ESTIMATION IN THE SIMPLE LINEAR REGRESSION MODEL WITH ONE-FOLD NESTED ERROR

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OF

MIDDLE EAST TECHNICAL UNIVERSITY

BY

BURÇİN EMRE ÜLGEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR

THE DEGREE OF MASTER OF SCIENCE

IN

STATISTICS

MAY 2005

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. H. Öztaş AYHAN
Head of the Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Bilgehan GÜVEN
Supervisor

Examining Committee Members

Prof. Dr. H. Öztaş AYHAN	(METU, STAT)	
Assoc. Prof. Dr. Bilgehan GÜVEN	(METU, STAT)	
Prof. Dr. Ali UZUN	(METU,STAT)	
Prof. Dr. Kerim DEMİRBAŞ	(METU, EEE)	
Dr. Özlem İLK	(METU, STAT)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name :

Signature :

ABSTRACT

ESTIMATION IN THE SIMPLE LINEAR REGRESSION MODEL WITH ONE-FOLD NESTED ERROR

Ülgen, Burçin Emre
M.S., Department of Statistics
Supervisor : Assoc. Prof. Dr. Bilgehan Güven

May 2005, 71 pages

In this thesis, estimation in simple linear regression model with one-fold nested error is studied.

To estimate the fixed effect parameters, generalized least squares and maximum likelihood estimation procedures are reviewed. Moreover, Minimum Norm Quadratic Estimator (MINQE), Almost Unbiased Estimator (AUE) and Restricted Maximum Likelihood Estimator (REML) of variance of primary units are derived.

Also, confidence intervals for the fixed effect parameters and the variance components are studied. Finally, the aforesaid estimation techniques and confidence intervals are applied to a real-life data and the results are presented.

Keywords: Least Squares Estimation, Maximum Likelihood Estimation, Estimation for Variance Components, Confidence Intervals

ÖZ

İÇ İÇE GEÇMİŞ HATA TERİMLİ BASİT DOĞRUSAL REGRESYON MODELİNDE TAHMİN YÖNTEMLERİ

Ülgen, Burçin Emre
Yüksek Lisans, İstatistik Bölümü
Tez Yöneticisi: Doç. Dr. Bilgehan Güven

Mayıs 2005, 71 sayfa

Bu tezde, iç içe geçmiş hata terimli doğrusal regresyon modelinde tahmin yöntemleri esas olarak çalışılmıştır.

Sabit etki parametrelerini tahmin etmek için, Genelleştirilmiş En Küçük Kareler Yöntemi ve En Çok Olabilirlik Tahmin Yöntemi prosedürleri sunulmuştur. Ayrıca, birinci aşama birimlerin varyansının En Küçük Norm Karesel Tahmin Edicisi, Hemen Hemen Yansız Tahmin Edicisi ve Kısıtlı En Çok Olabilirlik Tahmin Edicisi bulunmuştur.

Sabit etki parametreleri ve varyans bileşenleri için güven aralıkları da çalışılmıştır. Son olarak, adı geçen tahmin yöntemleri ve güven aralıkları bir veri setine uygulanmış ve bu uygulamanın sonuçları sunulmuştur.

Anahtar Kelimeler: En Küçük Kareler Tahmini, En Çok Olabilirlik Tahmini, Varyans Bileşenlerinin Tahmini, Güven Aralıkları

To My Parents

ACKNOWLEDGMENTS

First, I would like to express my sincere appreciation to my supervisor Assoc. Prof. Dr. Bilgehan Güven for his integral view on research, wise direction and support in so many aspects. He could not even realize how much I have learned from him.

I would also like to thank to my committee members: Prof. Dr. H. Öztaş Ayhan, Prof. Dr. Ali Uzun, Prof. Dr. Kerim Demirbaş and Dr. Özlem İlk for their valuable suggestions which has a direct impact on the final form of this thesis.

Special thanks go to my dearest friends Burcin (you are an original like I've never seen before) and Sibel (you're just too good to be true) for their help, care and love. I really hope that we'll always be here and you'll always be near.

I feel deep sense of gratitude for my family who formed my vision and taught me that "There's more to life than this".

Finally, nothing would have been possible without the blessing of above.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

The one-fold nested error structure is appropriate when there is some form of subsampling from primary sampling, i.e., observations followed by randomly selected primary groups. Particular cases in which the data sets arise are : (i) observational studies with two-stage sampling structure, i.e., random selection of whole units followed by random selection of several subunits within each of the whole units, and the researcher is interested in the regression of correlated response variables (ii) experiments designed as split-plot or repeated measures designs and (iii) experiments designed as incomplete blocks where the block effect is random.

Suppose for each observation Y_{jk}, we have a covariate x_{jk} and JK observations are collected by selecting J primary units followed by taking K measurements from each primary unit. Since we look for regression of observations on the covariate and assume that there is only one observational error e_{jk} independently and identically distributed, we have J regression equations such that each of them has the same intercept parameter but different slope parameters. However, when we fit the same regression equation to all J primary units, a new error term which we call a primary unit error a_j arises. In this situation our model which is called the simple linear regression model with one-fold nested error structure is

$$Y_{jk} = \mu + \beta x_{jk} + a_j + e_{jk}\,,\ j = 1,2,\ldots,J\,;\ k = 1,2,\ldots,K$$

where a_j and e_{jk} are independently distributed with zero means and variances σ_a^2 and σ_e^2 respectively (these variances are also referred as variance components). The term a_j represents an error component associated with primary units and e_{jk} represents an error component associated with the subunit. A classical example of the simple linear regression model with one-fold nested error structure is two stage sampling with random selection of J primary units followed by the selection of K subunits within each primary unit, i.e., an experiment on J selected individuals with K measurements taken on each individual.

A factor is defined as fixed if its levels consist of the entire population of possible levels. Otherwise, i.e., if its levels consist of a random sample of levels from a population of possible levels, a factor is called random. In a classification by Eisenhart (1947), a model is defined as :

- fixed or fixed effects model if all the factors in the treatment structure are fixed effects,
- random or random effects model if all the factors in the treatment structure are random effects and
- mixed or mixed effects model if some of the factors in the treatment structure are fixed effects and some are random effects.

Since the simple linear regression model with one-fold nested error structure consists of two fixed effects μ, β and two random effects a_j, e_{jk}, the model is a linear mixed model which is a type of the mixed analysis of variance models. Thus, the estimation techniques for mixed analysis of variance models can be used for the estimation of unknown parameters of a simple linear regression model with one-fold nested error structure.

1.2 Review of Literature

Mixed models are an extension of linear models for which covariance structure is based on random effects and their covariance parameters. The use of such models has increased dramatically in the past decade. For example, they are now the main vehicle for analysis of longitudinal data (e.g. Diggle, Heagarty, and Zeger, 2002). Their use in semiparametric regression is advocated in some recent literature (e.g. Ruppert, Wand & Carroll, 2003). The popularity of linear mixed models has been accompanied by vigorous research on analytic results and computational methods.

After Eisenhart (1947) distinguished fixed, random and mixed models, Henderson (1953) proposed a method for estimating variance components by equating each computed reduction in sums of squares (due to fitting different subgroups of factors in the model) to its expected value. Hartley and Rao (1967) described the maximum likelihood estimation of the fixed effects and variance components in the general mixed analysis of variance model and established some properties of these estimators. Corbeil and Searle (1976) modified Hartley and Rao's solution by applying a transformation to the likelihood function partitioned into two parts, one being free of fixed effects. Maximizing this part yields what are called the restricted maximum likelihood estimates (REML). A comprehensive survey of maximum likelihood approaches to variance component estimation were given by Harville (1977).

Rao (1970) proposes minimum norm estimation technique for variance components. The resulting estimator is the so-called Minimum Norm Quadratic Unbiased Estimators (MINQUE) for a model with unequal error variances. Moreover, Rao (1972) describes niminum norm and minimum variance quadratic unbiased estimators (MINQUE and MIVQUE) in a general setting. Henderson (1975) also gave the method of obtaining generalized least squares estimates in a general mixed model and described the mixed model and developed best linear unbiased estimators.

Rao and Chaubey (1978) obtained nonnegative estimators (including Minimum Norm Quadratic Estimator) for the variance components by ignoring the condition for unbiasedness in the principle of Minimum Norm Quadratic Unbiased Estimation. Moreover, Horn, Horn and Duncan (1975) and Horn and Horn (1975) compared the estimators of variance components along with the estimator they obtained which is called Almost Unbiased Estimator (AUE). Searle (1977) gave a comprehensive review on variance components estimation. An excellent summary of developments concerning the mixed models is provided by McCulloch and Searle (2000).

Having reviewed more cited studies about the mixed models, we now summarize the former studies concerning the simple linear regression model with one-fold nested error structure as follows.

Tong and Cornelius (1989, 1991) treat the simple linear regression model with one-fold nested error as an analysis of a one way classification, in which treatment effects are random, and compared the four estimators of the slope β (ordinary least squares, maximum likelihood, estimated generalized least squares and the covariance estimators) with respect to their mean squared errors in a Monte Carlo simulation study. They also constructed and compared six hypothesis testing procedures for the slope β with respect to Type I error and power of test in a Monte Carlo simulation study.

Park and Burdick (1994) derived exact and approximate confidence intervals for the regression coefficient in the simple linear regression model with the one-fold nested error structure and compared these confidence intervals using computer simulation. Park and Hwang (2002) also derived the exact and approximate confidence intervals for the mean response for a given level of the independent variable in the simple linear regression with one-fold nested error structure and compared them by simulation.

The generalized least squares (GLS) estimator for the linear model with a nested error is given by Fuller and Battese (1973). They suggest a transformation depending on the variance components for the nested error model. The transformation is simply based on multiplying an observation vector by the square root of its covariance matrix and the transformed observations are uncorrelated. In order to get the GLS estimator of the fixed effect parameters, ordinary least squares (OLS) estimation is applied to the transformed observations. They use Henderson's Method 3 (Henderson, 1953) to estimate the variance components since it seems better than both of Henderson's other two methods since it provides estimators of the variance components which are invariant to the design matrix.

Restricted maximum likelihood estimation (Corbeil and Searle, 1976) was been applied to simple linear regression model with one-fold nested error structure by Tong and Cornelius (1989).

Güven (1995) studied the maximum likelihood estimation in simple linear regression with one-fold nested error structure differently than restricted maximum likelihood estimation and derived the exact maximum likelihood estimates of the four unknown parameters together. Moreover, for this model Güven (1998) combined two independent unbiased estimators for the slope β and obtained the unbiased estimator of β whose variance is less than the variances of either, which is called uniformly better unbiased estimator.

1.2 Aims and Scope of the Study

The primary aim of this thesis study is to review the estimation of fixed-effect parameters and variance components of the simple linear regression model with one-fold nested error. We present a comprehensive summary of what has been done up to date along and apply them into a real-life data.

The definition and a brief introduction of simple linear regression models with one-fold nested error are presented in this chapter.

Chapter 2 focuses on estimation techniques for fixed effect parameters in the simple linear regression model with one-fold nested error. It covers the estimated generalized least squares and maximum likelihood estimation of fixed-effect parameters.

Chapter 3 gives the estimation of the primary unit variance and involves Minimum Norm Quadratic Estimator (MINQE), Almost Unbiased Estimator (AUE) and Restricted Maximum Likelihood Estimator (REML). Among the different variance component estimators (summarized by Searle, 1977) we choose these four estimators since the exact mean squared errors (MSEs) of them can be derived. In addition to the derivations of these estimators, we also obtained the exact mean squared errors (MSEs) of them in order to compare these variance component estimators with respect to their MSEs in the last section of this chapter.

Chapter 4 covers the exact and approximate confidence intervals for the fixed effect parameters and the variance components of the model.

In Chapter 5, the application of all of these estimation techniques and confidence intervals to a real life data taken from Vonesh and Carter (1987) is presented.

Finally, in Chapter 6, summary and conclusions on the findings are given.

1.4 Simple Linear Regression Model with One-Fold Nested Error

We consider the model :

$$Y_{jk} = \mu + \beta x_{jk} + a_j + e_{jk} \qquad j = 1,2,...,J, \quad k = 1,2,...,K$$

where Y_{jk} denotes the observation from the k^{th} second stage sampling unit in the j^{th} first-stage unit; x_{jk} denotes the value of a nonstochastic regressor variable measured on the k^{th} second stage sampling unit in the j^{th} first-stage unit; μ, β are parameters to be estimated; a_j and e_{jk} are unobservable random effects, where a_j is an error associated with the j^{th} first-stage unit and e_{jk} with the k^{th} second-stage sampling unit in the j^{th} first-stage unit. The "errors" a_j and e_{jk} are independent normal random variables with zero mean and variances σ_a^2 and σ_e^2, respectively, where $\sigma_a^2 \geq 0$, and $\sigma_e^2 > 0$. Under these assumptions, the covariance structure for the observed variables Y_{jk} is as follows :

$$Cov(Y_{jk}, Y_{j'k'}) = \begin{cases} \sigma_a^2 + \sigma_e^2 & if \quad j = j', \quad k = k' \\ \sigma_a^2 & if \quad j = j', \quad k \neq k' \\ 0 & if \quad j \neq j' \end{cases}$$

The observation Y_{jk} for the k^{th} element of the j^{th} unit can be written as follows :

$$Y_{jk} = E[y_{jk}] + a_j + e_{jk} \text{ where } E[Y_{jk}] = \mu + \beta x_{jk} \text{ .}$$

We can write the linear model, in matrix form, as follows :

$$\mathbf{Y} = \mathbf{X\alpha} + \mathbf{Ua} + \mathbf{e}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y_1} \\ \mathbf{Y_2} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{Y_J} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \mathbf{1_K} & \mathbf{X_1} \\ \mathbf{1_K} & \mathbf{X_2} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \mathbf{1_K} & \mathbf{X_K} \end{bmatrix}, \ \mathbf{a} = \begin{bmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{a_J} \end{bmatrix}$$

with $\mathbf{Y_j} = (\mathbf{Y_{j1}}, \mathbf{Y_{j2}}, ..., \mathbf{Y_{jK}})^\mathbf{T}$, $\mathbf{X_j} = (\mathbf{x_{j1}}, \mathbf{x_{j2}}, ..., \mathbf{x_{jK}})^\mathbf{T}$, $\mathbf{e_j} = (\mathbf{e_{j1}}, \mathbf{e_{j2}}, ..., \mathbf{e_{jK}})^\mathbf{T}$ for $j = 1,2,...J$, $\mathbf{U} = \mathbf{I_J} \otimes \mathbf{1_K}$, $\mathbf{1_K}$ is a $K \times 1$ vector whose elements are all 1, $\otimes$ denotes the Kronecker matrix product, $\boldsymbol{\alpha} = (\boldsymbol{\mu}, \boldsymbol{\beta})^\mathbf{T}$, the vector $\mathbf{e}$ is defined similar to $\mathbf{Y}$.

It follows from the independence of $\mathbf{Y_1}, \mathbf{Y_2}, ..., \mathbf{Y_J}$ that $\mathbf{Y} \sim N_{JK}(\mathbf{X\alpha}, \mathbf{\Lambda})$ where

$$\mathbf{\Lambda} = \mathbf{I_J} \otimes \mathbf{\Sigma} = \mathbf{\Lambda} = \sigma_a^2(\mathbf{I_J} \otimes \mathbf{1_K}\mathbf{1_K^T}) + \sigma_e^2\mathbf{I_{JK}} \text{ ,} \qquad (1.4.1)$$

and $\mathbf{\Sigma}$ is the variance-covariance matrix of $\mathbf{Y_L}$ for $L = 1,2,...,J$ which is given by $\mathbf{\Sigma} = \sigma_a^2\mathbf{1_K}\mathbf{1_K^T} + \sigma_e^2\mathbf{I_K}$.

Let us define some quantities which will occur in formulas for estimation of parameters and variance components. All of these quantities are easily obtained from a one-way analysis of covariance of **x** and **Y**. Let S_{xxa}, S_{xya} and S_{yya} be the among primary unit sum of squares and cross products, defined as,

$$S_{xxa} = K\sum_{j=1}^{J}(\bar{x}_{j.} - \bar{x}_{..})^2 \text{ ,} \qquad (1.4.2)$$

$$S_{xya} = K\sum_{j=1}^{J}(\bar{x}_{j.} - \bar{x}_{..})(\bar{Y}_{j.} - \bar{Y}_{..}), \tag{1.4.3}$$

and

$$S_{yya} = K\sum_{j=1}^{J}(\bar{Y}_{j.} - \bar{Y}_{..})^2 , \tag{1.4.4}$$

where $\bar{Y}_{j.} = \sum_{k=1}^{K} Y_{jk} / K$, $\bar{x}_{j.} = \sum_{k=1}^{K} x_{jk} / K$, $\bar{Y}_{..} = \sum_{j=1}^{J}\sum_{k=1}^{K} Y_{jk} /(JK)$ and $\bar{x}_{..} = \sum_{j=1}^{J}\sum_{k=1}^{K} x_{jk} /(JK)$.

