac{1}{n} \sum_{i=1}^{n} \log f_i(x_i, \beta_0) + \left( 0, \frac{a}{\sqrt{n}} \right)' \sum_{i=1}^{n} \frac{\partial \log f_i(x_i, \beta_0)}{\partial \beta_0} +
\]
\[ + \left( 0 \ a_n \right) \gamma + \sum_{i=1}^{n} \frac{\partial^2 \log f_i(x_i, \beta_0)}{\partial \beta_i \partial \beta_i^t} \left( 0 \ a_n \right) + o_p(1). \]

Let \((0 \ a_n) = a_0\) where 0 is a 2p x 1 vector of zeros, a is a 2p x 1 vector and

\[ U_n^0 = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \log f_i(x_i, \beta_0)}{\partial \beta_i^t} \quad \text{and} \quad V_n^0 = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f_i(x_i, \beta_0)}{\partial \beta_i \partial \beta_i^t}. \]

Then we can write the log of the likelihood as \( \log L_n = a_0^t U_n^0 + \frac{1}{2} a_0^t V_n^0 a_0 + o_p(1). \)

Since \( U_n^0 \sim N(0, I(\beta_0)) \) and \( V_n^0 \sim -I(\beta_0) \) by Slutzky's theorem, we conclude that

\[ \log L \sim N \left( -\frac{1}{2} a_0^t I(\beta_0) a_0, a_0^t I(\beta_0) a_0 \right) \]

Let \( a_0^t I(\beta_0) a_0 = \sigma \), we have

\[ \log L \sim N \left( -\frac{1}{2} \sigma^2, \sigma^2 \right). \]

Using Corollary VI. 1. 2 in Hajek and Sidak, we conclude that the densities \( f(x, \beta_a) \) are contiguous to the densities \( f(x, \beta_0) \). By Le Cam's third Lemma, we conclude that under \( K_n \),

\[ \sqrt{n} (\beta_a - \beta) \sim N (a_0, I(\beta_0)^{-1}) \quad (2.1.19) \]

where \( \beta_a = (\gamma, a) \) is the p.m.l.e. of \( \beta \) under \( K_n \).

To establish the relationship between \( \hat{\gamma} \) and \( \tilde{\gamma} \) consider now the vector of the partial scores under \( H_0 \)

\[ \tilde{U} = \left( \tilde{U}_\gamma \ \tilde{U}_\Delta \right) \]

where

\[ \tilde{U}_\gamma = \left[ \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \gamma} \right]_{\Delta = 0, \gamma = \hat{\gamma}} = 0 \]
\[ \hat{U}_\Delta = \left[ \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \Delta} \right]_{\Delta = 0, \gamma = \hat{\gamma}} \]

Expanding the scores \( \hat{U} \) in a Taylor series about the true parameter value \( \beta = (\gamma, 0) \) we have

\[ \hat{U}(\hat{\beta}) = U(\beta) + \sqrt{n}(\hat{\beta} - \beta) U(\beta^*)' \]

where \( |\beta^* - \beta| < |\hat{\beta} - \beta| \). Now, we can write the expression above as

\[
\begin{pmatrix}
0 & \hat{U}_\Delta
\end{pmatrix} =
\begin{pmatrix}
U_\gamma^0 & U_\Delta^0
\end{pmatrix} - \sqrt{n}(\hat{\gamma} - \gamma) I(\beta) + o_p(1)
\]

since \( I(\beta) = -\left( \frac{\partial^2 \log L(\beta)}{\partial \beta \partial \beta} \right) \). Then

\[ \hat{U}_\Delta = U_\Delta^0 - I_{\Delta \gamma} \sqrt{n}(\hat{\gamma} - \gamma) + o_p(1) \]

and

\[ U_\gamma^0 = \sqrt{n}(\hat{\gamma} - \gamma) I_{\gamma \gamma} + o_p(1) \]

\[ \sqrt{n}(\hat{\gamma} - \gamma) = I_{\gamma \gamma} U_\gamma^0 + o_p(1) \quad (2.1.20) \]

Then

\[ \hat{U}_\Delta = U_\Delta^0 - I_{\Delta \gamma} \frac{I_{\gamma \gamma} U_\gamma^0}{\sqrt{n}} \sim N(0, I_{\Delta \Delta; \gamma}) \]

since \( E(U^*) = 0 \) and \( V(U^0) = I(\beta) \). Consider now the vector of partial scores under \( H_a \)

\[ \hat{U} = (\hat{U}_\gamma, \hat{U}_\Delta) \] where

\[ \hat{U}_\gamma = \left[ \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \gamma} \right]_{\Delta = \frac{\beta}{n}, \gamma = \hat{\gamma}} = 0 \]
\[ \hat{U}_\Delta = \left[ \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \Delta} \right]_{\Delta = \frac{a}{\sqrt{n}}, \gamma = \hat{\gamma}} \]

Expanding the scores \( \hat{U} \) in a Taylor series about the true parameter value \( \beta = (\gamma, \frac{a}{\sqrt{n}}) \), we have

\[ \hat{U}(\hat{\beta}) = U(\beta) + \sqrt{n}(\hat{\beta} - \beta) U(\beta^*)' \]

where \( |\beta^* - \beta| < |\hat{\beta} - \beta| \). Now, we can write the expression above as

\[ \begin{pmatrix} 0 & \hat{U}_\Delta \end{pmatrix} = \begin{pmatrix} U_{\gamma} & U_{\Delta} \end{pmatrix} - \sqrt{n}(\hat{\gamma} - \gamma \ 0) I(\beta) + o_p(1) \]

Since \( \hat{U}_\gamma = 0 \)

\[ U_{\gamma} = \sqrt{n}(\hat{\gamma} - \gamma) I_{\gamma\gamma} + o_p(1) \]

\[ \sqrt{n}(\hat{\gamma} - \gamma) = I_{\gamma\gamma}^{1/2} U_{\gamma} + o_p(1) \quad (2.1.21) \]

\[ \hat{U}_\Delta = U_{\Delta} - I_{\Delta\gamma} \sqrt{n}(\hat{\gamma} - \gamma) + o_p(1) \]

Now expanding the true scores \( U^0 \) in a Taylor series about \( \beta^0 \) we have

\[ U(\beta^0) = U(\beta^0) + \sqrt{n}(\beta^0 - \beta^0) U(\beta^*)' \quad (2.1.22) \]

where \( |\beta^* - \beta^0| < |\beta^0 - \beta^0| \)

\[ U(\beta^0) = U(\beta^0) + \sqrt{n}(0 \ \frac{a}{\sqrt{n}} I(\beta) + o_p(1) \]

\[ \]
Since $U(\beta^0) \sim N(0, \mathbf{I}(\beta))$ it follows that

$$U(\beta^\circ) \sim N(\mathbf{a}^0 \mathbf{I}(\beta), \mathbf{I}(\beta))$$

where $\mathbf{a}^0=(0 \ a)$. Then

$$\hat{U}_\Delta = U^\circ_\Delta - I^0_{\Delta \gamma} \hat{I}^{-1}_{\gamma \gamma} + o_p(1) \sim N(\mathbf{a} I_{\Delta \Delta; \gamma}, I_{\Delta \Delta; \gamma}).$$

Now from (2.1.30) we have

$$U^0_\gamma = U^0_\gamma + \frac{a}{\sqrt{n}} I_{\Delta \gamma} + o_p(1)$$

(2.1.22)

$$U^\circ_\Delta = U^\circ_\Delta + \frac{a}{\sqrt{n}} I_{\Delta \Delta} + o_p(1)$$

(2.1.23)

Using (2.1.22) in (2.1.20) we have

$$\sqrt{n}(\hat{\gamma} - \gamma) = I^1_{\gamma \gamma} U^\circ_\gamma + I^1_{\gamma \gamma} I_{\gamma \Delta} \frac{a}{\sqrt{n}} + o_p(1)$$

Using (2.1.21) in the expression above we have

$$\sqrt{n}(\hat{\gamma} - \gamma) = \sqrt{n}(\hat{\gamma} - \gamma) + I^1_{\gamma \gamma} I_{\gamma \Delta} \frac{a}{\sqrt{n}} + o_p(1)$$

$$\sqrt{n}(\hat{\gamma} - \gamma) - \sqrt{n}(\hat{\gamma} - \gamma) = I^1_{\gamma \gamma} I_{\gamma \Delta} \frac{a}{\sqrt{n}} + o_p(1)$$

(2.1.24)

2.2.2. The Union-Intersection Test

In this section we derive the union-intersection test for the hypothesis in (1.7) $H_0$: $\Delta=0$ against the restricted alternative $H_a$: $\Delta > 0$, that is based on partial maximum likelihood estimators and the application of Roy's union-intersection principle. Define
\[ \Omega_0 = \{ \Delta: \Delta = 0 \} \quad \text{and} \quad \Omega_1 = \{ \Delta: \Delta \geq 0 \} \] such that \( \Omega = \Omega_0 \cup \Omega_1 \). Define also \( a = (a_1, a_2) \), where \( a_i = (a_{i1}, a_{i2}, ..., a_{ip}) \), \( i = 1, 2 \) such that \( a \geq 0 \). Assume that \( \Omega \) is positively homogeneous and for each \( \Delta \in \Omega \) let \( \Delta = \frac{a}{\sqrt{n}} \). For a fixed \( a \in \mathbb{R}^{+2p} \) construct

\[ \Omega(a) = \left\{ \Delta: \Delta \geq 0, \Delta = \frac{a}{\sqrt{n}} \right\} \tag{2.2.1} \]

and consider the alternative hypothesis

\[ H^*_a : \Delta = \frac{a}{\sqrt{n}} \tag{2.2.2} \]

If we define the set \( A \) such that

\[ A = \left\{ a: a \geq 0 \right\} \tag{2.2.3} \]

we have

\[ \Omega \subseteq \bigcup_{a \in A} \Omega(a) \]

since \( \Omega \) is positively homogeneous. Thus according to the union-intersection principle

\[ l_1^* = \sup_{a \in A} 2 \log l_1 \tag{2.2.4} \]

where

\[ l_1 = \frac{L(\hat{\gamma}, \frac{a}{\sqrt{n}})}{L(\hat{\gamma}, 0)} \tag{2.2.5} \]

where \( \hat{\gamma} \) is the p.m.l.e of \( \gamma \) given \( \Delta = \frac{a}{\sqrt{n}} \), \( \tilde{\gamma} \) is the p.m.l.e of \( \gamma \) given \( \Delta = 0 \) and \( L(\cdot, \cdot) \) is the likelihood function as defined in (2.3). Then

\[ \log l_1 = \sum_{l=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ (\gamma_i + \frac{a_i}{\sqrt{n}}w_{i(l)})z_{i(l)} - \tilde{\gamma}_i z_{i(l)} \right\} \]

\[ - \frac{\sum_{u} \exp \left( \gamma'_i + \frac{a_i}{\sqrt{n}}w_{i(l)}z_{i(l)} \right)}{\sum_{u} \exp \left( \gamma'_i + \frac{a_i}{\sqrt{n}}w_{i(l)}z_{i(l)} \right)} \tag{2.2.6} \]
Since from (2.1.25)

\[ \sqrt{n}(\hat{\gamma}_i - \gamma_i) = -(I_{i\gamma}^{-1}) I_{i\gamma} \Delta a_i \]

we have

\[
\log l_i = \sum_{l=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ \left( -\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + \hat{\gamma}_i' + a_i' \frac{w_{i(l)}}{\sqrt{n}} \right) z_{i(l)} - \hat{\gamma}_i^{0'} z_{i(l)} \right\}
\]

\[
- \log \sum_u \exp \left[ \frac{-\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + \hat{\gamma}_i^{0'} + a_i' \frac{w_{i(l)}}{\sqrt{n}}}{\sum \exp (\hat{\gamma}_i^{0'} z_{i(l)})} z_{i(l)} \right]
\]

\[= \sum_{l=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ \left( -\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + a_i' \frac{w_{i(l)}}{\sqrt{n}} \right) z_{i(l)} \right\}
\]

\[- \log \frac{\sum_u \exp \hat{\gamma}_i^{0'} z_{i(l)} \exp \left[ \frac{-\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + a_i' \frac{w_{i(l)}}{\sqrt{n}}}{\sum \exp (\hat{\gamma}_i^{0'} z_{i(l)})} z_{i(l)} \right]}{\sum \exp (\hat{\gamma}_i^{0'} z_{i(l)})}
\]

Expanding \( \exp \left[ \frac{-\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + a_i' \frac{w_{i(l)}}{\sqrt{n}}}{\sum \exp (\hat{\gamma}_i^{0'} z_{i(l)})} z_{i(l)} \right] \) in a Taylor series up to the second order around zero, we have

\[
\log l_i = \sum_{l=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ \left( -\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + a_i' \frac{w_{i(l)}}{\sqrt{n}} \right) z_{i(l)} \right\}
\]

\[- \log \frac{\sum_u \exp \hat{\gamma}_i^{0'} z_{i(l)} \sum_u \exp (\hat{\gamma}_i^{0'} z_{i(l)}) \left\{ 1 + \left( -\frac{a_i'}{\sqrt{n}} I_{i\gamma} \Delta (I_{i\gamma}^{-1}) + a_i' \frac{w_{i(l)}}{\sqrt{n}} \right) z_{i(l)} - \hat{\gamma}_i^{0'} z_{i(l)} \right\}}{\sum \exp (\hat{\gamma}_i^{0'} z_{i(l)})}
\]
\begin{align*}
&+ \frac{1}{2} \left[ \left( - \frac{a'_i}{\sqrt{n}} \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} + a'_i \frac{w_i(l)}{\sqrt{n}} \right) z_i(l) \right]^2 \right) \right] \\
\log l_1 &= \sum_{i=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ \left( - \frac{a'_i}{\sqrt{n}} \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} + a'_i \frac{w_i(l)}{\sqrt{n}} \right) z_i(l) \right\} \\
&- \log \left( 1 + \frac{1}{\sum_{u} \exp \gamma_i^{0} z_i(l)} \left( \sum_{u} \exp \gamma_i^{0} z_i(l) \left( - \frac{a'_i}{\sqrt{n}} \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} + a'_i \frac{w_i(l)}{\sqrt{n}} \right) z_i(l) \right)^2 \right) \right) \right) \\
\log l_1 &= \sum_{i=1}^{n} \sum_{i=1}^{2} \delta_{i(l)} \left\{ \left( - \frac{a'_i}{\sqrt{n}} \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} + a'_i \frac{w_i(l)}{\sqrt{n}} \right) z_i(l) \right\} \\
&- \frac{1}{\sqrt{n}} \frac{a'_i \sum_{u} \exp \gamma_i^{0} z_i(l) w_i(l) z_i(l)}{\sum_{u} \exp \gamma_i^{0} z_i(l)} \\
&+ \frac{1}{\sqrt{n}} \frac{a'_i \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} \sum_{u} z_i(l) \exp \gamma_i^{0} z_i(l)}{\sum_{u} \exp \gamma_i^{0} z_i(l)} \\
&- \frac{1}{2} \frac{a'_i \sum_{u} \exp (\gamma_i^{0} z_i(l)) w_i^2(l) z_i(l) a_i}{\sum_{u} \exp \gamma_i^{0} z_i(l)} \\
&+ \frac{1}{\sqrt{n}} \frac{a'_i \mathbf{I}_n \Delta \gamma (\mathbf{I}_n \gamma)^{-1} \left( \sum_{u} \exp (\gamma_i^{0} z_i(l)) z_i(l) z_i(l)^{-1} \right) a_i}{\sum_{u} \exp \gamma_i^{0} z_i(l)}
\end{align*}
\[ + \frac{1}{2n} \left\{ \sum_{i=1}^{n} \exp(\gamma_{i}^{0}z_{i(l)}) \left( -a_{i}^{\prime} I_{\Delta \gamma} (I_{\gamma}^{l})^{-1} + a_{i}^{\prime} w_{i(l)} z_{i(l)} \right) \right\}^{2} \frac{2}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} + o_p(1). \]

\[
\log \lambda_{1} = \frac{2}{2n} \sum_{i=1}^{2} a_{i} \left\{ \frac{1}{n} \sum_{l=1}^{n} \delta_{i(l)} \left( \sum_{i}^{n} z_{i(l)} \exp(\gamma_{i}^{0}z_{i(l)}) \right) \right\} - \frac{2}{n} \sum_{i=1}^{2} a_{i} \sum_{l=1}^{n} \delta_{i(l)} \left\{ \sum_{i}^{n} \exp(\gamma_{i}^{0}z_{i(l)}) \frac{z_{i(l)}^{2} z_{i(l)}^{l}}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} \right\} \]

\[
- \frac{1}{2n} \sum_{i=1}^{2} a_{i} \sum_{l=1}^{n} \delta_{i(l)} \left\{ \sum_{i}^{n} \exp(\gamma_{i}^{0}z_{i(l)}) \frac{z_{i(l)}^{l}}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} \right\} \left\{ \frac{\sum \exp(\gamma_{i}^{0}z_{i(l)}) \sum z_{i(l)}^{l} w_{i(l)} \exp(\gamma_{i}^{0}z_{i(l)})}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} \right\} \}
\]

\[
+ \frac{1}{n} \sum_{i=1}^{2} a_{i}^{\prime} I_{\Delta \gamma} (I_{\gamma}^{l})^{-1} \sum_{l=1}^{n} \delta_{i(l)} \left\{ \sum_{i}^{n} \exp(\gamma_{i}^{0}z_{i(l)}) \frac{z_{i(l)}^{l} z_{i(l)}^{l}}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} \right\} \]

\[
+ \frac{1}{n} \sum_{i=1}^{2} a_{i}^{\prime} I_{\Delta \gamma} (I_{\gamma}^{l})^{-1} \sum_{l=1}^{n} \delta_{i(l)} \left\{ \sum_{i}^{n} \exp(\gamma_{i}^{0}z_{i(l)}) \frac{z_{i(l)}^{l} z_{i(l)}^{l}}{\sum \exp(\gamma_{i}^{0}z_{i(l)})} \right\} \]
\[
\sum \frac{z_{i(t)} \exp(\gamma_{0}^{0} z_{i(i(t)}) \sum \frac{z_{i(t)} \exp(\gamma_{0}^{0} z_{i(t)})}{\sum \exp(\gamma_{0}^{0} z_{i(t)})} (I_{\Delta\Delta})^{-1} I_{\Delta\gamma} a_{i}}{\sum \exp(\gamma_{0}^{0} z_{i(t)})^2} \right) (I_{\gamma\gamma})^{-1} I_{\Delta\gamma} a_{i}
\]

+ \text{op}(1).

\[
\log l_1 = \sum a_i \hat{U}_\Delta - \sum \frac{a'_i I_{\Delta\gamma} (I_{\gamma\gamma})^{-1} \hat{U}_\gamma}{2} - \frac{1}{2} \sum a'_i I_{\Delta\Delta} a_i \\
+ \sum a'_i I_{\Delta\gamma} (I_{\gamma\gamma})^{-1} I_{\Delta\gamma} a_i - \frac{1}{2} \sum a'_i I_{\Delta\gamma} (I_{\gamma\gamma})^{-1} (I_{\gamma\gamma})^{-1} I_{\Delta\gamma} a_i \\
+ \text{op}(1).
\]

\[
= a' \hat{U}_\Delta - a' I_{\Delta\gamma} I_{\gamma\gamma} \hat{U}_\gamma - \frac{1}{2} a' I_{\Delta\Delta} a + a' I_{\Delta\gamma} I_{\gamma\gamma} I_{\Delta\gamma} a \\
- \frac{1}{2} a' I_{\Delta\gamma} I_{\gamma\gamma} I_{\gamma\Delta} a + \text{op}(1)
\]

\[
= a' \hat{U}_\Delta - a' I_{\Delta\gamma} I_{\gamma\gamma} \hat{U}_\gamma - \frac{1}{2} a' \left( I_{\Delta\Delta} - I_{\Delta\gamma} I_{\gamma\gamma} I_{\gamma\Delta} \right) a + \text{op}(1)
\]

\[
\log l_1 = a' \hat{U}_\Delta - \frac{1}{2} I_{\Delta\Delta} \gamma + \text{op}(1) \tag{2.2.6}
\]

since \( \hat{U}_\gamma = 0 \) and \( I_{\Delta\Delta} \gamma = I_{\Delta\Delta} - I_{\Delta\gamma} I_{\gamma\gamma} I_{\gamma\Delta} \). Then the union-intersection test in (2.2.4) requires the maximization of \( 2 \log l_1 \) over the set \( A \).

The problem of determining the maximum of \( 2 \log l_1 = a' \hat{U}_\Delta - I_{\Delta\Delta} \gamma + \text{op}(1) \) which is a convex quadratic form in \( a \), over the set \( A \) is equivalent to finding the

\[- \inf(-2a'\hat{U}_\Delta + I_{\Delta\Delta} \gamma) \text{ over } A. \]

In non-linear programming terminology, the problem is
to determine the vector \( a \in A \) such that
\[ h(a^*) = - \inf \left\{ h(a), \ a \in A \right\} \]  \hspace{1cm} (2.2.8)

where \( h(a) \) is a scalar-valued function of \( a \) and

\[ A = \left\{ a: h_1(a) \leq 0, \ h_2(a) = 0 \right\} \]

where \( h_1(a) \) and \( h_2(a) \) are functions of \( a \). The Kuhn-Tucker-Lagrange (KTL) minimization technique can be applied here in order to minimize \( h(a) \), considering \( h_1(a) = -a \leq 0 \) and the equality constraint missing. The solutions for the Lagrange function

\[ L(a, t_1) = h(a) + t_1^i h_1(a) = -2 a' U_\Delta + a' I_{\Delta \Delta} : \gamma a - t_1 a \]  \hspace{1cm} (2.2.9)

where \( t_1 \) is a 2p vector values, is any point \( (a^*, t_1^*) \) which satisfies the system of conditions:

i) \( t_1 \geq 0 \)

ii) \( a \geq 0 \)

iii) \( a' t_1 = 0 \)

iv) \[ \frac{\partial L(a, t_1)}{\partial a} = -2 \tilde{U}_\Delta + 2 I_{\Delta \Delta} : \gamma a - t_1 = 0 \]  \hspace{1cm} (2.2.10)

From iv), we have

\[ t_1 = -2 \tilde{U}_\Delta + 2 I_{\Delta \Delta} : \gamma a \]

and

\[ a' t_1 = -2 a' \tilde{U}_\Delta + 2 a' I_{\Delta \Delta} : \gamma a = 0 \] (from iii)

Then \( a' \tilde{U}_\Delta = a' I_{\Delta \Delta} : \gamma a \). Therefore

\[ h(a) = -2 a' U_\Delta + a' I_{\Delta \Delta} : \gamma a = - a' I_{\Delta \Delta} : \gamma a \]
\[ I^* = -\inf h(a) = a' I_{\Delta:\gamma} a \]

From (iv) we have
\[ t_1 + 2 \bar{U}_\Delta = 2 I_{\Delta:\gamma} a \]

\[ a = \frac{1}{2} I_{\Delta:\gamma}^{-1} t_1 + I_{\Delta:\gamma}^{-1} \bar{U}_\Delta \]  \hspace{1cm} (2.2.11)

From (i) and (ii) we see that both \( t_1 \) and \( a \) as given above are non-negative, yet their inner product is zero by (iii) in (2.2.10). This is possible if for some \( \emptyset \subseteq J \subseteq A \), \( t_{(j)} = 0 \) and \( a_{(j)} > 0 \), that is, we partition \( a \) and \( t_1 \) as \( a = (a_{(j)}, a_{(j)'}) \) and \( t = (t_{(j)}, t_{(j)}) \). Also for each \( a \) we partition \( \bar{U}_\Delta \) and \( I_{\Delta:\gamma} \) as

\[ \bar{U}'_\Delta = (\bar{U}'_{\Delta(j)}, \bar{U}'_{\Delta(j)}) \]

and

\[ I_{\Delta:\gamma} = \begin{bmatrix} I_{\Delta:\gamma}^{jj} & I_{\Delta:\gamma}^{ij} \\ I_{\Delta:\gamma}^{ij} & I_{\Delta:\gamma}^{jj} \end{bmatrix} \]

We denote by
\[ \bar{W}_\Delta = I_{\Delta:\gamma}^{-1} \bar{U}_\Delta, \]  \hspace{1cm} (2.2.12)

where

\[ I_{\Delta:\gamma}^{ij} = \Gamma = \begin{bmatrix} \Gamma_{jj}^{ij} & \Gamma_{jj}^{ij} I_{\Delta:\gamma}^{jj} I_{\Delta:\gamma}^{-1} I_{\Delta:\gamma}^{ij} \\ \Gamma_{jj}^{jj} I_{\Delta:\gamma}^{-1} I_{\Delta:\gamma}^{ij} & \Gamma_{jj}^{jj} \end{bmatrix} \]  \hspace{1cm} (2.2.13)

and

\[ \Gamma_{jj}^{jj} = \begin{bmatrix} I_{\Delta:\gamma}^{jj} - I_{\Delta:\gamma}^{jj} (I_{\Delta:\gamma}^{jj})^{-1} I_{\Delta:\gamma}^{jj} \end{bmatrix}^{-1} \]

\[ \Gamma_{jj}^{ij} = \begin{bmatrix} I_{\Delta:\gamma}^{ij} - I_{\Delta:\gamma}^{ij} (I_{\Delta:\gamma}^{ij})^{-1} I_{\Delta:\gamma}^{ij} \end{bmatrix}^{-1} \]
Then from (2.2.11) and using (2.2.12) and (2.2.13) we have

\[
a = \begin{bmatrix} a_j \\ a_j^* \end{bmatrix} = \begin{bmatrix} \tilde{W}'_{\Delta(j)} \\ \tilde{W}'_{\Delta(j)} \end{bmatrix} + \frac{1}{2} \Gamma_{\Delta\Delta;\gamma}^1 \begin{bmatrix} t_j \\ t_j^* \end{bmatrix}
\]

If we let \( t_j = 0, a_j > 0 \) and \( a_j^* = 0 \), we have

\[
a_j = \tilde{W}'_{\Delta(j)} + \frac{1}{2} \Gamma_{jj;\gamma}^{-1} t_j = 0
\]

and

\[
a_j = \tilde{W}'_{\Delta(j)} + \frac{1}{2} \Gamma_{jj;\gamma}^{-1} \Gamma_{jj} (\Gamma_{\Delta\Delta;\gamma}^{-1})^{-1} t_j > 0
\]

From (2.2.14) we have

\[
t_j = -2 \Gamma_{jj} \tilde{W}'_{\Delta(j)} \geq 0
\]

and from (2.2.15) we have

\[
\tilde{W}'_{\Delta(j)} + \frac{1}{2} \Gamma_{jj;\gamma}^{-1} \Gamma_{jj} (\Gamma_{\Delta\Delta;\gamma}^{-1})^{-1} t_j
\]

\[
= \tilde{W}_{\Delta(j)} - \Gamma_{jj} (\Gamma_{jj}^{-1})^{-1} \tilde{W}_{\Delta(j)} = \tilde{W}^*_{\Delta(j)} > 0
\]

Kudo (1963) has shown that the collection of all \( 2^r \) sets, \( r = 2(p+1) \)

\[
R = \{ \tilde{W}_{\Delta(j)}: \tilde{W}^*_{\Delta(j)} \geq 0, \Gamma_{jj} \tilde{W}'_{\Delta(j)} \leq 0 \}
\]

\( \emptyset \subseteq J \subseteq A = \{1, 2, \ldots, r = 2p\} \) is a disjoint and exhaustive partitioning of \( r \)-th dimensional Euclidean space, and for a fixed \( a \), \( (a^*, t^*_j) \) is a KTL point, where
\[ t_i^* = \begin{bmatrix} 0 \\ -2 \Gamma_{jj} \tilde{W}_{\Delta(j)} \end{bmatrix} \]  

(2.2.18)

and

\[ a^* = \begin{bmatrix} \tilde{W}_{\Delta(j)} \\ 0 \end{bmatrix} \]  

(2.2.19)

Then

\[ h(a^*) = -a^* \cdot I_{\Delta:q} a^* \]

\[ = -\left[ \begin{bmatrix} \tilde{W}_{\Delta(j)}^* \\ 0 \end{bmatrix} \right] I_{\Delta:q} \begin{bmatrix} \tilde{W}_{\Delta(j)}^* \\ 0 \end{bmatrix} = -\tilde{W}_{\Delta(j)}^* \cdot I_{\Delta:q} \tilde{W}_{\Delta(j)}^* \]

so that the union-intersection test statistic for the hypothesis (2.3.1) can be written as

\[ I_i^* = \sum_{\emptyset \subseteq J \subseteq P} \left\{ \tilde{W}_{\Delta(j)}^* \cdot I_{jj:q} \tilde{W}_{\Delta(j)}^* \right\} I\{ \tilde{W}_{\Delta(j)}^* > 0 \} \cdot I\{ \Gamma_{jj} \tilde{W}_{\Delta(j)} < 0 \} \]

(2.2.20)

In terms of \( \tilde{U}_\Delta \), the statistic test (2.3.20) can be written as

\[ I_i^* = \sum_{\emptyset \subseteq J \subseteq P} \left\{ \tilde{U}_{\Delta(j)}' \cdot (I_{jj:q}^{-1} \tilde{U}_{\Delta(j)}) \right\} I\{ (I_{jj:q}^{-1} \tilde{U}_{\Delta(j)}) > 0 \} \cdot I\{ \tilde{U}_{\Delta(j)}^* \leq 0 \} \]

or

\[ I_i^* = \sum_{\emptyset \subseteq J \subseteq P} \left\{ \tilde{U}_{\Delta(j)}' \cdot (I_{jj:q}^{-1} \tilde{U}_{\Delta(j)}) \right\} I\{ (I_{jj:q}^{-1} \tilde{U}_{\Delta(j)}) > 0 \}

\times I\{ \tilde{U}_{\Delta(j)}^* \leq 0 \}, \]

as from (2.2.17)

\[ \tilde{W}_{\Delta(j)}^* = \tilde{W}_{\Delta(j)} - r_{ij}^{-1} \tilde{W}_{\Delta(j)} - r_{jj}^{-1} \tilde{U}_{\Delta(j)}, \]

(2.2.21)
where

\[ \Gamma^{\prime \prime} = (\Gamma - \Gamma^{\prime \prime} (\Gamma^{\prime \prime})^{-1} \Gamma) \]

and from (2.2.11)

\[ \tilde{\mathbf{W}}_{\Delta(j)} = (\Gamma^{\prime \prime})^{-1} \tilde{\mathbf{U}}_{\Delta(j)} - \Gamma^{\prime \prime} \tilde{\mathbf{U}}_{\Delta(j)} \Delta : \gamma \right(\Gamma^{\prime \prime})^{-1} \tilde{\mathbf{U}}_{\Delta(j)} \]

2.3. Asymptotic Distribution of the Union-Intersection Test Statistic \( l_1^* \)

In this section, we study the asymptotic distribution of the UI test \( l_1^* \) under \( H_0^* : \Delta = 0 \) and under \( H_a^* : a = 0 \).

