THE ESTIMATION OF STABILITY PARAMETERS

By

CHUNG-HSIEN SUNG

Bachelor of Science
Tamkang College
Taipei, Taiwan ROC
1974

Master of Arts
Northeast Missouri State University
Kirksville, Missouri
1979

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1983

THE ESTIMATION OF STABILITY PARAMETERS

Thesis Approved:

Ronald W. McNew
Thesis Adviser

William H. Stewart

M. Palmer Terrell

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my thanks and appreciation to my adviser, Dr. Ronald W. McNew, for suggesting this topic and for his valuable advice.

I also wish to thank Dr. Nitis Mukhopadhyay, Dr. William H. Stewart and Dr. Marvin Palmer Terrell for serving on my advisory committee.

Special gratitude is extended to my parents for their dedicated support and encouragement.

Finally, the patience and untold sacrifices of my wife, Mei-Mei, are deeply appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In plant breeding experiments, many genotypes are usually evaluated in different environments (such as location, years) before selecting a desirable genotype. Genotypes tested in different environments almost invariably show genotype-environment interaction; that is, the relative phenotypic performances of the genotypes vary from one environment to another. Such differential response of genotypes in different environments makes it difficult for breeders to decide which genotypes should be selected.

Different attempts have been made to solve the problems created by genotype-environment interaction. Stratification of environments has been used to reduce the genotype-environment interaction. That is, the region for which a breeder is developing improved genotypes is subdivided so that within a subregion the interaction is reduced. The stratification usually is based on such macro-environmental factors as temperature, water availability, and soil type. Even with this refinement of technique, the interaction of genotype with location in a subregion and with environments encountered at the same location in different years may remain unacceptedly large.

Since little additional progress can be expected in reducing genotype-environment interaction by the stratification of environments, other methods need to be investigated. One such method is to select stable genotypes that interact less with the environments in which they are to be grown. If the stability of performance, or the ability to show a minimum of interaction with the environment, is a genetic characteristic, then preliminary evaluation could be planned to identify the stable genotypes. With only more stable genotypes remaining for the final stages of testing, a breeder would be greatly aided in selecting superior genotypes.

The regression of each genotype in an experiment on an environmental index, originally proposed by Yates and Cochran (1938) and later used by Finlay and Wilkinson (1963), has been used to estimate the regression coefficient which is defined as the stability parameter. Regression coefficients measure phenotypic stability; that is, genotypes with regression coefficients of 1.0 have an average stability, whereas coefficients less than or greater than 1.0 indicate above average or below average stability.

In the regression approach, a desirable index, independent of experimental genotypes, is one obtained from environmental factors such as temperature, rainfall, and soil fertility. Current knowledge of the relationship of these factors and phenotypic performance does not permit the computation of such an index. The commonly used substitute for this index is the average phenotypic performance of genotypes in a particular environment.

Several researchers (Eberhart and Russell, 1966; Miezen, Milliken, and Liang, 1979) have evaluated the environmental index by the mean performance of all genotypes grown in that environment. However, estimation and testing have been carried out as if this index had been determined independently of the data. Although some difficulties have been acknowledged (Freeman, 1973), no general study has been undertaken.

The purposes of the research reported herein are to obtain the estimators of stability parameters and derive the distributions of the estimators. Attention will be restricted to two simple designs, namely, the Completely Randomized Disign (CRD) within each environment and the Randomized Complete Block Design (RCBD) within each environment.

In general, we assume that there are n genotypes and r replications within each of the m environments. That is, there are nr observations in each of the m environments. A general statistical model to be used herein is the following:

$$Y_{ijk} = \mu + \tau_i + \beta_i EV_j + e_{ijk} \qquad (1.1)$$

$$i = 1,\ldots,n; \; j = 1,\ldots,m; \; k = 1,\ldots,r.$$

where

Y_{ijk} is the phenotypic performance of the ith genotype at the jth environment on the kth replicate;

μ is the overall mean of the population;

τ_i is the effect of the ith genotype;

β_i is the stability parameter which measures the response of the ith genotype to varying environments;

EV_j is the effect of the jth environment;

e_{ijk} is the random error associated with ith genotype at the jth environment on the kth replication.