Similarly, let S_{xxe}, S_{xye} and S_{yye} be the within primary unit sums of squares and cross products, i.e.,

$$S_{xxe} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{j.})^2 , \tag{1.4.5}$$

$$S_{xye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{j.})(Y_{jk} - \bar{Y}_{j.}), \tag{1.4.6}$$

and

$$S_{yye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \bar{Y}_{j.})^2 . \tag{1.4.7}$$

It follows that the the total sum of squares and cross products S_{xxt}, S_{xyt} and S_{yyt} can be written as :

$$S_{xxt} = S_{xxa} + S_{xxe} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{..})^2 , \tag{1.4.8}$$

$$S_{xyt} = S_{xya} + S_{xye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{..})(Y_{jk} - \bar{Y}_{..}), \tag{1.4.9}$$

and

$$S_{yyt} = S_{yya} + S_{yye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \bar{Y}_{..})^2 . \tag{1.4.10}$$

Moreover S_{xxt}, S_{xyt}, S_{yyt} S_{xxa}, S_{xya} , S_{yya}, S_{xxe}, S_{xye} and S_{yye} can be rewritten, in a matrix form, as follows :

$$S_{xxt} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{..})^2 = \left(\mathbf{X} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\,\mathbf{X}\right)^T\left(\mathbf{X} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\,\mathbf{X}\right)$$

$$= \mathbf{X^T}\left(\mathbf{I_{JK}} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right)\mathbf{X}, \tag{1.4.11}$$

$$S_{xyt} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{..})(Y_{jk} - \bar{Y}_{..}) = \mathbf{X^T}\left(\mathbf{I_{JK}} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right)\mathbf{Y}, \tag{1.4.12}$$

$$S_{yyt} = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \bar{Y}_{..})^2 = \mathbf{Y^T}\left(\mathbf{I_{JK}} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right)\mathbf{Y}, \tag{1.4.13}$$

$$S_{xxa} = K\sum_{j=1}^{J}(\bar{x}_{j.} - \bar{x}_{..})^2 = \mathbf{X^T}\left[\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right]\mathbf{X}, \tag{1.4.14}$$

$$S_{xya} = K\sum_{j=1}^{J}(\bar{x}_{j.} - \bar{x}_{..})(\bar{Y}_{j.} - \bar{Y}..) = \mathbf{X^T}\left[\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right]\mathbf{Y}, \tag{1.4.15}$$

$$S_{yya} = K\sum_{j=1}^{J}(\bar{Y}_{j.} - \bar{Y}_{..})^2 = \mathbf{Y^T}\left[\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right]\mathbf{Y} \;, \tag{1.4.16}$$

$$S_{xxe} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{j.})^2 = \mathbf{X^T}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{X} = \mathbf{X^T}\left(\mathbf{I_J} \otimes \mathbf{P}\right)\mathbf{X} \tag{1.4.17}$$

with $\mathbf{P} = \mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}$,

$$S_{xye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{j.})(Y_{jk} - \bar{Y}_{j.}) = \mathbf{X^T}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{Y} = \mathbf{X^T}\left(\mathbf{I_J} \otimes \mathbf{P}\right)\mathbf{Y} \tag{1.4.18}$$

and

$$S_{yye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \bar{Y}_{j.})^2 = \mathbf{Y^T}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{Y} = \mathbf{Y^T}\left(\mathbf{I_J} \otimes \mathbf{P}\right)\mathbf{Y}. \tag{1.4.19}$$

Having found the matrix representations of S_{xya} and S_{xye}, it can easily be shown that, under our model assumptions :

$$S_{xya} \sim N(S_{xxa}\beta, \quad S_{xxa}(\sigma_e^2 + K\sigma_a^2)) \text{ and } \quad S_{xye} \sim N(S_{xxe}\beta, \quad S_{xxe}\sigma_e^2)$$

since

$$E(S_{xya}) = \mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right)E(\mathbf{Y})$$

$$= \mathbf{X}^{\mathbf{T}}\left(\mathbf{I}_J \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_K}\mathbf{1_K^T}\right)\mu\mathbf{1_K} + \mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T}\right)\mathbf{X}\beta$$

$$= S_{xxa}\beta \ ,$$

$$V(S_{xya}) = V\left[\mathbf{X}^{\mathbf{T}}(\mathbf{I_J} \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T} - \frac{1}{JK}\mathbf{1_{JK}}\mathbf{1_{JK}^T})\mathbf{Y}\right]$$

$$= \sigma_e^2 S_{xxa} + K\sigma_a^2\mathbf{X}^{\mathbf{T}}\left(\left(\mathbf{I_J} - \frac{1}{J}\mathbf{1_J}\mathbf{1_J^T}\right) \otimes \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\mathbf{X}$$

$$= S_{xxa}(\sigma_e^2 + K\sigma_a^2),$$

$$E(S_{xye}) = \mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)E(\mathbf{Y}) = \mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{X}\beta$$

$$= S_{xxe}\beta \, ,$$

and

$$V(S_{xye}) = V\left[\mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{Y}\right]$$

$$= \sigma_e^2\mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\right)\mathbf{X} + \sigma_a^2\mathbf{X}^{\mathbf{T}}\left(\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K}\mathbf{1_K}\mathbf{1_K^T}\right)\mathbf{1_K}\mathbf{1_K^T}\right)\mathbf{X}$$

$$= S_{xxe}\sigma_e^2 .$$

CHAPTER 2

ESTIMATION TECHNIQUES FOR FIXED EFFECT PARAMETERS

In this chapter, we review the estimated generalized least squares estimation of the unknown parameters of the model given by Fuller and Battese (1973), Tong and Cornelius (1991) and maximum likelihood estimation of the unknown parameters of the model given by Güven (1995).

2.1 Estimated Generalized Least Squares Estimation

Fuller and Battese (1973) obtained the estimated generalized least squares estimators of the fixed effect parameters after they find the estimators of the variance components. Since the model can be considered as an analysis of covariance model of the one way classification model in which the treatment (primary) effects are random, Henderson's method 3 (Henderson, 1953) can be used for obtaining the variance components. This method uses reduction in sum of squares due to fitting different subgroup of factors in the model, i.e., it estimates the variance components by equating each computed reduction to its expected value. For our model, we use;

(i) the sum of squares from the primary units adjusted for regression on X, $R(a \mid \mu, \beta)$, to estimate σ_a^2

(ii) the residual sum of squares in the one way covariance model, RSS to estimate σ_e^2 .

We have

$$R(a \mid \mu,\beta) = R(\mu,\beta,a) - R(\mu,\beta) \tag{2.1.1}$$

$$= R(\mu,\beta,a) - \frac{(S_{xya}+S_{xye})^2}{(S_{xxa}+S_{xxe})}$$

and

$$R(\mu,\beta,a) = R(\mu,a) + R(\beta) \tag{2.1.2}$$

$$= S_{yya} + \frac{S_{xye}^2}{S_{xxe}}$$

where $R(\mu,a)$ is the sum of squares from the one way analysis of variance model and $R(\beta)$ is the sum of squares of covariance. Substituting (2.1.2) into (2.1.1), we have

$$R(a \mid \mu,\beta) = R(\mu,a) + R(\beta) - R(\mu,\beta) \tag{2.1.3}$$

where $R(\mu,\beta)$ is the sum of squares for the primary units. Then we have,

$$R(a \mid \mu,\beta) = S_{yya} + \frac{S_{xye}^2}{S_{xxe}} - \frac{(S_{xya}+S_{xye})^2}{(S_{xxa}+S_{xxe})}$$

$$= S_{yya} - \frac{S_{xya}^2}{S_{xxa}} + \frac{S_{xxa}}{S_{xxe}(S_{xxa}+S_{xxe})}\left(\frac{S_{xxe}}{S_{xxa}}S_{xya} - S_{xye}\right)^2 \tag{2.1.4}$$

where S_{xxa}, S_{xya}, S_{yya}, S_{xxe}, S_{xye} are given in (1.4.2)-(1.4.6)

We can rewrite (2.1.4) as follows :

$$R(a \mid \mu,\beta) = S_a^2(\hat{\beta}_a) + \frac{(\hat{\beta}_e-\hat{\beta}_a)^2 S_{xxa}S_{xxe}}{S_{xxa}+S_{xxe}} = S_a^2(\hat{\beta}_a) + S_d^2 \tag{2.1.5}$$

where

$$S_a^2(\hat{\beta}_a) = S_{yya} - \frac{S_{xya}^2}{S_{xxa}}, \tag{2.1.6}$$

$$S_d^2 = \frac{S_{xxa}}{S_{xxe}(S_{xxa}+S_{xxe})}\left(\frac{S_{xxe}}{S_{xxa}}S_{xya} - S_{xye}\right)^2$$

$$= \frac{(\hat{\beta}_e-\hat{\beta}_a)^2 S_{xxa}S_{xxe}}{S_{xxa}+S_{xxe}}, \tag{2.1.7}$$

$$\hat{\beta}_a = \frac{S_{xya}}{S_{xxa}}, \tag{2.1.8}$$

$$\hat{\beta}_e = \frac{S_{xye}}{S_{xxe}}. \tag{2.1.9}$$

$\hat{\beta}_a$ is distributed according to a normal distribution with the mean β and the variance $(\sigma_e^2 + K\sigma_a^2)/S_{xxa}$ since $S_{xya} \sim N\left(S_{xxa}\beta,\ \ S_{xxa}(\sigma_e^2 + K\sigma_a^2)\right)$. Also, we can rewrite $\hat{\beta}_a$ and $S_a^2(\hat{\beta}_a)$ as follows :

$$\hat{\beta}_a = \frac{\mathbf{X_{J.}^T P Y_{J.}}}{\mathbf{X_{J.}^T P X_{J.}}} \tag{2.1.10}$$

$$S_a^2(\hat{\beta}_a) = K\sum_{j=1}^{J}(\overline{Y}_{j.} - \overline{Y}_{..} - \hat{\beta}_a(\overline{x}_{j.} - \overline{x}_{..}))^2$$

$$= K(\mathbf{Y_{J.}} - \hat{\beta}_a \mathbf{X_{J.}})^T \mathbf{P}(\mathbf{Y_{J.}} - \hat{\beta}_a \mathbf{X_{J.}}) \tag{2.1.11}$$

where the symmetric and idempotent matrix P is defined as :

$$\mathbf{P} = \left[\mathbf{I_J} - \frac{1}{J}\mathbf{1_J 1_J^T}\right] \tag{2.1.12}$$

and $\mathbf{X_{j.}} = (\overline{x}_{1.}, \overline{x}_{2.}, ..., \overline{x}_{J.})^T$, $\mathbf{Y_{j.}} = (\overline{Y}_{1.}, \overline{Y}_{2.}, ..., \overline{Y}_{J.})^T$

Substituting the matrix representation of $\hat{\beta}_a$ (2.1.10) into $S_a^2(\hat{\beta}_a)$ (2.1.11) , we obtain :

$$S_a^2(\hat{\beta}_a) = K\left(\mathbf{Y_{J.}} - \frac{\mathbf{X_{J.}^T P Y_{J.}}}{\mathbf{X_{J.}^T P X_{J.}}}\mathbf{X_{J.}}\right)^T \mathbf{P}\left(\mathbf{Y_{J.}} - \frac{\mathbf{X_{J.}^T P Y_{J.}}}{\mathbf{X_{J.}^T P X_{J.}}}\mathbf{X_{J.}}\right)$$

$$= K\mathbf{Y_{J.}^T}\left(\mathbf{P} - \frac{\mathbf{P X_{J.} X_{J.}^T P}}{\mathbf{X_{J.}^T P X_{J.}}}\right)\mathbf{Y_{J.}} \tag{2.1.13}$$

where $\mathbf{P}$ is given in (2.1.12)

Let $\mathbf{u_{J.}^T} = \mathbf{(e_{J.} + a)^T} = \mathbf{(e_{1.} + a_1, e_{2.} + a_2, ..., e_{J.} + a_j)}$, then $\mathbf{Y_{J.}} = \mu\mathbf{1_J} + \beta\mathbf{X_{J.}} + \mathbf{u_{J.}}$ Then the quadratic form $S_a^2(\hat{\beta}_a)$ becomes

$$S_a^2(\hat{\beta}_a) = K\mathbf{u_{J.}^T}\left(\mathbf{P} - \frac{\mathbf{P X_{J.} X_{J.}^T P}}{\mathbf{X_{J.}^T P X_{J.}}}\right)\mathbf{u_{J.}} \tag{2.1.14}$$

since $\mathbf{P1_J} = 0$, $\left(\mathbf{P} - \frac{\mathbf{P X_{J.} X_{J.}^T P}}{\mathbf{X_{J.}^T P X_{J.}}}\right)\mathbf{1_J} = 0$ and $\left(\mathbf{P} - \frac{\mathbf{P X_{J.} X_{J.}^T P}}{\mathbf{X_{J.}^T P X_{J.}}}\right)\mathbf{X_{J.}} = 0$

Since the quadratic form $S_a^2(\hat{\beta}_a)$ is of the form

$$S_a^2(\hat{\beta}_a) = \mathbf{u}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{R}\mathbf{u}_{\mathbf{J.}}$$

where the symmetric and idempotent matrix R is defined as,

$$\mathbf{R} = \left(\mathbf{P} - \frac{\mathbf{P}\mathbf{X}_{\mathbf{J.}}\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}}{\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\mathbf{X}_{\mathbf{J.}}} \right) \tag{2.1.15}$$

We can use the theory of the distribution of the quadratic forms to find the distribution of $S_a^2(\hat{\beta}_a) = \mathbf{u}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{R}\mathbf{u}_{\mathbf{J.}}$. It can be proved that $S_a^2(\hat{\beta}_a)\ /(\sigma_e^2 + K\sigma_a^2)$ is distributed as a chi square with degrees of freedom $tr(\mathbf{R})$ where

$$tr(\mathbf{R}) = tr\left(\mathbf{P} - \frac{\mathbf{P}\mathbf{X}_{\mathbf{J.}}\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}}{\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\mathbf{X}_{\mathbf{J.}}} \right) = tr(\mathbf{P}) - \frac{1}{\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\mathbf{X}_{\mathbf{J.}}} tr\left(\mathbf{P}\mathbf{X}_{\mathbf{J.}}\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\right)$$

$$= tr(\mathbf{P}) - \frac{1}{\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\mathbf{X}_{\mathbf{J.}}} tr\left(\mathbf{X}_{\mathbf{J.}}^{\mathbf{T}}\mathbf{P}\mathbf{X}_{\mathbf{J.}}\right) = J - 2$$

Also, it can easily be showed that S_d^2 in (2.1.7) is distributed as $\left(\sigma_e^2 + \frac{KS_{xxe}\sigma_a^2}{S_{xxa} + S_{xxe}} \right)\chi_1^2$ as follows. Since $\hat{\beta}_e \sim N(\beta,\ \ \sigma_e^2 / S_{xxe})$ and $\hat{\beta}_a \sim N(\beta,\ \ (\sigma_e^2 + K\sigma_a^2)/S_{xxa}))$, it follows that :

$$(\hat{\beta}_a - \hat{\beta}_e)^2 \sim \left(\frac{\sigma_e^2}{S_{xxe}} + \frac{\sigma_e^2 + K\sigma_a^2}{S_{xxa}} \right)\chi_1^2$$

$$S_d^2 = \frac{(\hat{\beta}_e - \hat{\beta}_a)^2 S_{xxa}S_{xxe}}{S_{xxa} + S_{xxe}} \sim \left(\sigma_e^2 + \frac{KS_{xxe}\sigma_a^2}{S_{xxa} + S_{xxe}} \right)\chi_1^2$$

where $\hat{\beta}_a$ and $\hat{\beta}_e$ are given in (2.1.8) and (2.1.9) respectively. Hence the expectation of $R(a \mid \mu, \beta)$ is ;

$$E[R(a \mid \mu, \beta)] = (J-1)\sigma_e^2 + K\sigma_a^2\left((J-2) + \frac{S_{xxe}}{S_{xxa} + S_{xxe}} \right) \tag{2.1.16}$$

Also, we can rewrite $\hat{\beta}_e$ and $S_e^2(\hat{\beta}_e)$ as follows:

$$\hat{\beta}_e = \frac{\mathbf{X}^{\mathbf{T}}\mathbf{Q}\mathbf{Y}}{\mathbf{X}^{\mathbf{T}}\mathbf{Q}\mathbf{X}} \tag{2.1.17}$$

$$S_e^2(\hat{\beta}_e) = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \overline{Y}_{j.} - \hat{\beta}_e(x_{jk} - \overline{x}_{j.}))^2$$

$$= (\mathbf{Y} - \hat{\beta}_e \mathbf{X})^T \mathbf{Q}(\mathbf{Y} - \hat{\beta}_e \mathbf{X}) \qquad (2.1.18)$$

where the symmetric and idempotent matrix $\mathbf{Q}$ is defined as :

$$\mathbf{Q} = \mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{1}{K} \mathbf{1_K} \mathbf{1_K^T} \right) \qquad (2.1.19)$$

Substituting the matrix representation of $\hat{\beta}_e$ (2.1.17) into $S_e^2(\hat{\beta}_e)$ (2.1.18), we obtain

$$S_e^2(\hat{\beta}_e) = \left(\mathbf{Y} - \frac{\mathbf{X^T Q Y}}{\mathbf{X^T Q X}} \mathbf{X} \right)^{\mathbf{T}} \mathbf{Q} \left(\mathbf{Y} - \frac{\mathbf{X^T Q Y}}{\mathbf{X^T Q X}} \mathbf{X} \right)$$

$$= \mathbf{Y^T} \left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right) \mathbf{Y} \qquad (2.1.20)$$

where $\mathbf{Y} = \mu \mathbf{1_{JK}} + \beta \mathbf{X} + (\mathbf{I_J} \otimes \mathbf{1_K})\mathbf{a} + \mathbf{e}$. Then the quadratic form $S_e^2(\hat{\beta}_e)$ becomes

$$S_e^2(\hat{\beta}_e) = \mathbf{e^T} \left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X' Q X}} \right) \mathbf{e}, \qquad (2.1.21)$$

where $\left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right) \mathbf{1_{JK}} = 0$ and then $\left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right) (\mathbf{I_J} \otimes \mathbf{1_K}) = 0$, $\left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right) \mathbf{X} = 0$ since $\mathbf{Q 1_{JK}} = 0$ and $\mathbf{Q}(\mathbf{I_J} \otimes \mathbf{1_K}) = 0$.

Since the quadratic form $S_e^2(\hat{\beta}_e)$ is of the form ,

$$S_e^2(\hat{\beta}_e) = \mathbf{e^T S e},$$

where the symmetric and idempotent matrix **S** is defined as,

$$\mathbf{S} = \left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right). \qquad (2.1.22)$$

We can use the theory of the distribution of the quadratic forms to find the distribution of $S_e^2(\hat{\beta}_e) = \mathbf{e^T S e}$. It can be proved that $S_e^2(\hat{\beta}_e)/\sigma_e^2$ is distributed as a chi square variable with degrees of freedom $tr(\mathbf{S})$ where

$$tr(\mathbf{S}) = tr\left(\mathbf{Q} - \frac{\mathbf{Q X X^T Q}}{\mathbf{X^T Q X}} \right) = tr(\mathbf{Q}) - \frac{1}{\mathbf{X^T Q X}} tr\left(\mathbf{Q X X^T Q} \right)$$

$$= tr(\mathbf{Q}) - 1 = J(K-1) - 1.$$

Hence the expectation of RSS is ;

$$E[RSS] = (J(K-1)-1)\sigma_e^2. \tag{2.1.23}$$

By equating each sum of squares to its expected value, we have the following estimator equations for the variance components (Fuller and Battese, 1973):

$$RSS = (J(K-1)-1)\tilde{\sigma}_e^2 \tag{2.1.24}$$

$$R(a \mid \mu, \beta) = (J-1)\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2\left((J-2) + \frac{S_{xxe}}{S_{xxa} + S_{xxe}}\right) \tag{2.1.25}$$

Solving (2.1.24) and (2.1.25), the following variance component estimators can be obtained :

$$\tilde{\sigma}_e^2 = \frac{RSS}{J(K-1)-1}, \tag{2.1.26}$$

and

$$\tilde{\sigma}_a^2 = \max\left(0,\quad \frac{R(a \mid \mu, \beta) - (J-1)\tilde{\sigma}_e^2}{K\left((J-2) + \frac{S_{xxe}}{S_{xxe} + S_{xxa}}\right)}\right), \tag{2.1.27}$$

where the variances of $\tilde{\sigma}_e^2$ and $\tilde{\sigma}_a^2$ are calculated as follows :

$$V(\tilde{\sigma}_e^2) = \frac{1}{(J(K-1)-1)^2} V(S_e^2(\hat{\beta}_e)) = \frac{2\sigma_e^4}{J(K-1)-1}, \tag{2.1.28}$$

and

$$V(\tilde{\sigma}_a^2) = \frac{1}{\left[K\left((J-2) + \frac{S_{xxe}}{S_{xxa} + S_{xxe}}\right)\right]^2} V\left(R(a \mid \mu, \beta) - (J-1)\tilde{\sigma}_e^2\right). \tag{2.1.29}$$

From the distributional properties of our model, we know that $S_a^2(\hat{\beta}_a)$ and $S_e^2(\hat{\beta}_e)$ are independent. Moreover, it is obvious that S_d^2 and $S_e^2(\hat{\beta}_e)$ are also independent since $\hat{\beta}_a$, $\hat{\beta}_e$ and $S_e^2(\hat{\beta}_e)$ independent. As a result, $S_a^2(\hat{\beta}_a)$, $S_e^2(\hat{\beta}_e)$ and S_d^2 are independent, thus we can rewrite (2.1.29) as follows:

$$V(\tilde{\sigma}_e^2) = c.(V(S_a^2(\hat{\beta}_a)) + V(S_d^2) + (J-1)^2 V(\tilde{\sigma}_e^2)), \tag{2.1.30}$$

where $c = \dfrac{1}{\left[K\left((J-2) + \dfrac{S_{xxe}}{S_{xxa} + S_{xxe}}\right)\right]^2}$.