First we want to show the following.

1. Under \( \{ H_0 : \Delta = 0 \} \) there exist \( c > 0 \) such that

\[ P\{ l_1^* \leq c \mid H_0 \} = \sum_{i=0}^{2p} w_i \ P\{ \chi_i^2 \leq c \} \]

where the \( w_i \) are nonnegative quantities with \( \sum_{i=0}^{2p} w_i = 1 \) and \( \chi_i^2 \) is a Chi-squared random variable with \( k \) degree of freedom and \( \chi_0^2 = 1 \) with probability 1. The distribution above is referred as chi-bar distribution. The weights \( w_i \) can be computed as

\[ w_i = \sum_{J \subseteq P} P\{ \mathbf{W}_{\Delta(j)}^* > 0 \mid H_0 \} \ P\{ \Gamma_{\Delta(j)} \tilde{\mathbf{W}}_{\Delta(j)} \leq 0 \mid H_0 \} \]

where \( \sum_{J \subseteq P} \) denotes the sum over all sets \( J, \emptyset \subseteq J \subseteq P \), such that \( J \) has \( s \) elements, \( \tilde{\mathbf{W}}_{\Delta(j)}^* \) is \( N_{sj}(0, (\Gamma^{\prime \prime})^{-1}) \) under \( H_0 \) and \( \tilde{\mathbf{W}}_{\Delta(j)} \) is \( N_{sj} (0, \Gamma_{\Delta(j)}^{-1}) \), under \( H_0 \).

**Proof:** From (2.2.20) we have for \( c > 0 \),
\[ P\{\xi_i^* \leq c\} = \sum_{\emptyset \subseteq J \subseteq P} P\left\{ \left( \hat{W}_{\Delta(j)}^*, I_{\Delta^\gamma}^{jj} \hat{W}_{\Delta(j)}^* \right) \leq c, \hat{W}_{\Delta(j)}^* > 0, \Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0 \mid E \hat{W}_\Delta = 0 \right\} \]  

(2.3.3) 

as the distribution of \( \xi_i^* \). First, we need to show that the events \( I(\hat{W}_{\Delta(j)}^* > 0) \) and \( I(\Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0) \) are independent under \( H_0 \). Second, we need independence of \( I(\hat{W}_{\Delta(j)}^*, I_{\Delta^\gamma}^{jj} \hat{W}_{\Delta(j)}^*) \) and \( I(\hat{W}_{\Delta(j)}^* > 0) \).

The first independence can be shown as follows. The \( r \)-variate normal density function for \( \hat{W}_\Delta \) with mean \( \mu \) and dispersion matrix \( \Gamma_{\Delta^\gamma}^1 \), say \( g(\hat{W}_\Delta, \mu, \Gamma_{\Delta^\gamma}^1) \) can be factored into

\[ g(\hat{W}_\Delta, \mu, \Gamma_{\Delta^\gamma}^1) = g(\Gamma_{jj}^j \hat{W}_{\Delta(j)}; \mu_j; \Gamma_{jj}^j) g(\hat{W}_{\Delta(j)}^*, \mu_j^*, (1_{\Delta^\gamma})^{-1}) \]

where \( \emptyset \subseteq J \subseteq P = \{1, 2, \ldots, r\} \). Then

\[ P\{I(\hat{W}_{\Delta(j)}^* > 0) \cap I(\Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0)\} = P(\hat{W}_{\Delta(j)}^* > 0) \cdot P(\Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0). \]

Then, \( \{\hat{W}_{\Delta(j)}^* > 0\} \) and \( \{\Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0\} \) are independent, regardless whether or not \( E(\hat{W}_\Delta) = 0 \).

The second part can be shown by using Lemma 3.2 of Kudo (1963). Kudo has shown that under \( H_0 \), if \( X \) is an \( a \)-variate, null mean normal random vector with covariance \( \Delta \), that \( P(\Delta^{-1}X \leq x, Bx \geq 0) = P(X'\Delta^{-1}X \leq x) P(Bx \geq 0) \), where \( B \) is a (\( b \times a \)) matrix of full rank \( b \) (\( \leq a \)). If we take \( B \) to be the identity matrix, then the two events \( I(\hat{W}_{\Delta(j)}^*, I_{\Delta^\gamma}^{jj} \hat{W}_{\Delta(j)}^* \leq c) \) and \( I(\hat{W}_{\Delta(j)}^* > 0) \) are independent under \( H_0 \).

Thus, from the independence above, we conclude that the events \( I(\hat{W}_{\Delta(j)}^*, I_{\Delta^\gamma}^{jj} \hat{W}_{\Delta(j)}^* \leq c) \), \( I(\hat{W}_{\Delta(j)}^* > 0) \) and \( I(\Gamma_{jj}^j \hat{W}_{\Delta(j)} < 0) \) are mutually
independent, under $H_0$.

From (2.2.11) we have that under $H_0$, $\tilde{W}_\Delta$ is $r$-variate normally distributed with null mean and dispersion matrix $I^{ij}_{\Delta \Delta : \gamma}$. From (2.2.21), we can see that $\tilde{W}^*_\Delta(j)$ is normally distributed with null mean and dispersion matrix $(I^{ij}_{\Delta \Delta : \gamma})^{-1}$. It follows that $\tilde{W}^*_\Delta(j) I^{ij}_{\Delta \Delta : \gamma} \tilde{W}^*_\Delta(j)$ is $\chi^2_{s_j}$ where $s_j$ is the degree of freedom and $s$ is the dimension of set $J, J \subseteq P$.

Then (2.3.3) can be written as

$$P\{l_1^* \leq c \} = \sum_{\emptyset \subseteq J \subseteq P} p \{ \chi^2_{s_j} \leq x \} P\{ \tilde{W}^*_\Delta(j) > 0 \} P\{ \Gamma_{jj} \tilde{W}^*_\Delta(j) < 0 \}$$

Regrouping the $2^r$ terms ($r=2p$) in terms of the cardinality $k$ ($0 \leq k \leq 2p$), we obtain

$$P\{l_1^* \leq c \} = \sum_{s=0}^{2p} w_s P\{ \chi^2_{s_j} \leq x \}$$

The second part of this section studies the distribution of $l_1^*$ under the set of local alternatives $H_a: \Delta = \frac{A}{\sqrt{n}}$. We want to show that

II. Under $\{H_a: \Delta = \frac{A}{\sqrt{n}}\}$ there exist $c > 0$ such that

$$\lim P\{l_1^* \leq c \} = \sum_{\emptyset \subseteq J \subseteq P} \int_{C_{1(j)}} d N_{s_j}(t_1; a_j^*; (I^{ij}_{\Delta \Delta : \gamma})^{-1})$$

$$\times \int_{C_{2(j)}} d N_{s_j}(t_2; a_j; (\Gamma_{jj})^{-1})$$

(2.3.4)

where

$N_k(X, \mu, \Sigma)$ is the $k$-variate normal distribution function with mean $\mu$ and dispersion matrix $\Sigma$.

$$C_{1(j)} = \left\{ t_1 : (t_1' I^{ij}_{\Delta \Delta : \gamma} t_1 \leq r) \cap (t_1 > 0) \right\}$$

(2.3.5)

and
\[ C_{2(j)} = \{ t_2 : \Gamma_{jj} t_2 \leq 0 \} \]  

(2.3.6)

**Proof:** From (2.2.20), we have for \( c > 0 \),

\[
P\{ l^*_1 \leq c | H_a \} = \sum_{\emptyset \subseteq J \subseteq \mathcal{P}} \sum_{\emptyset \subseteq I \subseteq \mathcal{P}} \prod \left\{ \hat{W}_{\Delta(j)}^{*, \gamma} I^{ij}_{\Delta; \gamma} \hat{W}_{\Delta(j)}^{*} \right\} I\{ \hat{W}_{\Delta(j)} > 0 \} \times I\{ \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \} \leq c | H_a \}
\]

(2.3.7)

as the distribution of \( l^*_1 \) under \( H_a; \Delta = \frac{a_{j_*}}{n} \). Following the same arguments used to show the distribution under \( H_0 \), we can see that the events \( I\{ \hat{W}_{\Delta(j)}^{*} > 0 \} \) and \( I\{ \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \} \) are independent regardless of whether or not \( E\{ \hat{W}_{\Delta(j)}^{*} \} \) is non null. In fact, we can see from (2.2.21) that \( \hat{W}_{\Delta(j)}^{*} \sim N_{\gamma}(a^*_j; (I^{ij}_{\Delta; \gamma})^{-1}) \), since \( \hat{W}_{\Delta} \sim N_{\gamma}(a, I^{ij}_{\Delta; \gamma}) \) under \( H_a \). Thus, (2.3.7) becomes

\[
P\{ l^*_1 \leq c | H_a \} = \sum_{\emptyset \subseteq J \subseteq \mathcal{P}} \sum_{\emptyset \subseteq I \subseteq \mathcal{P}} \prod \left\{ \hat{W}_{\Delta(j)}^{*, \gamma} I^{ij}_{\Delta; \gamma} \hat{W}_{\Delta(j)}^{*} \leq c, \hat{W}_{\Delta(j)}^{*} > 0 \right\} I\{ \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \}
\]

(2.3.8)

Therefore, the probabilities in (2.3.8) are expressed as follows.

\[
P\{ \hat{W}_{\Delta(j)}^{*, \gamma} I^{ij}_{\Delta; \gamma} \hat{W}_{\Delta(j)}^{*} \leq c, \hat{W}_{\Delta(j)}^{*} > 0 \} = \int_{C_{1(j)}} d N_{\gamma}(t_1; a^*_j; (I^{ij}_{\Delta; \gamma})^{-1})
\]

\[
= \int_{C_{1(j)}} (2\pi)^{-\frac{1}{2}} |I^{ij}_{\Delta; \gamma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\hat{W}_{\Delta(j)}^{*} - a^*_j)^{\gamma} (\hat{W}_{\Delta(j)}^{*} - a^*_j)^{\gamma} \right\} d \hat{W}_{\Delta(j)}
\]

where \( C_{1(j)} \) is given in (2.3.5).

\[
P\{ \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \} = \int_{C_{2(j)}} d N_{\gamma}(t_2; a_j; (\Gamma_{jj})^{-1})
\]

\[
= \int_{C_{2(j)}} (2\pi)^{-\frac{1}{2}} |\Gamma_{jj}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\hat{W}_{\Delta(j)} - a_j)^{\gamma} (\hat{W}_{\Delta(j)} - a_j)^{\gamma} \right\} d \hat{W}_{\Delta(j)}
\]
2.4. Local Asymptotic Power and Efficiency

In this section, we present our study of local asymptotic power of the union-intersection test for the hypothesis given in (1.7). The power function \( \beta(\theta) \) can be expressed as a function of \( a \) and examined when \( a \) tends to zero. The comparison of the two tests may be performed in terms of the slopes of their power functions at a point \( a = 0 \).

1. The power function for the UI test of \( H_0: \Delta = 0 \) against \( H_a: \Delta > 0 \) is given by

\[
\beta_1(a) = \alpha + a \left\{ \sum_{\emptyset \subseteq J \subseteq C_{2(j)}} P \left\{ \begin{array}{c} \hat{W}_{\Delta(j)}^\star, I_{\Delta \Delta: \gamma}^{jj} \hat{W}_{\Delta(j)}^\star \geq c, \hat{W}_{\Delta(j)}^\star > 0 \end{array} \right\} \right. \\
\times \left. \int_{C_{2(j)}} \Gamma_{jj} t_2 \ d N_{jj}(t_2; 0; (\Gamma_{jj}^{-1})^{\ast}) + P \left\{ \begin{array}{c} \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \end{array} \right\} \right. \\
\times \left. \int_{C_{1(j)}} I_{\Delta \Delta: \gamma}^{jj} t_1 \ d N_{jj}(t_1; 0; (I_{\Delta \Delta: \gamma}^{jj})^{-1}) \right\} 
\]

(2.4.1)

where the expression in the brackets is a 2p-vector of non-negative elements and \( \alpha \) is the level of significance of the test. In the particular case of \( J = P \) and \( J = \emptyset \), the power can be expressed as

\[
\beta_1(a) = \alpha + a \left[ \int_{C_{(j)}} \hat{W}_{\Delta} d N_f(\hat{W}_{\Delta}; 0; \Gamma_{\Delta \Delta: \gamma}^{\ast}) \right] 
\]

(2.4.2)

where the expression in the square brackets is a 2p vector of non-negative elements and \( C_{(j)} = \{ \hat{W}_{\Delta}: \hat{W}_{\Delta} > 0 \} \). From the distribution of the UI test under the alternative, given in (2.3.4), the power of the test can be expressed as

\[
\beta_1(a) = \sum_{\emptyset \subseteq J \subseteq C} P \left\{ \begin{array}{c} \hat{W}_{\Delta(j)}^\star, I_{\Delta \Delta: \gamma}^{jj} \hat{W}_{\Delta(j)}^\star \geq c, \hat{W}_{\Delta(j)}^\star > 0 \end{array} \right\} \times P \left\{ \begin{array}{c} \Gamma_{jj} \hat{W}_{\Delta(j)} < 0 \end{array} \right\}
\]
Expanding $\beta_1(a)$ in a Taylor series around $a=0$, we have

$$\beta_1(a) = \beta_1(a) \bigg|_{a=0} + a \frac{\partial \beta_1(a)}{\partial a} \bigg|_{a=0} + o(|a|) = \alpha + a \frac{\partial \beta_1(a)}{\partial a} \bigg|_{a=0} + o(|a|)$$

where

$$\frac{\partial \beta_1(a)}{\partial a} = \sum_{0 \leq J \leq p} \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} < 0 \right\} + \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} > 0 \right\}$$

$$\times \frac{\partial}{\partial a} \mathbb{P}_H \left\{ \check{W}^*_{\Delta(j)} , I_{\Delta(j)} \check{W}^*_{\Delta(j)} \geq c, \check{W}^*_{\Delta(j)} > 0 \right\}$$

$$\times \frac{\partial}{\partial a} \mathbb{P}_H \left\{ \check{W}^*_{\Delta(j)} , I_{\Delta(j)} \check{W}^*_{\Delta(j)} \geq c, \check{W}^*_{\Delta(j)} > 0 \right\}$$

(2.4.3)

The derivatives of the probabilities above can be calculated as:

$$\frac{\partial}{\partial a} \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} < 0 \right\} = - \frac{\partial}{\partial a} \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} > 0 \right\}$$

(2.4.4)

which is a $2p$ vector whose $s_j$ elements are zero, where $s_j$ is the dimension of $J$, that is,

$$\frac{\partial}{\partial a_j} \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} > 0 \right\} = 0$$

and

$$\frac{\partial}{\partial a_j} \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} > 0 \right\} = - \int_{C_2(j)} \Gamma_{jj} (t_2 - a_j) \ d N_{s_j}(t_2; a_j; (\Gamma_{jj}^{-1})$$

$$= - \Gamma_{jj} \int_{C_2(j)} t_2 \ d N_{s_j}(t_2; a_j; (\Gamma_{jj}^{-1})$$

$$+ a_j \mathbb{P}_H \left\{ \Gamma_{jj} \check{W}_{\Delta(j)} > 0 \right\}$$

(2.4.5)
which is also a \(2p\) vector whose \(s_j\) elements are zero, where \(s_j\) is the dimension of \(j\), that is,

\[
\frac{\partial}{\partial a_j^*} P_{H_a} \left\{ \bar{W}^{\star}_{\Delta(j)} \left< \bar{I}^{ij}_{\Delta;\gamma}, \bar{W}^{\star}_{\Delta(j)} \geq c, \bar{W}^{\star}_{\Delta(j)} > 0 \right\} = 0
\]

and

\[
\frac{\partial}{\partial a_j^*} P_{H_a} \left\{ \bar{W}^{\star}_{\Delta(j)} \left< \bar{I}^{ij}_{\Delta;\gamma}, \bar{W}^{\star}_{\Delta(j)} \geq c, \bar{W}^{\star}_{\Delta(j)} > 0 \right\} \times \frac{\partial a_j^*}{\partial a_j}
\]

\[
= \int_{C_1(j)} \bar{I}^{ij}_{\Delta;\gamma} (t_1 - a_j^*) \, d N_s_j(t_1; a_j^*; (\bar{I}^{ij}_{\Delta;\gamma})^{-1})
\]

\[
= \bar{I}^{ij}_{\Delta;\gamma} \int_{C_1(j)} t_1 \, d N_s_j(t_1; a_j^*; (\bar{I}^{ij}_{\Delta;\gamma})^{-1})
\]

\[
- a_j^* \bar{I}^{ij}_{\Delta;\gamma} P_{H_a} \left\{ \bar{W}^{\star}_{\Delta(j)} \left< \bar{I}^{ij}_{\Delta;\gamma}, \bar{W}^{\star}_{\Delta(j)} \geq c, \bar{W}^{\star}_{\Delta(j)} > 0 \right\}
\]

where \(a_j^* = a_j - \bar{I}^{ij}_{\Delta;\gamma} (\bar{I}^{ij}_{\Delta;\gamma})^{-1} a_j\). Then

\[
\frac{\partial \beta_1(a)}{\partial a} \bigg|_{a=0} = - \sum_{\emptyset \subseteq J \subseteq \mathbb{P}} P_{H_0} \left\{ \bar{W}^{\star}_{\Delta(j)} \left< \bar{I}^{ij}_{\Delta;\gamma}, \bar{W}^{\star}_{\Delta(j)} \geq c, \bar{W}^{\star}_{\Delta(j)} > 0 \right\}
\]

\[
\times \Gamma_{jj} \int_{C_2(j)} t_2 \, d N_s_j(t_2; 0; (\Gamma_{jj})^{-1}) + P_{H_0} \left\{ \Gamma_{jj} \bar{W}_{\Delta(j)} < 0 \right\}
\]

\[
\times \bar{I}^{ij}_{\Delta;\gamma} \int_{C_1(j)} t_1 \, d N_s_j(t_1; 0; (\bar{I}^{ij}_{\Delta;\gamma})^{-1})
\]

\[
= \sum_{\emptyset \subseteq J \subseteq \mathbb{P}} \left\{ \int_{C_1(j)} d N_s_j(t_1; 0; (\bar{I}^{ij}_{\Delta;\gamma})^{-1})
\right\}
\]

\[
\times \Gamma_{jj} \int_{C_2(j)} t_2 \, d N_s_j(t_2; 0; (\Gamma_{jj})^{-1})
\]

\[
+ \int_{C_2(j)} d N_s_j(t_2; a_j; (\Gamma_{jj})^{-1})
\]
\[ x \mathbf{I}^{jj}_{\Delta \Delta : \gamma} \int_{C_{1(j)}} t_i \, d \mathbf{N}_{t_i}(0; (\mathbf{I}^{jj}_{\Delta \Delta : \gamma})^{-1}) \] 

which is non-negative, since \( \mathbf{W}_{\Delta(j)}^* > 0 \), \( \mathbf{I}^{jj}_{\Delta \Delta : \gamma} \) is p.d. and \( -\mathbf{I}^{jj}_{\Delta \Delta : \gamma} \mathbf{W}_{\Delta(j)} > 0 \). We would like to compare the test with restricted alternative with the test with unrestricted alternative, that is, with \( H_a : \Delta \neq 0 \). For the test with unrestricted alternative, the test statistic will be

\[ Q = \mathbf{W}'_{\Delta} \mathbf{I}^{1}_{\Delta \Delta : \gamma} \mathbf{W}_{\Delta} = \mathbf{U}'_{\Delta} \mathbf{I}^{1}_{\Delta \Delta : \gamma} \mathbf{U}_{\Delta} \tag{2.4.6} \]

which under \( H_0 \) is a \( \chi^2 \) with 2p degrees of freedom and the power function will be a non-central chi-squared with 2p degrees of freedom and non-centrality parameter \( \theta = a' \mathbf{I}^{1}_{\Delta \Delta : \gamma} a \), that is,

\[ \beta_2(\theta) = P_{H_a}\{ \mathbf{W}'_{\Delta} \mathbf{I}^{1}_{\Delta \Delta : \gamma} \mathbf{W}_{\Delta} \geq x \} = e^{-\frac{1}{2} \sum_{r=0}^{\infty} \frac{\theta^r}{r!}} P\{ \chi^2_{r+2} \geq x \} \]

where \( r = 2p \). When \( H_0 \) is true, \( \beta_2(\theta) = a = P\{ \chi^2 \geq x \} \). In order to compare the two sets, we need to express \( \beta_2(\theta) \) as a function of \( a \) or as a function of \( \theta \). Thus, we expand \( \beta_2(\theta) \) around \( \theta = 0 \) in a Taylor series

\[ \beta_2(\theta) = \beta_2(\theta) \bigg|_{\theta = 0} + \theta \frac{\partial \beta(\theta)}{\partial \theta} \bigg|_{\theta = 0} + o(\theta) \]

The derivative of the power function is given by

\[ \frac{\partial \beta(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta} P_{H_a}\{ \mathbf{W}'_{\Delta} \mathbf{I}^{1}_{\Delta \Delta : \gamma} \mathbf{W}_{\Delta} \geq x \} = \frac{\partial}{\partial \theta} P_{H_a}\{ \chi^2 \geq x \} \]

\[ = \frac{1}{2} \left\{ P\{ \chi^2_{r+2} \geq x \} - P\{ \chi^2 \geq x \} \right\} = f_{r+2}(\mathbf{W}_\Delta; \theta) \]
which is the density function of a non-central chi-squared with \( r+2 \) degrees of freedom and non-centrality parameter \( \theta \). When \( \theta = 0 \) we have

\[
\frac{\partial}{\partial \theta} P_{H_0}\{\chi_r^2 \geq x\} = \frac{1}{2} \left\{ P\{\chi_{r+2}^2 \geq x\} - P\{\chi_r^2 \geq x\} \right\}
\]

which is non-negative since it is a p.d.f. Thus,

\[
\beta_2(\theta) = \alpha + \frac{1}{2} a' I_{\Delta:\gamma} a \left\{ P\{\chi_{r+2}^2 \geq x\} - P\{\chi_r^2 \geq x\} \right\} + o(|\theta|) \tag{2.4.7}
\]

is a function of \( \theta = a' I_{\Delta:\gamma} a \), a quadratic form in \( a \). Since \( a' I_{\Delta:\gamma} a \leq k \), where \( k = \text{ch}_{\text{max}} I_{\Delta:\gamma} \) (Rao, page 50, 1f2.1), we have \( \left| a' I_{\Delta:\gamma} a \right| \leq k \) for every \( a \). Then,

\[
\left| a' I_{\Delta:\gamma} a \right| \leq O(|a|^2) = o(|a|). \]

The power function \( \beta_2 \) can be expressed as

\[
\beta_2(\theta) = \alpha + \left( \frac{k_2}{2} \right) a' I_{\Delta:\gamma} a + o(|a|^2) \leq \alpha + \left( \frac{k_2}{2} \right) k (|a|^2) + o(|a|^2) \tag{2.4.8}
\]

Now, for the comparison of the power functions given in (2.4.1) and (2.4.8), we can calculate

\[
\beta_1(a) - \beta_2(a) \geq k_1 a + o(|a|) - \left( \frac{k_2}{2} \right) k (|a|^2) - o(|a|^2)
\]

\[
= k_1 a - \left( \frac{k_2}{2} \right) o(|a|) + o(|a|^2) = a \left( k_1 - \frac{1}{2} k_2 (o(1)) \right),
\]

For larger values of \( a \) the quantity \( \frac{1}{2} k_2 (o(1)) = 0 \) and \( \beta_1(a) - \beta_2(a) \) is positive. Therefore, the test statistic \( \lambda^* \) defined in (3.2.20) is more powerful than \( Q \) defined in (2.4.6).
CHAPTER III

RESTRICTED ALTERNATIVES TEST IN A PARAMETRIC MODEL
AND IN A LOGISTIC MODEL WITH COMPETING RISK DATA

3.1 Introduction

In this Chapter we derive the union-intersection test for comparison of survival distributions from two independent subpopulations in a competing risk situation with two causes of failure using two different approaches: the first, when a parametric model is known, and the second, using the logistic model.

Since independence of causes of failures is not realistic in many competing risk situation we may consider dependent causes of failure, that is, the theoretical times of an individual failing from one cause may be correlated with the theoretical times of the same individual failing from the other cause. Thus, a dependent parametric model may be considered. The exponential survival distribution has found considerable acceptance in survival analysis for a single cause of failure. Hence it would be desirable to work with some bivariate distribution whose marginals are exponential within the context of dependent competing cause of failure. The bivariate exponential of Sarkar (1987) appear to meet the requirements cited above, that is, it is a dependent model with exponential marginals. Furthermore, it is an absolutely continuous distribution, which is adequate for the situation where failure do not occur for both causes simultaneously.

With data of type \((T,I)\), where \(T\) is the random variable corresponding to the observed failure time and \(I\) is the random index indicating failure type, it may be possible to estimate all the parameters of the dependent underlying distribution. However, the estimated value of the parameter in a dependent model measuring the dependence between the hypothetical times \(X_i, i=1, 2, \ldots, k\), should not be taken as an
indicator of dependence. The reason for this lies in the fact that the cause-specific hazard function from a dependent model may arise an independent model with the same cause-specific hazard function (Prentice et al., 1980).

The logistic linear model has been an alternative to the proportional hazard models, in the analysis of the data from prospective studies. The data from those studies consist basically of values on a dichotomized dependent variable, failure or not, and of values on concomitant variables observed at entry time. In a competing risk situation, the cause of failure among the k causes is also observed. Others data may be incorporated to the model: entry time y, follow-up time τ. This information is useful in situations when subjects enter the study at different times or withdrawal before its termination. The concomitant variables can be time dependent.

Some studies have appeared comparing logistic linear models and proportional hazard models in prospective studies; among others, Green and Symons(1983) and Elandt-Johnson (1983). It appears that the two models are equivalent for studies when the disease is rare and/or follow-up time is short.

For cohort data, when all individuals enter the study at the same time and there are no withdrawals before the termination of the fixed time t, no knowledge of the underlying survival distribution is needed. When individuals enter the study at different times or censoring occurs before the termination, the underlying parametric survival distribution may be specified.

3.2. Union-Intersection Test for a Parametric Model

In this section we derive a union-intersection test for the comparison of two parametric distributions within Sarkar's family of the absolutely continuous bivariate exponential distribution, in a competing risk situation.

The absolutely continuous bivariate exponential distribution of Sarkar (1987) has a survival distribution of the form
\[ P(X_1 > t_1, X_2 > t_2) = \exp\{- (\lambda_2 + \lambda_{12})t_2\} \left\{ 1 - \left[A(\lambda_1 t_1)\right]^{-\gamma} [A(\lambda_1 t_1)]^{1+\gamma} \right\} \]

if \(0 < t_1 < t_2\)

\[ = \exp\{- (\lambda_1 + \lambda_{12})t_1\} \left\{ 1 - \left[A(\lambda_2 t_2)\right]^{-\gamma} [A(\lambda_2 t_2)]^{1+\gamma} \right\} \]

if \(t_1 > t_2 > 0\)

(2.1)

where \(\lambda_1 > 0, \lambda_2 > 0\) and \(\lambda_{12} \geq 0, \gamma = \frac{\lambda_{12}}{\lambda_1 + \lambda_2}\) and \(A(z) = 1 - \exp(-z)\) for \(z > 0\). Note that \(X_1\) and \(X_2\) are independent when \(\lambda_{12} = 0\). The marginals are exponential,

\[ S_{X_1}(t_1,0) = \exp\{- (\lambda_1 + \lambda_{12}) t_1\}, \ t_1 > 0, \]

(2.2)

\[ S_{X_2}(0,t_2) = \exp\{- (\lambda_2 + \lambda_{12}) t_2\}, \ t_2 > 0, \]

(2.3)

and the density of \(\min(X_1, X_2) = T\) is exponential with parameter \(\lambda\), that is,

\[ S_T(t,0) = S(t,t) = \exp\{- (\lambda_1 + \lambda_2 + \lambda_{12}) t\} = \exp\{-t\lambda\}. \]

(2.4)

The hazard functions are

\[ \lambda_i(t) = \frac{- \frac{\partial S(t_1, t_2)}{\partial t_i}}{S_T(t)} \bigg|_{t_1 = t_2 = t} = \frac{\lambda_i t}{\lambda_1 + \lambda_2} = (1 + \gamma) \lambda_i, \ i = 1, 2, \]

(2.5)

since

\[ - \frac{\partial S(t_1, t_2)}{\partial t_1} = (1 + \gamma) \lambda_1 \exp\{-\lambda_1 t_1\} \exp\{-(\lambda_2 + \lambda_{12})t_2\} \times [A(\lambda_1 t_1)]^{-\gamma} [A(\lambda_1 t_1)]^{-\gamma} \]

if \(0 < t_1 \leq t_2\)

\[ - \frac{\partial S(t_1, t_2)}{\partial t_2} = (1 + \gamma) \lambda_2 \exp\{-\lambda_{12} t_2\} \exp\{-(\lambda_1 + \lambda_{12})t_1\} \times [A(\lambda_2 t_2)]^{-\gamma} [A(\lambda_2 t_2)]^{-\gamma} \]
if \( t_1 \geq t_2 > 0 \)

and

\[
- \frac{\partial S(t_1,t_2)}{\partial t_1} \bigg|_{t_1=t_2=t} = (1+\gamma) \lambda_1 \exp\{-\lambda t\}, \text{ if } t_1 \leq t_2,
\]

\[
- \frac{\partial S(t_1,t_2)}{\partial t_2} \bigg|_{t_1=t_2=t} = (1+\gamma) \lambda_2 \exp\{-\lambda t\}, \text{ if } t_2 \leq t_1,
\]

\[S_T(t) = \exp\{-\lambda t\}.\]

The probability distribution of the time of failure for cause \( C_i \) are

\[
Q_i(t) = \int_0^t \lambda_i(u) S_T(u) \, du = \int_0^t (1+\gamma) \lambda_i \exp\{-\lambda u\} \, du
\]

\[
= (1+\gamma) \frac{\lambda_i}{\lambda_1+\lambda_2} \left( \exp\{-\lambda t\} \right) = \frac{\lambda_i}{\lambda_1+\lambda_2} (1 - \exp\{-\lambda t\}), \quad i = 1, 2
\]  

(2.6)

and the proportions of failure due to \( C_i \) are

\[
\Pi_i = \frac{\lambda_i}{\lambda_1+\lambda_2}, \quad i = 1, 2.
\]  

(2.7)

The conditional survival function due to cause \( C_i \) in the presence of each other cause is also exponential with parameter \( \lambda \)

\[
S_i^*(t) = \frac{Q_i^*(\infty) - Q_i^*(t)}{Q_i^*(\infty)} = \exp\{-\lambda t\}, \quad i = 1, 2
\]  

(2.8)

with density given by \( f_i^*(t) = \lambda \exp\{-\lambda t\}, \quad i = 1, 2 \). Since \( S_i^*(t) = S_T(t) = \exp\{-\lambda t\} \), the overall hazard rate and the hazard functions are proportional, that is, \( \frac{\lambda_i}{\lambda_T(t)} = c_i \), where \( 0 < c_i < 1 \). Thus, the distribution associated with cause \( C_i \) alone (Elandt-Johnson, 1980) is
a fractional power (power is $c_i$) of $S_T$,  

$$G_i(t) = \{S_T(t)\}^{c_i} = \exp\left\{-\frac{\lambda \lambda_i}{\lambda_1 + \lambda_2} \ t\right\}, i = 1, 2. \quad (2.9)$$

The unconditional probability of failure from $C_i$ in an interval of time $[t_i, t_i + \Delta]$ is given by

$$-\frac{\partial S(t_1, t_2)}{\partial t_i}_{|t_i=t} = \lambda_i(t) \ S_T(t) = f_i^*(t) \ \prod_i = (1 + \gamma) \ \lambda_i \exp\{-\lambda \ t\} \quad (2.10)$$

and the likelihood for $\theta' = (\lambda_1, \lambda_2, \lambda_{12})$ can be written as

$$L(\theta, t) = \prod_{l=1}^{n} \ \prod_{i=1}^{2} \ \{\lambda_i(t_i; \theta)\} \ \delta_{i(l)} \ S_T(t_l)$$

$$= \prod_{l=1}^{n} \ \left\{\frac{\lambda \lambda_i}{\lambda_1 + \lambda_2}\right\} \ \delta_{i(l)} \ \exp\{-\lambda t l\} \quad (2.11)$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$ and $\delta_{i(l)}$ is the indicator which assumes the value 1 if cause of failure for the $l$-th subject is $C_i$ and zero otherwise.