The organization of this thesis is as follows: The relevant literature is reviewed in Chapter II. In Chapter III, the Maximum Likelihood method is employed to obtain the estimators of stability parameters when the design is a CRD within each environment. There are two sections in this chapter. A fixed model will be considered in the first section and a mixed model in the second section. In both sections, the model will be described along the basic restrictions and assumptions. After obtaining the estimators of all the parameters, the distribution of the estimators for the stability parameters will be derived. In the end of each section, the sample size for which the estimators will have approximately a normal distribution will be discussed. In Chapter IV, the Restricted Maximum Likelihood method is employed. Again, the model will be described, the estimators for stability parameters will be obtained and the distributions of the estimators will be derived. In Chapter V, we will consider the estimation of the stability parameters when the design is a RCBD within each environment. In the first part of this chapter, the block effects are assumed to be random. In the second part, the block effects are assumed to be fixed. In Chapter VI, a Ratio method and Least Squares will be used to estimate the stability parameters in the case that the distribution of the random errors is not normal. In Chapter VII, we compare M.L.E., R.M.L.E., and Ratio Estimators for the stability parameters via approximate mean squared errors and computer simulation. Chapter VIII contains a brief summary of the thesis and recommendations for further research.

CHAPTER II

REVIEW OF LITERATURE

The existence of interactions between genotypes and environmental factors has long been recognized, the earliest reference, which indeed precedes the analysis of variances, being Fisher and Mackenzie (1923). In considering the manurial responses of different potato varieties they concluded that "the yields of different varieties under different manurial treatments are better fitted by a product formula than by a sum formula".

The idea of breaking down an interaction into several parts was given by Yates and Cochran (1938). In their words, "the degree of association between varietal differences and general fertility can be further investigated by calculating the regression of the yields of all varieties". That is, in terms of a statistical model, the yield Y_{ijk} of the kth replicate of the ith genotype in the jth environment is regarded as made up of a general mean μ, a genotype effect τ_i, an environmental effect EV_j, an interaction effect ξ_{ij}, and a random error ε_{ijk}, i.e.,

$$Y_{ijk} = \mu + \tau_i + EV_j + \xi_{ij} + \varepsilon_{ijk}. \quad (2.1)$$

The ξ_{ij} in equation (2.1) is regressed on EV_j, i.e.,

$$\xi_{ij} = \alpha_i EV_j + \varepsilon^*_{ij} \quad (2.2)$$

where α_i is a linear regression coefficient for the ith genotype and ε^*_{ij}

is a deviation.

Using (2.2), we can rewrite (2.1) as

$$Y_{ijk} = \mu + \tau_i + (1+\alpha_i)EV_j + (\varepsilon^*_{ij} + \varepsilon_{ijk}). \tag{2.3}$$

Yates and Cochran showed that this regression accounted for a large part of interaction in a set of barley trials, but their ideas were not really taken up until Finlay and Wilkinson (1963) rediscovered them. The same method was also used by Perkins and Jinks (1968), who used it for estimating parameters in a biometrical genetical model.

Regression methods were also considered by Rowe and Andrew (1964) and Eberhart and Russell (1966), who defined the environmental index to be the mean performance of all genotypes grown in that environment. In their analysis, they repartitioned the sum of squares for environments and genotype-environment interactions (see Table I). Their partitioning is into a linear component between environments with 1 degree of freedom, a linear component of genotype-environment interaction with n-1 degrees of freedom, and deviations from regression, the deviations being found separately for each of the n genotypes with m-2 degree of freedom each.

The trouble with this approach, as pointed out by Freeman and Perkins (1971), is that the sum of squares for the linear component between environments, which is allocated one degree of freedom, is the same as the sum of squares for environments with m-1 degree of freedom. Perkins and Jinks (1971) recognized that the environmental sum of squares is the same as that for the combined regression overall genotypes but did not use it, thus avoiding this difficulty.