Using the distributional properties of $S_a^2(\hat{\beta}_a)$ and $S_e^2(\hat{\beta}_e)$, substituting (2.1.28) into (2.1.30), we have :

$$V(\tilde{\sigma}_e^2) = 2c\left[(\sigma_e^2 + K\sigma_a^2)(J-2) + \left(\sigma_e^2 + \frac{K\sigma_a^2 S_{xxe}}{S_{xxe} + S_{xxa}}\right)^2 + \frac{(J-1)^2\sigma_e^4}{J(K-1)-1}\right].$$

Having obtained the estimators of variance components, Fuller and Batesse (1973) found the estimated generalized least squares of fixed effect parameters. In a vector form, we can write,

$$\tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \tilde{\mu} \\ \tilde{\beta} \end{pmatrix} = \left(\mathbf{X}^{\mathbf{T}}\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathbf{T}}\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{Y},$$

where $\tilde{\boldsymbol{\Sigma}} = \tilde{\sigma}_e^2\mathbf{I}_{\mathbf{JK}} + \tilde{\sigma}_a^2(\mathbf{I}_{\mathbf{J}} \otimes \mathbf{1}_{\mathbf{K}}\mathbf{1}_{\mathbf{K}}^{\mathbf{T}})$, $\tilde{\sigma}_e^2$ and $\tilde{\sigma}_a^2$ are given in (2.1.14) and (2.1.15) respectively.

$$\left(\mathbf{X}^{\mathbf{T}}\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{X}\right)^{-1} = \frac{\tilde{\sigma}_e^2(\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2)}{\tilde{\sigma}_e^2 S_{xxa} + (\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2)S_{xxe}}\begin{pmatrix} \dfrac{\tilde{\sigma}_e^2\sum_{j=1}^{J}\sum_{k=1}^{K}x_{jk}^2 + K\tilde{\sigma}_a^2 S_{xxe}}{JK\tilde{\sigma}_e^2} & -x_{..} \\ -x_{..} & 1 \end{pmatrix}$$

$$\mathbf{X}^{\mathbf{T}}\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{Y} = \frac{1}{\tilde{\sigma}_e^2(\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2)}\begin{pmatrix} \tilde{\sigma}_e^2\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{jk} \\ \tilde{\sigma}_e^2\sum_{j=1}^{J}\sum_{k=1}^{K}x_{jk}Y_{jk} + K\tilde{\sigma}_a^2 S_{xye} \end{pmatrix}.$$

Then, we have

$$\tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \tilde{\mu} \\ \tilde{\beta} \end{pmatrix} = \begin{pmatrix} \overline{Y}_{..} - \tilde{\beta}\overline{x}_{..} \\ \dfrac{\tilde{\sigma}_e^2 SS_{xya} + (\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2)S_{xye}}{\tilde{\sigma}_e^2 SS_{xxa} + (\tilde{\sigma}_e^2 + K\tilde{\sigma}_a^2)S_{xxe}} \end{pmatrix}. \qquad (2.1.31)$$

2.2 Maximum Likelihood Estimation

Güven (1995) derived the exact maximum likelihood estimates of unknown parameters of the simple linear regression with the one-fold nested error structure. The procedure can be summarized as follows.

Consider the column vector $\mathbf{Y_j} = (Y_{j1}, Y_{j2}, ..., Y_{jK})^T$, $j = 1,2,...,J$. Then $\mathbf{Y_j}$ is distributed as a K-dimensional multivariate normal with mean vector $\mu\mathbf{1_K} + \beta\mathbf{X_j}$ where $\mathbf{X_j} = (x_{j1}, x_{j2}, ..., x_{jK})^T$ and covariance matrix $\Sigma = \sigma_e^2\mathbf{I_K} + \sigma_a^2\mathbf{J_K}$ where $\mathbf{J_K} = \mathbf{1_K}\mathbf{1_K^T}$. So, we have an independent but not identically distributed sample of size J. The likelihood function of $\mathbf{Y_1}, \mathbf{Y_2}, ..., \mathbf{Y_J}$ is given by

$$L = (2\pi)^{-JK/2}|\Sigma|^{-J/2} \exp\left(-(1/2)\sum_{j=1}^{J}(\mathbf{Y_j} - \mu\mathbf{1_K} - \beta\mathbf{X_j})^T \Sigma^{-1}(\mathbf{Y_j} - \mu\mathbf{1_K} - \beta\mathbf{X_j}) \right), \tag{2.2.1}$$

We can rewrite Σ as follows after reparametrizing the variance components by $\sigma^2 = \sigma_a^2 + \sigma_e^2$ and $\rho = \sigma_a^2 / \sigma^2$

$$\Sigma = \sigma_e^2\mathbf{I_K} + \sigma_a^2\mathbf{J_K} = \sigma^2\left((1-\rho)\mathbf{I_K} + \rho\mathbf{J_K}\right)$$

$$= \sigma^2\left((1-\rho)\left(\mathbf{I_K} - \frac{1}{K}\mathbf{J_K}\right) + (1-(K-1)\rho)\frac{1}{K}\mathbf{J_K} \right)$$

where $\left(\mathbf{I_K} - \frac{1}{K}\mathbf{J_K}\right)$ and $\frac{1}{K}\mathbf{J_K}$ are symmetric, idempotent matrices that are mutually orthogonal. Thus,

$$\Sigma^{-1} = \sigma^{-2}\left[\frac{\left(\mathbf{I_K} - \frac{1}{K}\mathbf{J_K}\right)}{(1-\rho)} + \frac{\frac{1}{K}\mathbf{J_K}}{(1+(K-1)\rho)} \right],$$

which is equal to :

$$\Sigma^{-1} = \frac{1}{(1-\rho)\sigma^2}\left(\mathbf{I_K} - \frac{\rho}{(1+(K-1)\rho)}\mathbf{J_K} \right).$$

So, the exponent term can be rewritten as :

$$\frac{1}{(1-\rho)\sigma^2}\sum_{j=1}^{J}(\mathbf{Y_j}-\mu\mathbf{1_K}-\beta\mathbf{X_j})^T\left(\mathbf{I_K}-\frac{\rho}{1+(K-1)\rho}\mathbf{J_K}\right)(\mathbf{Y_j}-\mu\mathbf{1_K}-\beta\mathbf{X_j})=$$

$$=\frac{1}{(1-\rho)\sigma^2}\left[\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk}-\mu-\beta x_{jk})^2-\frac{\rho}{(1+(K-1)\rho)}K^2\sum_{j=1}^{J}(\overline{Y}_{j.}-\mu-\beta x_{j.})^2\right]$$

(2.2.2)

Since the two sum of squares in the exponent term (2.2.2) can be expanded as follows :

$$\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk}-\mu-\beta x_{jk})^2=\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk}-\overline{Y}_{j.}-\beta(x_{jk}-\overline{x}_{j.}))^2+K\sum_{j=1}^{J}(y_{j.}-\mu-\beta x_{j.})^2$$

and

$$\frac{\rho}{(1+(K-1)\rho}K^2\sum_{j=1}^{J}(\overline{Y}_{j.}-\mu-\beta\overline{x}_{j.})^2=$$

$$=\frac{\rho}{(1+(K-1)\rho)}K^2\sum_{j=1}^{J}((\overline{Y}_{j.}-\overline{Y}_{..})-\beta(\overline{x}_{j.}-\overline{x}_{..}))^2+\frac{\rho}{(1+(K-1)\rho}JK^2(\overline{Y}_{..}-\mu-\beta\overline{x}_{..})^2$$

The exponent term (2.2.2) can be simplified to

$$\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{(Y_{jk}-\overline{Y}_{j.}-\beta(x_{jk}-\overline{x}_{j.}))^2}{(1-\rho)\sigma^2}+K\sum_{j=1}^{J}\frac{(\overline{Y}_{j.}-\overline{Y}_{..}-\beta(\overline{x}_{j.}-\overline{x}_{..}))^2}{(1+(K-1)\rho)\sigma^2}+\frac{JK(\overline{Y}_{..}-\mu-\beta\overline{x}_{..})^2}{(1+(K-1)\rho)\sigma^2}$$

(2.2.3)

Substituting (2.2.3) into it, the likelihood function given in (2.2.1) becomes :

$$L=(2\pi)^{-JK/2}(\sigma^2)^{-JK/2}(1-\rho)^{-J(K-1)/2}(1+(K-1)\rho)^{-J/2}$$

$$\times\exp\left\{(-1/2\sigma^2)\left(\frac{S_e^2(\beta)}{(1-\rho)}+\frac{S_a^2(\beta)}{(1+(K-1)\rho)}+\frac{JK(\overline{Y}_{..}-\mu-\beta\overline{x}_{..})^2}{(1+(K-1)\rho)}\right)\right\},$$

(2.2.4)

where

$$S_e^2(\beta)=\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk}-\overline{Y}_{j.}-\beta(x_{jk}-\overline{x}_{j.}))^2\ , \qquad (2.2.5)$$

and

$$S_a^2(\beta) = K\sum_{j=1}^{J}(\overline{Y}_{j.} - \overline{Y}_{..} - \beta(\overline{x}_{j.} - \overline{x}_{..}))^2 \ . \tag{2.2.6}$$

Hence, the log likelihood function $\ln L$ of the observations $\mathbf{Y_1, Y_2, ..., Y_J}$ is

$$\ln L = -(JK/2)\ln 2\pi - (JK/2)\ln\sigma^2 - (J(K-1)/2)\ln(1-\rho)$$
$$-(J/2)\ln(1+(K-1)\rho) + (-1/2\sigma^2)\left\{\frac{S_e^2(\beta)}{(1-\rho)} + \frac{S_a^2(\beta)}{(1+(K-1)\rho)} + \frac{JK(y_{..} - \mu - \beta\overline{x}_{..})^2}{(1+(K-1)\rho)}\right\}$$

.

(2.2.7)

Fixing β, ρ, and maximizing $\ln L$ with respect to μ, the maximum likelihood estimate of μ is :

$$\hat{\mu}(\beta) = \overline{Y}_{..} - \beta\overline{x}_{..} \ . \tag{2.2.8}$$

Substituting $\hat{\mu}(\beta)$ into $\ln L$ yields :

$$\ln L = -(JK/2)\ln 2\pi - (JK/2)\ln\sigma^2 - (J(K-1)/2)\ln(1-\rho)$$
$$-(J/2)\ln(1+(K-1)\rho) + (-1/2\sigma^2)\left\{\frac{S_e^2(\beta)}{(1-\rho)} + \frac{S_a^2(\beta)}{(1+(K-1)\rho)}\right\} \ .$$

Fixing β, ρ, and maximizing $\ln L$ with respect to σ^2 gives :

$$\hat{\sigma}^2(\beta,\rho) = \frac{1}{JK}\left[\frac{S_e^2(\beta)}{(1-\rho)} + \frac{S_a^2(\beta)}{(1+(K-1)\rho)}\right] \ . \tag{2.2.9}$$

Finally, substituting $\hat{\mu}(\beta)$ and $\hat{\sigma}^2(\beta)$ (given in (2.2.8) and (2.2.9)) into $\ln L$, the likelihood function becomes:

$$\ln L = -(JK/2)\ln 2\pi - (JK/2)\ln\left\{\frac{1}{JK}\left(\frac{S_e^2(\beta)}{1-\rho} + \frac{S_a^2(\beta)}{(1+(K-1)\rho)}\right)\right\}$$
$$-(J(K-1)/2)\ln(1-\rho) - (J/2)\ln(1+(K-1)\rho) - JK/2 \ .$$

Fixing β and maximizing the log likelihood function $\ln L$ with respect to ρ, after substituting $\hat{\mu}(\beta)$ and $\hat{\sigma}^2(\beta,\rho)$ into $\ln L$ yields:

$$\hat{\rho}(\beta) = \frac{S_a^2(\beta) - \dfrac{S_e^2(\beta)}{K-1}}{S_T^2(\beta)} \tag{2.2.10}$$

where $S_T^2(\beta) = S_a^2(\beta) + S_e^2(\beta) = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \overline{Y}_{..} - \beta(x_{jk} - \overline{x}_{..}))^2$, $S_a^2(\beta)$ and $S_e^2(\beta)$ are given in (2.2.5) and (2.2.6).

When $0 < \hat{\rho}(\beta) < 1$ we take $\hat{\rho}(\beta)$ as the maximum likelihood estimate of ρ instead of 0 since $\ln L$ is greater than 0 at $\hat{\rho}(\beta)$, which can be shown as follows;

$$\ln L = -(JK/2)\ln 2\pi - (JK/2)\ln\left(\frac{S_T^2(\beta)}{JK}\right) - (JK/2) \quad \text{when } \hat{\rho}(\beta) = 0$$

$$\ln L = -(JK/2)\ln 2\pi - (JK/2)\ln\left(\frac{S_T^2(\beta)}{JK}\right) - (JK/2) - \frac{J(K-1)}{2}\ln(1-\hat{\rho}(\beta))$$

$$-\frac{J}{2}\ln(1+(K-1)\hat{\rho}(\beta)) \quad \text{when } \hat{\rho}(\beta) > 0\,.$$

The difference between the likelihood function $\ln L$ at $\hat{\rho}(\beta)$ and the likelihood function $\ln L$ at 0 is

$$-\frac{J(K-1)}{2}\ln(1-\hat{\rho}(\beta)) - \frac{J}{2}\ln(1+(K-1)\hat{\rho}(\beta))\,.$$

Consider a function

$$d(\rho) = -\frac{J(K-1)}{2}\ln(1-\rho) - \frac{J}{2}\ln(1+(K-1)\rho) \qquad 0 < \rho < 1,$$

The derivative of $d(\rho)$ is

$$d'(\rho) = -\frac{J(K-1)}{2}\left(\frac{K\rho}{(1-\rho)(1+(K-1)\rho)}\right) > 0 \qquad \text{for } 0 < \rho < 1\,.$$

So it has been shown that the difference between the likelihood function $\ln L$ at $\hat{\rho}(\beta)$ and the likelihood function $\ln L$ at 0 is an increasing function of $\hat{\rho}(\beta)$, indicating that $\ln L(0) \le \ln L(\hat{\rho}(\beta))$. Since the likelihood function $\ln L$ is maximized at $\rho = 0$ when $\hat{\rho}(\beta) < 0$, Güven (1995) took $\hat{\rho}(\beta)$ as follows :

$$\hat{\rho}(\beta) = \max\left(0, \frac{S_a^2(\beta) - \frac{S_e^2(\beta)}{K-1}}{S_T^2(\beta)}\right). \qquad (2.2.11)$$

By substituting $\hat{\rho}(\beta)$ into $\hat{\sigma}^2(\beta,\rho)$ given in (2.2.9) when $\hat{\rho}$ is zero will yield :

$$\hat{\sigma}^2(\beta) = \frac{S_T^2(\beta)}{JK}$$

It is obvious that if $\hat{\rho} > 0$, substituting $\hat{\rho}(\beta)$ into $\hat{\sigma}^2(\beta,\rho)$ will yield the same result since;

$$\hat{\sigma}^2(\beta) = \frac{1}{JK}\left(\frac{S_e^2(\beta)}{1-\left(\frac{S_a^2(\beta)-S_e^2(\beta)/(K-1)}{S_T^2(\beta)}\right)} + \frac{S_a^2(\beta)}{1+(K-1)\left(\frac{S_a^2(\beta)-S_e^2(\beta)/(K-1)}{S_T^2(\beta)}\right)} \right)$$

$$= \frac{S_e^2(\beta)S_T^2(\beta)}{JKS_e^2(\beta)(1+1/(K-1))} + \frac{S_a^2(\beta)S_T^2(\beta)}{JK^2S_a^2(\beta)} = \frac{S_T^2(\beta)}{JK} \ .$$

Utilizing the information that $\sigma_a^2 = \rho\sigma^2$ and $\sigma_e^2 = (1-\rho)\sigma^2$, Güven (1995) derived the maximum likelihood estimators of σ_a^2 and $\sigma_e^2 = (1-\rho)\sigma^2$ by using the invariance property of the maximum likelihood estimation :

$$\hat{\sigma}_a^2(\beta) = \hat{\rho}(\beta)\hat{\sigma}^2(\beta) = \begin{cases} 0 & \text{if} \quad S_a^2(\beta) \le S_e^2/K-1 \\ (1/JK)\left[S_a^2(\beta)-S_e^2(\beta)/(K-1)\right] & \text{if} \quad S_a^2(\beta) > S_e^2/K-1 \end{cases} \tag{2.2.12}$$

$$\hat{\sigma}_e^2(\beta) = (1-\hat{\rho}(\beta))\hat{\sigma}^2(\beta) = \begin{cases} S_T^2(\beta)/JK & \text{if} \quad \hat{\sigma}_a^2(\beta) = 0 \\ S_e^2(\beta)/(J(K-1)) & \text{if} \quad \hat{\sigma}_a^2(\beta) > 0 \end{cases} \tag{2.2.13}$$

Substituting $\hat{\mu}(\beta)$, $\hat{\rho}(\beta)$ and $\hat{\sigma}^2(\beta)$ into the log likelihood function $\ln L$ given in (2.2.7) will give :

$$L_1(\beta) = -(JK/2)(1+\ln 2\pi) - (J/2)\ln\left(\frac{S_a^2(\beta)}{J}\right) - (J(K-1)/2)\ln\left(\frac{S_e^2(\beta)}{J(K-1)}\right)$$

if $S_a^2(\beta) > S_e^2(\beta)/(K-1)$

$$L_2(\beta) = -(JK/2)(1+\ln 2\pi) - (JK/2)\ln\left(\frac{S_T^2(\beta)}{JK}\right)$$

if $S_a^2(\beta) \le S_e^2(\beta)/(K-1)$

(2.2.14)

When $S_a^2(\beta) \le S_e^2(\beta)/(K-1)$, maximizing $L_2(\beta)$ with respect to β yields;

$$\hat{\beta}_T = \frac{S_{xyt}}{S_{xxt}}$$

where $S_{xyt} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk}-\bar{x}_{..})(Y_{jk}-\bar{Y}_{..})$, $S_{xxt} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk}-\bar{x}_{..})^2$

When $S_a^2(\beta) > S_e^2(\beta)/(K-1)$, maximizing $L_1(\beta)$ with respect to β gives the following derivative equation;

$$f(\beta) = \frac{J(S_{xya} - \beta S_{xxa})}{S_a^2(\beta)} + \frac{J(K-1)(S_{xye} - \beta S_{xxe})}{S_e^2(\beta)} = 0,$$

where $S_{xya} = K\sum_{j=1}^{J}(\bar{x}_{j.} - \bar{x}_{..})(\bar{Y}_{j.} - \bar{Y}_{..})$, $S_{xye} = \sum_{j=1}^{J}\sum_{k=1}^{K}(x_{jk} - \bar{x}_{j.})(\bar{Y}_{jk} - \bar{Y}_{j.})$.