3.2.1. The Problem and Assumptions.

For the comparison of survival distributions for two independent subpopulations, the statistical formulation of the hypothesis involve the equality of the parameters of the corresponding distributions, or in a reparameterized form, the difference between the parameters corresponding to the two distributions. That is, if the underlying distribution is the absolutely continuous bivariate exponential of Sarkar, the hypothesis of interest is

$$H_0: \lambda^{(1)}_1 = \lambda^{(2)}_1, \lambda^{(1)}_2 = \lambda^{(2)}_2, \lambda^{(1)}_{12} = \lambda^{(2)}_{12} \quad (2.1.1)$$
which may be equivalent to

$$H_0: \Delta_1 = 0, \Delta_2 = 0, \Delta_{12} = 0$$  \hspace{1cm} (2.1.2)$$

where $\Delta=(\Delta_1, \Delta_2, \Delta_{12})$ is the vector of differences between the parameters of the distributions of the two subpopulations, that is,

$$\lambda_{s(t)}^{(1)} = \lambda_{s(t)} + \Delta_{s(t)} \text{ and } \lambda_{s(t)}^{(2)} = \lambda_{s(t)}, s=1, 2, 12,$$

where the superscripts (1) and (2) indicate the subpopulations. The alternative of interest is

$$H_a: \Delta_1 > 0, \Delta_2 > 0, \Delta_{12} > 0.$$  \hspace{1cm} (2.1.3)$$

The parameter vector is $\theta=(\lambda_1, \lambda_2, \lambda_{12}, \Delta_1, \Delta_2, \Delta_{12})$. For testing $H_0$ in (2.1.2) against the restricted alternative in (2.1.3), $\lambda=(\lambda_1, \lambda_2, \lambda_{12})$ will be considered a vector of the nuisance parameters. The likelihood function (2.11) may be written as

$$L(\theta, t) = \prod_{l=1}^n \left\{ \frac{(\lambda + w_{il}\Delta)}{(\lambda_1 + w_{il}(\Delta_1 + \Delta_2))} \right\}^{\delta_{i(l)}}$$

$$+ \left\{ \frac{(\lambda_2 + w_{il}\Delta_2)}{(\lambda_1 + \lambda_2 + w_{il}(\Delta_1 + \Delta_2))} \right\}^{\delta_{i(l)}} \times \exp \left\{ - (\lambda + w_{il}\Delta) t_i \right\}$$

$$= \prod_{l=1}^n f(t_i; \theta)$$  \hspace{1cm} (2.1.4)$$

where $w_{il}$ is an indicator function which assumes the value 1 if the subject comes from subpopulation 1 and zero otherwise. The log of the likelihood (2.1.4) can be written as

$$\log L(\theta, t) = \sum_{l=1}^n \left\{ \delta_{i(l)} \left( \log (\lambda_1 + w_{il}\Delta_1) + \log(\lambda + w_{il}\Delta) \right) \right\}$$
\[
- \log((\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)) + \delta_{2(i)} \left( \log \left( \frac{1}{\lambda_1 + w_i(\Delta_1 + \Delta_2)} \right) + \log \left( \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} \right) - (\lambda + w_i(\Delta_1 + \Delta_2)) \right) = 0.
\]

The maximum likelihood estimators of \( \lambda \) and \( \Delta \) are the solutions of the system of equations \( \frac{\partial \log L(\theta, t)}{\partial \lambda} = 0 \) and \( \frac{\partial \log L(\theta, t)}{\partial \Delta} = 0 \), respectively.

For the study, the regularity conditions for the large sample derivations of MLE must be satisfied. They are assumptions A1-A3 stated in (2.13)-(2.15), in Chapter II.

1. The first derivatives of \( \log f(t_i; \theta) \) exist, where \( f(t_i; \theta) \) are the unconditional probabilities of failing due to \( C_i \) for the \( l \)-th subject (2.1.4). They are:

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \lambda_1} = \delta_{1(i)} \left[ \frac{1}{\lambda_1 + w_i(\Delta_1 + \Delta_2)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] + \delta_{2(i)} \left[ \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] - t_i
\]

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \lambda_2} = \delta_{1(i)} \left[ \frac{1}{\lambda_2 + w_i(\Delta_2)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] + \delta_{2(i)} \left[ \frac{1}{\lambda_2 + w_i(\Delta_2)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] - t_i
\]

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \lambda_{12}} = \delta_{1(i)} \left[ \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} \right] + \delta_{2(i)} \left[ \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} \right] - t_i
\]

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \Delta_1} = \delta_{1(i)} \left[ \frac{1}{\lambda_1 + w_i(\Delta_1)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] + \delta_{2(i)} \left[ \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] - w_i t_i = w_i \frac{\partial f(t_i; \theta)}{\partial \lambda_1}
\]

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \Delta_2} = \delta_{1(i)} \left[ \frac{1}{\lambda_2 + w_i(\Delta_2)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] + \delta_{2(i)} \left[ \frac{1}{\lambda_2 + w_i(\Delta_2)} + \frac{1}{\lambda + w_i(\Delta_1 + \Delta_2)} - \frac{1}{\lambda_1 + \lambda_2 + w_i(\Delta_1 + \Delta_2)} \right] - w_i t_i = w_i \frac{\partial f(t_i; \theta)}{\partial \lambda_2}
\]
\[
\frac{\partial \ln f(t_i; \theta)}{\partial \Delta_2} = w_i \frac{\partial \ln f(t_i; \theta)}{\partial \lambda_2}
\]

\[
\frac{\partial \ln f(t_i; \theta)}{\partial \Delta_{12}} = w_i \frac{\partial \ln f(t_i; \theta)}{\partial \lambda_{12}}
\]

(2.1.6)

The second derivatives also exist and they are:

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \lambda_2} = \delta_{1(1)} \left[ -\frac{1}{(\lambda_1 + w_i \Delta_1)^2} - \frac{1}{(\lambda_1 + w_i \Delta)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right] + \delta_{2(1)} \left[ -\frac{1}{(\lambda_1 + w_i \Delta)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right]
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \lambda_{12}} = \delta_{1(1)} \left[ -\frac{1}{(\lambda_1 + w_i \Delta)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right] + \delta_{2(1)} \left[ -\frac{1}{(\lambda_1 + w_i \Delta)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right]
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \lambda_{12}} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_{12} \partial \lambda_{12}} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_{12} \partial \lambda_{12}}
\]

\[
= \delta_{1(1)} \left[ -\frac{1}{(\lambda + w_i \Delta)^2} \right] + \delta_{2(1)} \left[ -\frac{1}{(\lambda + w_i \Delta)^2} \right]
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \lambda_2} = \delta_{1(1)} \left[ -\frac{1}{(\lambda_2 + w_i \Delta_2)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right] + \delta_{2(1)} \left[ -\frac{1}{(\lambda_2 + w_i \Delta_2)^2} + \frac{1}{[(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)]^2} \right]
\]
\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_1} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \lambda_1}, \text{ etc.}
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_1} = \delta_{1(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta_1)} - \frac{w_i}{(\lambda + w_i \Delta_1)^2} + \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} \right\} 
\]

\[-\frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} + \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} \right\} 
\]

\[+ \delta_{2(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} - \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]} \right\} 
\]

\[+ \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} \right\} 
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_2} = \delta_{1(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} - \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]} \right\} 
\]

\[+ \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} \right\} + \delta_{2(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} \right\} 
\]

\[+ \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} \right\} 
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_{12}} = \delta_{1(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} \right\} + \delta_{2(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} \right\} 
\]

\[
\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_1} = \delta_{1(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} - \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]} \right\} 
\]

\[+ \frac{w_i}{[(\lambda_1 + \lambda_2) + w_i(\Delta_1 + \Delta_2)]^2} \right\} + \delta_{2(i)} \left\{ \frac{w_i}{(\lambda + w_i \Delta)} - \frac{w_i}{(\lambda + w_i \Delta)^2} \right\} 
\]
\[ \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_2} = \delta_{i(1)} \left\{ \frac{w_i}{(\lambda+w_i \Delta)} - \frac{w_i}{(\lambda+w_i \Delta)^2} - \frac{w_i}{[(\lambda_1+\lambda_2) + w_i(\Delta_1+\Delta_2)]} \right\} + \delta_{2(1)} \left\{ \frac{w_i}{(\lambda+w_i \Delta)} - \frac{w_i}{(\lambda_2+w_i \Delta_2)^2} + \frac{w_i}{(\lambda+w_i \Delta)} \right\} \]

\[ = \frac{w_i}{(\lambda+w_i \Delta)^2} - \frac{w_i}{[(\lambda_1+\lambda_2) + w_i(\Delta_1+\Delta_2)]} + \frac{w_i}{[(\lambda_1+\lambda_2) + w_i(\Delta_1+\Delta_2)]^2} \]

\[ \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_{12}} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_{12}} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_1} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_2} = \frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_1 \partial \Delta_{12}} \]

It is easy to verify that the third derivatives exist for all \( \theta \in \Theta, r=1, 2, 12 \). Then assumption A1 holds.

2. We may write \( f(t_i; \theta) \) from (2.11) as

\[ f(t_i; \theta) = \prod_{i=1}^{2} \left\{ \frac{(\lambda+w_i \Delta)(\lambda_1+w_i \Delta_1)}{(\lambda_1+\lambda_2)+w_i(\Delta_1+\Delta_2)} \right\}^{\delta_{i(1)}} \times \exp \left\{ - (\lambda+w_i \Delta) t_i \right\} \]

\[ = a_i \exp \left\{ (\lambda+w_i \Delta) t_i \right\} \]

where \( a_i = \prod_{i=1}^{2} \left\{ \frac{(\lambda+w_i \Delta)(\lambda_1+w_i \Delta_1)}{(\lambda_1+\lambda_2)+w_i(\Delta_1+\Delta_2)} \right\}^{\delta_{i(1)}} \) does not depend on \( t \) and \( b_i = (\lambda+w_i \Delta) \).

Then, for \( \theta \in \Theta \),
\[
\frac{\partial f(t_i; \theta)}{\partial \theta_s} = a_i \frac{\partial b_i}{\partial \theta_s} \exp\{-b_i t_i\} + \exp\{-b_i t_i\} \frac{\partial b_i}{\partial \theta_s} t_i
\]

\[
= \exp\{-b_i t_i\} \left(-a_i \frac{\partial b_i}{\partial \theta_s} t_i + \frac{\partial a_i}{\partial \theta_s}\right)
\]

\[
= \exp\{-b_i t_i\} \left(-c_i t_i + d_i\right)
\]

where \(c_i = -a_i \frac{\partial b_i}{\partial \theta_s}\) and \(d_i = \frac{\partial a_i}{\partial \theta_s}\). Then,

\[
\left|\frac{\partial f(t_i; \theta)}{\partial \theta_s}\right| \leq |c_i| t_i \exp\{-b_i t_i\} + |d_i| \exp\{-b_i t_i\} = g(t_i)
\]

Also, for \(\theta_r \in \theta, \theta_r \notin \theta\), we have

\[
\frac{\partial^2 f(t_i; \theta)}{\partial \theta_s \partial \theta_r} = (c_i t_i + d_i) \exp\{-b_i t_i\} \left(-t_i \frac{\partial b_i}{\partial \theta_r} + (t_i \frac{\partial c_i}{\partial \theta_r} + \frac{\partial d_i}{\partial \theta_r}) \exp\{-b_i t_i\}\right)
\]

\[
= -c_i t_i^2 \frac{\partial b_i}{\partial \theta_r} \exp\{-b_i t_i\} - d_i \frac{\partial b_i}{\partial \theta_r} t_i \exp\{-b_i t_i\}
\]

\[
+ \frac{\partial c_i}{\partial \theta_r} t_i \exp\{-b_i t_i\} + \frac{\partial d_i}{\partial \theta_r} \exp\{-b_i t_i\}
\]

\[
= e_i t_i^2 \exp\{-b_i t_i\} + f_i t_i \exp\{-b_i t_i\} + g_i \exp\{-b_i t_i\}
\]

where \(e_i = -c_i \frac{\partial b_i}{\partial \theta_r}\), \(f_i = d_i \frac{\partial d_i}{\partial \theta_r} + \frac{\partial c_i}{\partial \theta_r}\), and \(g_i = \frac{\partial d_i}{\partial \theta_r}\). Then,

\[
\left|\frac{\partial^2 f(t_i; \theta)}{\partial \theta_s \partial \theta_r}\right| \leq |e_i| t_i^2 \exp\{-b_i t_i\} + |f_i| t_i \exp\{-b_i t_i\} + |g_i| \exp\{-b_i t_i\} = h(t_i)
\]

For all \(i = 1, 2, \ldots\), \(b_i = \lambda + w_i \Delta = (\lambda_1 + \lambda_2 + \lambda_1) > 0\). Then, \(\exp\{-b_i t_i\}, t_i \exp\{-b_i t_i\}\) and \(t_i^2 \exp\{-b_i t_i\}\) are integrable on \((0, \infty)\) and so are \(\int g(t_i) \, dt_i\) and \(\int h(t_i) \, dt_i\). Thus assumption \(A_2\) holds.
3. Consider now the first derivatives given in (2.16). For each subject who fails from the $ith$ cause, we can write those derivatives as

$$\frac{\partial \ln f^{(i)}(t_i; \theta)}{\partial \theta_i} = \delta_{i(i)} f(\theta) - t_i = \delta_{i(i)} f(\theta) \left[ 1 - \frac{t_i}{\delta_{i(i)} f(\theta)} \right]$$

where $f(\theta)$ is a function of the parameters. Then, for a positive $\eta$,

$$E \left| \frac{\partial \ln f^{(i)}(t_i; \theta)}{\partial \theta_i} \right|^{2+\eta} < \delta_{i(i)} f(\theta)^{2+\eta} \left[ 1 - \frac{t_i}{\delta_{i(i)} f(\theta)} \right]$$

since $t_i$ is exponentially distributed, all its moments are finite and the last term of the inequality above is positive and finite, that is,

$$E \left| \frac{\partial \ln f^{(i)}(t_i; \theta)}{\partial \theta_i} \right|^{2+\eta} < K < \infty \text{ for all } \theta \in \Theta.$$

Under the regularity conditions above, consistency and asymptotic normality of M.L.E. are verified. The M.L.E. of $\lambda=(\lambda_1, \lambda_2, \lambda_{12})$, the vector of nuisance parameters, is calculated under $H_0$, solving the following equations.

$$\begin{align*}
\frac{\partial \ln L(t_i; \theta)}{\partial \lambda_1} \big|_{\Delta=0} &= \sum_{i=1}^{\delta_{1(i)}} \left\{ \sum_{j=1}^{\delta_{1(i)}} \frac{1}{\lambda_1 + w_j \Delta_1} + \frac{1}{\lambda + w_i \Delta} - \frac{1}{(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)} \right\} - t_i \big|_{\Delta=0} (2.1.8) \\
\frac{\partial \log L(t_i; \theta)}{\partial \lambda_2} \big|_{\Delta=0} &= \sum_{i=1}^{\delta_{1(i)}} \left\{ \sum_{j=1}^{\delta_{1(i)}} \frac{1}{\lambda + w_i \Delta} - \frac{1}{(\lambda_1 + \lambda_2) + w_i (\Delta_1 + \Delta_2)} \right\} - t_i \big|_{\Delta=0} (2.1.9)
\end{align*}$$
\frac{\partial \log L(t_i; \theta)}{\partial \lambda_{12}} \bigg|_{\Delta = 0} = \sum_{l=1}^{n} \left\{ \delta_{i(l)} \left[ \frac{1}{\lambda + w_i \Delta} \right] + \delta_{2(l)} \left[ \frac{1}{\lambda + w_i \Delta} \right] - t_l \right\} \bigg|_{\Delta = 0} = 0. \tag{2.1.10}

Let \( n_i = \sum_{l=1}^{n} \delta_{i(l)} \) indicate the number of subjects who die due to \( C_i \), \( i = 1, 2 \), \( \sum_{i=1}^{n} \delta_{i(l)} = n \) and \( \lambda^0 = (\lambda^0_1, \lambda^0_2, \lambda^0_{12}) \) be the vector of MLE of \( \lambda \) and \( \lambda^0 = \lambda^0_1 + \lambda^0_2 + \lambda^0_{12} \). Consider the score functions

\[ \hat{U}^0_{\Delta} = (\hat{U}^0_{\Delta_1}, \hat{U}^0_{\Delta_2}, \hat{U}^0_{\Delta_{12}}) \]

where

\[
\hat{U}^0_{\Delta_1} = \frac{\partial \log L(t_i; \theta)}{\partial \lambda_1} \bigg|_{\Delta = 0, \lambda = \hat{\lambda}} = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} w_i \left\{ \delta_{1(l)} \frac{1}{\lambda^0_1} + \sum_{i=1}^{n} \delta_{i(l)} \left[ \frac{1}{\lambda^0} - \frac{1}{\lambda^0_1 + \lambda^0_2} \right] - t_i \right\} \tag{2.1.11} \]

\[
\hat{U}^0_{\Delta_2} = \frac{\partial \log L(t_i; \theta)}{\partial \lambda_2} \bigg|_{\Delta = 0, \lambda = \hat{\lambda}} = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} w_i \left\{ \delta_{2(l)} \frac{1}{\lambda^0_2} + \sum_{i=1}^{n} \delta_{i(l)} \left[ \frac{1}{\lambda^0} - \frac{1}{\lambda^0_1 + \lambda^0_2} \right] - t_i \right\} \tag{2.1.12} \]

and

\[
\hat{U}^0_{\Delta_{12}} = \frac{\partial \log L(t_i; \theta)}{\partial \lambda_{12}} \bigg|_{\Delta = 0, \lambda = \hat{\lambda}} = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} w_i \left\{ \delta_{1(l)} \frac{1}{\lambda^0_1} + \delta_{2(l)} \frac{1}{\lambda^0_2} - t_i \right\}. \tag{2.1.13} \]

and the sample variance-covariance matrix

\[
\begin{bmatrix}
I(\hat{\theta})_{6 \times 6} & = & \begin{bmatrix}
I_{\lambda\lambda} & I_{\lambda \Delta} \\
I_{\Delta\lambda} & I_{\Delta \Delta}
\end{bmatrix}
\tag{2.1.14}
\end{bmatrix}
\]

where \( I_{\lambda\lambda} \) is a 3 \times 3 submatrix with elements
\[
I_{\lambda_1 \lambda_1} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \lambda_1 \partial \lambda_1} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} \left[ \frac{1}{(\lambda_1^0)^2} + \frac{1}{(\lambda_2^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] + \delta_{2(l)} \left[ \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

\[
I_{\lambda_1 \lambda_2} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \lambda_1 \partial \lambda_2} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} \left[ \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] + \delta_{2(l)} \left[ \frac{1}{(\lambda_2^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

\[
I_{\lambda_1 \lambda_{12}} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \lambda_1 \partial \lambda_{12}} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} \left[ \frac{1}{(\lambda_1^0)^2} + \delta_{2(l)} \frac{1}{(\bar{\lambda}_2^0)^2} \right] + \delta_{2(l)} \left[ \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

\[
I_{\lambda_2 \lambda_2} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \lambda_2 \partial \lambda_2} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} \left[ \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] + \delta_{2(l)} \left[ \frac{1}{(\lambda_2^0)^2} + \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

\[
I_{\lambda_2 \lambda_{12}} = I_{\lambda_1 \lambda_1} = I_{\lambda_1 \lambda_{12}}, \quad I_{\Delta \Delta} \text{ is a 3 x 3 submatrix with elements}
\]

\[
I_{\Delta_1 \Delta_1} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \Delta_1 \partial \Delta_1} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} w_l^2 \left[ \frac{1}{(\lambda_1^0)^2} + \frac{1}{(\lambda_2^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] + \delta_{2(l)} w_l^2 \left[ \frac{1}{(\lambda_1^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

\[
I_{\Delta_1 \Delta_2} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \Delta_1 \partial \Delta_2} \Big|_{\Delta=0, \lambda = \bar{\lambda}} = \frac{1}{n} \sum_{l=1} \left\{ \delta_{1(l)} w_l^2 \left[ \frac{1}{(\lambda_1^0)^2} + \frac{1}{(\lambda_2^0)^2} - \frac{1}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]
\[ I_{\Delta_1 \Delta_2} = \frac{\partial^2 \ln L(t_1; \theta)}{\partial \Delta_1 \partial \Delta_2} \bigg|_{\Delta = 0, \lambda = \lambda} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} w_i \left[ \frac{1}{(\lambda_i^0)^2} - \frac{1}{((\lambda_i^0 + \lambda_2^0)^2)} \right] \right. \\
\left. + \delta_{2(i)} w_i \left[ -\frac{1}{(\lambda_i^0)^2} - \frac{1}{((\lambda_i^0 + \lambda_2^0)^2)} \right] \right\} \]

\[ I_{\Delta_2 \Delta_1} = I_{\Delta_2 \Delta_12} = I_{\Delta_12 \Delta_2} \]

\[ I_{\Delta_2 \Delta_2} = -\frac{\partial^2 \ln L(t_1; \theta)}{\partial \Delta_2 \partial \Delta_2} \bigg|_{\Delta = 0, \lambda = \lambda} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} w_i \left[ \frac{1}{(\lambda_i^0)^2} - \frac{1}{((\lambda_i^0 + \lambda_2^0)^2)} \right] \right. \\
\left. + \delta_{2(i)} w_i \left[ -\frac{1}{(\lambda_i^0)^2} - \frac{1}{((\lambda_i^0 + \lambda_2^0)^2)} \right] \right\} \]

and finally \( I_{\gamma \Delta} \) and \( I_{\Delta \gamma} \) are 3 x 3 submatrices with elements given below.

\[ I_{\lambda_1 \Delta_1} = -\frac{\partial^2 \ln f(t_1; \theta)}{\partial \lambda_1 \partial \Delta_1} \bigg|_{\Delta = 0, \gamma = \gamma} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} \left( \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} + \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} \right) \right. \\
\left. - \frac{w_i}{[(\lambda_i^0 + \lambda_2^0)]^2} + \frac{w_i}{[(\lambda_i^0 + \lambda_2^0)]^2} \right\} + \delta_{2(i)} \left( \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} \right) + \frac{w_i}{[(\lambda_i^0 + \lambda_2^0)]^2} \}

\[ I_{\lambda_1 \Delta_2} = -\frac{\partial^2 \ln f(t_1; \theta)}{\partial \lambda_1 \partial \Delta_2} \bigg|_{\Delta = 0, \gamma = \gamma} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} \left( \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} \right) \right. \\
\left. + \frac{w_i}{[(\lambda_i^0 + \lambda_2^0)]^2} \right\} + \delta_{2(i)} \left( \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} + \frac{w_i}{(\lambda_i^0)^2} \right) + \frac{w_i}{[(\lambda_i^0 + \lambda_2^0)]^2} \} = I_{\lambda_2 \Delta_1} \]

\[ I_{\lambda_1 \Delta_{12}} = -\frac{\partial^2 \ln f(t_1; \theta)}{\partial \lambda_1 \partial \Delta_{12}} \bigg|_{\Delta = 0, \gamma = \gamma} = \frac{1}{n} \sum_{i=1}^{n} \delta_{1(i)} \left( \frac{w_i}{(\lambda_i^0)^2} - \frac{w_i}{(\lambda_i^0)^2} \right) \]
\[ + \delta_{2(i)} \left\{ \frac{w_i}{(\lambda)} - \frac{w_j}{(\lambda)}^2 \right\} = I_{\lambda_2} \Delta_{12} \]

\[ I_{\Delta_2} = -\frac{\partial^2 \ln f(t_i; \theta)}{\partial \lambda_2 \partial \Delta_2} \bigg|_{\Delta=0, \gamma = \hat{\gamma}} = \frac{1}{h} \sum_{i=1}^n \delta_{1(i)} \left\{ \frac{w_i}{(\lambda)} - \frac{w_j}{(\lambda)}^2 \right\} \left\{ \frac{w_i}{(\lambda_1 + \lambda_2)} + \frac{w_j}{(\lambda_1 + \lambda_2)^2} \right\} + \delta_{2(i)} \left\{ \frac{w_i}{(\lambda_2)} - \frac{w_j}{(\lambda_2)}^2 + \frac{w_i}{(\lambda)} - \frac{w_j}{(\lambda)}^2 - \frac{w_i}{((\lambda_1 + \lambda_2) + (\lambda_1 + \lambda_2)^2} + \frac{w_j}{((\lambda_1 + \lambda_2) + (\lambda_1 + \lambda_2)^2} \right\} \]

Under the regularity conditions, we have the following properties of \( \hat{\gamma} \), under \( H_0 \).

1. There exists a vector \( \hat{\lambda} \) of solutions to the likelihood equations (2.19)-(2.21), which converges in probability to \( \gamma \).

2. \( \sqrt{n}(\hat{\lambda} - \lambda) \) has a limiting multivariate normal distribution with mean 0 and variance-covariance \( \Gamma_{\hat{\lambda} \lambda} \).