Multiplying both sides of the derivative equation by $\dfrac{S_a^2(\beta)S_e^2(\beta)}{J}$, Güven(1995) obtained;

$$\begin{aligned} g(\beta) &= S_a^2 S_e^2 f(\beta)/J = (S_{xya} - \beta S_{xxa})S_e^2(\beta) + (K-1)(S_{xye} - \beta S_{xxe})S_a^2(\beta) \\ &= -KS_{xxe}S_{xxa}\beta^3 + \big((K+1)S_{xye}S_{xxa} + (2K-1)S_{xxe}S_{xya}\big)\beta^2 \\ &\quad - \big(S_{xxa}S_{yye} + 2KS_{xya}S_{xye} + (K-1)S_{xxe}S_{yya}\big)\beta + \big(S_{yye}S_{xya} + (K-1)S_{yya}S_{xye}\big) = 0 \end{aligned} \tag{2.2.15}$$

The previous equation is a third degree polynomial equation in β. One of the real roots of this equation which maximizes $L_1(\beta)$ is the maximum likelihood estimate of β. When $S_{xxe}S_{xxa} > 0$, dividing the polynomial equation into the coefficient of β^3 and multiplying it by –1, obtain the following polynomial equation can be obtained :

$$\begin{aligned} P(\beta) &= -\frac{g(\beta)}{KS_{xxa}S_{xxe}} = \beta^3 - \left(\frac{(K+1)\hat{\beta}_e + (2K-1)\hat{\beta}_a}{K}\right)\beta^2 \\ &+ \left(\frac{\gamma_e + 2K\hat{\beta}_e\hat{\beta}_a + (K-1)\gamma_a}{K}\right)\beta - \left(\frac{\gamma_e\hat{\beta}_a + (K-1)\gamma_a\hat{\beta}_e}{K}\right) = 0 \end{aligned} \tag{2.2.16}$$

where $\hat{\beta}_a = \dfrac{S_{xya}}{S_{xxa}}$, $\hat{\beta}_e = \dfrac{S_{xye}}{S_{xxe}}$, $\gamma_a = \dfrac{S_{yya}}{S_{xxa}}$ and $\gamma_e = \dfrac{S_{yye}}{S_{xxe}}$.

Güven (1995) developed an algorithm to sort out to the roots of $P(\beta)$ in order to find the maximum likelihood estimate of β.

CHAPTER 3

ESTIMATION TECHNIQUES FOR THE VARIANCE OF PRIMARY UNITS

In this chapter, Minimum Norm Quadratic Estimator (MINQE) (Rao & Chaubey, 1978), Almost Unbiased Estimator (AUE) (Horn & Horn, 1975) and Restricted Maximum Likelihood Estimator (REML) (Corbeil & Searle, 1976) of the variance of primary units, σ_a^2 are derived in the simple linear regression model with the one-fold nested error structure. Among the different variance component estimators summarized by Searle (1977), we choose these four estimators since the exact mean squared errors (MSEs) of them can be derived. In the last section of this chapter, these estimators are compared with respect to their Mean Squared Errors (MSEs).

3.1 Minimum Norm Quadratic Estimator (MINQE)

Rao and Chaubey (1978) introduced the principle of MINQE for the variance components of a linear model, where the model is

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

and $\mathbf{Y}$ is an n - vector of observations, $\mathbf{X}$ is a known nxm matrix, $\boldsymbol{\beta}$ is an m -vector of fixed-effect parameters. Here $\boldsymbol{\varepsilon}$ is of the form :

$$\boldsymbol{\varepsilon} = \mathbf{U_1}\boldsymbol{\xi_1} + \mathbf{U_2}\boldsymbol{\xi_2} + \ldots + \mathbf{U_K}\boldsymbol{\xi_K}.$$

and $\mathbf{U_i}$ is an nxn_i matrix of known constants, $\boldsymbol{\xi_i}$ is an n_i -vector with mean zero and dispersion $\sigma_i^2\mathbf{I}$. Thus, $\boldsymbol{\varepsilon}$ has zero mean and dispersion :

$$\mathbf{\Lambda} = \sigma_1^2 \mathbf{V}_1 + \sigma_2^2 \mathbf{V_2} + ... + \sigma_k^2 \mathbf{V_k} \text{ with } \mathbf{V_i} = \mathbf{U_i} \mathbf{U}_i'.$$

Rao and Chaubey (1978) derived the MINQE for the i-th variance component σ_i^2 in the following form :

$$\hat{\sigma}_i^2 = \frac{\alpha_i^4 \mathbf{e}' \mathbf{V_i} \mathbf{e}}{n_i}, \tag{3.1.1}$$

where α_i is a priori weight, $\mathbf{e} = \mathbf{RY}$, $\mathbf{R} = \mathbf{WQ}$, $\mathbf{W} = \mathbf{V}_*^{-1}$, $\mathbf{V}_* = \alpha_1^2 \mathbf{V}_1 + \alpha_2^2 \mathbf{V_2} + ... + \alpha_k^2 \mathbf{V_k}$ and $\mathbf{Q} = \mathbf{I} - \mathbf{P}$, $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^- \mathbf{X}'\mathbf{W}$.

For our model, we can write ε as follows :

$$\boldsymbol{\varepsilon} = \mathbf{U_1} \boldsymbol{\xi_1} + \mathbf{U_2} \boldsymbol{\xi_2} = (\mathbf{I_J} \otimes \mathbf{1_K}) \boldsymbol{\xi_1} + \mathbf{I_{JK}} \boldsymbol{\xi_2}.$$

We know that ε has zero mean and dispersion

$$\mathbf{\Lambda} = \sigma_a^2 \mathbf{V_1} + \sigma_e^2 \mathbf{V_2} = \sigma_a^2 (\mathbf{I_J} \otimes \mathbf{J_K}) + \sigma_e^2 \mathbf{I_{JK}}.$$

Thus, using (3.1.1), we can derive the MINQE of σ_a^2 given by :

$$\hat{\sigma}_a^2 = \frac{\sigma_{a,0}^4}{J} \mathbf{Y}' \mathbf{R}' (\mathbf{I_J} \otimes \mathbf{J_K}) \mathbf{RY}, \tag{3.1.2}$$

where $\sigma_{a,0}^2$ is the a priori information and $\mathbf{R} = \mathbf{\Lambda_0^{-1}} (\mathbf{I} - \mathbf{P})$ with,

$$\mathbf{\Lambda_0^{-1}} = \frac{1}{\sigma_{e,0}^2} \mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{\mathbf{1}}{\mathbf{K}} \mathbf{J_K} \right) + \frac{1}{\sigma_{e,0}^2 (1 + K\rho)} \mathbf{I_J} \otimes \frac{\mathbf{1}}{\mathbf{K}} \mathbf{J_K} \text{ and}$$

$\mathbf{P} = \mathbf{X}(\mathbf{X}' \mathbf{\Lambda_0^{-1}} \mathbf{X})^- \mathbf{X}' \mathbf{\Lambda_0^{-1}}$. Here ρ is the ratio of the variance components defined as $\rho = \dfrac{\sigma_{a,0}^2}{\sigma_{e,0}^2}$.

After simplifying (3.1.2), we find the MINQE of σ_a^2 as follows:

$$\hat{\sigma}_a^2 = \frac{\rho^2}{J(1+K\rho)^2} K^2 \sum_{j=1}^{J} (Y_{J.} - \hat{Y}_{J.})^2 = \frac{K\rho^2}{J(1+K\rho)^2} S_a^2(\hat{\beta}_{GLS}) \tag{3.1.3}$$

where $\hat{\beta}_{GLS}$ is the generalized least squares estimator of β given in (2.1.31) and $S_a^2(\hat{\beta}_{GLS}) = K \sum_{j=1}^{J} (\overline{Y}_{j.} - \overline{Y}_{..} + \hat{\beta}_{GLS} (\overline{x}_{j.} - \overline{x}_{..}))^2$.

To evaluate the expected value and the variance of $\hat{\sigma}_a^2$ for finding the MSE of it, we use a decomposition of $S_a^2(\hat{\beta}_{GLS})$. It is possible to write $S_a^2(\hat{\beta}_{GLS})$ as a linear combination of R_A and R_1 as follows :

$$S_a^2(\hat{\beta}_{GLS}) = R_A + g(\rho)R_1 \tag{3.1.4}$$

where $R_A = S_{yya} - S_{xya}^2 / S_{xxa}$, $R_1 = S_{xxa} / S_{xxe}\left[(S_{xxe} / S_{xxa})S_{xya} - S_{xye}\right]^2$, $g(\rho) = (1+K\rho)^2 S_{xxe} /(S_{xxa} + (1+K\rho)S_{xxe})^2$.

The assumption that the error terms a_j and e_{jk} are independent normal variables with zero means and variances σ_a^2 and σ_e^2 respectively yields that $R_A \sim (\sigma_e^2 + K\sigma_a^2)\chi_{J-2}^2$ and $R_1 \sim [(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_e^2 S_{xxa}]\chi_1^2$. As a result of the two facts, the expectation and variance of $S_a^2(\hat{\beta}_{GLS})$ are :

$$E(S_a^2(\hat{\beta}_{GLS})) = E(R_A) + g(\rho)E(R_1), \tag{3.1.5}$$

$$V(S_a^2(\hat{\beta}_{GLS})) = V(R_A) + (g(\rho))^2 V(R_1) \tag{3.1.6}$$

where $E(R_A) = (\sigma_e^2 + K\sigma_a^2)(J-2)$, $V(R_A) = (\sigma_e^2 + K\sigma_a^2)^2 2(J-2)$,

$$E(R_1) = [(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_e^2 S_{xxa}],\ V(R_1) = 2[(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_e^2 S_{xxa}]^2 . \tag{3.1.7}$$

Thus, we can state the MSE of $\hat{\sigma}_a^2$ as follows :

$$MSE(\hat{\sigma}_a^2) = V(\hat{\sigma}_a^2) + (E(\hat{\sigma}_a^2) - \sigma_a^2)^2$$

$$= \frac{K^2\rho^4}{J^2(1+K\rho)^4}(V(R_A) + (g(\rho))^2 V(R_1)) + \left[\frac{K\rho^2}{J(1+K\rho)^2}\left((E(R_A) + g(\rho)E(R_1)\right) - \sigma_a^2\right]^2 \tag{3.1.8}$$

where $g(\rho) = (1+K\rho)^2 S_{xxe} /(S_{xxa} + (1+K\rho)S_{xxe})^2$, $E(R_A)$, $V(R_A)$, $E(R_1)$ and $V(R_1)$ are given in (3.1.7).

3.2 Almost Unbiased Estimator (AUE)

Horn, Horn and Duncan (1975) and Horn and Horn (1975) defined the AUE for the variance component σ_i^2 as :

$$\tilde{\sigma}_i^2 = \frac{\alpha_i}{\mathbf{trRV_i}} \mathbf{Y'RV_iRY} \tag{3.2.1}$$

where α_i is a priori weight, $\mathbf{e} = \mathbf{RY}$, $\mathbf{R} = \mathbf{WQ}$, $\mathbf{W} = \mathbf{V}_*^{-1}$, $\mathbf{V}_* = \alpha_1^2\mathbf{V_1} + \alpha_2^2\mathbf{V_2} + ... + \alpha_k^2\mathbf{V_k}$, $\mathbf{Q} = \mathbf{I} - \mathbf{P}$, $\mathbf{P} = \mathbf{X(X'WX)^-X'W}$, $\mathbf{V_i}$ is defined as in the previous section. The superscript "-" denotes the generalized inverse of any matrix. For our model, we can derive the AUE of the variance of primary units. It is given by:

$$\tilde{\sigma}_a^2 = \frac{\sigma_{a,0}^2}{\mathbf{trR(I_J \otimes J_K)}} \mathbf{Y'RI_J \otimes J_K RY} \tag{3.2.2}$$

where $\sigma_{a,0}^2$ is the a priori information to σ_a^2, and $\mathbf{R} = \mathbf{\Lambda_0^{-1}(I - P)}$ with

$$\mathbf{\Lambda_0^{-1}} = \frac{1}{\sigma_{e,0}^2}\mathbf{I_J} \otimes \left(\mathbf{I_K} - \frac{\mathbf{1}}{\mathbf{K}}\mathbf{J_K}\right) + \frac{1}{\sigma_{e,0}^2(1+K\rho)}\mathbf{I_J} \otimes \frac{\mathbf{1}}{\mathbf{K}}\mathbf{J_K}\,,\ \rho = \frac{\sigma_{a,0}^2}{\sigma_{e,0}^2}$$

and $\mathbf{P} = \mathbf{X(X'\Lambda_0^{-1}X)^-X'\Lambda_0^{-1}}$.

After some algebra, the expression (3.2.2) is simplified to :

$$\tilde{\sigma}_a^2 = \frac{\rho}{1+K\rho}\frac{S_a^2(\hat{\beta}_{GLS})}{J-1-C} \tag{3.2.3}$$

where $C = \dfrac{S_{xxa}}{S_{xxa} + (1+K\rho)S_{xxe}}$, $\hat{\beta}_{GLS}$ is the generalized least squares estimator of β given in (2.1.31) and $S_a^2(\hat{\beta}_{GLS}) = K\sum_{j=1}^{J}(\bar{Y}_{j.} - \bar{Y}_{..} + \hat{\beta}_{GLS}(\bar{x}_{j.} - \bar{x}_{..}))^2$.

To evaluate the expected value and the MSE of $\tilde{\sigma}_a^2$, we use a decomposition of $S_a^2(\hat{\beta}_{GLS})$ given in (3.1.4). Using (3.1.4) with (3.1.5) and (3.1.6), the $MSE(\tilde{\sigma}_a^2)$ is written as

$$MSE(\tilde{\sigma}_a^2) = V(\tilde{\sigma}_a^2) + (E(\tilde{\sigma}_a^2) - \sigma_a^2)^2 = \frac{\rho^2}{(1+K\rho)^2 (J-1-C)^2}(V(R_A) + (g(\rho))^2 V(R_1)) +$$

$$+\left[\frac{\rho}{(1+K\rho)(J-1-C)}\left((E(R_A) + g(\rho)E(R_1)\right) - \sigma_a^2\right]^2, \tag{3.2.4}$$

where $g(\rho) = (1+K\rho)^2 S_{xxe} / (S_{xxa} + (1+K\rho)S_{xxe})^2$, $C = \dfrac{S_{xxa}}{S_{xxa} + (1+K\rho)S_{xxe}}$, and $E(R_A)$, $V(R_A)$, $E(R_1)$ and $V(R_1)$ are given in (3.1.7).

3.3 Restricted Maximum Likelihood Estimator (REML)

Corbeil and Searle(1976) studied the restricted maximum likelihood estimation of variance components in the mixed model and obtained the following estimator equations for r-variance components :

$$\left\{\mathbf{tr}\left(\mathbf{\Lambda}^{-1}\mathbf{V_i V_i'}\right)\right\}_{\mathbf{i=0}}^{\mathbf{r}} = \left\{\mathbf{Y'RV_i V_i'RY}\right\}_{\mathbf{i=0}}^{\mathbf{r}} \tag{3.3.1}$$

where $\mathbf{\Lambda}$ is the variance structure of $\mathbf{Y}$, r is the number of variance components, $\mathbf{R} = \mathbf{\Lambda}^{-1}(\mathbf{I} - \mathbf{P})$, $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{\Lambda}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{\Lambda}^{-1}$ and V_i is defined in section 3.1. We have extended these results for our model as follows :

If we write the estimator equations given in (3.3.1) for our model, they will be

$$tr(\mathbf{\Lambda}^{-1}\mathbf{I_J} \otimes \mathbf{J_K}) = \mathbf{Y'RI_J} \otimes \mathbf{J_K RY} \tag{3.3.2}$$

$$tr(\mathbf{\Lambda}^{-1}\mathbf{I_{JK}}) = \mathbf{Y'RI_{JK}RY} \tag{3.3.3}$$

The expressions (3.3.2) and (3.3.3) are simplified to:

$$\frac{JK}{\hat{\sigma}_e^2 + \hat{\sigma}_a^2} = \frac{K}{(\hat{\sigma}_e^2 + K\hat{\sigma}_a^2)} S_a^2(\hat{\beta}_{GLS}), \tag{3.3.4}$$

and

$$\frac{J(K-1)}{\hat{\sigma}_e^2} + \frac{J}{\hat{\sigma}_e^2 + K\hat{\sigma}_a^2} = \frac{1}{\hat{\sigma}_e^4} S_e^2(\hat{\beta}_{GLS}) + \frac{1}{(\hat{\sigma}_e^2 + K\hat{\sigma}_a^2)^2} S_a^2(\hat{\beta}_{GLS}), \tag{3.3.5}$$

respectively where $S_a^2(\hat{\beta}_{GLS}) = K\sum_{j=1}^{J}(\overline{Y}_{j.} - \overline{Y}_{..} + \hat{\beta}_{GLS}(\overline{x}_{j.} - \overline{x}_{..}))^2$,

$S_e^2(\hat{\beta}_{GLS}) = \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{jk} - \overline{Y}_{j.} - \hat{\beta}_{GLS}(x_{jk} - \overline{x}_{j.}))^2$ and $\hat{\beta}_{GLS}$ is the generalized least squares estimator of β given in (2.1.31). Solving these equations, we find the REML of σ_a^2 as follows :

$$\hat{\sigma}_a^2 = \frac{1}{K}\left(\frac{S_a^2(\hat{\beta}_{GLS})}{J} - \frac{S_e^2(\hat{\beta}_{GLS})}{J(K-1)}\right). \qquad (3.3.6)$$

To evaluate the expected value and the MSE of $\tilde{\sigma}_a^2$, we use the decompositions of $S_a^2(\hat{\beta}_{GLS})$ and $S_e^2(\hat{\beta}_{GLS})$ where the decomposition of $S_a^2(\hat{\beta}_{GLS})$ is given in (3.1.4). Moreover, it is possible to write $S_e^2(\hat{\beta}_{GLS})$ as a linear combination of R_E and R_1 as shown :

$$S_e^2(\hat{\beta}_{GLS}) = R_E + f(\rho)R_1 \qquad (3.3.7)$$

where $R_E = S_{yye} - S_{xye}^2 / S_{xxe}$, $R_1 = S_{xxa} / S_{xxe}\left[(S_{xxe} / S_{xxa})S_{xya} - S_{xye}\right]^2$ and $f(\rho) = S_{xxa} / (S_{xxa} + (1 + K\rho)S_{xxe})^2$.

The assumption that the error terms a_j and e_{jk} are independent normal variables with zero means and variances σ_a^2 and σ_e^2 respectively yields that $R_E \sim \sigma_e^2\chi_{JK-J-1}^2$ and $R_1 \sim [(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_a^2 S_{xxa}]\chi_1^2$. As a result of the two facts, the expectation and variance of $S_e^2(\hat{\beta}_{GLS})$ will be :

$$E(S_e^2(\hat{\beta}_{GLS})) = E(R_E) + f(\rho)E(R_1), \qquad (3.3.8)$$

$$V(S_E^2(\hat{\beta}_{GLS})) = V(R_E) + (f(\rho))^2 V(R_1) \qquad (3.3.9)$$

where $E(R_E) = \sigma_e^2(JK - J - 1)$, $V(R_E) = 2(JK - J - 1)\sigma_e^4$ (3.3.10)

$$E(R_1) = [(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_e^2 S_{xxa}],\ V(R_1) = 2[(\sigma_e^2 + K\sigma_a^2)S_{xxe} + \sigma_e^2 S_{xxa}]^2$$

Thus, we obtain the MSE (of $\hat{\sigma}_a^2$) where it is :

$$MSE(\hat{\sigma}_a^2) = V(\hat{\sigma}_a^2) + (E(\hat{\sigma}_a^2) - \sigma_a^2)^2 \qquad (3.3.11)$$

with

$$V(\hat{\sigma}_a^2) = \frac{V(S_a^2(\hat{\beta}_{GLS}))}{J^2K^2} + \frac{V(S_e^2(\hat{\beta}_{GLS}))}{J^2K^2(K-1)^2} - \frac{2}{J^2K^2(K-1)} Cov(S_a^2(\hat{\beta}_{GLS}), \quad S_e^2(\hat{\beta}_{GLS}))$$

$$= \frac{1}{J^2K^2(K-1)^2}\left[(K-1)^2 V(R_A) + V(R_E) + ((K-1)g(\rho) - f(\rho))^2 V(R_1)\right],$$

and

$$E(\hat{\sigma}_a^2) = \frac{E(R_A) + g(\rho)E(R_1)}{JK} - \frac{E(R_E) + f(\rho)E(R_1)}{JK(K-1)}.$$

where $E(R_A)$, $V(R_A)$, $E(R_1)$ and $V(R_1)$ are given in (3.1.7), $E(R_E)$ and $V(R_E)$ are given in (3.3.10).

3.4 Comparison of the Estimators for the Primary Unit Variance

To compare the three estimators (MINQE, AUE and REML) of the variance of primary units, we calculated the exact MSE's given in (3.1.8), (3.2.4) and (3.3.11) along with the ANOVA estimator of σ_a^2 , $\left(\frac{R_A}{J-2} - \frac{R_E}{JK-J-1}\right)$ under various values of the following quantities by the computer programs listed in Appendix B:

1) Values considered for (S_{xxa}, S_{xxe}) pairs were (0.83,0.17), (0.66,0.34), (0.34,0.66) and (0.17,0.83).

2) The ratio of variance components $\rho = \frac{\sigma_a^2}{\sigma_e^2}$ were taken as 0.1, 0.5, 1, 2 and 5.

3) The pair (J,K) indicating both the number of primary sampling units and the number of secondary sampling units is (4,7) for all cases.

As it is expected, Table 3.4.1 indicates that the MINQE of σ_a^2 has the superiority over the other presented three estimators of σ_a^2. Both the REML estimator and the AUE of σ_a^2 have smaller MSE than the ANOVA estimator of σ_a^2. Therefore, it is concluded that both the REML estimator and the AUE have the superority over the ANOVA estimator.

When ρ is equal to 0.1 and 0.5, the MSE of the AUE of σ_a^2 is smaller than the MSE of the REML estimator of σ_a^2. However this statement is reversed when ρ is equal to 1, 2 and 5. It yields that the AUE has the superiority over the REML estimator for small values of ρ. But for large values of ρ, the REML has the superiority over the AUE.

Table 3.4.1 : MSE values of ANOVA, MINQE, REML and AUE estimates of σ_a^2 for given values of ρ , S_{xxa} and S_{xxe} .

		ρ =0,1	ρ =0,5	ρ =1	ρ =2	ρ =5
S_{xxa} =0,83 S_{xxe} =0,17	ANOVA	0.0607	0.4150	1.3078	4.5936	26.4507
	MINQE	0.0063	0.1091	0.4103	1.6296	10.4459
	REML	0.0274	0.1742	0.5418	1.9131	11.2306
	AUE	0.0081	0.1713	0.6945	2.7090	16.7239
S_{xxa} =0,66 S_{xxe} =0,34	ANOVA	0.0607	0.4150	1.3078	4.5936	26.4507
	MINQE	0.0060	0.1035	0.4027	1.6395	10.5729
	REML	0.0257	0.1713	0.5462	1.9485	11.4001
	AUE	0.0072	0.1708	0.6731	2.6750	16.6767
S_{xxa} =0,34 S_{xxe} =0,66	ANOVA	0.0607	0.4150	1.3078	4.5936	26.4507
	MINQE	0.0056	0.1006	0.4033	1.6570	10.6591
	REML	0.0253	0.1757	0.5603	1.9857	11.5117
	AUE	0.0067	0.1670	0.6672	2.6673	16.6674
S_{xxa} =0,17 S_{xxe} =0,83	ANOVA	0.0607	0.4150	1.3078	4.5936	26.4507
	MINQE	0.0055	0.1002	0.4043	1.66259	10.6805
	REML	0.0259	0.1782	0.5656	1.9964	11.5390
	AUE	0.0066	0.1667	0.6667	2.6667	16.6667

CHAPTER 4

CONFIDENCE INTERVALS

In this chapter, we describe the exact and approximate confidence intervals for the four unknown parameters $(\mu, \beta, \sigma_a^2, \sigma_e^2)$ of the model. In addition to the two exact confidence intervals given by Park and Burdick (1994), an exact confidence interval for β is constructed. Also, we constructed one approximate and one exact confidence interval for μ by obtaining two unbiased estimators of μ where one is depending on $\hat{\beta}_a$ and the other is depending on $\hat{\beta}_e$, and $\hat{\beta}_a$ and $\hat{\beta}_e$ are given in (2.1.8) and (2.1.9) respectively . Also, the confidence intervals for σ_a^2 and σ_e^2 are given.

4.1 Confidence Intervals for β

Park and Burdick (1994) derived the exact confidence interval for β by usng the best linear unbiased estimator (BLUE) of β, which was obtained by Tong and Cornelius (1989).

$$\hat{\beta}_{BLU} = \frac{\phi S_{xya} + S_{xye}}{\phi S_{xxa} + S_{xxe}}$$

where $\phi = \sigma_e^2 / (\sigma_e^2 + K\sigma_a^2)$, S_{xxa}, S_{xya}, S_{xxe} and S_{xye} are given in (1.2.2)-(1.2.6). Let $\omega = \phi S_{xxa} / (\phi S_{xxa} + S_{xxe})$, then $\hat{\beta}_{BLU}$ can be written as the convex combination of two unbiased estimators, $\hat{\beta}_a$ and $\hat{\beta}_e$ of $\hat{\beta}$ as follows :

$$\hat{\beta}_{BLU} = \omega\hat{\beta}_a + (1-\omega)\hat{\beta}_e$$

Since $\hat{\beta}_a \sim N\left(\beta,\ \ (\sigma_e^2 + K\sigma_a^2)/S_{xxa}\right)$ and $\hat{\beta}_e \sim N(\beta,\ \ \sigma_e^2/S_{xxe})$ where $\hat{\beta}_a$ and $\hat{\beta}_e$ are independent, it is obvious that $\hat{\beta}_{BLU}$ is distributed as normal distribution with mean β and the following variance :

$$V(\hat{\beta}_{BLU}) = \omega^2\left(\frac{\sigma_e^2}{\phi S_{xxa}}\right) + (1-\omega)^2\left(\frac{\sigma_e^2}{S_{xxe}}\right) = \frac{\sigma_e^2}{\phi S_{xxa} + S_{xxe}}.$$

Assuming that ϕ is known and utilizing $S_B^2 = \dfrac{R_B}{JK-2}$ as the estimator of σ_e^2, where $R_B = \phi S_a^2(\hat{\beta}_a) + \left(\phi S_d^2/(r^2(\phi-1)+1)\right) + S_e^2(\hat{\beta}_e)$. Here r^2 is defined as $r^2 = S_{xxa}/(S_{xxa} + S_{xxe})$, $S_a^2(\hat{\beta}_a)$, S_d^2 and $S_e^2(\hat{\beta}_e)$ are given in (2.1.6), (2.1.7) and (2.1.18) respectively, Park and Burdick (1994) derived the following $(1-2\alpha)$ two-sided exact confidence interval

$$\hat{\beta}_{BLU} \pm t_{\alpha:JK-2}\sqrt{\frac{S_B^2}{\phi S_{xxa} + S_{xxe}}}. \tag{4.1.1}$$

Also, the length of this interval can be stated as follows :

$$L_1 = 2t_{\alpha:JK-2}\sqrt{\frac{S_B^2}{\phi S_{xxa} + S_{xxe}}}.$$

Using $\hat{\beta}_e \sim N(\beta, \sigma_e^2/S_{xxe})$ and $S_e^2(\hat{\beta}_e) \sim \sigma_e^2\chi^2_{J(K-1)-1}$, Park and Burdick (1994) derived another $(1-2\alpha)$ two-sided exact confidence interval by estimating σ_e^2 by S_E^2 :

$$\hat{\beta}_e \pm t_{\alpha:J(K-1)-1}\sqrt{\frac{S_E^2}{S_{xxe}}}\quad, \tag{4.1.2}$$

where $S_E^2 = \dfrac{S_e^2(\hat{\beta}_e)}{J(K-1)-1}$, $\hat{\beta}_e$ and $S_e^2(\hat{\beta}_e)$ are given in (2.1.9) and (2.1.18) respectively. The length of this confidence interval is as follows :

$$L_2 = 2t_{\alpha:J(K-1)-1}\sqrt{\frac{S_E^2}{S_{xxe}}}.$$

Moreover, we modify the preceding technique by using $\hat{\beta}_a \sim N(\beta, (\sigma_e^2 + K\sigma_a^2)/S_{xxa})$ and $S_a^2(\hat{\beta}_a) \sim (\sigma_e^2 + K\sigma_a^2)\chi^2_{J-2}$ and obtain the following interval is :

$$\hat{\beta}_a \pm t_{\alpha:J-2}\sqrt{\frac{S_A^2}{S_{xxa}}}\,, \tag{4.1.3}$$

where $S_A^2 = \dfrac{S_a^2(\hat{\beta}_a)}{J-2}$, where $\hat{\beta}_a$ and $S_a^2(\hat{\beta}_a)$ is given in (2.1.8) and (2.1.6) respectively. Also we can state the length of this interval as follows :

$$L_3 = 2t_{\alpha:J-2}\sqrt{\frac{S_A^2}{S_{xxa}}}$$

4.2 Confidence Intervals for μ

By using $\hat{\beta}_a$ and $\hat{\beta}_e$ of the estimator of β seperately, we can construct two different confidence intervals for μ.

4.2.1 By using $\hat{\beta}_e$ as the estimator of β

The unbiased estimator of μ, depending on $\hat{\beta}_e$, is of the form :

$$\hat{\mu}_e = \overline{Y}_{..} - \hat{\beta}_e \overline{x}_{..} \tag{4.2.1.1}$$

where $\hat{\beta}_e$ is given in (2.1.9). Observe that

$$Cov(\overline{Y}_{..}, \hat{\beta}_e) = Cov\left(\frac{1}{JK}\mathbf{1}_{\mathbf{JK}}^{\mathbf{T}}\mathbf{Y}, \frac{\mathbf{X}^{\mathbf{T}}\mathbf{QY}}{\mathbf{X}^{\mathbf{T}}\mathbf{QX}}\right) = \frac{1}{JK}(K\sigma_a^2 + \sigma_e^2)\frac{\mathbf{1}_{\mathbf{JK}}^{\mathbf{T}}\mathbf{QX}}{\mathbf{X}^{\mathbf{T}}\mathbf{QX}} = 0$$

where $\mathbf{Q} = \mathbf{I_J} \otimes \mathbf{P}$, $\mathbf{P} = \mathbf{I_K} - \dfrac{\mathbf{1}}{\mathbf{K}}\mathbf{J_K}$ and $\mathbf{1}_{\mathbf{JK}}^{\mathbf{T}}\mathbf{Q} = \mathbf{0}$. Thus, the variance of (4.2.1.1) is

$$V(\hat{\mu}_e) = \frac{\sigma_e^2 + K\sigma_a^2}{JK} + \overline{x}_{..}^2\,\frac{\sigma_e^2}{S_{xxe}},$$

since $V(\overline{Y}_{..}) = \dfrac{\sigma_e^2 + K\sigma_a^2}{JK}$ and $V(\hat{\beta}_e) = \dfrac{\sigma_e^2}{S_{xxe}}$.

We now find the standard error of $\hat{\mu}_e = \overline{Y}_{..} - \hat{\beta}_e \overline{x}_{..}$ Since $\dfrac{S_a^2(\hat{\beta}_a)}{J-2}$ and $\dfrac{S_e^2(\hat{\beta}_e)}{JK-J-1}$ are unbiased estimators of $\sigma_e^2 + K\sigma_a^2$ and σ_e^2 respectively, the unbiased estimator of $V(\hat{\mu}_e)$ is

$$\frac{1}{JK}\frac{S_a^2(\hat{\beta}_a)}{J-2}+x_{..}^2\frac{1}{S_{xxe}}\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}.$$

In order to construct a confidence interval for μ, depending on $\hat{\beta}_e$, we have to find an approximate distribution for the estimates of the variance components. Using Satterthwaite's approximation (Satterthwaite, (1946)) we can say that the distribution of

$$m\left[\left(\frac{1}{JK}\frac{S_a^2(\hat{\beta}_a)}{J-2}+\bar{x}_{..}^2\frac{1}{S_{xxe}}\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}\right)\Big/\left(\frac{1}{JK}(\sigma_e^2+K\sigma_a^2)+\bar{x}_{..}^2\frac{1}{S_{xxe}}\sigma_e^2\right)\right]$$

can be approximated by a chi-square distribution with m degrees of freedom where m is obtained by the following formula ;

$$m=\frac{\left(\frac{1}{JK}(\sigma_e^2+K\sigma_a^2)+\frac{\bar{x}_{..}^2}{S_{xxe}}\sigma_e^2\right)^2}{\frac{\left(\frac{1}{JK}(\sigma_e^2+K\sigma_a^2)\right)^2}{J-2}+\frac{\left(\frac{\bar{x}_{..}^2}{S_{xxe}}\sigma_e^2\right)^2}{JK-J-1}}.$$

Substituting the unbiased estimators $\frac{S_a^2(\hat{\beta}_a)}{J-2}$ and $\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}$ into $\sigma_e^2+K\sigma_a^2$ and σ_e^2 respectively in m, we'll obtain the estimated m which we denote it by $\hat{m}$. It is given by

$$\hat{m}=\frac{\left(\frac{1}{JK}\frac{S_a^2(\hat{\beta}_a)}{J-2}+\frac{\bar{x}_{..}^2}{S_{xxe}}\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}\right)^2}{\frac{\left(\frac{1}{JK}S_a^2(\hat{\beta}_a)\right)^2}{(J-2)^2}+\frac{\left(\frac{\bar{x}_{..}^2}{S_{xxe}}S_e^2(\hat{\beta}_e)\right)^2}{(JK-J-1)^2}}. \tag{4.2.1.2}$$

It follows that an approximate $1-2\alpha$ confidence interval for μ is

$$\hat{\mu}_e\pm t_{\alpha,\hat{m}}\sqrt{\frac{1}{JK}\frac{S_a^2(\hat{\beta}_a)}{J-2}+\frac{\bar{x}_{..}^2}{S_{xxe}}\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}}. \tag{4.2.1.3}$$

where $\hat{\mu}_e$, $\hat{m}$, $S_a^2(\hat{\beta}_a)$ and $S_e^2(\hat{\beta}_e)$ are given in (4.2.1.1), (4.2.1.2), (2.1.6) and (2.1.18) respectively. The length of this confidence interval can be stated as follows :

$$L_4 = 2t_{\alpha,\hat{m}}\sqrt{\frac{1}{JK}\frac{S_a^2(\hat{\beta}_a)}{J-2} + \frac{\bar{x}_{..}^2}{S_{xxe}}\frac{S_e^2(\hat{\beta}_e)}{JK-J-1}}$$

4.2.2 By using $\hat{\beta}_a$ as the estimator of β

The unbiased estimator of μ, depending on $\hat{\beta}_a$, is of the form :

$$\hat{\mu}_a = \bar{Y}_{..} - \hat{\beta}_a \bar{x}_{..} \tag{4.2.2.1}$$

where $\hat{\beta}_a$ is given in (2.1.8). We have

$$Cov(y_{..}, \hat{\beta}_a) = Cov\left(\frac{1}{JK}\mathbf{1}_{\mathbf{JK}}^{\mathbf{T}}\mathbf{Y}, \frac{\mathbf{X}^{\mathbf{T}}\mathbf{WY}}{\mathbf{X}^{\mathbf{T}}\mathbf{WX}}\right) = \frac{1}{JK}(K\sigma_a^2 + \sigma_e^2)\frac{\mathbf{1}_{\mathbf{JK}}^{\mathbf{T}}\mathbf{WX}}{\mathbf{X}^{'}\mathbf{QX}} = 0$$

where $\mathbf{W} = \left(\mathbf{I_J} - \frac{\mathbf{1}}{\mathbf{J}}\mathbf{J_J}\right) \otimes \frac{\mathbf{1}}{\mathbf{K}}\mathbf{J_K}$. Then, the variance of $\hat{\mu}_a$ is :

$$V(\hat{\mu}_a) = V(\bar{Y}_{..}) + \bar{x}_{..}^2 V(\hat{\beta}_a) = \frac{\sigma_e^2 + K\sigma_a^2(S_{xxa} + JK\bar{x}_{..}^2)}{S_{xxa}JK},$$

since $V(\bar{Y}_{..}) = \frac{\sigma_e^2 + K\sigma_a^2}{JK}$ and $V(\hat{\beta}_a) = \frac{\sigma_e^2 + K\sigma_a^2}{S_{xxa}}$.