3.2.2. The Union-Intersection Test

We derive the union-intersection test for the hypothesis

\[ H_0^* : \Delta_1 = 0, \Delta_2 = 0, \Delta_{12} = 0 \] \hspace{1cm} (2.2.1)

against

\[ H_1^* : \Delta_1 > 0, \Delta_2 > 0, \Delta_{12} > 0 \] \hspace{1cm} (2.2.2)

Define \( \Omega_0 = \{ \Delta : \Delta = 0 \} \) and \( \Omega_1 = \{ \Delta \in \Omega \subset \mathbb{R}^{+3} \} \) such that \( \Omega = \Omega_0 \cup \Omega_1 \). Define also \( a = (a_1, a_2, a_{12}) \) such that \( a \geq 0 \). Assume that \( \Omega \) is positively homogeneous and for each \( \Delta \in \Omega \) and \( \Delta \neq 0 \), let \( \Delta = \frac{a}{\sqrt{n}} \). For a fixed \( a \in \mathbb{R}^{+3} \), construct the parameter subspace

\[ \Omega(a) = \{ \Delta \in \mathbb{R}^{+3} : \Delta = \frac{a}{\sqrt{n}} \} \] \hspace{1cm} (2.2.3)
and construct the alternative hypothesis

\[ H^*_0: \Delta_i = \frac{a_i}{\sqrt{n}}, \quad i = 1, 2, 12. \]  

(2.2.4)

If we define the set \( A \) such that

\[ A = \{ a \in \mathbb{R}^+ \} \]  

(2.2.5)

and due to positive homogeneity of \( \Omega \),

\[ \Omega \subset \bigcup_{a \in A} \Omega(a) \]

Then, according to the UI principle,

\[ l^* = \sup_{a \in A} \log l, \]  

(2.2.6)

where

\[ l = \frac{L(\tilde{\lambda}_i^a, \frac{a}{\sqrt{n}})}{L(\lambda^0, 0)} \]  

(2.2.7)

where \( L(\theta) = L(\lambda, \Delta) \) is given in (2.1.4), \( \tilde{\lambda} = (\tilde{\lambda}_1^a, \tilde{\lambda}_2^a, \tilde{\lambda}_{12}^a) \) and \( \lambda^0 = (\lambda_1^0, \lambda_2^0, \lambda_{12}^0) \) are the vectors of MLE under \( \Omega_i \) and \( \Omega_0 \), respectively. Then we can write the log of (2.2.7) as

\[
\log l = \log L(\tilde{\lambda}_i^a, \frac{a}{\sqrt{n}}) - \log L(\lambda^0, 0)
\]

\[
= \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \delta_{i(l)} \left\{ \log (\tilde{\lambda}_1^a + \frac{w_i}{\sqrt{n}}a_1) + \log (\tilde{\lambda}_2^a + \tilde{\lambda}_{12}^a)ight. \\
+ \frac{w_i}{\sqrt{n}} (a_1 + a_2 + a_{12}) \right\} \\
- \log \left( (\tilde{\lambda}_1^a + \tilde{\lambda}_2^a) + \frac{w_i}{\sqrt{n}} (a_1 + a_2) \right) \left( (\tilde{\lambda}_1^a + \tilde{\lambda}_{12}^a) + \frac{w_i}{\sqrt{n}} (a_1 + a_{12}) \right)
\]

\[
+ \frac{w_i}{\sqrt{n}} (a_1 + a_2 + a_{12}) t_i
\]
\[- \delta_{i(t)} \left\{ \log \lambda_i^0 + \log (\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) - \log (\lambda_1^0 + \lambda_2^0) \right\} \]

\[- (\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) t_i \right\} \]

\[= \sum_{i=1}^{2} \sum_{l=1}^{n} \delta_{i(t)} \left\{ \log \frac{(\lambda_i^0 + \frac{w_{i(t)}}{\sqrt{n}} a_{l_i})}{\lambda_i^0} \right\} \]

\[+ \log \frac{(\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) + \frac{w_i}{\sqrt{n}} (a_{1} + a_{2} + a_{12}) t_i}{(\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0)} \]

\[- \log \left\{ \frac{(\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) + \frac{w_i}{\sqrt{n}} (a_{1} + a_{2})}{(\lambda_1^0 + \lambda_2^0)} \right\} - \left\{ (\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) + \frac{w_i}{\sqrt{n}} (a_{1} + a_{2} + a_{12}) \right\} \]

\[- (\lambda_1^0 + \lambda_2^0 + \lambda_{12}^0) t_i \right\} \]  \hspace{1cm} (2.2.8)

Using matrix notation

\[\tilde{\lambda}_i = \tilde{\lambda}' C_i, \text{ i = 1, 2, 12},\]

where

\[C_1 = (1 0 0)', \; C_2 = (0 1 0)', \; C_{12} = (0 0 1)',\]  \hspace{1cm} (2.2.9)

\[\tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_{12} = \tilde{\lambda}' C,\]  \hspace{1cm} where

\[C = (1 1 1) = C_1 + C_2 + C_{12}\]  \hspace{1cm} (2.2.10)

and

\[\tilde{\lambda}_1 + \tilde{\lambda}_2 = \tilde{\lambda} D,\]

where

\[D = (1 1 0) = C_1 + C_2.\]  \hspace{1cm} (2.2.11)
Similar notation holds for $a_i$, $i = 1, 2, 12$ and $a$. Then, the log of $l$ given in (2.2.8) can be written as

$$\log l = \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \gamma_{i(l)} \right\} \left\{ \log \frac{\lambda^0 + \frac{w_i}{\sqrt{n}} a'}{\lambda^0 C_i} + \log \frac{\tilde{\lambda}^0 + \frac{w_i}{\sqrt{n}} a'}{\tilde{\lambda}^0 C_i} \right\}$$

$$- \log \frac{\tilde{\lambda}^0 + \frac{w_i}{\sqrt{n}} a'}{\tilde{\lambda}^0 D} \right\} = (\tilde{\lambda}^0 + \frac{w_i}{\sqrt{n}} a' - \tilde{\lambda}^0')C$$

Since $\sqrt{n}(\tilde{\lambda}^0 - \lambda^0) = - \Gamma_{i\lambda}^1 I_{\lambda \Delta} a$, the log of $l$ becomes

$$\log l = \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \gamma_{i(l)} \right\} \left\{ \log \frac{\lambda^0 + \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a'}{\lambda^0 C_i} \right\}$$

$$+ \log \frac{\tilde{\lambda}^0 - \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a'}{\tilde{\lambda}^0 C_i}$$

$$- \log \frac{\tilde{\lambda}^0 - \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a'}{\tilde{\lambda}^0 D} \right\} \right\} = \sum_{i=1}^{2} \sum_{l=1}^{n}$$

$$\left\{ \gamma_{i(l)} \right\} \left\{ \log \left(1 + \frac{1}{\sqrt{n}} (\tilde{\lambda}^0 - \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a')C_i \right) \right\}$$

$$+ \log \left(1 + \frac{1}{\sqrt{n}} (\tilde{\lambda}^0 - \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a')C_i \right)$$

$$- \log \left(1 + \frac{1}{\sqrt{n}} (\tilde{\lambda}^0 - \frac{1}{\sqrt{n}} a' I_{\Delta \lambda} \Gamma_{i\lambda}^1 + \frac{w_i}{\sqrt{n}} a')D \right) \right\} \right\}$$
\[-\frac{1}{\sqrt{n}} \left( \vec{\lambda}' - \frac{1}{\sqrt{n}} \mathbf{a}' \mathbf{I}_{\Delta \mathbf{\lambda}} \mathbf{I}_{\Delta \mathbf{\lambda}}^\top + \frac{w_i}{\sqrt{n}} \mathbf{a}' - \vec{\lambda}' \right) \mathbf{C} \right] \}

Using Taylor's expansion up to the second term for \( \log(1+x) \), we can write the expression above as

\[
\log l = \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \delta_{i(l)} \left\{ \frac{1}{\sqrt{n}} \frac{B'_{i} C_i}{\vec{\lambda}' C_i} - \frac{1}{2n} \left( \frac{B'_{i} C_i}{\vec{\lambda}' C_i} \right)^2 + \frac{1}{\sqrt{n}} \frac{B'_{i} C}{\vec{\lambda}' C} \right. \\
- \frac{1}{2n} \left( \frac{B'_{i} C}{\vec{\lambda}' C} \right)^2 - \frac{1}{\sqrt{n}} \frac{B'_{i} D}{\vec{\lambda}' D} - \frac{1}{2n} \left( \frac{B'_{i} D}{\vec{\lambda}' D} \right)^2 \left. \right\} - \frac{1}{\sqrt{n}} \mathbf{B}' \mathbf{C} \right\}
\]

where

\[
B' = \mathbf{a}' w_i - \mathbf{a}' \mathbf{I}_{\Delta \mathbf{\lambda}} \mathbf{I}_{\Delta \mathbf{\lambda}}^\top = B'_1 - B'_2, \quad (2.2.12)
\]

\[
B'_1 = \mathbf{a}' w_i \quad \text{and} \quad B'_2 = \mathbf{a}' \mathbf{I}_{\Delta \mathbf{\lambda}} \mathbf{I}_{\Delta \mathbf{\lambda}}^\top. \quad (2.2.13)
\]

Rearranging the expression above, we have

\[
\log l = \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \delta_{i(l)} \left\{ \frac{B'_1 C_i}{\vec{\lambda}' C_i} + \frac{B'_1 C}{\vec{\lambda}' C} - \frac{B'_1 D}{\vec{\lambda}' D} \right. \right. \\
- \frac{1}{\sqrt{n}} \mathbf{B}' \mathbf{C} \right\}
\]

\[
- \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \delta_{i(l)} \left\{ \frac{B'_2 C_i}{\vec{\lambda}' C_i} + \frac{B'_2 C}{\vec{\lambda}' C} - \frac{B'_2 D}{\vec{\lambda}' D} \right. \right. \\
- \frac{1}{\sqrt{n}} \mathbf{B}' \mathbf{C} \right\}
\]

\[
- \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{2n} \left[ \left( \frac{B'_1 C_i}{\vec{\lambda}' C_i} \right)^2 + \left( \frac{B'_1 C}{\vec{\lambda}' C} \right)^2 - \left( \frac{B'_1 D}{\vec{\lambda}' D} \right)^2 \right] \right. \\
\left. - \frac{1}{\sqrt{n}} \mathbf{B}' \mathbf{C} \right\}
\]

\[
+ \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{n} \left[ \left( \frac{B'_1 C_i C'_2}{\vec{\lambda}' C_i} \right)^2 + \left( \frac{B'_1 C C'_2}{\vec{\lambda}' C} \right)^2 - \left( \frac{B'_1 D D'B'_2}{\vec{\lambda}' D} \right)^2 \right] \right. \\
\left. - \frac{1}{\sqrt{n}} \mathbf{B}' \mathbf{C} \right\}
\]
\[ \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{2n} \left[ \frac{(B_i^2 C_i)^2}{(\lambda^{0^i} C_i)^2} + \frac{(B_i^2 C)^2}{(\lambda^{0^l} C)^2} - \frac{(B_i^2 D)^2}{(\lambda^{0^l} D)^2} \right] \right\} \quad (2.2.14) \]

Each term of the expression above is calculated below. The first term can be calculated as:

\[ \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{4n} \left[ \frac{B_i^2 C_i}{\lambda^{0^i} C_i} + \frac{B_i^1 C}{\lambda^{0^l} C} - \frac{B_i^1 D}{\lambda^{0^l} D} \right] - \frac{1}{\sqrt{n}} B_i^1 C t_i \right\} \]

\[ = \sum_{l=1}^{n} \left\{ \frac{\delta_{1(l)}}{\sqrt{n}} a_1 \left[ \frac{1}{\lambda_1^0} + \frac{1}{\lambda_0^l} - \frac{1}{\lambda_1^0 + \lambda_2^0} \right] \right. \]

\[ + \frac{\delta_{2(l)}}{\sqrt{n}} \left[ \frac{1}{\lambda_0^l} - \frac{1}{\lambda_1^0 + \lambda_2^0} \right] - w_i t_i \right\} \]

\[ + \sum_{l=1}^{n} \left\{ \frac{\delta_{2(l)}}{\sqrt{n}} a_2 \left[ \frac{1}{\lambda_0^l} - \frac{1}{\lambda_1^0 + \lambda_2^0} \right] \right. \]

\[ + \frac{\delta_{2(l)}}{\sqrt{n}} \left[ \frac{1}{\lambda_0^l} + \frac{1}{\lambda_0^l} - \frac{1}{\lambda_1^0 + \lambda_2^0} \right] - w_i t_i \right\} \]

\[ + \frac{a_{12}}{\sqrt{n}} \sum_{l=1}^{n} \left\{ \frac{\delta_{1(l)}}{\lambda_0^l} + \frac{\delta_{2(l)}}{\lambda_0^l} - w_i t_i \right\} \quad (2.2.15) \]

using (2.2.9)-(2.2.13). Using the score functions given in (2.13)-(2.15) the first term (2.2.14) above can be summarized as:

\[ \sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{4n} \left[ \frac{B_i^2 C_i}{\lambda^{0^i} C_i} + \frac{B_i^1 C}{\lambda^{0^l} C} - \frac{B_i^1 D}{\lambda^{0^l} D} \right] - \frac{1}{\sqrt{n}} B_i^1 C t_i \right\} \]
\[= a_1 U_{\Delta_1} + a_2 U_{\Delta_2} + a_{12} U_{\Delta_{12}} = a U_{\Delta} \quad (2.2.16)\]

The second term of \(\log 1\) in \((2.2.14)\) can be developed as

\[
\sum_{i=1}^{\infty} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{\sqrt{n}} \left[ C_i + C_i - \frac{1}{\lambda_0} \right] - \frac{B_2}{\sqrt{n}} C_t \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ \frac{\delta_{1(l)}}{\sqrt{n}} b_1 \left[ \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2} \right] \right. \\
\left. + \frac{\delta_{2(l)}}{\sqrt{n}} \left[ \frac{1}{\lambda_0} - \frac{1}{\lambda_1 + \lambda_2} \right] - t_i \right\}
\]

\[
+ \sum_{i=1}^{n} \left\{ \frac{\delta_{2(l)}}{\sqrt{n}} b_2 \left[ \frac{1}{\lambda_1^0} + \frac{1}{\lambda_2^0} - \frac{1}{\lambda_1^0 + \lambda_2^0} \right] \right. \\
\left. + \frac{\delta_{2(l)}}{\sqrt{n}} \left[ \frac{1}{\lambda_0} - \frac{1}{\lambda_1 + \lambda_2} \right] - t_i \right\}
\]

\[
+ \frac{b_{12}}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{\delta_{1(l)}}{\lambda_0} + \frac{\delta_{2(l)}}{\lambda_0} - t_i \right\} \quad (2.2.17)
\]

where \(B_2\) given in \((2.2.12)\) can be written also as \(B_2 = (b_1, b_2, b_{12})'\). Rearranging the expression above using the score functions \((2.1.13)-(2.1.15)\), we have

\[
\sum_{i=1}^{\infty} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{\sqrt{n}} \left[ C_i + C_i - \frac{1}{\lambda_0} \right] - \frac{B_2}{\sqrt{n}} C_t \right\}
\]

\[
= b_1 U_{\lambda_1} + b_2 U_{\lambda_2} + b_3 U_{\lambda_{12}} = B_2 U_{\lambda} = a' I_{\Delta \lambda} \Gamma_{\lambda \lambda}^{-1} U_{\lambda} \quad (2.2.18)
\]
The third term of $\log I$ in (2.2.13) is calculated below.

\[
\sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{2n} \left[ \frac{(B_i C_i)^2}{(\lambda_0 C_i)^2} + \frac{(B_i C)^2}{(\lambda_0 C)^2} - \frac{(B_i D)^2}{(\lambda_0 D)^2} \right] \right\}
\]

\[
= \frac{1}{2n} \sum_{l=1}^{n} \left\{ \delta_{1(l)} \cdot w_i \left[ \frac{a_1}{(\lambda_1^0)^2} + \frac{(a_1 + a_2 + a_{12})^2}{(\lambda_0)^2} - \frac{(a_1 + a_2)^2}{(\lambda_1^0 + \lambda_2^0)^2} \right] + \delta_{2(l)} \cdot w_i \left[ \frac{a_1}{(\lambda_1^0)^2} + \frac{(a_1 + a_2 + a_{12})^2}{(\lambda_0)^2} - \frac{(a_1 + a_2)^2}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\}
\]

(2.2.19)

Rearranging the expression above and using the elements of (2.1.14), we have

\[
\sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{2n} \left[ \frac{(B_i C_i)^2}{(\lambda_0 C_i)^2} + \frac{(B_i C)^2}{(\lambda_0 C)^2} - \frac{(B_i D)^2}{(\lambda_0 D)^2} \right] \right\}
\]

\[
= \frac{1}{2n} \left\{ a_1 a_2 \Delta_1 \Delta_1 + a_1 a_2 \Delta_1 \Delta_2 + a_1 a_{12} \Delta_1 \Delta_{12} + a_2 a_1 \Delta_2 \Delta_1 + a_2 a_2 \Delta_2 \Delta_2 + a_2 a_{12} \Delta_2 \Delta_{12} + a_{12} a_1 \Delta_{12} \Delta_1 + a_{12} a_2 \Delta_{12} \Delta_2 + a_{12} a_{12} \Delta_{12} \Delta_{12} \right\}
\]

\[
= \frac{1}{2} \left( a_1 a_2 a_{12} \right) I_{\Delta} (a_1 a_2 a_{12})' = \frac{1}{2} a'_{I_{\Delta}} a
\]

(2.2.20)

Using (2.1.11)-(2.1.13) and also (2.2.13)-(2.2.14) the fourth term of $\log I$ in (2.2.13) can be written as above.

\[
\sum_{i=1}^{2} \sum_{l=1}^{n} \left\{ \frac{\delta_{i(l)}}{n} \left[ \frac{(B_i C_i C' B_2)^2}{(\lambda_0 C_i)^2} + \frac{(B_i C C' B_2)^2}{(\lambda_0 C)^2} - \frac{(B_i D D' B_2)^2}{(\lambda_0 D)^2} \right] \right\}
\]
\[
= \frac{1}{n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} w_i^2 \left[ \frac{a_1 b_1}{(\lambda_1^0)^2} + \frac{(a_1+a_2+a_{12})(b_1+b_2+b_{12})}{(\lambda_1^0)^2} - \frac{(a_1+a_2)(b_1+b_2)}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right. \\
+ \delta_{2(i)} w_i^2 \left[ \frac{a_1 b_1}{(\lambda_1^0)^2} + \frac{(a_1+a_2+a_{12})(b_1+b_2+b_{12})}{(\lambda_1^0)^2} - \frac{(a_1+a_2)(b_1+b_2)}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\} \\
= a' I_{\Delta\lambda} B_2 = a' I_{\Delta\lambda} \Gamma_{\lambda\lambda}^1 a 
\]

(2.2.21)

Finally, the fifth term of log \( l \) in (2.2.13) can be calculated similarly to the third term and expressed as

\[
\sum_{i=1}^{n} \left\{ \delta_{i(l)} w_i^2 \left[ \frac{(B_2 C_i)^2}{(\lambda_1^0 C_i)^2} + \frac{(B_2 C_i)^2}{(\lambda_1^0 C_i)^2} - \frac{(B_2 D_i)^2}{(\lambda_1^0 D_i)^2} \right] \right\} \\
= \frac{1}{2n} \sum_{i=1}^{n} \left\{ \delta_{1(i)} w_i^2 \left[ \frac{b_1}{(\lambda_1^0)^2} + \frac{(b_1+b_2+b_{12})^2}{(\lambda_1^0)^2} - \frac{(b_1+b_2)^2}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right. \\
+ \delta_{2(i)} w_i^2 \left[ \frac{b_1}{(\lambda_1^0)^2} + \frac{(b_1+b_2+b_{12})^2}{(\lambda_1^0)^2} - \frac{(b_1+b_2)^2}{(\lambda_1^0 + \lambda_2^0)^2} \right] \right\} \\
= \frac{1}{2} B_2 I_{\lambda\lambda} B_2 = \frac{1}{2} a' I_{\Delta\lambda} \Gamma_{\lambda\lambda}^1 I_{\lambda\Delta} a 
\]

(2.2.22)

Then, using (2.2.16), (2.2.18), (2.2.20), (2.2.21) and (2.2.22), the log of \( l \) from (2.2.13) can be expressed as

\[
\log l = a' U_{\Delta} - a' I_{\Delta\lambda} \Gamma_{\lambda\lambda}^1 U_{\lambda} + a' I_{\Delta\Delta} a + a' I_{\Delta\lambda} \Gamma_{\lambda\lambda}^1 I_{\lambda\Delta} a - \frac{1}{2} a' I_{\Delta\lambda} \Gamma_{\lambda\lambda}^1 I_{\lambda\Delta} a 
\]
\[ = a' U_\Delta + \frac{1}{2} a' I_{\Delta : \gamma} a + o_p(1). \] (2.2.23)

The derivation of the UI test statistic \( l^* \) as defined in (2.2.6) with \( \log l \) as above in (2.2.23), can be proceed in the same lines as in Chapter II, by utilization of the Kuhn-Tucker-Lagrange minimization technique. The test statistic for \( H_0 \) defined in (2.2.1) against the orthant alternative defined in (2.2.2) can be expressed by (2.2.20), in Chapter II.

3.3. Union-Intersection Test for a Logistic Model in Competing Risk Situation.

In this section we study the comparison of two survival distributions from two independent subpopulations in a competing risk situation with two causes of failures and with a logistic model using union-intersection test.

We consider cohort data, where all \( n \) subjects enter the study at the same time (\( y \)) and all are followed over a fixed period of time \( t \). The response variable, failure or not, is observed and also the cause of failure, say \( C_i, i=1, 2, \ldots, k \). Also, a vector of concomitant variables \( z'=(z_1, \ldots, z_p) \), which may be considered as predicting factors of failure and is observed at entry times for each subject and may be also time dependent.

Following Elandt-Johnson (1983), let us denote the conditional probability of failure from cause \( C_i \) over the period \( t \) in the presence of all causes as \( Q_i^*(t|y,z) \) and

\[ \sum_{i=1}^{k} Q_i^*(t|y,z) = F_T(t|y,z) = 1 - S_T(t|y,z), \] (3.1)

where \( S_T(t|y,z) \) is the conditional probability of surviving this period, for a subject whose vector of concomitant variables \( z \) was observed at time of entry. A logistic model over a period \( t \) can be defined as
\[
\log \frac{Q_i^*(t|y,z)}{S_T(t|y,z)} = \gamma_i(t) + \alpha_i y + \beta_i' z
\]  

(3.2)

where \( \gamma_i = (\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{ip}) \) and \( \beta_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ip}) \) are the parameter vectors for \( i = 1, 2, \ldots, k \). Note that \( \gamma \) may be a function of the period \( t \) and \( z \) may be time dependent. From (3.1) we have

\[
S_T(t|y,z) = 1 - \sum_{i=1}^{k} Q_i^*(t|y,z)
\]  

(3.3)

and from (3.2),

\[
\frac{Q_i^*(t|y,z)}{S_T(t|y,z)} = \exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}
\]

Thus,

\[
Q_i^*(t|y,z) = S_T(t|y,z) \exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}.
\]  

(3.4)

Using (3.4) in (3.3), we have

\[
S_T(t|y,z) = 1 - S_T(t|y,z) \sum_{i=1}^{k} \exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}
\]

or

\[
S_T(t|y,z) = \frac{1}{1 + \sum_{i=1}^{k} \exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}}
\]  

(3.5)

Thus, using (3.5) in (3.4),

\[
Q_i^*(t|y,z) = \frac{\exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}}{1 + \sum_{i=1}^{k} \exp\{\gamma_i(t) + \alpha_i y + \beta_i' z\}}
\]  

(3.6)

The likelihood functions for \( n \) subjects in the study may be written as

\[
L(\gamma, \beta) = \prod_{l=1}^{n} \left\{ \prod_{i=1}^{k} \left[ Q_i^*(t_i|y_i,z_i) \right]^{\delta_{i(l)}} \left[ S_T(t_i|y_i,z_i) \right]^{-1} \sum_{i=1}^{k} \delta_{i(l)} \right\}
\]
\[
\prod_{i=1}^{n} \left\{ \prod_{i=1}^{k} \exp\left( \gamma_i(t) + \alpha_i y + \beta_i' z \right) \left[ 1 + \sum_{i=1}^{k} \exp\left( \gamma_i(t) + \alpha_i y + \beta_i' z \right) \right] \right\}^{-1}
\]

(3.7)

where \( \delta_{i(t)} \) is 1 if the \( l \)-th subject fails from \( C_i \) and 0, otherwise.

### 3.3.1. Notation and Assumptions

Consider a follow-up study in which the time of entry is the same for all subjects and they are followed for length of time \( t \). The response variable (dead, alive) and the cause of death (the cause of interest or other causes) are then observed. The vector of concomitant variables \( z \) is then observed at entry time and considered here not dependent on time. Also, a particular variable \( w_i \) indicate which one the two subpopulations the \( l \)-th subject belongs to. Thus, the logistic model (3.1) can be written as

\[
\log \frac{Q_i(t|y,z)}{S_T(t|y,z)} = (\alpha_{0i} + w_i \alpha^*_0) + (\beta_i' + w_i \beta^*_i) z_i
\]

(3.1.1)

where \( \alpha^*_0 \) and \( \beta^*_i \) indicate the difference between the subpopulation 1 and 2, \( \beta^*_i = (\beta^*_{i1}, \beta^*_{i2}, \ldots, \beta^*_{ip}) \), \( i = 1, 2 \), \( z_i = (z_1, z_2, \ldots, z_p) \). The logistic model can be expressed as

\[
\log \frac{Q_i(t|y,z)}{S_T(t|y,z)} = (\alpha_{0i} + \beta_i' z_i) + w_i (\alpha^*_0 + \beta^*_i' z_i) = \gamma_i' x_i + w_i \Delta_i' x_i
\]

(3.1.2)

where \( \gamma_i' = (\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{ip}) \), \( \Delta_i' = (\Delta_{i1}, \Delta_{i2}, \ldots, \Delta_{ip}) \), \( x_i = (1, z_i) \) are \( (p+1) \) vectors. We are interested in testing

\[
H_0: \Delta = 0
\]

(3.1.3)

against the alternative

\[
H_a: \Delta > 0
\]

(3.1.4)
considering \( \gamma \) as a vector of nuisance parameters. The likelihood function (3.7) can be written as

\[
L(\gamma, \Delta) = \prod_{i=1}^{n} \exp\left[\delta_{i(l)}(\gamma_i' + w_i \Delta_i')x_i\right]
\]

\[
\times \prod_{l=1}^{n} \left[1 + \sum_{i=1}^{k} \exp\{\gamma_i' + w_i \Delta_i'x_i\}\right]^{-1}
\]

(3.1.5)

Since \( L(\gamma, \Delta) = \prod_{i=1}^{n} f(\gamma, \Delta, x_i) \), we can write \( f(\gamma, \Delta, x_i) \) as

\[
f(\gamma, \Delta, x_i) = \left\{\prod_{i=1}^{2} \exp\left[\delta_{i(l)}(\gamma_i' + w_i \Delta_i')x_i\right]\left[1 + \sum_{i=1}^{k} \exp\{\gamma_i' + w_i \Delta_i'x_i\}\right]^{-1}\right\}
\]

(3.1.6)

and

\[
\log f(\gamma, \Delta, x_i) = \sum_{i=1}^{2} \delta_{i(l)}(\gamma_i' + w_i \Delta_i')x_i - \log \left[1 + \sum_{i=1}^{k} \exp\{\gamma_i' + w_i \Delta_i'x_i\}\right]^{-1}
\]

(3.1.7)

In order to estimate the parameters by the maximum likelihood method and derive the large sample test for the hypothesis (3.1.3), some regularity conditions have to hold.

1. The first derivatives of \( \log f(\gamma, \Delta, x_i) \) with respect to \( \gamma \) and \( \Delta \) exist. They are:

\[
\frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \gamma_i} = \left\{\delta_{i(l)} x_i - \frac{x_i' \exp[\delta_{i(l)}(\gamma_i' + w_i \Delta_i')x_i]}{1 + \sum_{i=1}^{k} \exp\{\gamma_i' + w_i \Delta_i'x_i\}}\right\}
\]

(3.1.8)

and

\[
\frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \Delta_i} = \left\{\delta_{i(l)} w_i x_i - \frac{w_i x_i' \exp[\delta_{i(l)}(\gamma_i' + w_i \Delta_i')x_i]}{1 + \sum_{i=1}^{k} \exp\{\gamma_i' + w_i \Delta_i'x_i\}}\right\}
\]

(3.1.9)

The second derivatives of \( \log f(\gamma, \Delta, x_i) \) with respect to \( \gamma \) and \( \Delta \) exist and are:
\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_i} = - \left\{ \frac{x_i x_i \exp[(\gamma'_i + w_i \Delta'_i) x_i]}{1 + \sum_{i=1}^{2} \exp(\gamma'_i + w_i \Delta'_i) x_i} \right\}^2 \text{ for } i = 1, 2, \\
- \frac{x_i x_i \exp[(\gamma'_i + w_i \Delta'_i) x_i]}{\left[1 + \sum_{i=1}^{2} \exp(\gamma'_i + w_i \Delta'_i) x_i \right]^2} \text{ for } i = 1, 2.
\]

(3.1.10)

\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_j} = \\
= \frac{x_i x_i \exp[(\gamma'_i + w_i \Delta'_i) x_i] \exp[(\gamma'_j + w_i \Delta'_j) x_i]}{\left[1 + \sum_{i=1}^{2} \exp(\gamma'_i + w_i \Delta'_i) x_i \right]^2} \text{ for } i \neq j, i, j = 1, 2
\]

(3.1.11)

\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \Delta'_i} = w_i \frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_i} \text{ for } i = 1, 2,
\]

(3.1.12)

\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \Delta'_j} = w_i \frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_j} \text{ for } i \neq j, i, j = 1, 2,
\]

(3.1.13)

\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \Delta_i \partial \Delta'_i} = w_i^2 \frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_i} \text{ for } i = 1, 2,
\]

(3.1.14)

\[
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \Delta_i \partial \Delta'_j} = w_i^2 \frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma'_j} \text{ for } i \neq j, i, j = 1, 2,
\]

(3.1.15)
2. Now consider the $f(\beta; x_i)$ as given in (3.1.6). If we write $f(\beta; x_i) = a_i x_i b_i$ with

$$a_i = \prod_{i=1}^{2} \exp\left[\delta_{i(1)}(\gamma'_i + w_i \Delta'_i) x_i\right]$$

$$b_i = \left[1 + \sum_{i=1}^{k} \exp(\gamma'_i + w_i \Delta'_i) x_i\right]^{-1}$$

$$\frac{\partial f(\beta; x_i)}{\partial \beta_{ir}} = a_i \frac{\partial b_i}{\partial \beta_{ir}} + b_i \frac{\partial a_i}{\partial \beta_{ir}}$$

and for a positive $K$, $\left|\frac{\partial f(\beta; x_i)}{\partial \beta_{ir}}\right| \leq a_i \left|\frac{\partial b_i}{\partial \beta_{ir}}\right| + b_i \left|\frac{\partial a_i}{\partial \beta_{ir}}\right| = K < \infty$

which happens for all $i=1, 2$ and $r=1, 2, \ldots, p$. In the same way, we can show that

$$\left|\frac{\partial^2 f(\beta; x_i)}{\partial \beta_{ir} \partial \beta_{js}}\right| \leq K < \infty$$

for $i, j = 1, 2, r, s = 1, 2, \ldots, p$.

Thus, assumption A2 holds.

3. Finally, consider the first derivatives in (3.1.7). Then, for a positive number $\eta$ and $K$, we have

$$\left|\frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \gamma_{ir}}\right|^{2+\eta} \leq \left|\delta_{i(1)}\right|^{2+\eta} + \frac{-x_i \exp(\gamma'_i + w_i \Delta'_i) x_i}{\log \left[1 + \sum_{i=1}^{k} \exp(\gamma'_i + w_i \Delta'_i) x_i\right]}$$

$$\left|\frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \gamma_{ir}}\right|^{2+\eta} = K < \infty$$

since $\gamma > 0$, $\Delta > 0$, $\exp(\gamma'_i + w_i \Delta'_i) x_i > 0$. Then Assumption A3 holds. Consider the vector of efficient scores

$$\bar{U} = (\bar{U}_\gamma, \bar{U}_\Delta),$$

where $\bar{U}_\gamma = (U_{\gamma_1}, U_{\gamma_2})$ and $\bar{U}_\Delta = (U_{\Delta_1}, U_{\Delta_2})$ are $2(p+1)$ vectors such that

$$U_{\gamma_i} = \frac{1}{\sqrt{n}} \frac{\partial \log L(\gamma, \Delta, x_i)}{\partial \gamma_i} \bigg|_{\Delta = 0, \gamma = \tilde{\gamma}}$$
\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \delta_{i(i)} x_i - \frac{x_i' \exp(\gamma_i x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)} \right\} \tag{3.1.16}
\]

\[
U_{\Delta_i} = \frac{1}{\sqrt{n}} \frac{\partial \log L(\gamma, \Delta, x_i)}{\partial \Delta_i} \bigg|_{\Delta = 0, \gamma = \bar{\gamma}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \delta_{i(i)} \omega_i x_i - \frac{\omega_i x_i' \exp(\gamma_i x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)} \right\} \tag{3.1.17}
\]

The sample efficient Fisher information matrix is given by

\[
I(\bar{\gamma}, \bar{\Delta}, x_i)_{4(p+1) \times 4(p+1)} = \begin{bmatrix} I_{\gamma \gamma} & I_{\gamma \Delta} \\ I_{\Delta \gamma} & I_{\Delta \Delta} \end{bmatrix}
\]

where each element (vector) of the matrix above can be calculated using (3.1.10)-(3.1.15):

\[
I_{\gamma_i, \gamma_i} = -\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma_i} \bigg|_{\Delta = 0, \gamma = \bar{\gamma}} = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{x_i' x_i \exp(\gamma_i x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)} - \frac{x_i' x_i \exp(\gamma_i x_i)^2}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)^2} \right\} \text{ for } i = 1, 2, \tag{3.1.18}
\]

\[
I_{\gamma_i, \gamma_j} = -\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \gamma_j} \bigg|_{\Delta = 0, \gamma = \bar{\gamma}} = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{x_i' x_i \exp(\gamma_i x_i) \exp(\gamma_j x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)^2} \right\} \text{ for } i \neq j, i, j = 1, 2 \tag{3.1.19}
\]

\[
I_{\gamma_i, \Delta_i} = -\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \Delta_i} \bigg|_{\Delta = 0, \gamma = \bar{\gamma}} = \]

\[=-\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{x_i' x_i \exp(\gamma_i x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma_i x_i)^2} \right\} \bigg|_{\Delta = 0, \gamma = \bar{\gamma}}
\]
\[ \begin{align*}
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \gamma_i \partial \Delta_j} & \bigg|_{\Delta = 0, \gamma = \tilde{\gamma}} \\
& = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{w_i x_i^2 \exp(\gamma^\prime x_i) \exp(\gamma^\prime x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma^\prime x_i)^2} \right\} \quad \text{for } i \neq j, \ i, j = 1, 2 \\
& = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{w_i^2 x_i^2 \exp(\gamma^\prime x_i) \exp(\gamma^\prime x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma^\prime x_i)^2} \right\} \quad \text{for } i = 1, 2 \\
\frac{\partial^2 \log f(\gamma, \Delta, x_i)}{\partial \Delta_i \partial \Delta_j} & \bigg|_{\Delta = 0, \gamma = \tilde{\gamma}} \\
& = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{w_i x_i^2 \exp(\gamma^\prime x_i) \exp(\gamma^\prime x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma^\prime x_i)^2} \right\} \quad \text{for } i \neq j, \ i, j = 1, 2 \\
& = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{w_i^2 x_i^2 \exp(\gamma^\prime x_i) \exp(\gamma^\prime x_i)}{1 + \sum_{i=1}^{2} \exp(\gamma^\prime x_i)^2} \right\} \quad \text{for } i = 1, 2 \\
\end{align*} \]

The maximum likelihood estimators of \( \gamma \) under \( H_0 \), say \( \tilde{\gamma} \), are found as solutions of the likelihood equations

\[ \sum_{i=1}^{n} \frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \gamma_i} = 0 \quad \text{and} \quad \sum_{i=1}^{n} \frac{\partial \log f(\gamma, \Delta, x_i)}{\partial \Delta_i} = 0, \]

where the derivatives are given in (3.1.8) and (3.1.9), respectively. Under regularity conditions, consistency and asymptotic normality of \( \tilde{\gamma} \) hold:
1. There exists $\tilde{\gamma}$ such that under $H_0$, $\tilde{\gamma}$ converges in probability to $\gamma$.

2. $\sqrt{n} (\tilde{\gamma} - \gamma)$ has a limiting multivariate normal distribution with mean 0 and covariance matrix $I_{\tilde{\gamma} \gamma}$.

3.3.2. The Union-Intersection Test

In this section we derive the union-intersection test for the hypothesis $H_0: \Delta = 0$ against the alternative $H_a: \Delta > 0$. Define $\Omega_0 = \{ \Delta: \Delta = 0 \}$ and $\Omega_1 = \{ \Delta: \Delta \geq 0 \}$ such that $\Omega = \Omega_0 \cup \Omega_1$. Define also $\mathbf{a} = (a_1, a_2)$, $\mathbf{a}_i = (a_{i0}, a_{i1}, \ldots, a_{ip})$, such that $\mathbf{a} \geq 0$. Assume that $\Omega$ is positively homogeneous and for each $\Delta \in \Omega$ let $\Delta = \frac{a}{\sqrt{n}}$. For a fixed $a \in \mathbb{R}^{+3}$ construct

$$\Omega(a) = \{ \Delta: \Delta = \frac{a}{\sqrt{n}} \} \quad (3.2.1)$$

and consider the alternative hypothesis

$$H^*_{a}: \Delta = \frac{a}{\sqrt{n}}. \quad (3.2.2)$$

If we define the set $A$ such that

$$A = \{ a: a \geq 0 \} \quad (3.2.3)$$

we have

$$\Omega \subseteq \bigcup_{a \in A} \Omega(a)$$

since $\Omega$ is positively homogeneous. Thus, according to the UI principle

$$l^* = \sup_{a \in A} \log l \quad (3.2.4)$$

where

$$l_{s} = \frac{L(\tilde{\gamma}, \frac{a}{\sqrt{n}})}{L(\tilde{\gamma}, 0)} \quad (3.2.5)$$

where $\tilde{\gamma}$ is the maximum likelihood estimator of $\gamma$ given $\Delta = \frac{a}{\sqrt{n}}$, $\gamma$ is the maximum likelihood estimator of $\gamma$ given $\Delta = 0$ and $L(\cdot, \cdot)$ is defined in (3.1.5). Then,
\[
\log l = \sum_{i=1}^{n} \left\{ \frac{2}{\sum_{i=1}^{2} \delta_i(i) (\tilde{\gamma}_i' + \frac{w_i}{\sqrt{n}} a_i') x_i} - \log \left[ 1 + \sum_{i=1}^{2} \exp(\tilde{\gamma}_i' + w_i a_i') x_i \right] \right\}
- \sum_{i=1}^{2} \delta_i(i) \tilde{\gamma}_i' x_i + \log \left[ 1 + \sum_{i=1}^{2} \exp(\tilde{\gamma}_i' x_i) \right] \}
= \sum_{i=1}^{n} \left\{ \sum_{i=1}^{2} \delta_i(i) \left[ (\tilde{\gamma}_i' - \tilde{\gamma}_i) x_i' + \frac{w_i}{\sqrt{n}} a_i' x_i \right]
- \log \frac{1 + \sum_{i=1}^{2} \exp(\tilde{\gamma}_i' + w_i a_i') x_i}{1 + \sum_{i=1}^{2} \exp(\tilde{\gamma}_i' x_i)} \right\}
\]

(3.2.6)

Now, consider a 2(p+1) \times (p+1) matrix \( C_i \), which is partitioned into two submatrices (p+1) \times (p+1), \( C_{ij}, j = 1, 2 \) in such way that whenever \( i=j \), the submatrix \( C_{ij} \) is identity matrix and whenever \( i \neq j \), the submatrix is zero. Then, we can write

\[
\tilde{\gamma}_i - \tilde{\gamma}_i = (\tilde{\gamma} - \tilde{\gamma}) C_i
\]

(3.2.7)

Also, we have the relation between \( \tilde{\gamma} \) and \( \tilde{\gamma} \):

\[
\sqrt{n} (\tilde{\gamma}_i - \tilde{\gamma}_1)' = -\Gamma_{\gamma' \gamma} I_{\gamma \Delta} a = -(B_1' B_2')
\]

(3.2.8)

Then, using (3.2.7) and (3.2.8) we have

\[
\sqrt{n} (\tilde{\gamma}_i - \tilde{\gamma}_1)' = \sqrt{n} (\tilde{\gamma} - \tilde{\gamma}) C_i = -\Gamma_{\gamma' \gamma} I_{\gamma \Delta} a C_i = -B'C_i
\]

(3.2.9)

Then, the log of \( l \) in (3.2.6) becomes

\[
\log l = \sum_{i=1}^{n} \left\{ \frac{2}{\sum_{i=1}^{2} \delta_i(i) \left[ \frac{B'_i}{\sqrt{n}} C_i x_i' + \frac{w_i}{\sqrt{n}} a_i' x_i \right]} -
\]


\[
1 + \sum_{i=1}^{2} \exp(\gamma_i' - \frac{B'C_i}{\sqrt{n}} + \frac{w_i a_i'}{\sqrt{n}}) x_i
\]
\[
-\log \left( 1 + \sum_{i=1}^{2} \exp(\gamma_i' x_i) \right) \}
\]

\begin{equation}
(3.2.10)
\end{equation}

Expanding \( \exp(x) \) in a Taylor series up to the second order, we have

\[
\exp \left[ \frac{1}{\sqrt{n}} (w_i a_i' - B'C_i) x_i \right]
\]

\[
= 1 + \frac{1}{\sqrt{n}} (w_i a_i' - B'C_i) x_i
\]

\[
+ \frac{1}{2n} (w_i^2 a_i' x_i a_i - 2w_i a_i' x_i x_i' C_i B + B'C_i x_i x_i' C_i B + o_p(1)).
\]

Using this expansion in the log of \( l_5 \) given in (3.2.10), we have

\[
\log l = \sum_{i=1}^{n} \left\{ \sum_{i=1}^{2} \delta_i(l) \left[ \frac{B_i'}{\sqrt{n}} C_i x_i + \frac{w_i}{\sqrt{n}} a_i' x_i \right] - \log \left( 1 + \sum_{i=1}^{2} \exp(\gamma_i' x_i) \right)
\]

\[
+ \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i a_i' x_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\gamma_i' x_i) B' C_i x_i
\]

\[
+ \frac{1}{2n} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i^2 a_i' x_i x_i' a_i - \frac{1}{n} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i a_i' x_i x_i' C_i B
\]

\[
+ \frac{1}{2n} \sum_{i=1}^{2} \exp(\gamma_i' x_i) B'C_i x_i x_i' C_i B + o_p(1) \}
\]

\[
+ \log \left( 1 + \sum_{i=1}^{2} \exp(\gamma_i' x_i) \right)
\]

\[
\log l_5 = \sum_{i=1}^{n} \left\{ \sum_{i=1}^{2} \delta_i(l) \left[ \frac{B_i'}{\sqrt{n}} C_i x_i' + \frac{w_i}{\sqrt{n}} a_i' x_i \right]
\]

\[
- \log \left( 1 + \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i a_i' x_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\gamma_i' x_i) B' C_i x_i
\]

\[
+ \frac{1}{2n} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i^2 a_i' x_i x_i' a_i - \frac{1}{n} \sum_{i=1}^{2} \exp(\gamma_i' x_i) w_i a_i' x_i x_i' C_i B
\]
\[ + \frac{1}{2n} \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) B' C_i x_i x_i' C_i B \left\{ 1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) \right\} \] (3.2.11)

Expanding \(\log(1+x)\) in a Taylor series, we have

\[ \log 1 = \sum_{i=1}^{n} \left\{ \sum_{i=1}^{2} \delta_{i(l)} \left[ \frac{B'}{\sqrt{n}} C_i x_i + \frac{w_i}{\sqrt{n}} a_i' x_i \right] - \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \left( \exp(\tilde{r}_i x_i) w_i a_i' x_i' x_i' a_i \right) \right\} \]

\[ \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \left( \exp(\tilde{r}_i x_i) w_i a_i' x_i x_i' x_i' a_i \right) \right) \left( 1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) \right) \]

\[ - \frac{1}{n} \sum_{i=1}^{2} \left( \exp(\tilde{r}_i x_i) w_i a_i' x_i x_i' C_i B + \frac{1}{2} \exp(\tilde{r}_i x_i) B' C_i x_i' x_i' C_i B \right) \right\} \left( 1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) \right) \]

\[ \left. - \frac{1}{2n} \left( \sum_{i=1}^{2} \left( \exp(\tilde{r}_i x_i) w_i a_i' x_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) B' C_i x_i \right)^2 \right) \right\} + o_p(1). \]

Developing the last term and rearranging the terms, we end up with the following expression:

\[ \log 1 = \sum_{i=1}^{2} \left\{ \delta_{i(l)} \left[ \frac{w_i}{\sqrt{n}} x_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) w_i x_i \right] \right\} \]

\[ - \frac{B'}{\sqrt{n}} C_i \sum_{i=1}^{n} \left[ \delta_{i(l)} x_i' - \frac{\exp(\tilde{r}_i x_i) x_i}{1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i)} \right] \]

\[ - \frac{1}{2n} a_i' \sum_{i=1}^{n} \left[ \frac{\exp(\tilde{r}_i x_i) w_i x_i' x_i' - \exp(2\tilde{r}_i x_i) w_i x_i x_i'}{1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) \left[ 1 + \sum_{i=1}^{2} \exp(\tilde{r}_i x_i) \right]^2} \right] a_i. \]
\[ + \frac{1}{\sqrt{n}} a'_i \left[ \frac{\exp(\gamma'_i x_i) \exp(\gamma'_2 x_i) w_i x_i'}{\left(1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)\right)^2} \right] a_2 \]

\[ - \frac{B'_i C_i}{2n} \sum_{i=1}^{n} \left[ \frac{\exp(\gamma'_i x_i) x_i x'_i}{\left(1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)\right)} - \frac{\exp(2\gamma'_i x_i) x_i x'_i}{1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)} \right] C'_i B \]

\[ + \frac{B'_i C_i}{2n} \sum_{i=1}^{n} \left[ \frac{\exp(\gamma'_i x_i) \exp(\gamma'_2 x_i) x_i x'_i}{\left(1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)\right)^2} \right] C'_i B \]

\[ - \frac{a'_i}{\sqrt{n}} \sum_{i=1}^{n} \left[ \frac{\exp(\gamma'_i x_i) \exp(\gamma'_2 x_i) w_i x_i'}{\left(1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)\right)} \right] C'_i B \]

\[ - \frac{a'_i}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{\exp(\gamma'_i x_i) \exp(\gamma'_2 x_i) w_i x_i'}{\left(1 + \sum_{i=1}^{2} \exp(\gamma'_i x_i)\right)^2} \right\} C'_i B + o_p(1). \] (3.2.12)

Therefore, using (3.1.16)-(3.1.23), we have

\[
\log l = \sum_{i=1}^{2} a'_i U_{\Delta_i} - \sum_{i=1}^{2} B' C_i U_{\gamma_i} - \frac{1}{2} \sum_{i=1}^{2} a'_i I_{\Delta_i \Delta_i} a_i - \frac{1}{2} a'_i I_{\Delta_2 \Delta_2} a_i \\
- \frac{1}{2} \sum_{i=1}^{2} B'C_i I_{\gamma_i \gamma_i} C'_i B - B'C'_i I_{\gamma_1 \gamma_2} C'^2_i B + \frac{2}{2} a'_i I_{\Delta_i \Delta_i} a_i - a'_i I_{\Delta_1 \Delta_2} C_2 B \\
- a'_2 I_{\gamma_2 \Delta_1} C'_2 B = a' U_{\Delta} - a' I_{\Delta \gamma} \Gamma_{\gamma \gamma} U_{\gamma} + \frac{1}{2} a' I_{\Delta \Delta} a - \frac{1}{2} a' I_{\Delta \gamma} \Gamma_{\gamma \gamma} \Gamma_{\gamma \Delta} \\
= a' U_{\Delta} + \frac{1}{2} a' \left( I_{\Delta \Delta} - I_{\Delta \gamma} \Gamma_{\gamma \gamma} I_{\gamma \Delta} \right) a + o_p(1). \]
\[ a'_\Delta U_\Delta + \frac{1}{2} a'_\gamma (I_{\Delta\Delta;\gamma}) a + o_p(1). \] (3.2.13)

The derivation of the UI test statistic \( l^* \) as defined in (3.2.5) with \( \log l \) as above in (3.2.13), can be proceed in the same lines as in Chapter II, by utilization of the Kuhn-Tucker-Lagrange minimization technique. The test statistic for \( H_0: \Delta = 0 \) against the orthant alternative defined in (3.2.2) can be expressed by (2.2.20), in Chapter II.
CHAPTER IV
ORDERS (ORTHANT) ALTERNATIVE TESTS IN A PROPORTIONAL
HAZARD MODEL WITH COMPETING RISK DATA

4.1. Introduction

In this Chapter we study the multisample situation, aiming to compare survival
distributions in a competing risk situation with k causes of failure. Ordered alternatives
are considered and union-intersection test statistic based on partial maximum likelihood
estimators is derived. The asymptotic distribution of the statistic under the null
hypothesis is also studied, as well as the efficiency of the test when compared to one
with an unrestricted alternative.

Section 4.2 present the problem and the hypothesis. Some notations are also
explained in this section. Section 4.3 concerns with the union-intersection test for the
restricted orthant (ordered) alternative. Section 4.4 studies the union-intersection test
statistic for the ordered alternative.

4.2. Problem and Notation

Suppose we have the following situation: r+1 groups of subjects from r+1
subpopulations are followed from a starting point and their survival times (T) are
observed as well as their causes of death C_i, i=1,2,..., k. A vector of covariates z_i=(z_1,
z_2, ..., z_k) is also observed for the i-th individual. The cause-specific hazard function for
the i-th subject having survival time T_i failing from cause C_i is

\[ \lambda_i(t_i | z) = \lambda_{0i}(t) \exp (\gamma_i' + \theta_i' c_i(t_i)) z_i(t_i) \]  

(2.1)
where \( \lambda_\odot(t) \), the nuisance underlying hazard rate, an unknown, arbitrary nonnegative function, \( \beta'=(\gamma' \theta') \) is the \( k \times k \) matrix of unknown parameters, \( \beta'_i=(\gamma'_i \theta'_i) \) is the \( k \) vector of the parameters for the cause \( C_i \), \( i=1, 2, \ldots, k \), \( \gamma'_i=(\gamma'_{i1}, \gamma'_{i2}, \ldots, \gamma'_{ip}) \) and \( \theta'_i=(\theta'_{i1}, \theta'_{i2}, \ldots, \theta'_{ip}) \) where \( \theta'_{ij}=(\theta'_{ij1}, \theta'_{ij2}, \ldots, \theta'_{ijp}) \). The parameter \( \gamma \) applies to the control subgroup and \( \theta_j \) is the differential effect of the \( j \)-th subgroup over the control, \( j=1, 2, \ldots, r \),

\[
c_{i(l)}' = (c_1, c_2, \ldots, c_r)' \tag{2.2}
\]

is an \( r \times p \) matrix such that \( c_j=1 \) if the \( l \)-th subject is from the \( j \)-th subpopulation and \( c_j=0 \) otherwise.

Suppose we are interested in the null hypothesis of no treatment effect, that is,

\[ H_0: \theta = 0 \tag{2.3} \]

against the alternative that none of the treatments is inferior to the control, that is,

\[ H_1: \theta_j \geq 0, j = 1, 2, \ldots, r \tag{2.4} \]

with at least one strict inequality. This situation is a generalization of the problem studied in Chapter II for the two sample case with control and treatment subpopulations. A second hypothesis of interest may be the ordering of the treatment effects when \( H_0 \) may not hold, that is,

\[ H_2: 0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_r \tag{2.5} \]

with at least one strict inequality; this is termed the ordered orthant alternative.
Finally, the ordered alternative

$$H_3: \theta_1 \leq \theta_2 \leq \ldots \leq \theta_r$$  \hspace{1cm} (2.6)

may be of interest.

For the hypothesis $H_0$ and the alternative $H_1$, the cause-specific hazard function was defined in (2.1). For the alternative hypothesis $H_2$ and $H_3$ there is a need to make a transformation which allow us to write these alternatives as the orthant alternative given in (2.4).

Let us consider

$$\Delta_j = \theta_j - \theta_{(j-1)} \quad j = 1, 2, \ldots, r, \theta_0 = 0$$  \hspace{1cm} (2.7)

so that

$$\Delta'_j = (\Delta'_{i1} \quad \Delta'_{i2} \ldots \Delta'_{ik})$$  \hspace{1cm} (2.8)

is a $p \times k$ matrix of the parameters for the $j$-th subpopulation, $j = 1, 2, \ldots, r$.

$\Delta'_{ij} = (\Delta_{ij1}, \Delta_{ij2}, \ldots, \Delta_{ijp})$ is a $p$-vector of the parameters for each cause and each subpopulation. The entire matrix of parameters is an $rp \times k$ matrix $\Delta' = (\Delta'_1 \quad \Delta'_2 \ldots \Delta'_r)$.

$$\begin{bmatrix} \theta_{i1} \\ \theta_{i2} \\ \vdots \\ \theta_{ir} \end{bmatrix}_{rp \times 1} = \begin{bmatrix} \Delta_{i1} \\ \Delta_{i1+\Delta_{i2}} \\ \vdots \\ \Delta_{i1+\Delta_{i2}+\ldots+\Delta_{ir}} \end{bmatrix}$$  \hspace{1cm} (2.9)

$$= \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \Delta_{i1} \\ \Delta_{i2} \\ \vdots \\ \Delta_{ir} \end{bmatrix} = D \Delta_i$$
The hypothesis $H_0: \theta = 0$ and the alternative $H_2$ defined in (2.5) can be written for an orthant test on $\Delta$:

$$H_0: \Delta = 0 \quad \text{against} \quad H_2: \Delta > 0.$$  \hfill (2.10)

The cause-specific hazard function (2.1) can be expressed as

$$\lambda_i(t|z) = \lambda_{0i}(t) \exp(\gamma_i' + \Delta_i' d_{i(t)}) z_{i(t)}$$  \hfill (2.11)

where

$$\Delta_i' = (\Delta_{i1}' \Delta_{i2}' \ldots \Delta_{ir}').$$  \hfill (2.12)

$$d_{i(l)} = \begin{bmatrix} D_{r_pXp} & c_{i(l)} \end{bmatrix} = (d_{1(l)} \ldots d_{r(l)})'$$  \hfill (2.13)

with

$$d_{j(l)} = \sum_{k=j}^{r} c_{ii}, \; j = 1, 2, \ldots, r, \; i = 1, 2, \ldots, n$$  \hfill (2.14)

with $c_{i(l)}$ defined in (2.2)

For the hypothesis $H_0$ defined in (2.3) and the alternative $H_3$ in (2.6), we can also utilize the transformation defined in (2.9). The null hypothesis can be stated as $H_0: \Delta = 0$; however, the alternative hypothesis consider the restriction of $(\Delta_{2}' \Delta_{3}' \ldots \Delta_{r}')$ leaving $\Delta_{1}'$ unspecified. Then, if we define the $(r-1)p \times pr$ matrix

$$A = \begin{bmatrix} 0 & I \end{bmatrix}$$  \hfill (2.15)

with 0 a $(r-1)p \times p$ matrix of zeros and I a $(r-1)p \times (r-1)p$ matrix of ones, we can write the alternative $H_3$ as

$$H_3^* : A\Delta \geq 0, \; \Delta_1 \neq 0$$  \hfill (2.16)
with \( A \) as defined in (2.15).

### 4.3. Union-Intersection Test for Orthant (Ordered) Alternative

In this section we study the union-intersection test statistic based on the p.m.l.e. for the hypothesis \( H_0: \theta = 0 \) against the orthant alternative \( H_1: \theta > 0 \) and against the orthant ordered alternative \( H_2: 0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_r \) with at least one strict inequality.

Let us first consider the hypothesis \( H_0: \theta = 0 \) and the alternative \( H_1: \theta > 0 \). The cause-specific hazard function is defined in (2.1) and the likelihood function can be written as

\[
L(\beta) = \prod_{i=1}^{k} \prod_{l=1}^{n} \left\{ \frac{\exp(\gamma'_{i,l} + \theta'_{i,l}z_{i,l})}{\sum_{u \in R \{t_{i,l}\}} \exp(\gamma'_{i,u} + \theta'_{i,u}z_{i,u})} \right\}^{\delta_{i,l}}
\]  

(3.1)

Now, assume that the \( z \)'s are random vectors independent and identically distributed with some mean \( \mu \) and \( \tau = E(z - \mu)(z - \mu)' \) positive definite and \( \det(\tau) < \infty \).

Under the conditions studied in section (2.2.1), in Chapter II, \( \gamma \), the p.m.l.e. of \( \gamma \), under \( H_0 \), has the properties of consistency and asymptotic normality, as we saw in Chapter II in (2.1.16). Also, under a sequence of local alternative \( H_{1n}: \theta = \frac{\delta}{\sqrt{n}} \), \( \hat{\beta} = (\hat{\gamma}, a) \) also has a consistency property and by contiguity it is asymptotically normality. Therefore, the score function

\[
\tilde{U}_{\theta} = \left. \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \theta} \right|_{\theta = 0, \hat{\gamma} = \gamma}
\]

(3.2)

is asymptotically \( N(0, I_{\Delta \Delta: \gamma}) \) under \( H_0 \), and

\[
\hat{U}_{\theta} = \left. \frac{1}{\sqrt{n}} \frac{\partial \log L(\beta)}{\partial \theta} \right|_{\theta = \frac{\delta}{\sqrt{n}}, \hat{\gamma} = \gamma}
\]

(3.3)

is asymptotically \( N(a I_{\Delta \Delta: \gamma}, I_{\Delta \Delta: \gamma}) \) under \( H_{1n} \). \( \hat{U}_{\theta} \) and \( \tilde{U}_{\theta} \) are kpr vectors, and
\( I_{\Delta \gamma} \) is a kprx kpr submatrix from

\[
I(\beta)^{-1} = \begin{bmatrix}
(I_{\gamma \gamma} - I_{\gamma \theta}^1 I_{\theta \gamma}^1 I_{\theta \theta})^{-1} & -\Gamma_{\gamma \gamma : \theta}^1 I_{\gamma \theta}^1 \\
-\Gamma_{\theta \theta}^1 I_{\theta \gamma}^1 & I_{\Delta \gamma}
\end{bmatrix}
\]

(3.4)

where \( \Gamma_{\gamma \gamma : \theta}^1 = (I_{\gamma \gamma} - I_{\gamma \theta}^1 I_{\theta \gamma}^1 I_{\theta \theta})^{-1} \) and \( I_{\gamma \gamma}, I_{\gamma \theta}, I_{\theta \gamma} \) and \( I_{\theta \theta} \) are submatrices of

\[
I(\beta) = \begin{bmatrix}
I_{\gamma \gamma} & I_{\gamma \theta} \\
I_{\theta \gamma} & I_{\gamma \gamma}
\end{bmatrix}
\]

with dimensions kpr x kpr, kp x kpr, kpr x kp and kpr x kpr respectively.

The derivation of the UI test statistic follows the same steps taken in Chapter II for the two sample problem. The test statistic resulting is the same as in (2.2.20), Chapter II, differing only in the dimensions of the vectors and matrices involved. Then

\[
l_2^* = \sum_{\emptyset \subseteq J \subseteq P} \left\{ \left\{ \mathbf{W}_{\Delta(j)}^{*} \Gamma_{\theta \gamma}^{ij} \mathbf{W}_{\Delta(j)}^{*} \right\} I\left\{ \mathbf{W}_{\Delta(j)}^{*} \right\} I\left\{ \Gamma_{\theta \gamma}^{ij} \mathbf{W}_{\Delta(j)} \right\} \right\} (3.5)
\]

where \( P = \{1, 2, ..., kpr\} \) and \( J \) is any subset of \( P \) (there are \( 2^{kpr} \) sets). Likewise in (2.3.1) in Chapter II, the statistic \( l_2^* \) is distributed under \( H_0 \) as a chi-bar, that is,

\[
P\{l_2^* \leq c| H_0\} = \sum_{s=0}^{kpr} w_s P\{\chi_2^2 \leq x\} (3.6)
\]

where \( w_s \) are nonnegative weights with \( \sum_{s=0}^{kpr} w_s = 1 \), \( w_s \) are given by

\[
w_s = \sum_{\{j\}} P\{\mathbf{W}_{\Delta(j)}^{*} > 0| H_0\} P\{\Gamma_{\theta \gamma}^{ij} \mathbf{W}_{\Delta(j)} \leq 0| H_0\} ,
\]

\[\mathbf{W}_{\Delta(j)}^{*} \sim \mathcal{N}(0, (I_{\theta \theta}^{ij})^{-1}) \text{ under } H_0,\]
\( \mathbf{W}_{\Delta(j)} \sim N_{n_0}(0, (\Gamma_{jj})^{-1}) \) under \( H_0 \) and finally

\[ (3.7) \]

\( \chi^2 \) represents the chi-squared random variable with \( s \) d.f.

The asymptotic distribution of \( l^*_0 \) is similar to that in (2.3.4) and the power study also follows the same results as obtained in section (2.4), Chapter II.

Now, let us consider the hypothesis \( H_0: \theta = 0 \) against the ordered orthant alternative \( H_2: 0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_r \). The transformation defined in (2.9) allow us to write the hypothesis \( H_0 \) versus \( H_2 \) in terms of \( \Delta \), that is, \( H_0: \Delta = 0 \) against \( H_2: \Delta > 0 \). Therefore, the cause-specific hazard function could be written as in (2.16) and the likelihood function is similar to that in (3.1) with \( \theta_i \) changed to \( \Delta_i \) and \( w_{i(l)} \) by \( d_{i(l)} \) as defined in (2.18). Then, the UI test statistic for \( H_0: \Delta = 0 \) against \( H_2: 0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_r \) is the same as given in (3.6).

\[ 4.4. \text{ Union-Intersection Test for Ordered Alternative} \]

In this section we study the UI test for \( H_0: \theta = 0 \) against \( H_3: A\Delta > 0 \), with \( A \) as defined in (2.15) and \( \Delta \) as defined in (2.9). Thus, the cause-specific hazard function is that defined in (2.11). The likelihood function, the vector of scores functions and the sample variance-covariance matrix are as defined in (3.1) to (3.5) with \( \Delta \) instead of \( \theta \) and \( d_{i(l)} \) instead of \( c_{i(l)} \).