$V(\hat{\mu}_a)$ can be estimated by using the distributional property that $S_a^2(\hat{\beta}_a) \sim (\sigma_e^2 + K\sigma_a^2)\chi_{J-2}^2$. Then, we can construct an exact $1-2\alpha$ confidence interval for μ. It is given by

$$\hat{\mu}_a \pm t_{\alpha,J-2}\sqrt{\frac{(S_{xxa} + JK\bar{x}_{..}^2)S_a^2(\hat{\beta}_a)}{JK(J-2)S_{xxa}}} \tag{4.2.2.2}$$

where $S_a^2(\hat{\beta}_a)$ is given in (2.1.6) . The length of this confidence interval can be found by :

$$L_5 = 2t_{\alpha,J-2}\sqrt{\frac{(S_{xxa} + JK\bar{x}_{..}^2)S_a^2(\hat{\beta}_a)}{JK(J-2)S_{xxa}}}$$

4.3 Confidence interval for σ_a^2

In order to construct a confidence interval for σ_a^2, we use a method suggested by Williams (1962) which constructs a confidence interval from experimental data by combining two or more intervals about the functions of the parameter of interest and nuisance parameters.

Since $S_a^2(\hat{\beta}_a) \sim (\sigma_e^2 + K\sigma_a^2)\chi_{J-2}^2$ as proved in section 2.1, we can state that;

$$P\left(\chi_{J-2,L}^2 \le \frac{S_a^2(\hat{\beta}_a)}{\sigma_e^2 + K\sigma_a^2} \le \chi_{J-2,U}^2\right) = 1-\alpha,$$

and it follows from that :

$$P\left(\frac{S_a^2(\hat{\beta}_a)}{\chi_{J-2,U}^2} \le \sigma_e^2 + K\sigma_a^2 \le \frac{S_a^2(\hat{\beta}_a)}{\chi_{J-2,L}^2}\right) = 1-\alpha \qquad (4.3.1)$$

where $S_a^2(\hat{\beta}_a)$ is given in (2.1.6) .

Let $V_1 = J-2$ and $V_2 = J(K-1)-1$. Then $\dfrac{V_2\sigma_e^2}{V_1(\sigma_e^2 + K\sigma_a^2)}\dfrac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)}$ is distributed according to the F-distribution with degrees of freedom V_1 and V_2 since it is shown that $S_e^2(\hat{\beta}_e) \sim \sigma_e^2\chi_{J(K-1)-1}^2$ in section 2.1, where $S_e^2(\hat{\beta}_e)$ is given in (2.1.18).

Hence, we have the following confidence interval :

$$P\left(F_{V_1,V_2,L} \le \frac{V_2}{V_1}\frac{\sigma_e^2}{(\sigma_e^2 + K\sigma_a^2)}\frac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)} \le F_{V_1,V_2,U}\right) = 1-\alpha$$

which can be rewritten as :

$$P\left(\frac{1}{K}\left(\frac{V_2}{V_1}\frac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)}\frac{1}{F_{V_1,V_2,U}} - 1\right) \le \frac{\sigma_a^2}{\sigma_e^2} \le \frac{1}{K}\left(\frac{V_2}{V_1}\frac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)}\frac{1}{F_{V_1,V_2,L}} - 1\right)\right) = 1-\alpha .$$

(4.3.2)

For any fixed σ_e^2, the two confidence intervals given in (4.3.1) and (4.3.2) together yield :

$$\frac{1}{K}\left(\frac{S_a^2(\hat{\beta}_a)}{\chi_{J-2,U}^2} - \sigma_e^2\right) \le \sigma_a^2 \le \frac{1}{K}\left(\frac{S_a^2(\hat{\beta}_a)}{\chi_{J-2,L}^2} - \sigma_e^2\right), \qquad (4.3.3)$$

$$\frac{1}{K}\left(\frac{V_2}{V_1}\frac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)}\frac{1}{F_{V_1,V_2,U}}-1\right)\sigma_e^2 \le \sigma_a^2 \le \frac{1}{K}\left(\frac{V_2}{V_1}\frac{S_a^2(\hat{\beta}_a)}{S_e^2(\hat{\beta}_e)}\frac{1}{F_{V_1,V_2,L}}-1\right)\sigma_e^2 .$$

(4.3.4)

Equating the left and right hand sides of (4.3.3) and (4.3.4) to each other, we have the following ;

$$\sigma_e^2 = \frac{V_1}{V_2}\frac{S_e^2(\hat{\beta}_e)F_{V_1,V_2}}{\chi^2_{V_1}} . \tag{4.3.5}$$

Substituting (4.3.5) into (4.3.4), we obtain the following two-sided confidence interval for σ_a^2 with probability $(1-\alpha)$;

$$\frac{1}{K\chi^2_{v_1,U}}\left(S_a^2(\hat{\beta}_a)-\frac{V_1S_e^2(\hat{\beta}_e)F_{V_1,V_2,U}}{V_2}\right)\le \sigma_a^2 \le \frac{1}{K\chi^2_{v_1,L}}\left(S_a^2(\hat{\beta}_a)-\frac{V_1S_e^2(\hat{\beta}_e)F_{V_1,V_2,L}}{V_2}\right)$$

.

(4.3.6)

where the length of the interval is :

$$L_6 = \frac{1}{K\chi^2_{v_1,U}}\left(S_a^2(\hat{\beta}_a)-\frac{V_1S_e^2(\hat{\beta}_e)F_{V_1,V_2,U}}{V_2}\right)+\frac{1}{K\chi^2_{v_1,L}}\left(S_a^2(\hat{\beta}_a)-\frac{V_1S_e^2(\hat{\beta}_e)F_{V_1,V_2,L}}{V_2}\right)$$

4.4 Confidence interval for σ_e^2

We derive an exact confidence interval for σ_e^2 by using the distributional property that $\frac{S_e^2(\beta_e)}{\sigma_e^2} \sim \chi^2_{J(K-1)-1}$ (shown in section 2.1) where $S_e^2(\hat{\beta}_e)$ is (2.1.18). Thus, we can state :

$$P\left(\chi^2_{J(K-1)-1,L} \le \frac{S_e^2(\hat{\beta}_e)}{\sigma_e^2} \le \chi^2_{J(K-1)-1,U}\right) = 1-\alpha$$

which can be written as a $(1-2\alpha)$ two sided exact confidence interval for σ_a^2 as follows :

$$P\left(\frac{S_e^2(\hat{\beta}_e)}{\chi^2_{J(K-1)-1,U}} \le \sigma_e^2 \le \frac{S_e^2(\hat{\beta}_e)}{\chi^2_{J(K-1)-1,L}}\right) = 1-\alpha . \tag{4.4.1}$$

where the length of the confidence interval can be stated as follows :

$$L_7 = \frac{S_e^2(\hat{\beta}_e)}{\chi^2_{J(K-1)-1,U}} + \frac{S_e^2(\hat{\beta}_e)}{\chi^2_{J(K-1)-1,L}}$$

CHAPTER 5

APPLICATION

For the application of all of the estimation techniques and confidence intervals we've mentioned in this study, we use the data from Vonesh and Carter (1987) to provide an example for the simple linear regression with one-fold nested error structure. A study was done to evaluate the in vivo ultrafiltration characteristics of a group of hollow fiber dialyzers. Ultrafiltration rates were measured at four different transmembrane pressures for each of 17 dialyzers. The actual data are shown in Table 5.1 where x_{jk} shows the transmembrane pressures and y_{jk} shows the ultrafiltration rates. Here, in addition to the observational error e_{jk}, dialyzer is the random effect since 17 dialyzers are randomly selected. Figure 1 shows the individual ultrafiltration rate (ml/hr) profiles for these 17 dialyzers.

Our goal is to estimate the linear relationship between ultrafiltration rate (ml/hr) and transmembrane pressure (mmHg). We apply the following model into the data :

$$Y_{jk} = \mu + \beta x_{jk} + a_j + e_{jk}, \qquad j = 1,2,\ldots,17, \quad k = 1,2,3,4$$

Quantities occuring in formulas for the estimation are computed through the computer programs listed in Appendix A.

The data summary is

$$S_{yya} = 158881.76, \qquad S_{xya} = 4096.27, \qquad S_{xxa} = 330.93,$$
$$S_{yye} = 16728030, \qquad S_{xye} = 3783722.25, \qquad S_{xxe} = 858087.62.$$

Table 5.1 : Ultrafiltration Data for 17 dialyzers

j \ k	1		2		3		4	
Dialyzer	x_{jk}	y_{jk}	x_{jk}	y_{jk}	x_{jk}	y_{jk}	x_{jk}	y_{jk}
1	160.0	600.0	265.0	1026.0	365.0	1470.0	454.0	1890.0
2	164.0	516.0	260.5	930.0	355.0	1380.0	451.0	1770.0
3	156.0	480.0	260.0	900.0	363.0	1380.0	466.0	1860.0
4	160.0	528.0	259.0	930.0	361.0	1410.0	462.0	1872.0
5	157.0	540.0	258.0	978.0	359.0	1410.0	471.0	1920.0
6	161.0	564.0	264.0	996.0	359.0	1422.0	466.0	1920.0
7	161.0	564.0	263.0	1062.0	363.0	1500.0	468.0	1980.0
8	158.0	492.0	255.0	900.0	360.0	1392.0	461.0	1860.0
9	161.0	1516.0	263.0	960.0	361.0	1380.0	462.0	1800.0
10	155.0	528.0	255.0	930.0	355.0	1356.0	455.0	1860.0
11	158.0	564.0	267.0	1020.0	360.0	1380.0	464.0	1884.0
12	165.0	618.0	263.0	1056.0	362.0	1500.0	461.0	1920.0
13	158.0	564.0	263.0	1038.0	367.0	1410.0	464.0	1770.0
14	162.0	552.0	268.0	1014.0	360.0	1440.0	465.0	1944.0
15	171.0	624.0	256.0	978.0	357.0	1440.0	466.0	1980.0
16	158.5	468.0	263.0	930.0	361.0	1332.0	460.0	1860.0
17	162.0	480.0	263.0	900.0	356.0	1272.0	463.0	1758.0

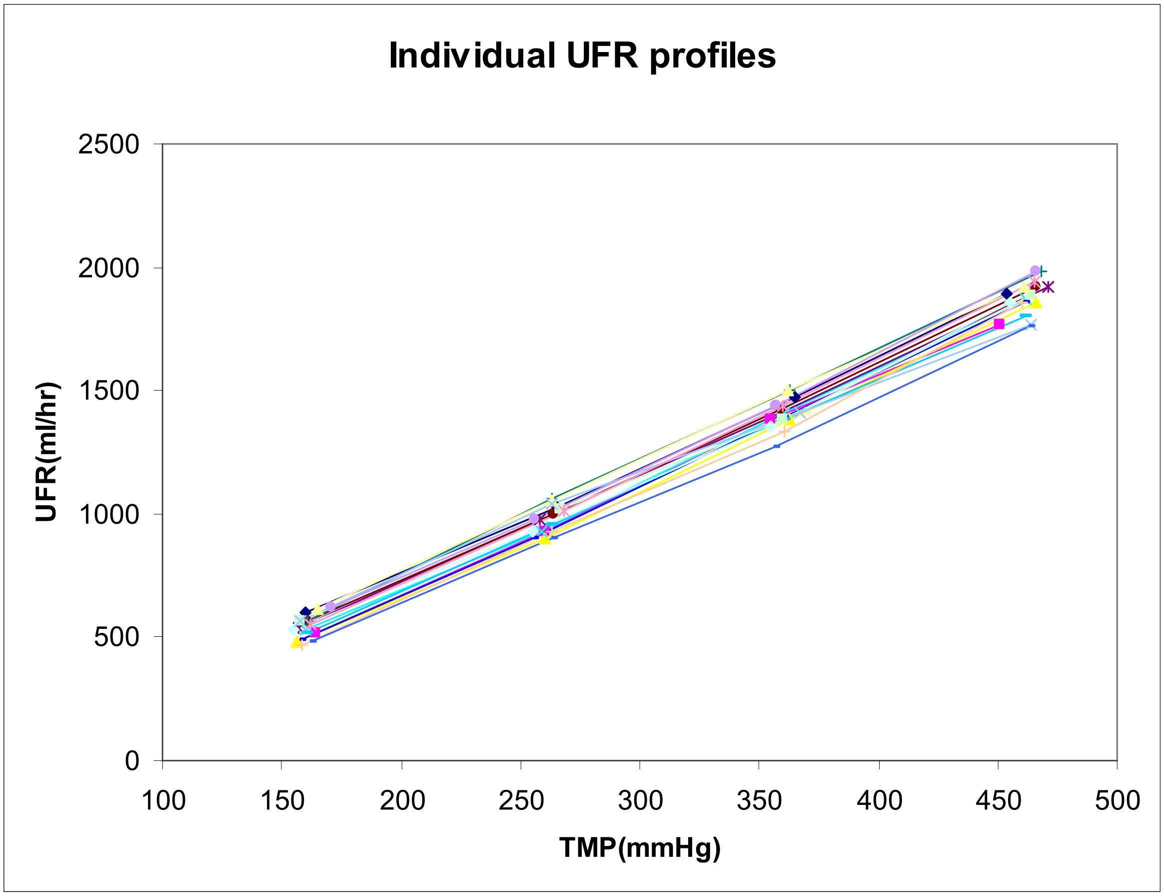

Figure 5.1 : Individual UFR Profiles for 17 dialyzers

5.1 Estimation of Fixed Effect Parameters

5.1.1 Maximum Likelihood Estimation

For this data, the second degree equation, given by (2.2.15) becomes,

$$-857094.79\beta^2 + 7542866.32\beta - 16251375.6 = 0$$

and has the roots 3.76 and 5.03. Thus the log likelihood function of β, given in (2.2.14) is as follows;

$$L(\beta) = \begin{cases} L_1(\beta) & if \quad \beta \in (3.76, 5.03) \\ L_2(\beta) & if \quad \beta \in (-\infty, 3.76] \cup [5.03, \infty) \end{cases}$$

The following figure is the log likelihood function L of β for this example when $-5 \le \beta \le 11$;

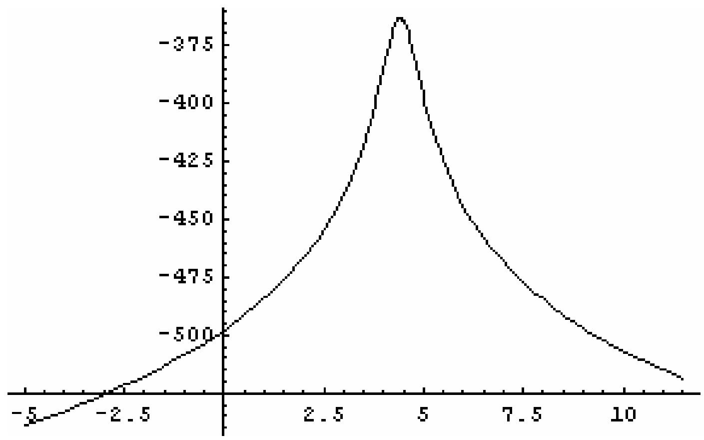

Figure 5.1.1 Graph of the log likelihood function of β when $-5 \le \beta \le 11$

For this data, the third degree equation given in (2.2.16) becomes

$$\beta^3 - 27.17\beta^2 + 474.11\beta - 1648.07 = 0$$

which has one real root with two complex roots. Its real root is $\hat{\beta}_M = 4.4098$. Since $S_a^2(\hat{\beta}_M) - S_e^2(\hat{\beta}_M)/(K-1) = 114598.41 > 0$, the maximum likelihood estimate of β is $\hat{\beta}_M = 4.40$. Then, $\hat{\mu}(\beta) = \overline{Y}_{..} - \hat{\beta}\overline{x}_{..}$ the maximum likelihood estimate of μ can be computed by using the previous formula and the estimated linear equation will be :

$$\hat{y}_{jk} = -173.91 + 4.40 x_{jk} .$$

5.1.2 Estimated Generalized Least Squares Estimation

The estimated generalized least squares estimates of μ and β given in (2.1.31) is computed through the program listed in Appendix A using   as follows :

$$\hat{\mu}_{GLS} = -173.91 ,\ \hat{\beta}_{GLS} = 4.40 .$$

Thus the estimated linear equation will be $\hat{y}_{jk} = -173.91 + 4.40x_{jk}$. The following four graphs give actual and estimated values of the first four UFR profiles for dialyzers.

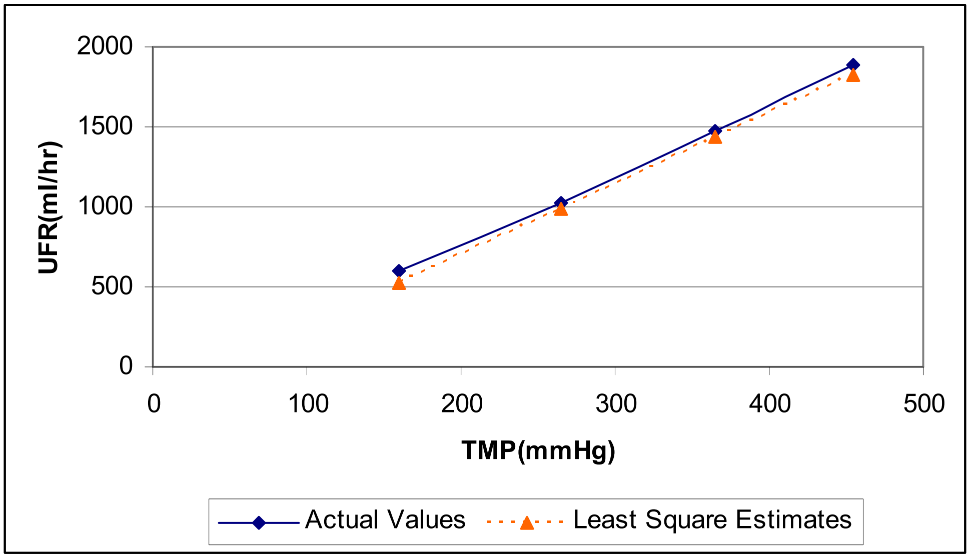

Figure 5.1.2 : Graph of actual and estimated values of the UFR profile of dialyzer 1

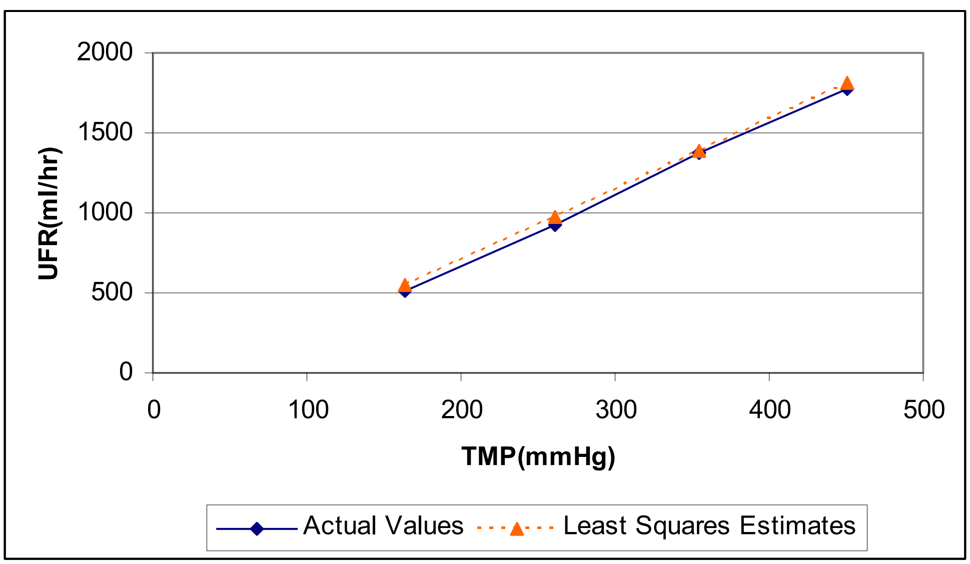

Figure 5.1.3 : Graph of actual and estimated values of the UFR profile of dialyzer 2

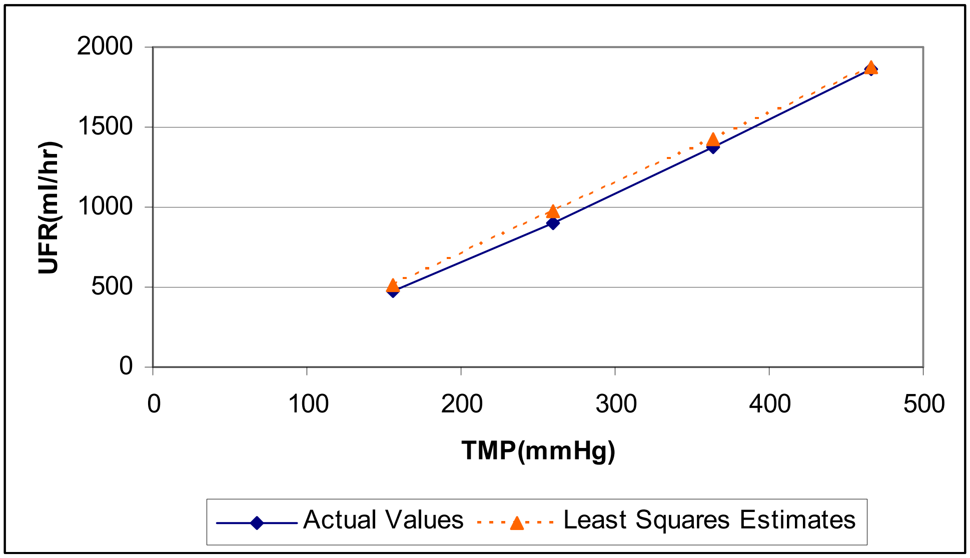

Figure 5.1.4 : Graph of actual and estimated values of the UFR profile of dialyzer 3

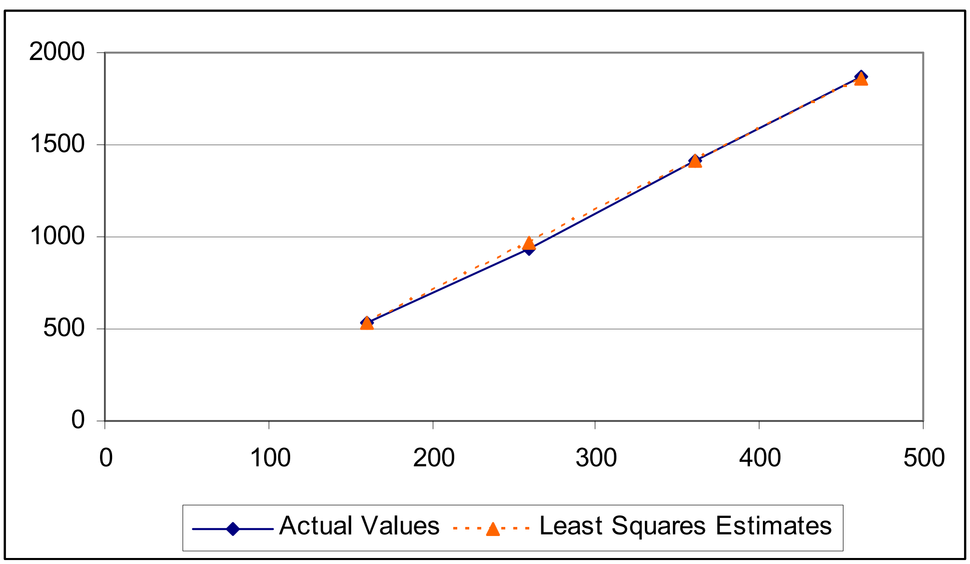

Figure 5.1.5 : Graph of actual and estimated values of the UFR profile of dialyzer 4

5.2 Estimation of Variance Components

5.2.1 Maximum Likelihood Estimation

Since $S_a^2(\hat{\beta}_M) - S_e^2(\hat{\beta}_M)/(K-1) = 114598.41 > 0$, the maximum likelihood estimate $\hat{\sigma}_a^2(\beta)$ of σ_a^2 given in (2.2.12) is ;

$$\hat{\sigma}_a^2(\beta) = 1685.72,$$

where $S_a^2(\hat{\beta}_M) = 129189.56$ $S_e^2(\hat{\beta}_M) = 43773.56$

We found that $\hat{\sigma}_a^2(\beta) > 0$, thus the maximum likelihood estimate $\hat{\sigma}_e^2(\beta)$ of σ_e^2 given in (2.2.13) is :

$$\hat{\sigma}_e^2(\beta) = 858.30.$$

5.2.2 Estimated Generalized Least Squares Estimation

The estimated generalized least squares estimates of σ_a^2 and σ_e^2 given in (2.1.26) and (2.1.27) are computed through the program listed in Appendix A as follows :

$$\hat{\sigma}_a^2 = 1799.66 \;,\; \hat{\sigma}_e^2 = 875.46$$

5.2.3 MINQE, AUE and REML

Utilizing the a priori information that $\rho = 2$, the MINQE, AUE and REML estimates of σ_a^2 given in (3.1.3), (3.2.3) and (3.3.6) respectively are computed as follows :

$$\hat{\sigma}_{a,MINQE}^2 = 1501.11,$$

$$\hat{\sigma}_{a,AUE}^2 = 1794.30,$$

and

$$\hat{\sigma}_{a,REML}^2 = 1685.27.$$

5.3 Confidence Intervals

5.3.1 Confidence Intervals for β

For this data, the exact % 95 confidence intervals for β given in (4.1.2) and (4.1.3) will become

$$4.34 \le \beta \le 4.47$$

and

$$2.42 \le \beta \le 22.32$$

respectively.

5.3.2 Confidence Intervals for μ

For this data, if we use $\hat{\beta}_e$ as the estimator of β, the % 95 confidence interval for μ given in (4.2.1.3) becomes

$$-235.36 \le \mu \le -112.20$$

When we use $\hat{\beta}_a$ as the estimator of β, the % 95 confidence interval for μ given in (4.2.2.2) becomes :

$$-5748.01 \le \mu \le 442.16$$

5.3.3 Confidence Intervals for σ_a^2 and σ_e^2

For this data, % 95 confidence intervals for σ_a^2 and σ_e^2 (given in 4.3.6 and 4.4.1) will yield :

$$1492.18 \le \sigma_a^2 \le 1801.02 \text{ and } 851.24 \le \sigma_e^2 \le 883.75 .$$

CHAPTER 6

SUMMARY AND CONCLUSIONS

Data sets with one-fold nested error structure occur in situations where there has been some form of subsampling from primary sampling units. An experiment from split-plot, one-restrictional lattice, incomplete block designs and cases of more than one "measurement" taken on each "subject" are examples of the simple linear regression model with one-fold nested error structure.

In this thesis, we mainly study the estimation techniques for fixed-effect parameters, μ and β and variance components, σ_a^2 and σ_e^2 for the simple linear regression model with one-fold nested error.

In the first part of the thesis, the definition and a brief introduction of simple linear regression models with one-fold nested error are presented. Chapter 2 gives the review of estimated generalized least squares estimation (Fuller and Battese, (1973)) and maximum likelihood estimation (Güven, (1995)) of fixed-effect parameters of the simple linear regression model with one-fold nested error.

Generalized least squares estimators are known as unbiased, more efficient than least squares estimators, asymptotically consistent and asymptotically normal. Concerning the optimality properties of maximum likelihood estimators, Güven (1995) showed the consistency, asymptotic normality and asymptotic efficiency properties of maximum likelihood estimation in the mixed analysis of variance models. Estimated generalized least squares estimators of fixed effect parameters are easy to compute and non-iterative in the nature, however they are not exact but "estimated" generalized least squares estimators since in order to find them, first we have to estimate the variance components. On the other hand, in the maximum likelihood estimation, sorting out the roots of the polynomial (given in 2.2.16) looks problematic but an algorithm to sort out the roots of this polynomial in order to find the maximum likelihood estimate of the slope term was developed which can be found in the unpublished doctorate thesis of Güven (1992). Also we have to note that the resulting maximum likelihood estimators are the exact roots of the likelihood function.

Chapter 3 considers the estimation of variance components and presents the Minimum Norm Quadratic Estimator (MINQE), the Almost Unbiased Estimator (AUE) and the Restricted Maximum Likelihood Estimator (REML) of the variance of primary units. Among the different variance component estimators summarized by Searle (1977) we choose these four estimators since they have exact MSEs.

After deriving the estimators and their MSEs, we computed the exact MSEs of these four estimators under various values of the pair (S_{xxa}, S_{xxe}) and the ratio of variance components $\rho = \frac{\sigma_a^2}{\sigma_e^2}$ along with the MSE of ANOVA estimator. The results are summarized in Table 3.4.1. It indicates that the MINQE of σ_a^2 has the superiority over the other presented three estimators of σ_a^2 where both the REML estimator and the AUE of σ_a^2 have smaller MSE than the MSE of the ANOVA estimator of σ_a^2. Hence, we conclude that both the REML estimator and the AUE have the superority over the ANOVA estimator. When we compare the REML estimator and AUE of σ_a^2, it can be seen that the MSE of the AUE of σ_a^2 is smaller than the MSE of the REML estimator when ρ is equal to 0,1 and 0,5. However this statement is reversed when ρ is equal to 1, 2 and 5. Thus, we can conclude that the AUE has the superiority over the REML estimator for small values of ρ whereas for large values of ρ, the REML has the superiority over the AUE.

We covered exact and approximate confidence intervals for the fixed effect parameters and the variance components of the model in Chapter 4. In addition to the two exact confidence intervals given by Park and Burdick (1994) , we constructed an exact confidence interval for β. When we compare these confidence intervals for β, we can infer that the ones obtained by Park and

Burdick are more reliable since their lengths are smaller than the one which we constructed but one of them given in (4.1.1) is generally not applicable since it assumes that $\phi = \sigma_e^2 / (\sigma_e^2 + K\sigma_a^2)$ is known. Moreover in Chapter 4, we constructed one approximate and one exact confidence interval for μ by obtaining two unbiased estimators of μ. One is depending on $\hat{\beta}_e$ whereas the other is depending on $\hat{\beta}_a$. By comparing their lengths, we can infer that the confidence interval using $\hat{\mu}_e = \overline{Y}_{..} - \hat{\beta}_e \overline{x}_{..}$ as the unbiased estimator of μ is more reliable than the one which employs $\hat{\mu}_a = \overline{Y}_{..} - \hat{\beta}_a \overline{x}_{..}$ as the unbiased estimator of μ. In addition to these, the last section of this chapter gives the confidence intervals for σ_a^2 and σ_e^2.

The last chapter of this thesis is devoted to the application of all of these estimation techniques and confidence intervals into a real life data we've covered in this study to a real life data set. Some of the conclusions that we mentioned in thiesis can be verified by the numerical results obtained in Chapter 5.

REFERENCES

- Corbeil, R.R., Searle, S.R. (1976). Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model, *Technometrics*, 18, 31-38.

- Diggle, P., Heagerty, P., Liang, K.L. (2002). Analysis of Longitudinal Data (Second Edition), Oxford : Oxford University Press.

- Eisenhart, C. (1947). The Assumptions Underlying the Analysis of Variance, *Biometrics,* 3, 1-21.

- Fuller, F.A., Battese, G.E. (1973). Transformation for Estimation of Linear Models with Nested Error, *J. Amer. Statist. Assoc.*, 68, 626-632.

- Güven, B. (1992). Estimation in Simple Linear Regression with A Nested Error, The Unpublished Ph.D. Thesis, University of Wisconsin - Madison.

- Güven, B. (1995). Maximum Likelihood Estimation in Simple Linear Regression with One Fold Nested Error, *Commun. Stat.-Theory Meth.*, 24, 121-130.

- Güven, B. (1995). Asymptotic Properties of Maximum Likelihood Estimation in Mixed Analysis of Variance Model, *Statistical Papers*, 36, 175-182.

- Güven, B. (1998). The Uniformly Better Estimator for the Slope Parameter in the Simple Linear Regression with One-Fold Nested Error, *Journal of the Turkish Statistical Association*, 1, 21-27.

- Hartley, H.O., Rao, J.N.K. (1967). Maximum Likelihood Estimation for the Mixed Analysis of Variance Model, *Biometrika*, 54, 93-108.

- Harville, D.A. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems, *J. Amer. Statist. Assoc.,* 72, 320-340.

- Henderson, C.R. (1953). Estimation of Variance and Covariance Components, *Biometrics*, 9, 226-252.

- Henderson, C.R. (1975). Best Linear Unbiased Estimation and Prediction Under a Selection Model, *Biometrics,* 31, 423-447.

- Horn, S.D., Horn, R.A., (1975). Comparison of Estimators of Heteroscedastic Variances in Linear Models, *J. Amer. Statist. Assoc.*, 70, 872-879.

- Horn, S.D., Horn, R.A., and Duncan, D.B. (1975). Estimating Heteroscedastic Variances in Linear Models, *J. Amer. Statist. Assoc.*, 70, 380-385.

- McCulloch, C.E., and Searle, S.R. (2000). Generalized, Linear and Mixed Models, New York : Wiley & Sons.

- Park, D.J., Burdick, R.K. (1994). Confidence Intervals on the Regression Coefficient in a Simple Linear Regression Model with a Balanced One-Fold Nested Error Structure, *Commun. Stat. – Simula.*, 23, 43-58.

- Park, D.J., Hwang, H.M. (2002). Confidence Intervals for the Mean Response in the Simple Linear Regression Model with Balanced Nested Error Structure, *Commun. Stat. – Theory Meth.* , 31, 107-118.

- Rao, C.R. (1970). Estimation of Heterogeneous Variances in Linear Models, *J. Amer. Statist. Assoc.,* 65, 161-172.

- Rao, C.R. (1971). Minimum Variance Quadratic Unbiased Estimation of Variance Components, *Journal of Multivariate Analysis*, 1, 445-456.

- Rao, C.R.. (1972) Estimation of Variance and Covariance Components in Linear Models, *J. Amer. Statist. Assoc.,* 67, 112-115..

- Rao, C.R., Chaubey, Y.P. (1978). Three Modifications of the Principle MINQUE, *Commun. Stat,-Theory Meth*., A7, 767-778.

- Rao, P.S.R.S. (2001). The MINQUE : Investigations, Modifications and Extensions, *Journal of Statistical Research*, 35, 29-43

- Rupper, D., Wand, M.P. and Carroll, R.J. (2003) Semiparametric Regression, New York : University Press.

- Satterthwaite, F.E. (1946). An Approximate Distribution of Estimates of Variance Components, *Biometrics Bull*., 2, 110-114.

- Searle, S.R. (1971). Topics in Variance Component Estimation, *Biometrics*, 27, 1-76.

- Searle, S.R. (1977). Variance Components Estimation : A Thumbnail Review, *Biometrics Unit Mimeo Series BU-612-M*, Cornell University, Ithaca.

- Tong, L., Cornelius, P.L. (1989). Studies on the Estimation of the Slope Parameter in the Simple Linear Regression Model with One-Fold Nested Error Structure, *Commun. Stat.-Simula*., 18, 220-225.

* Tong, L., Cornelius, P.L. (1991). Studies on the Hypothesis Testing of the Slope Parameter in the Simple Linear Regression Model with One-Fold Nested Error Structure, *Commun. Stat.-Theory Meth.*., 20, 2023-2043.

* Vonesh, E.F., Carter, R.L. (1987). Efficient Inference for Random-Coefficient Growth Curve Models with Unbalanced Data, *Biometrics*, 43, 617-628.

* Williams, J.S. (1962). A Confidence Interval for Variance Components, *Biometrika,* 49, 278-281.

APPENDIX A

COMPUTER PROGRAM FOR ESTIMATION PROCEDURES AND CONFIDENCE INTERVALS

Description :

This program computes the necessary quantities which will occur in the formulas for the estimation of fixed effect parameters and variance components in addition to the quantities which will occur in the formulas of the confidence intervals for the data taken from Vonesh & Carter (1987) mentioned in section 5.