In the subsequent subsections we present the derivation of the UI test statistic (in subsection 4.4.1), the study of the distribution under \( H_0 \) and \( H_3 \) (in subsection 4.4.2) and the power comparison with the unrestricted test (in subsection 4.4.3).

\[ 4.4.1. \text{ Union-Intersection Test Statistic Derivation} \]

Define the parameter spaces

\[ \Omega_0 = \{ \Delta : \Delta = 0 \} \]

such that
(4.1.2)

\[ \Delta = (\Delta_1 \Delta_2 \ldots \Delta_k) \]

is a \( p \times k \) matrix with \( \Delta'_i = (\Delta'_{i1} \Delta'_{i2} \ldots \Delta'_{ir}) \) a \( p \) vector, \( \Delta'_{ij} = (\Delta_{ij1}, \Delta_{ij2}, \ldots, \Delta_{ijp}) \) and

\[ \Omega_1 = \{ \Delta : A \Delta \geq 0 \} \quad (4.1.3) \]

where \( A \) is defined as in (2.15), \( A = [0 \ 1] \) such that \( \Omega = \Omega_0 \cup \Omega_1 \). Define similarly,

\[ a = (a_1 \ a_2 \ldots \ a_k) \quad (4.1.4) \]

a \( p \times k \) matrix with \( a'_i = (a'_{i1} \ a'_{i2} \ldots \ a'_{ir}) \) a \( p \) vector, \( a'_{ij} = (a_{ij1}, a_{ij2}, \ldots, a_{ijp}) \). We partition \( a \) in the same way with \( Aa \geq 0 \). Let us call \( a = (a^0 \ a^0) \), where \( a^0 = Aa \geq 0 \) is a \((r-1)p \times k \) matrix. Assume \( \Omega \) is positively homogeneous and for each \( \Delta \in \Omega \), let \( \Delta = \frac{a^0}{\sqrt{n}} \).

For a fixed \( a, a^0 \in \mathbb{R}^{krp} \) construct

\[ \Omega(a) = \{ \Delta : A \Delta \geq 0, \Delta \neq 0, \Delta = \frac{a^0}{\sqrt{n}} \} \quad (4.1.5) \]

and consider the alternative hypothesis

\[ H_3^* : A \Delta = \frac{Aa}{\sqrt{n}} = \frac{a^0}{\sqrt{n}}, \Delta \neq 0 \quad (4.1.6) \]

If we define the set \( B \) such that

\[ B = \{ a : Aa \geq 0, a^0 \ \text{unspecified} \} \quad (4.1.7) \]

we have

\[ \Omega \subseteq \bigcup_{a \in B} \Omega(a) \]

since \( \Omega \) is positively homogeneous. Thus, according to the UI principle
\[ l_3^* = \sup_{a \in B} 2 \log l_3 \] (4.1.8)

where

\[ l_3 = \frac{L(\hat{\gamma}, \frac{a}{\sqrt{n}})}{L(\hat{\gamma}, 0)} \] (4.1.9)

where \( \hat{\gamma} \) is the p.m.l.e of \( \gamma \) given \( A \Delta = \frac{\tilde{\theta}_0}{\sqrt{n}} \) and \( \tilde{\gamma} \) is the p.m.l.e of \( \gamma \) given \( \Delta = 0 \).

Thus, the \( l_2 \) resulting is the same as obtained in Chapter II in (2.2.6), that is

\[ \log l_3 = 2a' U_\Delta - a' I_{\Delta \Delta: \gamma} a + o_p(1) \] (4.1.10)

and now the problem is to find \( \sup_{a \in B} \{ 2a' U_\Delta - a' I_{\Delta \Delta: \gamma} a \} \). Then,

\[ l_3^* = - \inf_{Aa \geq 0} \{ 2a' U_\Delta - a' I_{\Delta \Delta: \gamma} a \} \] (4.1.11)

where \( I_{\Delta \Delta: \gamma} \) is defined in (3.4) and \( \tilde{U}_\Delta \) in (3.2), with \( \theta_i \) exchanged for \( \Delta_i \). The function to be minimized is a convex quadratic form in \( a \). The Kuhn-Tucker-Lagrange minimization technique can be applied here in order to minimize \( h(a) = 2a' U_\Delta - a' I_{\Delta \Delta: \gamma} a \) considering \( h_1(a) = -Aa < 0 \) and the equality constraint missing. The solution for the Lagrangean function

\[ L(a, t_1) = h(a) + t_1' h_1(a) \]

\[ = a' I_{\Delta \Delta: \gamma} a - 2a' \tilde{U}_\Delta - t_1' A a \]

is \( (a^*, t_1^*) \) which may satisfy the following assumptions:

a) \( t_1^* \geq 0, \ t_1^* \) is a kp(r-1) vector
b) \( A a \geq 0 \)

c) \( t_i' A a = 0 \)

d) \( \frac{\partial L(a, t_i)}{\partial a} = 2 I_{\Delta\Delta: \gamma} a - 2 \bar{U}_{\Delta} - A t_i' = 0 \) \hspace{1cm} (4.1.12)

Note that the vector of score functions \( \bar{U}_{\Delta} \) is partitioned as

\[
\bar{U}_{\Delta} = (U_{\Delta}^0 \quad \bar{U}_{\Delta}^0)
\] \hspace{1cm} (4.1.13)

where \( \bar{U}_{\Delta} \) is a kpr vector, \( U_{\Delta}^0 \) is a kp vector and \( \bar{U}_{\Delta}^0 \) is a kp\((r-1)\) vector. The variance-covariance matrix \( (kpr \times kpr) I_{\Delta\Delta: \gamma} \) is also partitioned in such way that the last kp\((r-1)\) columns and rows is denoted by \( I_{\Delta\Delta: \gamma}^{(11)} \), that is,

\[
I_{\Delta\Delta: \gamma} = \begin{bmatrix}
I_{\Delta\Delta: \gamma}^{(00)} & I_{\Delta\Delta: \gamma}^{(01)} \\
I_{\Delta\Delta: \gamma}^{(10)} & I_{\Delta\Delta: \gamma}^{(11)}
\end{bmatrix}
\] \hspace{1cm} (4.1.14)

\[
I_{\Delta\Delta: \gamma}^{(1)} = \begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix}
\] \hspace{1cm} (4.1.15)

where

\[
\Sigma_{11} = I_{\Delta\Delta: \gamma}^{(00)} - I_{\Delta\Delta: \gamma}^{(01)} (I_{\Delta\Delta: \gamma}^{(11)})^{-1} I_{\Delta\Delta: \gamma}^{(10)}
\]

\[
\Sigma_{12} = -\Sigma_{11} I_{\Delta\Delta: \gamma}^{(01)} (I_{\Delta\Delta: \gamma}^{(11)})^{-1} = \Sigma_{21}
\]

\[
\Sigma_{22} = \Gamma
\]

From d) in (4.1.12) we have
\[ A' t_1^* = 2 \tilde{U}_\Delta - 2 I_{\Delta \Delta : \gamma} a^* \]

\[ a^* A' t_1^* = 2 a^* \tilde{U}_\Delta - 2 a^* I_{\Delta \Delta : \gamma} a^* = 0 \]

\[ a^* I_{\Delta \Delta : \gamma} a^* = a^* \tilde{U}_\Delta \]

Then, we can write \( h(a) \) as

\[ h(a) = -a^* I_{\Delta \Delta : \gamma} a^* \quad (4.1.16) \]

From d) in (4.12) we can have also

\[ 2 I_{\Delta \Delta : \gamma} a^* = 2 \tilde{U}_\Delta + A' t_1^* \]

\[ a^* = \Gamma_{\Delta \Delta : \gamma}^{1} \tilde{U}_\Delta + \frac{1}{\sqrt{n}} \Gamma_{\Delta \Delta : \gamma}^{1} A' t_1^* \]

\[ A a^* = A \Gamma_{\Delta \Delta : \gamma}^{1} \tilde{U}_\Delta + \frac{1}{2} A \Gamma_{\Delta \Delta : \gamma}^{1} A' t_1^* \quad (4.1.17) \]

If we write \( \Gamma_{\Delta \Delta : \gamma}^{1} \tilde{U}_\Delta = \tilde{W}_\Delta \) and \( A \Gamma_{\Delta \Delta : \gamma}^{1} A' = \Gamma \) we have

\[ A a^* = A \tilde{W}_\Delta + \frac{1}{2} \Gamma t_1^* \]

\[ \tilde{a}^0 = \tilde{W}_\Delta^0 + \frac{1}{2} \Gamma t_1^* \quad (4.1.18) \]

where \( \tilde{W}_\Delta = (\tilde{W}_\Delta^0 \quad \tilde{W}_\Delta^0) \). Since \( A a^* \geq 0 \) and \( t_1^* \geq 0 \) but the inner product is zero, \( \tilde{a}^0 = A a^* \) and \( t_1^* \) can be partitioned as.
\[ \hat{a}^0 = (\hat{a}^0_j, \hat{a}^0_{\bar{j}}) \]

\[ t^*_i = i^0_i = (i^0_j, i^0_{\bar{j}}) \]

in such way that \( i^0_j = 0 \) and \( \hat{a}^0_j = 0 \), where \( j \) is any subset of \( P = 1, 2, \ldots, kp(r-1) \) and \( \bar{j} \) is the complementary subset (\( \emptyset \subseteq \bar{j} \subseteq P \)). Further partitions are necessary:

\[ \tilde{W}_\Delta^0 = (\hat{W}_{\Delta(j)}^0, \tilde{W}_{\Delta(\bar{j})}^0) \quad (4.1.19) \]

\[ \Gamma = \begin{bmatrix} \Gamma_{jj} & \Gamma_{j\bar{j}} \\ \Gamma_{\bar{j}j} & \Gamma_{\bar{j}\bar{j}} \end{bmatrix} \quad (4.1.20) \]

Then, (4.1.15) can be written as

\[ \begin{bmatrix} \hat{a}^0_j \\ \hat{a}^0_{\bar{j}} \end{bmatrix} = \begin{bmatrix} \tilde{W}_{\Delta(j)}^0 \\ \tilde{W}_{\Delta(\bar{j})}^0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Gamma_{jj} & \Gamma_{j\bar{j}} \\ \Gamma_{\bar{j}j} & \Gamma_{\bar{j}\bar{j}} \end{bmatrix} (i^0_j, i^0_{\bar{j}}) \quad (4.1.15) \]

Since \( i^0_j = 0 \) and \( \hat{a}^0_{\bar{j}} = 0 \), we have

\[ \begin{bmatrix} \hat{a}^0_j \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{W}_{\Delta(j)}^0 \\ \tilde{W}_{\Delta(\bar{j})}^0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Gamma_{jj} & \Gamma_{j\bar{j}} \\ \Gamma_{\bar{j}j} & \Gamma_{\bar{j}\bar{j}} \end{bmatrix} (0, i^0_{\bar{j}}) \quad (4.1.21) \]

From (4.1.21) we have

\[ 0 = \tilde{W}_{\Delta(\bar{j})}^0 + \frac{1}{2} \Gamma_{\bar{j}j} i^0_j \]

\[ \frac{1}{2} \Gamma_{j\bar{j}} i^0_j = -\tilde{W}_{\Delta(j)}^0 \]
\[ i^0_j = -2 \Gamma^1_{jj} \dot{W}^0_{\Delta(j)} > 0 \] (4.1.22)

and

\[ \ddot{a}^0_j = \ddot{W}^0_{\Delta(j)} + \frac{1}{2} \Gamma_{jj} \dot{i}^0_j \]

\[ = \ddot{W}^0_{\Delta(j)} + \frac{1}{2} \Gamma_{jj} (-2 \Gamma^1_{jj} \dot{W}^0_{\Delta(j)}) \]

\[ = \ddot{W}^0_{\Delta(j)} - \Gamma_{jj} \Gamma^1_{jj} \dot{W}^0_{\Delta(j)} \]

\[ = \ddot{W}^0_{\Delta(j)} \geq 0 \] (4.1.23)

Then,

\[ t^*_i = \begin{bmatrix} 0 \\ -2 \Gamma^1_{jj} \dot{W}^0_{\Delta(j)} \end{bmatrix} \] (4.2.24)

and

\[ \ddot{a}^* = \begin{bmatrix} \ddot{W}^0_{\Delta(j)} \\ 0 \end{bmatrix} \] (4.2.25)

From (4.1.16), we have

\[ I_{\Delta \Delta; \gamma} a^* = \ddot{U}_{\Delta} + \frac{1}{2} A' t^*_i \]

and using (4.1.14),

\[ \begin{bmatrix} I^{(00)}_{\Delta \Delta; \gamma} & I^{(01)}_{\Delta \Delta; \gamma} \\ I^{(10)}_{\Delta \Delta; \gamma} & I^{(11)}_{\Delta \Delta; \gamma} \end{bmatrix} \begin{bmatrix} a^0 \\ \ddot{a}^0 \end{bmatrix} = \begin{bmatrix} U^0_{\Delta} \\ \ddot{U}^0_{\Delta} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ I \end{bmatrix} t^*_i \] (4.1.26)
Therefore, from (4.1.26) we have

\[
I_{\Delta\Delta: \gamma}^{(0)} a^0 + I_{\Delta\Delta: \gamma}^{(01)} \tilde{a}^0 = U_\Delta^0
\]

\[
a^0 = (I_{\Delta\Delta: \gamma}^{(0)})^{-1} U_\Delta^0 - (I_{\Delta\Delta: \gamma}^{(0)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} \tilde{a}^0 \tag{4.1.27}
\]

\[
a^0 + (I_{\Delta\Delta: \gamma}^{(0)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} \tilde{a}^0 = (I_{\Delta\Delta: \gamma}^{(0)})^{-1} U_\Delta^0
\]

\[
\begin{bmatrix}
1 & (I_{\Delta\Delta: \gamma}^{(0)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} \\
(I_{\Delta\Delta: \gamma}^{(0)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} & I_{\Delta\Delta: \gamma}^{(0)}
\end{bmatrix}
\begin{bmatrix}
a^* \\
\tilde{a}^0
\end{bmatrix}
= (I_{\Delta\Delta: \gamma}^{(0)})^{-1} U_\Delta^0
\]

Now, we have from (4.1.16) that \( h(a) = -a^* I_{\Delta\Delta: \gamma} a^* \) which can be written as

\[
h(a) = -\begin{bmatrix} a^0 & \tilde{a}^0 \end{bmatrix}
\begin{bmatrix}
I_{\Delta\Delta: \gamma}^{(0)} & I_{\Delta\Delta: \gamma}^{(01)} \\
I_{\Delta\Delta: \gamma}^{(0)} & I_{\Delta\Delta: \gamma}^{(01)}
\end{bmatrix}
\begin{bmatrix} a^0 \\
\tilde{a}^0
\end{bmatrix}
\]

and using (4.1.27)

\[
h(a) = -\begin{bmatrix} U_\Delta^{0'} & I_{\Delta\Delta: \gamma}^{(00)} \end{bmatrix} - a^0 I_{\Delta\Delta: \gamma}^{(0)} (I_{\Delta\Delta: \gamma}^{(0)})^{-1} \tilde{a}^0
\]

\[
\times
\begin{bmatrix}
I_{\Delta\Delta: \gamma}^{(00)} & I_{\Delta\Delta: \gamma}^{(01)} \\
I_{\Delta\Delta: \gamma}^{(00)} & I_{\Delta\Delta: \gamma}^{(01)}
\end{bmatrix}
\begin{bmatrix} a^0 \\
\tilde{a}^0
\end{bmatrix}
\]

\[
\times
\begin{bmatrix}
U_\Delta^{0'} (I_{\Delta\Delta: \gamma}^{(00)})^{-1} - a^0 I_{\Delta\Delta: \gamma}^{(10)} (I_{\Delta\Delta: \gamma}^{(00)})^{-1} \\
\tilde{a}^0
\end{bmatrix}
\]

\[
= -\begin{bmatrix} U_\Delta^{0'} U_\Delta^{0'} (I_{\Delta\Delta: \gamma}^{(00)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} \tilde{a}^0 I_{\Delta\Delta: \gamma}^{(0)} (I_{\Delta\Delta: \gamma}^{(00)})^{-1} I_{\Delta\Delta: \gamma}^{(01)} + a^0 I_{\Delta\Delta: \gamma}^{(11)}
\end{bmatrix}
\]
\[
\begin{align*}
&\mathbf{x} = \left[ \begin{array}{c}
U_0' \left( 1^{(00)} \right)^{-1} - \hat{a}^0 \mathbf{I}_{\Delta \Delta;\gamma} \left( 1^{(00)} \right)^{-1} \\
\hat{a}^0 
\end{array} \right] \\
&= - \left( U_0' \left( 1^{(00)} \right)^{-1} U_0 - U_0' \left( 1^{(00)} \right)^{-1} 1^{(01)} \mathbf{I}_{\Delta \Delta;\gamma} \hat{a}^0 \right) \\
&\quad + U_0' \left( 1^{(00)} \right)^{-1} 1^{(01)} \mathbf{I}_{\Delta \Delta;\gamma} \hat{a}^0 - \hat{a}^0 \left( 1^{(10)} \right)^{-1} 1^{(01)} \mathbf{I}_{\Delta \Delta;\gamma} \hat{a}^0 + \hat{a}^0 \left( 1^{(11)} \right) \hat{a}^0 \\
&= - \left( U_0' \left( 1^{(00)} \right)^{-1} U_0 + \hat{a}^0 \left( 1^{(11)} \right) \right) \\
&\quad - \hat{a}^0 \left( 1^{(10)} \right)^{-1} \left( 1^{(01)} \right) \hat{a}^0 \\
&= - \left( U_0' \left( 1^{(00)} \right)^{-1} U_0 + \hat{a}^0 \mathbf{I}^{-1} \hat{a}^0 \right) \\
\end{align*}
\]

since \( \Gamma = \left( 1^{(11)} \right) - 1^{(10)} \left( 1^{(00)} \right)^{-1} 1^{(01)} \mathbf{I}_{\Delta \Delta;\gamma} \left( 1^{(00)} \right)^{-1} \).

Then,
\[
\mathbf{h}(\mathbf{a}^*) = \left( U_0' \left( 1^{(00)} \right)^{-1} U_0 + \left[ \hat{\mathbf{W}}^{0*} \right] \right) \left[ \begin{array}{c}
\Gamma_{jj} \\
\Gamma_{jj} \\
\Gamma_{jj} \\
\Gamma_{jj} 
\end{array} \right]^{-1} \left[ \begin{array}{c}
\hat{\mathbf{W}}_{\Delta(j)}^{0} \\
\Gamma_{jj} \\
\Gamma_{jj} \\
0 
\end{array} \right]
\]

and using
\[
\left[ \begin{array}{cc}
\Gamma_{jj} & \Gamma_{jj} \\
\Gamma_{jj} & \Gamma_{jj} 
\end{array} \right] = \left[ \begin{array}{cc}
(\Gamma_{jj,j})^{-1} & -\Gamma_{jj,j} \Gamma_{jj}(\Gamma_{jj,j})^{-1} \\
(\Gamma_{jj,j})^{-1} \Gamma_{jj} \Gamma_{jj} & (\Gamma_{jj,j})^{-1} 
\end{array} \right],
\]

where \( (\Gamma_{jj,j})^{-1} = (\Gamma_{jj} - \Gamma_{jj,j} \Gamma_{jj} \Gamma_{jj})^{-1} \), (4.29) can be expressed as
\[ h(a^*) = - \left( U_\Delta \left( I_{\Delta \Delta: \gamma}^{(00)} \right)^{-1} U_\Delta + \hat{W}_\Delta^{0*} \right) (\Gamma_{jj;j}^{1})^{-1} \hat{W}_\Delta^{0*} \Delta_{(j)} \]  

(4.1.31)

Since

\[ W_\Delta^{0*} = W_\Delta^{0} - \Sigma_0 \Sigma_1 \hat{W}_\Delta^{0} = (1 - \Sigma_0 \Sigma_1) W_\Delta^{0} = (1 - \Sigma_0 \Sigma_1) \Sigma \ U_\Delta = \Sigma_{00:1} \ U_\Delta \]  

(4.1.32)

and

\[ \Sigma_{00:1} = I_{\Delta \Delta: \gamma}^{(00)} \]

we can write (4.31) as

\[ h(a^*) = \left\{ W_\Delta^{0*} \Sigma_{00:1} \ W_\Delta^{0} + \hat{W}_\Delta^{0*} \ (\Gamma_{jj;j}^{1})^{-1} \hat{W}_\Delta^{0*} \right\} \]  

(4.1.33)

and the union-intersection test statistic can be written as

\[ I_5^* = \sum_{\emptyset \subseteq J \subseteq P} \left\{ W_\Delta^{0*} \Sigma_{00:1} \ W_\Delta^{0} + \hat{W}_\Delta^{0*} \ (\Gamma_{jj;j}^{1})^{-1} \hat{W}_\Delta^{0*} \right\} \]

\[ \times I \left\{ \Gamma_{jj;j}^{-1} \hat{W}_\Delta^{0} \Delta_{(j)} < 0 \right\} \times I \left\{ \hat{W}_\Delta^{0*} \Delta_{(j)} > 0 \right\} \]  

(4.1.34)

### 4.4.2. Distribution of the Test Statistic \( I_5^* \)

In this section we study the distribution of the statistic \( I_5^* \), given in (4.1.34), under the null hypothesis \( H_0 : \Delta = 0 \) and under a sequence of local alternatives \( H_{3n} : A \Delta \Delta_{\Delta \Delta}^{Aa} \), \( Aa > 0 \) as defined in (4.1.6). The distribution of \( I_5^* \) can be written as

\[ P \{ I_5^* \leq c \} = \sum_{\emptyset \subseteq J \subseteq P} P \left\{ \left[ W_\Delta^{0*} \Sigma_{00:1} \ W_\Delta^{0} + \hat{W}_\Delta^{0*} \ (\Gamma_{jj;j}^{1})^{-1} \hat{W}_\Delta^{0*} \right] \leq c, \]

\[ \hat{W}_\Delta^{0*} \Delta_{(j)} > 0, \Gamma_{jj;j}^{-1} \hat{W}_\Delta^{0} \Delta_{(j)} < 0 \right\} \]

(4.2.2)

Now, we want to derive the distribution of \( I_5^* \) in (4.2.2) under the null hypothesis
$H_0$: $\Delta = 0$. We want to show that, under $H_0$, the distribution in (4.2.2) is

$$P\left\{ l_s^* \leq c \mid H_0 \right\} = \sum_{s=0}^{kp(r-1)} w_s \ P\{\chi^2_{s+kp} \leq c\} \quad (4.2.3)$$

where

$$w_s = \sum_{\{j\}} P\{\hat{W}^0_{\Delta(j)} > 0\}, \ P\{\Gamma_{jj}^{-1} \hat{W}^0_{\Delta(j)} < 0\}, \ r(J) = s \quad (4.2.4)$$

(that is, $s$ is the cardinality of $J$), $\sum_{s=0}^{kp(r-1)} w_s = 1$ and $\chi^2_{s+kp}$, $0 \leq j \leq kp(r-1)$ represents a central chi-squared random variable with $s+kp$ degrees of freedom.

First, we need to demonstrate that for each $J$, $\emptyset \subseteq J \subseteq P$, the pair $l(\hat{W}^0_{\Delta(j)} > 0)$ and $l(\Gamma_{jj}^{-1} \hat{W}^0_{\Delta(j)} < 0)$ are asymptotically independent regardless of $H_0$. Let consider a $kp(r-1)$ variate $\hat{W}^0_{\Delta}$ with mean $\mu$ and dispersion matrix $\Gamma$ and partition $\hat{W}^0_{\Delta}$ according to sets $J$ and $\hat{J}$. Let $g(\hat{W}^0_{\Delta}, \mu, \Gamma)$ represent the normal density function of $\hat{W}^0_{\Delta}$. It follows that $g(\hat{W}^0_{\Delta}, \mu, \Gamma)$ factors into

$$g(\hat{W}^0_{\Delta}, \mu, \Gamma) = g(\Gamma_{jj}^{-1} \hat{W}^0_{\Delta(j)}, \Gamma_{jj}^{-1} \mu_j, \Gamma_{jj}^{-1}) \ g(\hat{W}^0_{\Delta(j)}, \mu_j, \Gamma_{jj}, \hat{J}) \quad (4.2.6)$$

Then

$$P\left\{\{\hat{W}^0_{\Delta(j)} > 0\} \cap \{\Gamma_{jj}^{-1} \hat{W}^0_{\Delta(j)} < 0\}\right\} = P\{\hat{W}^0_{\Delta(j)} > 0\} \ P\{\Gamma_{jj}^{-1} \hat{W}^0_{\Delta(j)} < 0\} \quad (4.2.7)$$

We need to show that $\hat{W}^0_{\Delta}$ and $\hat{W}^0_{\Delta(j)}$ are independent. In order to prove that, we need to show that $\hat{W}^0_{\Delta}$ is independent of $\hat{W}^0_{\Delta}$. Using the same argument as in first part, we consider a $kpr$ normal variate $W_{\Delta}$ with mean $\mu$ and dispersion matrix $I_{\Delta\Delta;\gamma}$. $W_{\Delta}$ and $\Gamma_{\Delta\Delta;\gamma}$ are partitioned according to the partition of restricted and unrestricted parameters, that is, according to $W' = (W^0_{\Delta} \ W^0_{\Delta})$ and $\Gamma_{\Delta\Delta;\gamma}$ as defined in (4.1.15).

Then,

$$g(W_{\Delta}, \mu, \Sigma) = g(\hat{W}^0_{\hat{\Delta}}; \mu_0; \Sigma_0) \ g(W^0_{\Delta}; \mu^*; \Sigma_{00};^{-1}).$$
Since $\Gamma^0_{j} \bar W_{\Delta(j)}^0$ is a function of $\bar W_\Delta^0$, we conclude for the independence between $W^0_{\Delta}$ and $\bar W_{\Delta(j)}^0$. Then we have that

$$I\{\left[ W^0_{\Delta}^0 \Sigma_{00:1} W^0_{\Delta}^0 + \bar W_{\Delta(j)}^0 (\Gamma^1_{jj;j})^{-1} \bar W_{\Delta(j)}^0 \right] \leq c\}$$

and $I\{\Gamma^{-1}_{jj} \bar W_{\Delta(j)}^0 < 0\}$ are independent.

Using the same argument as above, we can show that $\bar W_{\Delta}^0$ and $\bar W_{\Delta(j)}^0$ are independent. Therefore,

$$I\{\left[ W^0_{\Delta}^0 \Sigma_{00:1} W^0_{\Delta}^0 + \bar W_{\Delta(j)}^0 (\Gamma^1_{jj;j})^{-1} \bar W_{\Delta(j)}^0 \right] \leq c\} \cdot I\{\Gamma^{-1}_{jj} \bar W_{\Delta(j)}^0 < 0\}$$

and $I\{\bar W_{\Delta(j)}^0 > 0\}$ are mutually independents. Then,

$$P\{l^*_s \leq c\} = \sum_{\emptyset \subseteq J \subseteq P} P\left\{ W^0_{\Delta}^0 \Sigma_{00:1} W^0_{\Delta}^0 + \bar W_{\Delta(j)}^0 (\Gamma^1_{jj;j})^{-1} \bar W_{\Delta(j)}^0 \leq c\right\}$$

$$\times P\{\Gamma^{-1}_{jj} \bar W_{\Delta(j)}^0 < 0\} \cdot P\{\bar W_{\Delta(j)}^0 > 0\}$$

$$= \sum_{\emptyset \subseteq J \subseteq P} P\left\{(\chi^2_{p} + \chi^2_{r(j)}) \leq c\right\} \cdot P\{\Gamma^{-1}_{jj} \bar W_{\Delta(j)}^0 < 0\} \cdot P\{\bar W_{\Delta(j)}^0 > 0\}$$

(4.2.9)

since $\{W^0_{\Delta}^0 \Sigma_{00:1} W^0_{\Delta}^0\}$ is distributed as $\chi^2_p$ and $\{\bar W_{\Delta(j)}^0 (\Gamma^1_{jj;j})^{-1} \bar W_{\Delta(j)}^0\}$ as $\chi^2_{r(j)}$ under $H_0$, $s_j$ is the cardinality of $J$. This follows from the fact that $W^0_{\Delta}^0 \sim N(0, \Sigma_{00:1})$ and $\bar W_{\Delta(j)}^0 \sim N(0, \Gamma_{jj;j})$. Regrouping the $2^{kp(r-1)}$ terms in terms of the cardinality $s \leq kp(r-1)$, we have the distribution of $l^*_s$ under $H_0$ given as in (4.2.3) with $w_s$ given in (4.2.4).

Finally, we want to derive the distribution of $l^*_s$ in (4.2.2) under the local alternative.
\[
\{ H_{3n}: A \Delta = A \frac{\Delta}{\sqrt{n}}, A (r-1)p \times pr = \begin{bmatrix} 0 \\ (r-1) \times pr \end{bmatrix}, \Delta_i \neq 0 \}.
\]

Then, under \( H_{3n} \) there exists \( c > 0 \) such that

\[
P\{ I_3^* \leq c \} = \sum_{\emptyset \subseteq J \subseteq P} \left\{ \int_{C_{1(j)} \cap C_{2(j)}} g(\mathbf{W}_\Delta^{0*}, \mathbf{W}_{\Delta(j)}^{0*})d \mathbf{W}_\Delta^{0*}d \mathbf{W}_{\Delta(j)}^{0*} \int_{C_{3(j)}} dN(\mathbf{\bar{W}}_\Delta^0, \mathbf{\bar{\Delta}}_j^0, \mathbf{\Gamma}_{jj}^0) \right\}
\]

(4.2.10)

where

\[
C_{1(j)} = \left\{ \mathbf{W}_\Delta^{0*} \Sigma_{00:1} \mathbf{W}_\Delta^{0*} + \mathbf{W}_{\Delta(j)}^{0*} (\mathbf{\Gamma}_{jj}^{1})^{-1} \mathbf{W}_{\Delta(j)}^{0*} \leq c \right\},
\]

\[
C_{2(j)} = \left\{ \mathbf{W}_{\Delta(j)}^{0*} > 0 \right\}
\]

\[
C_{3(j)} = \left\{ \mathbf{\Gamma}_{jj}^{1-1} \mathbf{W}_\Delta^0 \mathbf{\Delta}(j) < 0 \right\}.
\]

**Proof:** The events \( \{ \mathbf{W}_{\Delta(j)}^{0*} > 0 \} \) and \( \{ \mathbf{\Gamma}_{jj}^{1-1} \mathbf{W}_\Delta^0 \mathbf{\Delta}(j) < 0 \} \) are independent regardless of whether or not \( \text{E}(\mathbf{W}_\Delta^0) = 0 \). However, the events \( \{ \mathbf{W}_{\Delta(j)}^{0*} (\mathbf{\Gamma}_{jj}^{1})^{-1} \mathbf{W}_{\Delta(j)}^{0*} \leq c \} \) and \( \{ \mathbf{W}_{\Delta(j)}^{0*} > 0 \} \) are not independent when \( \text{E}(\mathbf{W}_\Delta^0) \) is nonnull. Therefore, the distribution of \( I_3^* \) given in (4.2.2), under the alternative \( H_{3n} \) can be written as

\[
P\{ I_3^* \leq c \} = \sum_{\emptyset \subseteq J \subseteq P} P\left\{ \mathbf{W}_\Delta^{0*} \Sigma_{00:1} \mathbf{W}_\Delta^{0*} + \mathbf{W}_{\Delta(j)}^{0*} (\mathbf{\Gamma}_{jj}^{1})^{-1} \mathbf{W}_{\Delta(j)}^{0*} \leq c, \mathbf{W}_{\Delta(j)}^{0*} > 0 \} P\{ \mathbf{\Gamma}_{jj}^{1-1} \mathbf{W}_\Delta^0 \mathbf{\Delta}(j) < 0 \}.
\]

(4.2.11)

The probabilities in (4.2.11), under \( H_{3n} \), are given below:

\[
P\left\{ \mathbf{W}_\Delta^{0*} \Sigma_{00:1} \mathbf{W}_\Delta^{0*} + \mathbf{W}_{\Delta(j)}^{0*} (\mathbf{\Gamma}_{jj}^{1})^{-1} \mathbf{W}_{\Delta(j)}^{0*} \leq c, \mathbf{W}_{\Delta(j)}^{0*} > 0 \right\}
\]
\[ = \int_{C_{1(j) \cap C_{2(j)}}} g(W_{\Delta}^{0\star}, \hat{W}_{\Delta(j)}^{0\star}) \, dW_{\Delta}^{0\star} \, d\hat{W}_{\Delta(j)}^{0\star}, \]  
(4.2.12)

where

\[ C_{1(j)} = \left\{ \left[ W_{\Delta}^{0\star} \Sigma_{00:1} W_{\Delta}^{0\star} + \hat{W}_{\Delta(j)}^{0\star} (\Gamma_{jj,jj}^{1})^{-1} \hat{W}_{\Delta(j)}^{0\star} \right] \leq c \right\}, \]

\[ C_{2(j)} = \left\{ \hat{W}_{\Delta(j)}^{0\star} > 0 \right\}, \]

\[ \hat{W}_{\Delta(j)}^{0\star} \sim \mathcal{N}(\hat{a}_{j}^{0\star}, \Gamma_{jj,jj}^{1}), \text{ where } \hat{a}_{j}^{0\star} = a_{j}^{0} - \Gamma_{jj} \Gamma_{jj}^{-1} \hat{a}_{j}, \]

\[ W_{\Delta}^{0\star} \sim \mathcal{N}(a^{0\star}, \Sigma_{00:1}), \text{ where } a^{0\star} = a^{0} - \Sigma_{01} \Sigma_{11}^{-1} \hat{a}^{0} \]

and finally

\[ g(W_{\Delta}^{0\star}, \hat{W}_{\Delta(j)}^{0\star}) \]

is the joint normal probability density function of two independent multinormal variates

\[ W_{\Delta}^{0\star} \text{ and } \hat{W}_{\Delta(j)}^{0\star}. \]

\[ P\{\Gamma_{jj}^{-1} \hat{W}_{\Delta(j)}^{0} < 0\} = \int_{C_{2(j)}} dN(W_{\Delta(j)}^{0}; \hat{a}_{j}, \Gamma_{jj}) \]  
(4.1.14)

where

\[ C_{3(j)} = \left\{ \Gamma_{jj}^{-1} \hat{W}_{\Delta(j)}^{0} < 0 \right\}, \]

\[ \hat{W}_{\Delta(j)}^{0} \sim \mathcal{N}(\hat{a}_{j}^{0}; \Gamma_{jj}). \]

This follows from the fact that \( \hat{W}_{\Delta} = \Gamma_{\Delta\Delta:\gamma} \hat{U}_{\Delta} \) and \( \hat{U}_{\Delta} \sim \mathcal{N}(a_{\Delta\Delta:\gamma}; I_{\Delta\Delta:\gamma}) \) under \( H_{a} \).
4.4.3. Local Asymptotic Power and Efficiency

In this section we present the study of local asymptotic power of the union-intersection test for the ordered alternative given in (2.16).

From the distribution of the union-intersection test statistic under local alternatives (4.36), the power of the test can be expressed as

\[
\beta_1(a) = P \{ l^*_3 \geq c \}
\]

\[
= \sum_{\emptyset \subseteq J \subseteq P} P \left\{ \bar{W}_{\Delta}^{0*} \Sigma_{00:1} W_{\Delta}^{0*} + \bar{W}_{\Delta(j)}^{0*} (\Gamma_{jj;j}^{-1})^{-1} \bar{W}_{\Delta(j)}^{0*} \geq c, \bar{W}_{\Delta(j)}^{0*} > 0 \right\}
\times P \{ \Gamma_{jj; j}^{-1} W_{\Delta(j)}^{0} < 0 \}
\]

(4.3.1)

Expanding \( \beta_1(a) \) in a Taylor series around \( a=0 \), we have

\[
\beta_1(a) = \alpha + a \cdot \frac{\partial \beta_1(a)}{\partial a} \bigg|_{a=0} + o(|a|)
\]

(4.3.2)

where

\[
\frac{\partial \beta_1(a)}{\partial a} = \sum_{\emptyset \subseteq J \subseteq P} \left\{ P_{H_3} \left[ \bar{W}_{\Delta}^{0*} \Sigma_{00:1} W_{\Delta}^{0*} + \bar{W}_{\Delta(j)}^{0*} (\Gamma_{jj;j}^{-1})^{-1} \bar{W}_{\Delta(j)}^{0*} \right] \geq c, \bar{W}_{\Delta(j)}^{0*} > 0 \right\}
\times \frac{\partial}{\partial a} P_{H_3} \{ \Gamma_{jj; j}^{-1} W_{\Delta(j)}^{0} < 0 \}
\]

(4.3.3)

The two derivatives on the right hand side of (4.3.3) are calculated as follows.

\[
\frac{\partial}{\partial a} P_{H_3} \{ \Gamma_{jj; j}^{-1} W_{\Delta(j)}^{0} < 0 \} = - \frac{\partial}{\partial a} P_{H_3} \{ \Gamma_{jj; j}^{-1} W_{\Delta(j)}^{0} \geq 0 \}
\]
\[
= \left[ \frac{\partial}{\partial \hat{a}} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \} \right] \cdot \left[ \frac{\partial}{\partial \hat{a}} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \} \right]
\]

The first derivative is zero while the second can be partitioned into two parts since 
\[\hat{a}^0 = (\hat{a}^0_j, \hat{a}^\alpha_j).\] Further, \(\frac{\partial}{\partial \hat{a}_j} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \} = 0.\) Therefore,

\[
\frac{\partial}{\partial \hat{a}} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} < 0 \} = \begin{bmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{bmatrix}
\]

where \(\frac{\partial}{\partial \hat{a}^\alpha} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \}\) has k(J) elements zeros and k(\bar{J}) elements.

\[
\frac{\partial}{\partial \hat{a}^\alpha_j} P_{H_3} \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \}
\]

\[
= - \int_{C_{2(j)}} \Gamma^{-1}_{ij} (\hat{W}^0_{\Delta(j)} - \hat{a}^\alpha_j) \cdot dN(\hat{W}^0_{\Delta(j)}; \hat{a}^\alpha_j, \Gamma_{jj})
\]

\[
= - \Gamma^{-1}_{ij} \int_{C_{2(j)}} \hat{W}^0_{\Delta(j)} dN(\hat{W}^0_{\Delta(j)}; \hat{a}^\alpha_j, \Gamma_{jj}) - \hat{a}^\alpha_j \Gamma^{-1}_{jj} P \{ \Gamma^{-1}_{ij} \hat{W}^0_{\Delta(j)} \geq 0 \}.
\]

(4.3.5)

\[
\frac{\partial}{\partial \hat{a}} P_{H_3} \left[ \left[ \hat{W}^0_{\Delta(j)} \Sigma_{00;i} \hat{W}^0_{\Delta(j)} + \hat{W}^0_{\Delta(j)} (\Gamma^1_{ij;j})^{-1} \hat{W}^0_{\Delta(j)} \right] \geq c, \hat{W}^0_{\Delta(j)} > 0 \right]
\]

\[
= \left( \frac{\partial}{\partial \hat{a}} P_{H_3} \left[ \left[ \hat{W}^0_{\Delta(j)} \Sigma_{00;i} \hat{W}^0_{\Delta(j)} + \hat{W}^0_{\Delta(j)} (\Gamma^1_{ij;j})^{-1} \hat{W}^0_{\Delta(j)} \right] \geq c, \hat{W}^0_{\Delta(j)} > 0 \right] \right)
\]

\[
\frac{\partial}{\partial \hat{a}^\alpha_j} P_{H_3} \left[ \left[ \hat{W}^0_{\Delta(j)} \Sigma_{00;i} \hat{W}^0_{\Delta(j)} + \hat{W}^0_{\Delta(j)} (\Gamma^1_{ij;j})^{-1} \hat{W}^0_{\Delta(j)} \right] \geq c, \hat{W}^0_{\Delta(j)} > 0 \right]
\]

\[
= \left( \frac{\partial}{\partial \hat{a}^\alpha_j} \int_{C_{1(j)} \cap C_{2(j)}} g(\hat{W}^0_{\Delta}, \hat{W}^0_{\Delta}) \cdot dW^0_{\Delta} \cdot d\hat{W}^0_{\Delta(j)} \right)
\]

(4.3.6)
The first derivative can be calculated as

\[
\int_{C_{1(j)} \cap C_{2(j)}} \frac{\partial}{\partial \tilde{a}_j} \ g(W_\Delta^{0*}, \bar{W}_\Delta^{0*}(j)) \ dW_\Delta \ d\bar{W}_\Delta^{0*}(j)
\]

whereas the second derivative has \(k(j)\) elements zeros and \(k(j)\) elements calculated as

\[
\int_{C_{1(j)} \cap C_{2(j)}} \frac{\partial}{\partial \tilde{a}_j} \ g(W_\Delta^{0*}, \bar{W}_\Delta^{0*}(j)) \ dW_\Delta \ d\bar{W}_\Delta^{0*}(j) \times \frac{\partial \tilde{a}_j^{0*}}{\tilde{a}_j^{0*}}
\]

In order to evaluate each term of \(\frac{\partial \beta_1(a)}{\partial a}\) at \(a=0\) it is expressed as

\[
\frac{\partial \beta_1(a)}{\partial a} |_{a=0} = \sum_{\emptyset \subset J \subset \mathcal{P} \mathcal{P}} A_1 \begin{bmatrix} 0 \\ A_2 \\ A_5 \end{bmatrix} + A_3 \begin{bmatrix} A_4 \\ A_5 \end{bmatrix} \tag{4.3.8}
\]

where

\[
A_1 = P_{H_0} \left\{ W_\Delta^{0*} \sum_{00:1} W_\Delta^{0*} + \bar{W}_\Delta^{0*}(j) (\Gamma_{j;j}^{-1}) W_\Delta^{0*}(j) \geq c, \bar{W}_\Delta^{0*}(j) > 0 \right\}
\]

\(A_2\) has \(J\) elements zeros and the others \(\tilde{J}\) are given by

\[
A_{22} = \frac{\partial}{\partial \tilde{a}_j} \left. P_{H_0} \left\{ \Gamma_{j;j}^{-1} W_\Delta^{0*}(j) \geq 0 \right\} \right|_{a=0} = -\Gamma_{j;j}^{-1} \int_{C_{2(j)}} W_\Delta^{0*} \ dN(\bar{W}_\Delta^{0*}; 0; \Gamma_{j;j})
\]

from (4.3.5),
\[ A_3 = P_{H_0} \left\{ \Gamma^{-1}_{i;j} \bar{W}^0_{\Delta(j)} < 0 \right\}, \]
\[ A_4 = \frac{\partial}{\partial a} P_{H_0} \left\{ \left[ W^0_{\Delta} \Sigma_{00:1} W^0_{\Delta} + \bar{W}^0_{\Delta(j)} (\Gamma^1_{i;j})^{-1} \bar{W}^0_{\Delta(j)} \right] \geq c, \bar{W}^0_{\Delta(j)} > 0 \right\}_{a=0} \]
\[ = \int_{C_1(j) \cap C_2(j)} \Sigma^{-1}_{00:1} W^0_{\Delta(j)} \, g(W^0_{\Delta}, W^0_{\Delta(j)}) \, d W^0_{\Delta} \, d W^0_{\Delta(j)} \]

\[ A_5 \] also has \( J \) elements zeros and \( J \) elements calculated by

\[ A_{51} = \frac{\partial}{\partial a} P_{H_0} \left\{ \left[ W^0_{\Delta} \Sigma_{00:1} W^0_{\Delta} + \bar{W}^0_{\Delta(j)} (\Gamma^1_{i;j})^{-1} \bar{W}^0_{\Delta(j)} \right] \geq c, \bar{W}^0_{\Delta(j)} > 0 \right\}_{a=0} \]
\[ = \Gamma^1_{i;j} \int_{C_1(j)} \bar{W}^0_{\Delta(j)} \, d N(\bar{W}^0_{\Delta(j)} ; 0 ; \Gamma^1_{i;j}) \text{ from (4.3.6)}. \]

Since \( \bar{W}^0_{\Delta(j)} > 0 \), \( (\Gamma^1_{i;j})^{-1} \) is positive definite, \( \Gamma^{-1}_{i;j} \bar{W}^0_{\Delta(j)} \geq 0 \) then \( \frac{\partial \beta_1(a)}{\partial a} |_{a=0} \geq 0 \).

In a particular case of \( J=P \) and \( J=0 \),

\[ \frac{\partial \beta_1(a)}{\partial a} |_{a=0} = \begin{bmatrix} A_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ A_6 \end{bmatrix} \]

where

\[ A_1 = P_{H_3} \left\{ W^0_{\Delta} \Sigma_{00:1} W^0_{\Delta} \geq c, W^0_{\Delta} \geq 0 \right\}_{a=0} \]
\[ A_6 = \frac{\partial}{\partial a} P_{H_3} \left\{ W^0_{\Delta} \Sigma_{00:1} W^0_{\Delta} \geq c, W^0_{\Delta} \geq 0 \right\}_{a=0} = \]
\[ \left[ \frac{\partial}{\partial a} P_{H_3} \left\{ W^0_{\Delta} \Sigma_{00:1} W^0_{\Delta} \geq c, W^0_{\Delta} \geq 0 \right\} \right] \]

at \( a=0 \). The first derivative can be calculated using
\[ \frac{\partial}{\partial a} H_3 \left\{ W_{0\Delta}^0 \Sigma_{0:1} W_{0\Delta}^0 \geq c, W_{0\Delta}^0 \geq 0 \right\} \]

\[ = \int_{C_4} (W_{0\Delta}^0 - a_0^0) \ dN(W_{0\Delta}^0, a_0^0, \Sigma_{0:1}), \]

whereas the second derivative is zero. The set \( C_4 \) is defined as

\[ C_4 = \left\{ W_{0\Delta}^0 \Sigma_{0:1} W_{0\Delta}^0 \geq c, W_{0\Delta}^0 \geq 0 \right\} \]

Then, \( A_6 = \Gamma \int_{C_4} W_{0\Delta}^0 \ dN(W_{0\Delta}^0; 0, \Gamma) \). We have \( A_4 \geq 0 \) since \( \Sigma_{0:1}^1 \) is positive definite and \( W_{0\Delta}^0 \geq 0 \). Therefore, \( \frac{\partial \beta_1(a)}{\partial a} \mid_{a=0} \) is nonnegative.

We wish to compare the test statistic \( I_3^* \) with the test with unrestricted alternative, that is, when \( H_0: \Delta \neq 0 \). We know that the test statistic is given by

\[ R = W_{\Delta}^0 I_{\Delta: \gamma} W_{\Delta} \]

and the power function is a noncentral chi-squared random variable with \( kpr \) degrees of freedom and noncentrality parameter \( \theta = a' I_{\Delta: \gamma} a \). Following the same lines as in Chapter II, section 3.4, we can express the power function of this test as

\[ \beta_3(\theta) = \alpha + \frac{1}{2} a' I_{\Delta: \gamma} a \ f_{kpr}(W_{\Delta}) + o(|a|^2). \]

Then, using the same argument as in Chapter II, section 3.4, we conclude that \( I_3^* \) is locally more powerful than \( R \).
CHAPTER V
NUMERICAL ILLUSTRATION AND FURTHER RESEARCH

5.1. Introduction

In this chapter we consider a numerical applications of the UI test in a competing risk and a two-sample situation, with a proportional hazard model as studied in Chapter II. Also, discussion of the results and the conclusions of this study are presented here. In Section 5.2 we present the data and we discuss and list some the basic estimates needed for this numerical analysis. In Section 5.3 we present basically the results of the analysis using a proportional hazard model. In Section 5.4 we present some ideas for future research.

5.2. Numerical illustration

The following data, extracted from Lagakos (1977), are from a lung cancer clinical trial being conducted by the Eastern Cooperative Oncology Group at that time. Table 1 summarizes the results of 194 patients with squamous cell carcinoma. Eighty-three patients died with local spread of disease (cause 1: C=1), 44 died with metastatic spread of disease (cause 2: C=2) and 67 were censored (C=0). Two covariates were considered: \( Z_1 = \) performance status \( (\text{ambulatory}=0, \text{non-ambulatory}=1) \), \( Z_2 = \) age in years. A treatment was also considered: \( W = \) treatment \( (W=0, W=1) \). The failure times are indicated by \( T \). Lagakos considered an underlying exponential hazard function for each cause of failure as follows:

\[
\lambda_1(z) = \exp (\alpha + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3)
\]
and

\[ \lambda_2(z) = \exp (\alpha_1 + \beta_2^2 z_1 + \beta_2^2 z_2 + \beta_2^2 z_3) \]

and the estimated the parameters using the full likelihood function. Among other hypotheses, he considered the test of overall effect of the treatments on failure time expressed by \( H_0: \beta_2^1 = \beta_2^2 = 0 \). He utilized was a log-likelihood test and for the data and the particular hypothesis of no effect of treatment on failure times, the result was \( \chi^2 = 2.23 \) with 2 degree of freedom; thus not significant at \( \alpha = 5\% \).

Table 1: Outcome of 194 squamous cell patients

<table>
<thead>
<tr>
<th>Z1</th>
<th>W</th>
<th>Z2</th>
<th>T</th>
<th>C</th>
<th>Z1</th>
<th>W</th>
<th>Z2</th>
<th>T</th>
<th>C</th>
<th>Z1</th>
<th>W</th>
<th>Z2</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>61</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>69</td>
<td>34</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>52</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>60</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>76</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>49</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>52</td>
<td>39</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>54</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>58</td>
<td>24</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>55</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>59</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>43</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>68</td>
<td>12</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>65</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>64</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>51</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>68</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>66</td>
<td>61</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>65</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>47</td>
<td>59</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>48</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>69</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>67</td>
<td>22</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>57</td>
<td>30</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>63</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>44</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>57</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>73</td>
<td>41</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>67</td>
<td>24</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>51</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>49</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>63</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>43</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>46</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>52</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>69</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>60</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>70</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>70</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>12</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>71</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>60</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>54</td>
<td>18</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>55</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>67</td>
<td>40</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>62</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>57</td>
<td>56</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>58</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>49</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>68</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>68</td>
<td>15</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>62</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>67</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>58</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>64</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>68</td>
<td>27</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>53</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>50</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>58</td>
<td>34</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>48</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>60</td>
<td>20</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>64</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>54</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>57</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>39</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>74</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>54</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>74</td>
<td>20</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>39</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>65</td>
<td>14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>62</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>70</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>63</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>60</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>62</td>
<td>30</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>49</td>
<td>1010</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>55</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>57</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 1: Outcome of 194 squamous cell patients (Continued)

<table>
<thead>
<tr>
<th>Z₁</th>
<th>W₁</th>
<th>Z₂</th>
<th>T</th>
<th>C</th>
<th>Z₁</th>
<th>W₁</th>
<th>Z₂</th>
<th>T</th>
<th>C</th>
<th>Z₁</th>
<th>W₁</th>
<th>Z₂</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>62</td>
<td>21</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>52</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>52</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>48</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>64</td>
<td>17</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>66</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>53</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>79</td>
<td>36</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>60</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>58</td>
<td>25</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>65</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>65</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>57</td>
<td>34</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>41</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>44</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>64</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>73</td>
<td>21</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>62</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>43</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>54</td>
<td>78</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>65</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>43</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>57</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>60</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>64</td>
<td>62</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>72</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>55</td>
<td>24</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>42</td>
<td>33</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>70</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>74</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>46</td>
<td>84</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>59</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>58</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>67</td>
<td>16</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>58</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>59</td>
<td>28</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>71</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>49</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>62</td>
<td>29</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>44</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>58</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>66</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>48</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>50</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>59</td>
<td>41</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>73</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>56</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>76</td>
<td>34</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>61</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>44</td>
<td>38</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.3 Results

For the proportional hazard model (2.3) defined in Chapter II, we define for the data above \( n=194 \) and \( k=2 \). The parameters are then \( \beta' = (\gamma', \Delta') \), \( \gamma' = (\gamma_{10}, \gamma_{11}, \gamma_{12}) \) and \( \Delta' = (\Delta_{10}, \Delta_{11}, \Delta_{12}) \), and the hypotheses are \( H_0: \Delta = 0 \) against \( H_a: \Delta > 0 \).

Since the data present tied failure times, Breslow’s (1974) likelihood is used instead of the one defined in (1.2), Chapter II. The likelihood is

\[
L(\beta) = \prod_{j=1}^{r} \prod_{i=1}^{k} \left( \frac{\exp(\theta'_i s_{i(j)})}{\sum_{u \in R(t_{i(j)})} \exp(\theta'_i z_u)^{d_j}} \right)
\]

(5.1)

where \( s_{i(j)} \) indicates the sum of covariates corresponding to the \( i \)th cause of failure and \( j \)th failure time, that is \( s_{i(j)} = \sum_j z_{i(j)} \), \( d_j \) is the number of ties in the \( j \)th failure time.

For the study we reparameterize the model above considering
\[ \theta_i = \gamma_i + \Delta_i w_i \]

where \( \gamma'_i = (\gamma_{i0}, \gamma_{i1}, \gamma_{i2})' \) with \( \gamma_{i0} = 0 \) are the nuisance parameters related to performance status and age and \( \Delta'_i = (\Delta_{i0}, \Delta_{i1}, \Delta_{i2})' \) are the parameters related to the difference of the control and treatment groups and \( z' = (1, z_1, z_2)' \), where \( z_1 \) is the variable performance status and \( z_2 \) is age. Since the likelihood can be factored into different causes and there are no common parameters between the two causes, and there is no common parameter between the two causes, two separate analyses can be performed analyzing each cause of failure separately and treating other causes as censoring at that time.

The maximum likelihood estimates of \( \gamma \) under the null hypothesis \( H_0: \Delta = 0 \) are calculated solving the system of equations defined by the derivatives of the log of the likelihood above, similar to those defined in (2.1.1), in Chapter II, with \( \Delta = 0 \) by the iterative method of Newton-Raphson. SAS Proc PHGLM was utilized to calculated these estimators. They are:

\[ \bar{\gamma}' = (\bar{\gamma}_1', \bar{\gamma}_2') \]  

(5.2)

where

<table>
<thead>
<tr>
<th>Cause</th>
<th>Performance status</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local</td>
<td>( \bar{\gamma}_{11} = 0.5783 )</td>
<td>( \bar{\gamma}_{12} = -0.0198 )</td>
</tr>
<tr>
<td>Spread</td>
<td>( \bar{\gamma}_{21} = 0.3282 )</td>
<td>( \bar{\gamma}_{22} = 0.0161 )</td>
</tr>
</tbody>
</table>

The scores functions \( \hat{U}_\Delta \) and the information matrix \( I(\hat{\theta}) \) are calculated using the
likelihood above and are similar to those defined in (2.1.22) and (2.1.23), Chapter II. They are calculated by a program written in FORTRAN (Appendix), resulting in:

$$\tilde{U}_\Delta' = (5.0545 \ 4.1740 \ 265.6550 \ 2.6408 \ 3.4305 \ 170.3039)'$$  \hspace{1cm} (5.3)

$$I(\tilde{\beta}) = \begin{bmatrix} 1_{\gamma\gamma} & 1_{\gamma\Delta} \\ 1_{\Delta\gamma} & 1_{\Delta\Delta} \end{bmatrix}$$  \hspace{1cm} (5.4)

where the submatrices are given below:

$$I_{\gamma\gamma} = \begin{bmatrix} 19.3717 & 41.9168 & 0 & 0 \\ 41.9168 & 7924.4219 & 0 & 0 \\ 0 & 0 & 9.4544 & 25.8006 \\ 0 & 0 & 25.8006 & 3678.1194 \end{bmatrix}$$  \hspace{1cm} (5.5)

$I_{\Delta\Delta}$ is a block diagonal matrix with the following submatrices in the main diagonal (the superscript indicates cause of failure):

$$I^{(1)}_{\Delta\Delta} = \begin{bmatrix} 15.97 & 5.42 & 930.83 \\ 5.42 & 14.96 & 360.99 \\ 930.83 & 360.99 & 59988.43 \end{bmatrix}$$  \hspace{1cm} (5.6)

$$I^{(2)}_{\Delta\Delta} = \begin{bmatrix} 8.00 & 2.32 & 487.66 \\ 2.32 & 6.90 & 162.71 \\ 487.66 & 162.71 & 32365.12 \end{bmatrix}$$  \hspace{1cm} (5.7)

Also $I_{\gamma\Delta}$ is a block diagonal matrix with the following matrices in the diagonal:
\[ I^{(1)}_{\Delta \gamma} = \begin{bmatrix} -3.22 & 12.22 & -141.36 \\ 12.22 & 47.85 & 6434.86 \end{bmatrix}, \quad (5.8) \]

\[ I^{(2)}_{\Delta \gamma} = \begin{bmatrix} -1.79 & 5.84 & -84.69 \\ -2.17 & 20.83 & 2515.83 \end{bmatrix}, \quad (5.9) \]

and finally \( I_{\Delta \gamma} \) is a 6x4 submatrix transpose of \( I_{\gamma \Delta} \).

The next objective is to calculate the union-intersection statistic given in (2.2.20).