Outputs :

MEANX(I) : $\bar{x}_{i.} = \sum_{k=1}^{K} x_{ik} / K$

MEANY(I) : $\bar{Y}_{i.} = \sum_{k=1}^{K} Y_{ik} / K$

GRANDMEANX : $\bar{x}_{..} = \sum_{j=1}^{J} \sum_{k=1}^{K} x_{jk} /(JK)$

GRANDMEANY : $\bar{Y}_{..} = \sum_{j=1}^{J} \sum_{k=1}^{K} Y_{jk} /(JK)$

SXXA : S_{xxa} given in (1.2.2)

SYYA : S_{yya} given in (1.2.4)

SXYA : S_{xya} given in (1.2.3)

SXXE : S_{xxe} given in (1.2.5)

SYYE : S_{yye} given in (1.2.7)

SXYE : S_{xye} given in (1.2.6)

RA : $S_a^2(\hat{\beta}_a)$ given in (2.1.6)

R1 : S_d^2 given in (2.1.7)

RE : $S_e^2(\hat{\beta}_e)$ given in (2.1.18)

GLS_SIGMA_A_SQUARE : $\tilde{\sigma}_a^2$ given in (2.1.27)

GLS_SIGMA_E_SQUARE : $\tilde{\sigma}_e^2$ given in (2.1.26)

BETA_GLS : $\tilde{\beta}$ given in (2.1.31)

MU_GLS : $\tilde{\mu}$ given in (2.1.31)

SAB : $S_a^2(\beta)$ given in (2.2.6)

SEB : $S_e^2(\beta)$ given in (2.2.5)

P1 : Coefficient of second degree term in the polynomial (2.2.16)

P2 : Coefficient of first degree coefficient in the polynomial (2.2.16)

P3 : Constant term in the polynomial (2.2.16)

Program Listing :

THE OUTPUTS OF THIS PROGRAM ARE STORED IN FILE 'NESTED.txt'

```
PROGRAM NESTED
        PARAMETER (NCX=4,NCY=4,NRX=17,NRY=17)
        DOUBLE PRECISION
X(NRX,NCX),Y(NRY,NCY),SUMX(17),MEANX(17),SUMY(17)
        DOUBLE PRECISION
GRANDSUMX,GRANDMEANX,GRANDSUMY,GRANDMEANY,SXXA
```

```
      DOUBLE PRECISION
SYYA,SXYA,SXXE,MEANY(17),SYYE,SXYE,SA_SQUARE
      DOUBLE PRECISION
ANOVA_SIGMA_E_SQUARE,RO_HEAD,Y_N(17,4),X_N(17,4)
      DOUBLE PRECISION
GRANDSUMX_N,GRANDSUMY_N,GRANDMEANX_N,GRANDMEANY_N
      DOUBLE PRECISION SXYT_N,SXXT_N,BETA_N,MU_N,YFIT_N(17,4)
      DOUBLE PRECISION
RESID_N(17,4),BETA_A,BETA_E,GLS_SIGMA_A_SQUARE
      DOUBLE PRECISION
GLS_SIGMA_E_SQUARE,C,D,E,F,BETA_GLS,MU_GLS
      DOUBLE PRECISION
A,B,GAMMA_A,GAMMA_E,C1,C2,C3,P0,P1,P2,P3
      DOUBLE PRECISION
SE_SQUARE,ANOVA_SIGMA_A_SQUARE,SEB,SAB
      EXTERNAL WRRRN

      OPEN(2,FILE="NESTED.txt")

C     SET VALUES FOR X AND Y
      DATA ((X(I,J),J=1,4),I=1,17)/160.0,265.0,365.0,454.0,164.0,260.5,
     &  355.0,451.0,156.0,260.0,363.0,466.0,160.0,259.0,361.0,462.0,
     &  157.0,258.0,359.0,471.0,161.0,264.0,359.0,466.0,
     &  161.0,263.0,363.0,468.0,158.0,255.0,360.0,461.0,
     &  161.0,263.0,361.0,462.0,155.0,255.0,355.0,455.0,
     &  158.0,267.0,360.0,464.0,165.0,263.0,362.0,461.0,
     &  158.0,263.0,367.0,464.0,162.0,268.0,360.0,465.0,
     &  171.0,256.0,357.0,466.0,158.5,263.0,361.0,460.0,
     &  162.0,263.0,356.0,463.0/
      DATA ((Y(K,L),L=1,4),K=1,17)/600.0,1026.0,1470.0,1890.0,516.0,
```

```
     &    930.0,1380.0,1770.0,480.0,900.0,1380.0,1860.0,528.0,930.0,
     &    1410.0,1872.0,540.0,978.0,1410.0,1920.0,564.0,996.0,1422.0,
     &    1920.0,564.0,1062.0,1500.0,1980.0,492.0,900.0,1392.0,1860.0,
     &    516.0,960.0,1380.0,1800.0,528.0,930.0,1356.0,1860.0,
     &    564.0,1020.0,1380.0,1884.0,618.0,1056.0,1500.0,1920.0,
     &    564.0,1038.0,1410.0,1770.0,552.0,1014.0,1440.0,1944.0,
     &    624.0,978.0,1440.0,1980.0,468.0,930.0,1332.0,1860.0,
     &    480.0,900.0,1272.0,1758.0/
      CALL WRRRN('X',NRX,NCX,X,NRX,0)
      CALL WRRRN('Y',NRY,NCY,Y,NRY,0)

C     CALCULATING PRIMARY UNIT MEANS AND GRAND MEANS
      DO 10 I=1,17
      DO 20 J=1,4
      SUMX(I)=SUMX(I)+X(I,J)
      SUMY(I)=SUMY(I)+Y(I,J)
20    CONTINUE
      MEANX(I)=SUMX(I)/4
      MEANY(I)=SUMY(I)/4
10    CONTINUE
      PRINT*,'MEANX',MEANX
      WRITE(2,*)'MEANX',MEANX
      PRINT*,'MEANY',MEANY
      WRITE(2,*)'MEANY',MEANY
      DO 50 I=1,17
      DO 50 J=1,4
      GRANDSUMX=GRANDSUMX+X(I,J)
      GRANDSUMY=GRANDSUMY+Y(I,J)
50    CONTINUE
      GRANDMEANX=GRANDSUMX/68
```

```
      GRANDMEANY=GRANDSUMY/68
      PRINT*,'GRANDMEANX=',GRANDMEANX
      WRITE(2,*)'GRANDMEANX=',GRANDMEANX
      PRINT*,'GRANDMEANY=',GRANDMEANY
      WRITE(2,*)'GRANDMEANY=',GRANDMEANY

C     CALCULATING SXXA, SYYA, SXYA, SXXE, SYYE, SXYE
      DO 70 I=1,17
      SXXA=SXXA+((MEANX(I)-GRANDMEANX)**2)
      SYYA=SYYA+((MEANY(I)-GRANDMEANY)**2)
      SXYA=SXYA+((MEANX(I)-GRANDMEANX)*(MEANY(I)-
GRANDMEANY))
70    CONTINUE
      SXXA=4*SXXA
      PRINT*,'SXXA=',SXXA
      WRITE(2,*),'SXXA=',SXXA
      SYYA=4*SYYA
      PRINT*,'SYYA=',SYYA
      WRITE(2,*),'SYYA=',SYYA
      SXYA=4*SXYA
      PRINT*,'SXYA=',SXYA
      WRITE(2,*),'SXYA=',SXYA
      DO 100 I=1,17
      DO 100 J=1,4
      SXXE=SXXE+((X(I,J)-MEANX(I))**2)
      SYYE=SYYE+((Y(I,J)-MEANY(I))**2)
      SXYE=SXYE+((X(I,J)-MEANX(I))*(Y(I,J)-MEANY(I)))
100   CONTINUE
      PRINT*,'SXXE=',SXXE
      WRITE(2,*),'SXXE=',SXXE
```

```
      PRINT*,'SYYE=',SYYE
      WRITE(2,*),'SYYE=',SYYE
      PRINT*,'SXYE=',SXYE
      WRITE(2,*)'SXYE=',SXYE

C     CALCULATING RA, RE AND R1
      RA=SYYA-((SXYA**2)/SXXA)
      RE=SYYE-((SXYE**2)/SXXE)
      R1=(SXXA/(SXXE*(SXXE+SXXA)))*((((SXXE*SXYA)/SXXA)-
SXYE)**2)
      PRINT*,'RA=',RA
      WRITE(2,*)'RA=',RA
      PRINT*,'RE=',RE
      WRITE(2,*)'RE=',RE
      PRINT*,'R1=',R1
      WRITE(2,*)'R1=',R1

C     COMPUTING THE GENERALIZED LEAST SQUARES ESTIMATES
      GLS_SIGMA_E_SQUARE=RE/(68-17-1)
      BETA_A=SXYA/SXXA
      BETA_E=SXYE/SXXE
      RB=RA+R1
      A=RB-(16*GLS_SIGMA_E_SQUARE)
      B=4*(15+(SXXE/(SXXA+SXXE)))
      GLS_SIGMA_A_SQUARE=A/B
      PRINT*,'GLS_SIGMA_A_SQUARE=',GLS_SIGMA_A_SQUARE
      WRITE(2,*)'GLS_SIGMA_A_SQUARE=',GLS_SIGMA_A_SQUARE
      PRINT*,'GLS_SIGMA_E_SQUARE=',GLS_SIGMA_E_SQUARE
      WRITE(2,*)'GLS_SIGMA_E_SQUARE=',GLS_SIGMA_E_SQUARE
      C=GLS_SIGMA_E_SQUARE*SXYA
```

```
      D=(GLS_SIGMA_E_SQUARE+(4*GLS_SIGMA_A_SQUARE))*SXYE
      E=GLS_SIGMA_E_SQUARE*SXXA
      F=(GLS_SIGMA_E_SQUARE+(4*GLS_SIGMA_A_SQUARE))*SXXE
      BETA_GLS=(C+D)/(E+F)
      MU_GLS=GRANDMEANY-(BETA_GLS*GRANDMEANX)
      PRINT*,'BETA_GLS=',BETA_GLS
      WRITE(2,*)'BETA_GLS=',BETA_GLS
      PRINT*,'MU_GLS=',MU_GLS
      WRITE(2,*)'MU_GLS=',MU_GLS
```

C CALCULATIONS CONCERNING THE MAXIMUM LIKELIHOOD ESTIMATION

```
      GAMMA_A=SYYA/SXXA
      GAMMA_E=SYYE/SXXE
      P1=((5*BETA_E)+(7*BETA_A))/-4
      P2=(GAMMA_E+(8*BETA_A*BETA_E)+(3*GAMMA_A))/4
      P3=((GAMMA_E*BETA_A)+(3*GAMMA_A*BETA_E))/-4
      DO 180 I=1,17
      DO 180 J=1,4
      SEB=SEB+((Y(I,J)-MEANY(I)-(4.40982*(X(I,J)-MEANX(I))))**2)
180   CONTINUE
      DO 190 I=1,17
      SAB=SAB+((MEANY(I)-GRANDMEANY-(4.40982*(MEANX(I)-
GRANDMEANX)))**2)
190   CONTINUE
      SAB=4*SAB
      PRINT*,'P1=',P1
      WRITE(2,*)'P1=',P1
      PRINT*,'P2=',P2
```

```
WRITE(2,*)'P2=',P2
PRINT*,'P3=',P3
WRITE(2,*)'P3=',P3
PRINT*,'SEB=',SEB
WRITE(2,*)'SEB=',SEB
PRINT*,'SAB=',SAB
WRITE(2,*)'SAB=',SAB
END
```

APPENDIX B

COMPUTER PROGRAMS FOR CALCULATING MSE'S OF MINQE, AUE AND REML OF σ_a^2

Description :

This program computes the MSE of MINQE of σ_a^2 given in (3.1.8) under the conditions given in section 3.4.

Inputs :

SXXA : S_{xxa} given in (1.2.2)

SXXE : S_{xxe} given in (1.2.5)

RO : $\rho = \dfrac{\sigma_a^2}{\sigma_e^2}$

Outputs :

MSEMINQE : MSE of MINQE of σ_a^2 given in (3.1.8)

Program Listing

THE OUTPUTS OF THIS PROGRAM ARE STORED IN FILE 'MINQE.txt'

```
      PROGRAM MINQE
      DOUBLE PRECISION
SIGMA_A_SQUARE,SIGMA_E_SQUARE,RO,SXXA,SXXE,MSE1
      DOUBLE PRECISION J,K,A,B,C,D,E,F
      OPEN(2,FILE="MINQE.txt")
C     TAKING THE VALUES OF INPUTS
      J=4
      K=7
      PRINT*,'SXXA='
      READ*,SXXA
      PRINT*,'SXXE='
      READ*,SXXE
      PRINT*,'RO='
      READ*,RO
      SIGMA_E_SQUARE=1
      SIGMA_A_SQUARE=RO*SIGMA_E_SQUARE
C     COMPUTING THE MSE OF MINQE BY PARTITONING THE
EXPRESSION AND CALCULATING THESE PARTS
      A=(K*(RO**2)*SIGMA_E_SQUARE)/(J*(1+(K*RO)))
      B=J-2
      C=((1+(K*RO))*SXXE)/(SXXA+((1+(K*RO))*SXXE))
      D=(2*(A**2))*(B+(C**2))
      E=A*(B+C)
      F=(E-SIGMA_A_SQUARE)**2
      MSEMINQE=D+F
      PRINT*,'MSEMINQE=',MSEMINQE
      WRITE(2,*),'MSEMINQE=',MSEMINQE
      END
```

Description :

This program compute the MSE of AUE of σ_a^2 given in (3.2.4) under the conditions given in section 3.4.

Inputs :

SXXA : S_{xxa} given in (1.2.2)

SXXE : S_{xxe} given in (1.2.5)

RO : $\rho = \dfrac{\sigma_a^2}{\sigma_e^2}$

Outputs :

MSEAUE : MSE of AUE of σ_a^2 given in (3.2.4)

Program Listing :

THE OUTPUTS OF THIS PROGRAM ARE STORED IN FILE 'AUE.txt'

```
PROGRAM AUE
        DOUBLE PRECISION
SIGMA_A_SQUARE,SIGMA_E_SQUARE,RO,SXXA,SXXE
        DOUBLE PRECISION J,K,ERA,ER1,VRA,VR1,GR,EREML,VAUE
```

```
      DOUBLE PRECISION MSEAUE,SIGMA_B_SQUARE,EAUEN,EAUED
      OPEN(2,FILE="AUE.txt")
C     TAKING THE VALUES OF INPUTS
      J=4
      K=7
      PRINT*,'SXXA='
      READ*,SXXA
      PRINT*,'SXXE='
      READ*,SXXE
      PRINT*,'RO='
      READ*,RO
      SIGMA_E_SQUARE=1
      SIGMA_A_SQUARE=RO*SIGMA_E_SQUARE
      SIGMA_B_SQUARE=SIGMA_E_SQUARE+(K*SIGMA_A_SQUARE)
C     COMPUTING THE EXPECTED VALUES AND VARIANCES OF THE
RA AND R1
      ERA=SIGMA_B_SQUARE*(J-2)
      ER1=(SIGMA_B_SQUARE*SXXE)+(SIGMA_E_SQUARE*SXXA)
      VRA=(SIGMA_B_SQUARE**2)*2*(J-2)
      VR1=(((SIGMA_B_SQUARE*SXXE)+(SIGMA_E_SQUARE*SXXA))**2)
*2
C     COMPUTING THE MSE OF AUE BY PARTITIONING THE
EXPRESSION AND CALCULATING THESE PARTS SEPERATELY
      GR=(((1+(K*RO))**2)*SXXE)/((SXXA+((1+(K*RO))*SXXE))**2)
      C=SXXA/(SXXA+((1+(K*RO))*SXXE))
      EAUEN=(RO*(ERA+(GR*ER1)))
      EAUED=((1+(K*RO))*(J-1-C))
      EAUE=EAUEN/EAUED
      VAUE=((RO**2)*(VRA+((GR**2)*VR1)))/(((1+(K*RO))*(J-1-C))**2)
      MSEAUE=((EAUE-SIGMA_A_SQUARE)**2)+VAUE
```

```
PRINT*,'MSEAUE=',MSEAUE
WRITE(2,*),'MSEAUE=',MSEAUE
END
```

Description :

This programs computes the MSE of REML of σ_a^2 given in (3.3.1.1) under the conditions given in section 3.4.

Inputs :

SXXA : S_{xxa} given in (1.2.2)

SXXE : S_{xxe} given in (1.2.5)

RO : $\rho = \dfrac{\sigma_a^2}{\sigma_e^2}$

Outputs :

MSEREML : MSE of REML of σ_a^2 given in (3.3.1.1)

Program Listing :

THE OUTPUTS OF THIS PROGRAM ARE STORED IN FILE 'REML.txt'

```
PROGRAM REML
        DOUBLE PRECISION
SIGMA_A_SQUARE,SIGMA_E_SQUARE,RO,SXXA,SXXE
        DOUBLE PRECISION
J,K,ERA,ER1,ERE,VRA,VR1,VRE,FR,GR,EREML,VREMLN
```

```
      DOUBLE PRECISION
VREMLD,VREML,MSEREML,SIGMA_B_SQUARE
      OPEN(2,FILE="REML.txt")
C     TAKING THE VALUES OF INPUTS
      J=4
      K=7
      PRINT*,'SXXA='
      READ*,SXXA
      PRINT*,'SXXE='
      READ*,SXXE
      PRINT*,'RO='
      READ*,RO
      SIGMA_E_SQUARE=1
      SIGMA_A_SQUARE=RO*SIGMA_E_SQUARE
      SIGMA_B_SQUARE=SIGMA_E_SQUARE+(K*SIGMA_A_SQUARE)
C     COMPUTING THE EXPECTED VALUES AND VARIANCES OF RA,
RE AND R1
      ERA=SIGMA_B_SQUARE*(J-2)
      ER1=(SIGMA_B_SQUARE*SXXE)+(SIGMA_E_SQUARE*SXXA)
      ERE=SIGMA_E_SQUARE*((J*K)-J-1)
      VRA=(SIGMA_B_SQUARE**2)*2*(J-2)
      VR1=(((SIGMA_B_SQUARE*SXXE)+(SIGMA_E_SQUARE*SXXA))**2)
*2
      VRE=(SIGMA_E_SQUARE**2)*2*((J*K)-J-1)
C     COMPUTING THE MSE OF REML BY PARTITONING THE
EXPRESSION AND CALCULATING THESE PARTS SEPERATELY
      FR=SXXA/((SXXA+((1+(K*RO))*SXXE))**2)
      GR=(((1+(K*RO))**2)*SXXE)/((SXXA+((1+(K*RO))*SXXE))**2)
      EREML=(((K-1)*ERA)+((((K-1)*GR)-FR)*ER1)-ERE)/((J*K)*(K-1))
      VREMLN=(((K-1)**2)*VRA)+(((((K-1)*GR)-FR)**2)*VR1)+VRE
```

```
VREMLD=(((J*K)**2)*((K-1)**2))
VREML=VREMLN/VREMLD
MSEREML=((EREML-SIGMA_A_SQUARE)**2)+VREML
PRINT*,'MSEREML=',MSEREML
WRITE(2,*),'MSEREML=',MSEREML
END
```