Chapter II. First, the matrix \( I_{\Delta \Delta : \gamma} \) is calculated using the submatrices above as

\[ I_{\Delta \Delta : \gamma} = I_{\Delta \Delta} - I_{\Delta \gamma} I^{-1}_{\gamma \gamma} I_{\gamma \Delta} \]

which is a block diagonal matrix with the following submatrices in the main diagonal.

\[ I^{(1)}_{\Delta \Delta : \gamma} = \begin{bmatrix} 15.38 & 7.40 & 890.79 \\ 7.40 & 7.19 & 431.77 \\ 890.79 & 431.77 & 53156.6 \end{bmatrix} \]

and

\[ I^{(2)}_{\Delta \Delta : \gamma} = \begin{bmatrix} 7.66 & 3.42 & 469.53 \\ 3.42 & 3.29 & 211.33 \\ 469.53 & 211.33 & 29514.9 \end{bmatrix} \]

The scores \( \bar{W}_{\Delta} \) are calculated as defined in (2.2.11) in Chapter II and result in the following vector:

\[ \bar{W}'_{\Delta} = (1.13 \quad 0.49 \quad -0.018 \quad -0.74 \quad 1.27 \quad 0.0085)' \]

and

\[ \Gamma = \Gamma^{1}_{\Delta \Delta : \gamma} \]
Both $\mathbf{W}_\Delta$ and $\Gamma$ are partitioned according to $\theta \subseteq j \subseteq P$, where $P = \{1, 2, \ldots, p\}$, $p = k \times (r + 1) = 6$, the former as defined in (2.2.12). The vector of scores $\mathbf{W}^*_{\Delta(j)}$ is calculated as defined in (2.2.16) in Chapter II for each $j$, $\emptyset \subseteq j \subseteq P$. Observe that there exist $2^6 = 64$ such partitions and the proposed union-intersection test is calculated only on the term for which both $I(\mathbf{W}^*_{\Delta(j)} > 0)$ and $I(\Gamma_{jj} \mathbf{W}^*_{\Delta(j)} \leq 0)$ are one. The calculations of these two indicators are shown below:

<table>
<thead>
<tr>
<th>$j$</th>
<th>$j$</th>
<th>$\mathbf{W}^*_{\Delta(j)}$</th>
<th>$\Gamma_{jj} \mathbf{W}^*_{\Delta(j)}$</th>
<th>$\mathbf{W}^*_{\Delta(j)}$</th>
<th>$\Gamma_{jj} \mathbf{W}^*_{\Delta(j)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>{2,3,4,5,6}</td>
<td>1.74</td>
<td>-26.95</td>
<td>2.64</td>
<td>(2,3,5)</td>
</tr>
<tr>
<td></td>
<td>[0.33]</td>
<td>3.43</td>
<td>170.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>{2}</td>
<td>{1,3,4,5,6}</td>
<td>0.76</td>
<td>15.15</td>
<td>2.64</td>
<td>(2,3,6)</td>
</tr>
<tr>
<td></td>
<td>[0.058]</td>
<td>3.43</td>
<td>170.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>{3}</td>
<td>{1,2,4,5,6}</td>
<td>0.60</td>
<td>2.02</td>
<td>3.43</td>
<td>(2,4,5)</td>
</tr>
<tr>
<td></td>
<td>[0.0050]</td>
<td>170.3</td>
<td>-57.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>{4}</td>
<td>{1,2,3,5,6}</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>(2,4,6)</td>
</tr>
<tr>
<td></td>
<td>[4.80]</td>
<td></td>
<td></td>
<td>-0.174</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>( \tilde{w}_{\Delta(j)} )</td>
<td>( r_{ij} \tilde{w}_{\Delta(j)} )</td>
<td>j</td>
<td>( \tilde{w}_{\Delta(j)} )</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>----------------------------</td>
<td>----------------------------------</td>
<td>-----</td>
<td>----------------------------</td>
</tr>
</tbody>
</table>
| \{5\} | \{1,2,3,4,6\} | [1.04] | \[
\begin{bmatrix}
5.04 \\
4.17 \\
-0.93 \\
-50.27
\end{bmatrix}
\] | \{2,5,6\} | \{1,3,4\} | [0.58] | [0.76] |
| \{6\} | \{1,2,3,4,5\} | [0.0057] | \[
\begin{bmatrix}
5.05 \\
4.18 \\
265.66 \\
-0.068 \\
2.21
\end{bmatrix}
\] | \{3,4,5\} | \{1,2,6\} | [0.005] | [0.60] |
| \{1,2\} | \{3,4,5,6\} | [0.09] | [0.48] | \[
\begin{bmatrix}
-28.56 \\
2.64 \\
3.43 \\
170.3
\end{bmatrix}
\] | \{3,4,6\} | \{1,2,5\} | [0.005] | [0.60] |
| \{1,3\} | \{2,4,5,6\} | [1.31] | [0.017] | \[
\begin{bmatrix}
1.79 \\
2.64 \\
3.43 \\
170.3
\end{bmatrix}
\] | \{3,5,6\} | \{1,2,4\} | [0.005] | [0.60] |
| \{1,3\} | \{2,4,5,6\} | [1.31] | [0.017] | \[
\begin{bmatrix}
1.79 \\
2.64 \\
3.43 \\
170.3
\end{bmatrix}
\] | \{3,5,6\} | \{1,2,4\} | [0.005] | [0.60] |
| \{1,4\} | \{2,3,5,6\} | [0.32] | [0.34] | \[
\begin{bmatrix}
1.74 \\
-26.95 \\
2.25 \\
8.47
\end{bmatrix}
\] | \{4,5,6\} | \{1,2,3\} | [-0.74] | [5.05] |
| \{1,5\} | \{2,3,4,6\} | [0.32] | [1.05] | \[
\begin{bmatrix}
1.74 \\
-26.95 \\
-0.93 \\
-50.27
\end{bmatrix}
\] | \{1,2,3,4\} | \{5,6\} | [1.13] | [2.25] | [0.35] | [8.47] |
Table 3 (Continued)

<table>
<thead>
<tr>
<th>( j )</th>
<th>( j )</th>
<th>( \hat{W}_{\Delta(j)} )</th>
<th>( \Gamma_{jj} )</th>
<th>( \hat{W}_{\Delta(j)j} )</th>
<th>( j )</th>
<th>( \hat{W}_{\Delta(j)} )</th>
<th>( \Gamma_{jj} )</th>
<th>( \hat{W}_{\Delta(j)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1,6}</td>
<td>{2,3,4,5}</td>
<td>[0.32 \quad 0.006]</td>
<td>[1.74 \quad -26.95 \quad -0.07 \quad 2.21]</td>
<td>{1,2,3,5}</td>
<td>{4,6}</td>
<td>[1.13 \quad 0.50 \quad -0.018 \quad 1.04]</td>
<td>[-0.93 \quad -50.27]</td>
<td></td>
</tr>
<tr>
<td>{2,3}</td>
<td>{1,4,5,6}</td>
<td>[0.55 \quad 0.0006]</td>
<td>[0.51 \quad 2.64 \quad 3.43 \quad 170.3]</td>
<td>{1,2,3,6}</td>
<td>{4,5}</td>
<td>[1.13 \quad 0.50 \quad -0.018 \quad 0.006]</td>
<td>[-0.07 \quad 2.21]</td>
<td></td>
</tr>
<tr>
<td>{2,4}</td>
<td>{1,3,5,6}</td>
<td>[0.58 \quad 0.34]</td>
<td>[0.76 \quad 15.15 \quad 2.25 \quad 8.47]</td>
<td>{1,2,4,5}</td>
<td>{3,6}</td>
<td>[0.098 \quad 0.48 \quad -0.23 \quad 1.28]</td>
<td>[28.56 \quad 6.29]</td>
<td></td>
</tr>
<tr>
<td>{2,5}</td>
<td>{1,3,4,6}</td>
<td>[0.58 \quad 1.04]</td>
<td>[0.76 \quad 15.15 \quad -0.93 \quad -50.27]</td>
<td>{1,2,4,6}</td>
<td>{3,5}</td>
<td>[0.098 \quad 0.48 \quad -0.35 \quad 0.011]</td>
<td>[-28.56 \quad 2.23]</td>
<td></td>
</tr>
<tr>
<td>{2,6}</td>
<td>{1,3,4,5}</td>
<td>[0.58 \quad 0.058]</td>
<td>[0.76 \quad 15.15 \quad -0.07 \quad 2.21]</td>
<td>{1,2,5,6}</td>
<td>{3,4}</td>
<td>[0.098 \quad 0.48 \quad 1.25 \quad -0.003]</td>
<td>[-28.55 \quad -0.14]</td>
<td></td>
</tr>
<tr>
<td>{3,4}</td>
<td>{1,2,5,6}</td>
<td>[0.005 \quad 0.34]</td>
<td>[0.60 \quad 2.02 \quad 2.25 \quad 8.47]</td>
<td>{1,3,4,5}</td>
<td>{2,6}</td>
<td>[1.31 \quad -0.016 \quad -0.23 \quad 1.28]</td>
<td>[1.80 \quad 6.29]</td>
<td></td>
</tr>
<tr>
<td>{3,5}</td>
<td>{1,2,4,6}</td>
<td>[0.005 \quad 1.04]</td>
<td>[0.60 \quad 2.02 \quad -0.93 \quad -50.27]</td>
<td>{1,3,4,6}</td>
<td>{2,5}</td>
<td>[1.31 \quad -0.016 \quad -0.36 \quad 0.011]</td>
<td>[1.80 \quad 2.23]</td>
<td></td>
</tr>
</tbody>
</table>
Table 3 (Continued)

<table>
<thead>
<tr>
<th>j</th>
<th>j</th>
<th>$\hat{W}_{\Delta(j)}^*$</th>
<th>$r_{jj} \hat{W}_{\Delta(j)j}$</th>
<th>$\hat{W}_{\Delta(j)}^*$</th>
<th>$r_{jj} \hat{W}_{\Delta(j)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>{3,6} {1,2,4,5}</td>
<td>[0.005 2.02 0.07 2.21]</td>
<td>{1,3,5,6} {2,4}</td>
<td>1.31 -0.016 1.25 -0.003</td>
<td>1.80 -0.14</td>
<td></td>
</tr>
<tr>
<td>{4,5} {1,2,3,6}</td>
<td>[-0.23 4.17 265.7 6.29]</td>
<td>{1,4,5,6} {2,3}</td>
<td>1.36 -0.74 1.27 0.008</td>
<td>1.74 -26.95</td>
<td></td>
</tr>
<tr>
<td>{4,6} {1,2,3,5}</td>
<td>[-0.36 4.17 265.7 2.23]</td>
<td>{2,3,4,5} {1,6}</td>
<td>0.55 0.0006 -0.23 1.28</td>
<td>0.51 6.29</td>
<td></td>
</tr>
<tr>
<td>{5,6} {1,2,3,4}</td>
<td>[1.25 4.17 265.7 -0.14]</td>
<td>{2,3,4,6} {1,5}</td>
<td>0.55 0.0006 -0.36 0.011</td>
<td>0.51 2.23</td>
<td></td>
</tr>
<tr>
<td>{1,2,3} {4,5,6}</td>
<td>[5.24 2.64 3.43 170.3]</td>
<td>{2,3,5,6} {1,4}</td>
<td>0.55 0.0006 1.25 -0.003</td>
<td>0.51 -0.14</td>
<td></td>
</tr>
<tr>
<td>{1,2,4} {3,5,6}</td>
<td>[0.09 -28.56 0.50 2.25]</td>
<td>{2,4,5,6} {1,3}</td>
<td>0.58 -0.74 1.27 0.008</td>
<td>0.76 15.15</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>$\mathbf{W}^*$</td>
<td>$\mathbf{r}<em>{jj} \mathbf{W}</em>{\Delta(j)j}$</td>
<td>$\mathbf{W}^*$</td>
<td>$\mathbf{r}<em>{jj} \mathbf{W}</em>{\Delta(j)}$</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>---------------</td>
<td>---------------------------------</td>
<td>---------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td></td>
<td>(1,2,5)</td>
<td>(3,4,6)</td>
<td>0.097</td>
<td>-28.56</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.48</td>
<td>-0.93</td>
<td>-0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.04</td>
<td>-50.27</td>
<td>1.27</td>
</tr>
<tr>
<td>(1,2,6)</td>
<td>(3,4,5)</td>
<td></td>
<td>0.097</td>
<td>-28.56</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.48</td>
<td>-0.07</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.005</td>
<td>2.21</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.28</td>
</tr>
<tr>
<td>(1,3,4)</td>
<td>(2,5,6)</td>
<td></td>
<td>1.31</td>
<td>1.80</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.017</td>
<td>2.25</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.34</td>
<td>8.47</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.011</td>
</tr>
<tr>
<td>(1,3,5)</td>
<td>(2,4,6)</td>
<td></td>
<td>1.31</td>
<td>1.80</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.017</td>
<td>-0.93</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.05</td>
<td>-50.27</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.003</td>
</tr>
<tr>
<td>(1,3,6)</td>
<td>(2,4,5)</td>
<td></td>
<td>1.31</td>
<td>1.80</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.017</td>
<td>-0.07</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.006</td>
<td>2.21</td>
<td>-0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.008</td>
</tr>
<tr>
<td>(2,3,4)</td>
<td>(1,5,6)</td>
<td></td>
<td>0.061</td>
<td>0.17</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.068</td>
<td>50.83</td>
<td>-0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.80</td>
<td>-50.35</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.008</td>
</tr>
</tbody>
</table>
The union-intersection test statistic as given in (2.2.20) is calculated using only the term $J = \{1,2,5\}$ since the terms $I\{W_{\Delta(j)}^* > 0\}$ and $I\{r_{jj} W_{\Delta(j)} < 0\}$ are both equal to 1. Thus the test resulted $l_1^* = 6.07676$.

The next step is to calculate the P-value given in (2.3.1), Chapter II. The problem is to calculate the weights $w_k$, where $k$ is the cardinality of $J$, in our case, $k=1, \ldots, 6$. For $k$ up to 4, the weights of the chi-bar distribution can be calculated by formulas. For instance, Cramer (1946, page 290) give the expression for calculating the probability of both variates from a bivariate normal distribution being nonnegative ($J = \{1,2\}$) as

$$P(W > 0) = \int_{-\infty}^{0} \int_{-\infty}^{0} f(x,y) \, dx \, dy = \frac{1}{4} + \frac{1}{2n} \arcsin \rho,$$

where $\rho$ is the correlation between the two variates. The probability of both being
negative $P\{\Gamma_{jj} \bar{W}_j \leq 0\}$ is the same as above by symmetry. The probability of the first being positive and second being negative ($J=\{1\}$) is given by $\frac{1}{4} - \frac{1}{2n} \sin^{-1}(\rho)$. For $k=3$ and $k=4$, Barlow et al. (1972, page 134-142) present a summary discussion of the available methods for calculating the weights of the chi-bar distribution when $k \leq 4$.

For $k > 4$ the calculation of the weights has to be performed by utilizing simulation in a computer. For $\bar{W} \sim N_w (0, \Gamma_{\Delta} \Delta^{-1})$, with $w=6$, we generate a few thousand multivariate normal replicates and calculate $\bar{W}^*_\Delta$ and $\Gamma_{jj} \bar{W}_j$ for each normal replicate in order to estimate the orthant probabilities for these variables. The estimates are obtained by calculating the proportions which lie in the orthant. It was necessary to generate 5000 replicates to obtain the sum of the weights equal to one. The resulting weights were:

<table>
<thead>
<tr>
<th>k</th>
<th>weights $w_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1273</td>
</tr>
<tr>
<td>1</td>
<td>0.352209</td>
</tr>
<tr>
<td>2</td>
<td>0.346452</td>
</tr>
<tr>
<td>3</td>
<td>0.146520</td>
</tr>
<tr>
<td>4</td>
<td>0.026784</td>
</tr>
<tr>
<td>5</td>
<td>0.001190</td>
</tr>
<tr>
<td>6</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

Thus, the $P$-value is calculated as

$$P=1-\{1 \times 0.1273 + 0.986303 \times 0.352209 + 0.952088 \times 0.346452$$
$$+ 0.892065 \times 0.146520 + 0.806511 \times 0.026784 + 0.701179$$
$$\times 0.001190 + 0.585353 \times 0.00000 = 0.0418346$$

For the unrestricted alternative the score test resulted
\[ Q = \hat{W} \mathbf{I}_{\Delta \Delta: \gamma} \hat{W} = \bar{U} \mathbf{I}_{\Delta \Delta: \gamma} \bar{U} = 6.85497 \]

and the P-value was \( P = 0.335 \).

For the next step, the asymptotic power of the two tests was investigated. Since in the study of the power comparisons in Chapter II, section 2.4, the expression for the slope of the power curve of the proposed test is not tractable, we supplement the study with a numerical study. Utilizing the first 1000 from the 5000 six-variate normal replicates generated for calculating the weights, we calculate new vector by adding each time a combination of nonnegative numbers for each vector of the scores defined as \( \hat{W} \). The value of the union-intersection test statistic \( l_1^* \) and \( Q \) and their P-values are calculated and the proportions of times they exceed their respective critical value are observed.

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( l_1^* )</th>
<th>( Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1 0.0 0.0 0.1 0.0 0.0)</td>
<td>0.101</td>
<td>0.056</td>
</tr>
<tr>
<td>(0.3 0.0 0.0 0.3 0.0 0.0)</td>
<td>0.311</td>
<td>0.142</td>
</tr>
<tr>
<td>(0.5 0.0 0.0 0.5 0.0 0.0)</td>
<td>0.801</td>
<td>0.472</td>
</tr>
<tr>
<td>(1.0 0.0 0.0 1.0 0.0 0.0)</td>
<td>0.993</td>
<td>0.966</td>
</tr>
<tr>
<td>(1.5 0.0 0.0 1.5 0.0 0.0)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(0.0 0.1 0.0 0.0 0.1 0.0)</td>
<td>0.077</td>
<td>0.053</td>
</tr>
<tr>
<td>(0.0 0.3 0.0 0.0 0.3 0.0)</td>
<td>0.180</td>
<td>0.083</td>
</tr>
<tr>
<td>(0.0 0.5 0.0 0.0 0.5 0.0)</td>
<td>0.381</td>
<td>0.173</td>
</tr>
<tr>
<td>(0.0 1.0 0.0 0.0 1.0 0.0)</td>
<td>0.883</td>
<td>0.681</td>
</tr>
<tr>
<td>(0.0 1.5 0.0 0.0 1.5 0.0)</td>
<td>0.998</td>
<td>0.967</td>
</tr>
<tr>
<td>(0.0 2.0 0.0 0.0 2.0 0.0)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(0.0 0.0 0.01 0.0 0.0 0.01)</td>
<td>0.808</td>
<td>0.545</td>
</tr>
<tr>
<td>(0.0 0.0 0.1 0.0 0.0 0.1)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(0.1 0.1 0.0 0.1 0.1 0.0)</td>
<td>0.151</td>
<td>0.062</td>
</tr>
<tr>
<td>(0.3 0.3 0.0 0.3 0.3 0.0)</td>
<td>0.632</td>
<td>0.331</td>
</tr>
<tr>
<td>(0.5 0.5 0.0 0.5 0.5 0.0)</td>
<td>0.963</td>
<td>0.804</td>
</tr>
</tbody>
</table>
The proportions of rejections from $l_1^*$ are nearly twice than the proportions of rejections from Q in almost all combinations of $\Delta$ studied. Therefore, for $H_0: \Delta = 0$, the test statistic $l_1^*$ is more powerful than the test Q, that is, the test with restricted alternative is more powerful than the test with unrestricted alternative.

5.4. Suggestions for Further Research

In this dissertation, we have considered some comparative studies of two subpopulations in a competing risk situation using some well known analysis models, such as semiparametric and parametric model. Alternative models for such problems have also been considered. Extensions to more than two subpopulations have been studied. We have shown that a restricted alternative test is more powerfull than the corresponding unrestricted alternative test. There is a need for more research in this area, exploring both the applied and theoretical aspects. Following is a synopsis of further research we intend to undertake in the future:
1. Numerical comparisons of two survival functions related to the bivariate exponential model (Sarkar, 1987) and logistic model in a competing risk setting with two causes of failure are intended for future work. Also, numerical application for case of more than two subpopulations in a competing risk situation should be performed, using a proportional hazard model.

2. Derive a union-intersection test statistic for comparisons of survival functions of two subpopulations in a parametric regression model using full likelihood and considering a exponential family of distributions for the underlying survival distributions. Efficiency of proportional hazard regression model as compared with parametric regressions model can be of interest in assessing how semiparametric likelihood model behave compared with full likelihood model.

3. Derive union-intersection test statistic in the semiparametric Cox model as in Chapter II, for other hypotheses, in a competing risk situation with two causes of failure.

4. The estimation of the parameter in the case of rejection of the null hypothesis may be of interest. This problem is called "estimation after preliminary tests" and has been investigated in many studies. Basically, this study consists on the use of different estimation procedures for the parameter in question for the cases of significant or not significant parameters.
REFERENCES


APPENDIX

FORTRAN Program for calculating the Scores Functions $\hat{U}_\Delta$ and the Information Matrix $I(\hat{\delta})$ using Proportional Hazard Function as given in (2.2), Chapter II and using Breslow's Likelihood for Tied Data.

```
C CALCULATING $\hat{U}_\Delta$ AND $I(\hat{\delta})$
C
C X(1) PERFORMANCE STATUS
C X(2) AGE IN YEARS
C W TREATMENT
C T FAILURE TIME
C C CAUSE OF FAILURE
C X(3) TREATMENT
C X(4) PERFORMANCE STATUS X TREATMENT
C X(5) AGE X TREATMENT
C G(I) ESTIMATES OF $\gamma$ UNDER $H_0$
C DEL1 INDICATOR FOR CAUSE 1
C DEL2 INDICATOR FOR CAUSE 2
C T1 DESCENDING ORDERED FAILURE TIME FOR CAUSE 1

C
C IMPLICIT REAL*8 (A-Z)
C INTEGER I,J
C REAL X(5),G(10)
C DOUBLE PRECISION Q11,Q12,Q13,Q14,Q15,Q111,Q112,Q113,Q114,Q115,
C Q121,Q122,Q123,Q124,Q125,Q131,Q132,Q133,Q134,Q135,Q141,Q142,
C Q143,Q144,Q145,Q151,Q152,Q153,Q154,Q155,U11,U12,U13,U14,U5,U11,U12,
C 113,114,115,121,122,123,124,125,131,132,133,134,135,141,142,143,
C 4144,145,151,152,153,154,155

C
C SET VALUES FOR THE ESTIMATES G(I), I=3,4,5
C
C G(3)=0.0
C G(4)=0.0
C G(5)=0.0

C
C SET SAMPLE SIZE N
C
C N=194

C INITIALIZE HT=T1
C
C HT=-1.0

C INITIALIZE SUM(J), J=1,5, THE SUM OF X EXP{BZ} OVER ALL RISK SET.
C
C SUM11=0.0
C SUM12=0.0
C SUM13=0.0
C SUM14=0.0
```
SUM15 = 0.0

INITIALIZE SUM2, THE SUM OF EXP(BZ) OVER ALL RISK SET.

SUM2 = 0.0

INITIALIZE THE SCORES FUNCTIONS

U1 = 0.0
U2 = 0.0
U3 = 0.0
U4 = 0.0
U5 = 0.0

INITIALIZE THE ELEMENTS OF THE INFORMATION MATRIX.

I11 = 0.0
I12 = 0.0
I13 = 0.0
I14 = 0.0
I15 = 0.0
I21 = 0.0
I22 = 0.0
I23 = 0.0
I24 = 0.0
I25 = 0.0
I31 = 0.0
I32 = 0.0
I33 = 0.0
I34 = 0.0
I35 = 0.0
I41 = 0.0
I42 = 0.0
I43 = 0.0
I44 = 0.0
I45 = 0.0
I51 = 0.0
I52 = 0.0
I53 = 0.0
I54 = 0.0
I55 = 0.0

INITIALIZE SUMILM, L,M = 1,5, THE SUM OF X^2 EXP(BZ) OVER ALL RISK SET.

SUMI11 = 0.0
SUMI12 = 0.0
SUMI13 = 0.0
SUMI14 = 0.0
SUMI15 = 0.0
SUMI21 = 0.0
SUMI22 = 0.0
SUM123=0.0
SUM124=0.0
SUM125=0.0
SUM131=0.0
SUM132=0.0
SUM133=0.0
SUM134=0.0
SUM135=0.0
SUM141=0.0
SUM142=0.0
SUM143=0.0
SUM144=0.0
SUM145=0.0
SUM151=0.0
SUM152=0.0
SUM153=0.0
SUM154=0.0
SUM155=0.0
C
READ THE ESTIMATES \& OUTPUT FROM SAS PHGLM
C
20 READ (7,200) G(1),G(2)
200 FORMAT(2F11.8)
C
READ DATA AND CALCULATE THE SCORES UI, I=1,5 AND
THE INFORMATION MATRIX IMN, M,N=1,5.
C
DO 400 I=1, N
10 READ (5,100,END=990) X(1),W,X(2),T,C,X(3),X(4),X(5),
9DEL1,DEL2,T1
100 FORMAT (2F2.0,F3.0,F4.0,F2.0,F3.0,2F2.0,F6.1)
BZ=0
DO 300 J=1, 5
BZ=BZ+G(J)*X(J)
300 CONTINUE
SUM11=SUM11+X(1)*EXP(BZ)
SUM12=SUM12+X(2)*EXP(BZ)
SUM13=SUM13+X(3)*EXP(BZ)
SUM14=SUM14+X(4)*EXP(BZ)
SUM15=SUM15+X(5)*EXP(BZ)
SUM11=SUM11+X(1)*X(1)*EXP(BZ)
SUM12=SUM12+X(1)*X(2)*EXP(BZ)
SUM13=SUM13+X(1)*X(3)*EXP(BZ)
SUM14=SUM14+X(1)*X(4)*EXP(BZ)
SUM15=SUM15+X(1)*X(5)*EXP(BZ)
SUM121=SUM121+X(2)*X(1)*EXP(BZ)
SUM122=SUM122+X(2)*X(2)*EXP(BZ)
SUM123=SUM123+X(2)*X(3)*EXP(BZ)
SUM124=SUM124+X(2)*X(4)*EXP(BZ)
SUM125=SUM125+X(2)*X(5)*EXP(BZ)
SUM131=SUM131+X(3)*X(1)*EXP(BZ)
SUM132=SUM132+X(3)*X(2)*EXP(BZ)
SUM133=SUM133+X(3)*X(3)*EXP(BZ)
SUM134=SUM134+X(3)*X(4)*EXP(BZ)
SUM135=SUM135+X(3)•X(5)•EXP(BZ)
SUM141=SUM141+X(4)•X(1)•EXP(BZ)
SUM142=SUM142+X(4)•X(2)•EXP(BZ)
SUM143=SUM143+X(4)•X(3)•EXP(BZ)
SUM144=SUM144+X(4)•X(4)•EXP(BZ)
SUM145=SUM145+X(4)•X(5)•EXP(BZ)
SUM151=SUM151+X(5)•X(1)•EXP(BZ)
SUM152=SUM152+X(5)•X(2)•EXP(BZ)
SUM153=SUM153+X(5)•X(3)•EXP(BZ)
SUM154=SUM154+X(5)•X(4)•EXP(BZ)
SUM155=SUM155+X(5)•X(5)•EXP(BZ)
SUM2=SUM2+EXP(BZ)
IF (DEL1 .GE. 1.0) THEN
Q11=SUM11/SUM2
Q12=SUM12/SUM2
Q13=SUM13/SUM2
Q14=SUM14/SUM2
Q15=SUM15/SUM2
Q111=SUM111/SUM2
Q112=SUM112/SUM2
Q113=SUM113/SUM2
Q114=SUM114/SUM2
Q115=SUM115/SUM2
Q121=SUM121/SUM2
Q122=SUM122/SUM2
Q123=SUM123/SUM2
Q124=SUM124/SUM2
Q125=SUM125/SUM2
Q131=SUM131/SUM2
Q132=SUM132/SUM2
Q133=SUM133/SUM2
Q134=SUM134/SUM2
Q135=SUM135/SUM2
Q141=SUM141/SUM2
Q142=SUM142/SUM2
Q143=SUM143/SUM2
Q144=SUM144/SUM2
Q145=SUM145/SUM2
Q151=SUM151/SUM2
Q152=SUM152/SUM2
Q153=SUM153/SUM2
Q154=SUM154/SUM2
Q155=SUM155/SUM2
IF (T1.EQ.HT) THEN
DEL1=DEL1+HDEL1
ENDIF
U1=U1+X(1)-DEL1•Q11
U2=U2+X(2)-DEL1•Q12
U3=U3+X(3)-DEL1•Q13
U4=U4+X(4)-DEL1•Q14
U5=U5+X(5)-DEL1•Q15
I11=I11+DEL1•(QI11-Q11•Q11)
I12=I12+DEL1•(QI12-Q11•Q12)
I13 = I13 + DEL1 *(Q1I3-Q11*Q13)
I14 = I14 + DEL1 *(Q1I4-Q11*Q14)
I15 = I15 + DEL1 *(Q1I5-Q11*Q15)
I21 = I21 + DEL1 *(Q1I21-Q12*Q11)
I22 = I22 + DEL1 *(Q1I22-Q12*Q12)
I23 = I23 + DEL1 *(Q1I23-Q12*Q13)
I24 = I24 + DEL1 *(Q1I24-Q12*Q14)
I25 = I25 + DEL1 *(Q1I25-Q12*Q15)
I31 = I31 + DEL1 *(Q1I31-Q13*Q11)
I32 = I32 + DEL1 *(Q1I32-Q13*Q12)
I33 = I33 + DEL1 *(Q1I33-Q13*Q13)
I34 = I34 + DEL1 *(Q1I34-Q13*Q14)
I35 = I35 + DEL1 *(Q1I35-Q13*Q15)
I41 = I41 + DEL1 *(Q1I41-Q14*Q11)
I42 = I42 + DEL1 *(Q1I42-Q14*Q12)
I43 = I43 + DEL1 *(Q1I43-Q14*Q13)
I44 = I44 + DEL1 *(Q1I44-Q14*Q14)
I45 = I45 + DEL1 *(Q1I45-Q14*Q15)
I51 = I51 + DEL1 *(Q1I51-Q15*Q11)
I52 = I52 + DEL1 *(Q1I52-Q15*Q12)
I53 = I53 + DEL1 *(Q1I53-Q15*Q13)
I54 = I54 + DEL1 *(Q1I54-Q15*Q14)
I55 = I55 + DEL1 *(Q1I55-Q15*Q15)
IF (T1.EQ.HT) THEN
U1 = U1 + HDEL1 * HQ11
U2 = U2 + HDEL1 * HQ12
U3 = U3 + HDEL1 * HQ13
U4 = U4 + HDEL1 * HQ14
U5 = U5 + HDEL1 * HQ15
I11 = I11 - HDEL1 *(HQI11-HQ11*HQ11)
I12 = I12 - HDEL1 *(HQI12-HQ11*HQ12)
I13 = I13 - HDEL1 *(HQI13-HQ11*HQ13)
I14 = I14 - HDEL1 *(HQI14-HQ11*HQ14)
I15 = I15 - HDEL1 *(HQI15-HQ11*HQ15)
I21 = I21 - HDEL1 *(HQI21-HQ12*HQ11)
I22 = I22 - HDEL1 *(HQI22-HQ12*HQ12)
I23 = I23 - HDEL1 *(HQI23-HQ12*HQ13)
I24 = I24 - HDEL1 *(HQI24-HQ12*HQ14)
I25 = I25 - HDEL1 *(HQI25-HQ12*HQ15)
I31 = I31 - HDEL1 *(HQI31-HQ13*HQ11)
I32 = I32 - HDEL1 *(HQI32-HQ13*HQ12)
I33 = I33 - HDEL1 *(HQI33-HQ13*HQ13)
I34 = I34 - HDEL1 *(HQI34-HQ13*HQ14)
I35 = I35 - HDEL1 *(HQI35-HQ13*HQ15)
I41 = I41 - HDEL1 *(HQI41-HQ14*HQ11)
I42 = I42 - HDEL1 *(HQI42-HQ14*HQ12)
I43 = I43 - HDEL1 *(HQI43-HQ14*HQ13)
I44 = I44 - HDEL1 *(HQI44-HQ14*HQ14)
I45 = I45 - HDEL1 *(HQI45-HQ14*HQ15)
I51 = I51 - HDEL1 *(HQI51-HQ15*HQ11)
I52 = I52 - HDEL1 *(HQI52-HQ15*HQ12)
I53 = I53 - HDEL1 *(HQI53-HQ15*HQ13)
I54 = I54 - HDEL1 *(HQI54-HQ15*HQ14)
I55 = I55 - HDEL1 *(HQI55-HQ15*HQ15)
ENDIF
C  HOLD THE RATIOS FOR THE NEXT ITERATION
C
HT=T1
HQ11=Q11
HQ12=Q12
HQ13=Q13
HQ14=Q14
HQ15=Q15
HQ111=Q111
HQ112=Q112
HQ113=Q113
HQ114=Q114
HQ115=Q115
HQ121=Q121
HQ122=Q122
HQ123=Q123
HQ124=Q124
HQ125=Q125
HQ131=Q131
HQ132=Q132
HQ133=Q133
HQ134=Q134
HQ135=Q135
HQ141=Q141
HQ142=Q142
HQ143=Q143
HQ144=Q144
HQ145=Q145
HQ151=Q151
HQ152=Q152
HQ153=Q153
HQ154=Q154
HQ155=Q155
HDEL1=DEL1
ENDIF
400 CONTINUE
C  OUTPUT THE SCORES AND THE ELEMENTS OF INFORMATION
C  MATRIX
C
WRITE(15,*) U1,U2,U3,U4,U5
WRITE(16,*) I11,I12,I13,I14,I15,I21,I22,I23,I24,I25,I31,I32,I33,
990 STOP
END