Significance testing in these models is dependent upon even more assumptions than the usual analysis of variance. Eberhart and Russell

TABLE I

ANALYSIS OF VARIANCE WHEN STABILITY PARAMETERS ARE ESTIMATED

Source	D.F.	S.S.
Total	nmr - 1	$\sum_{ijk} Y_{ijk} - \frac{1}{nmr}Y^2_{\cdot\cdot\cdot}$
Genotypes (G)	n - 1	$\frac{1}{mr}\sum_{i} Y^2_{i\cdot\cdot} - \frac{1}{nmr}Y^2_{\cdot\cdot\cdot}$
Environments (E)	m - 1	$\frac{1}{nr}\sum_{j} Y^2_{\cdot j\cdot} - \frac{1}{nmr}Y^2_{\cdot\cdot\cdot}$
G × E	(n-1)(m-1)	$\frac{1}{r}\sum_{i,j} Y^2_{ij\cdot} - \frac{1}{mr}\sum_{i} Y^2_{i\cdot\cdot} - \frac{1}{nr}\sum_{j} Y^2_{\cdot j\cdot} + \frac{1}{nmr}Y^2_{\cdot\cdot\cdot}$
E (Linear)	1	$\frac{1}{nr}(\sum_{j} Y_{\cdot j\cdot}EV_j)^2/\sum_{j} EV_j^2$ *
G × E (Linear)	n - 1	$\frac{1}{r}\{\sum_{i}[(\sum_{j} Y_{ij\cdot}EV_j)^2)/\sum_{j} EV_j^2] - \frac{1}{n}(\sum_{j} Y_{\cdot j\cdot}EV_j)^2/\sum_{j} EV_j^2\}$
Pooled Deviation	n(m-2)	$\frac{1}{r}\sum_{i}\{[\sum_{j} Y^2_{ij\cdot} - \frac{1}{m}(Y_{i\cdot\cdot})^2] - (\sum_{j} Y_{ij\cdot}EV_j)^2/\sum_{j} EV_j^2\}$
G_1	m - 2	$\frac{1}{r}\{[\sum_{j} Y^2_{1j\cdot} - \frac{1}{m}(Y_{1\cdot\cdot})^2] - (\sum_{j} Y_{1j\cdot}EV_j)^2/\sum_{j} EV_j^2\}$
⋮		
G_n	m - 2	$\frac{1}{r}\{[\sum_{j} Y^2_{nj} - \frac{1}{m}(Y_{n\cdot\cdot})^2] - (\sum_{j} Y_{nj\cdot}EV_j)^2/\sum_{j} EV_j^2\}$
Error	nm(r-1)	$\sum_{i,j,k}\{\sum Y^2_{ijk} - \frac{1}{r}Y^2_{ij\cdot}\}$

*$EV_j = \frac{1}{nr}\sum_{ik} Y_{ijk} - \frac{1}{nmr}\sum_{ijk} Y_{ijk}$.

(1966) point out that in their approach the comparison of the linear component of the interaction against deviations from regression assumes that the deviations within the various genotypes are homogeneous. The same is true in the Yates and Cochran (1938) approach. For the same reason, pointed out by Freeman (1973), it is better to test the significance of the estimator of stability parameter for a particular genotype by comparing the appropriate sum of squares against the deviations for regression for that genotype rather than against the pooled deviation term.

Of the techniques discussed, there can be no doubt that the most fruitful has been the regression approach. For its success, a very high proportion of the interaction sum of squares should be explained by linear regression. When, as in the work of Bucio Alanis, Perkins, and Jinks (1969), very good linearity is found by regressing results from different generations of inbred lines on midparental means, the method is unchallengable, predictions across generations being remarkably good. The conditions making for success, i.e., linearity of regression, are very difficult to determine and one set of characteristics has frequently been found to give linear regressions, while other characteristics measured on the same set of genotypes have not.

CHAPTER III

MAXIMUM LIKELIHOOD METHOD

In this chapter, we will employ the Maximum Likelihood method to obtain the estimators of the stability parameters for the fixed model and mixed model. In this study, fixed model means that the environmental effects are assumed to be fixed and mixed model means that the environmental effects are assumed to be random.

3.1 Fixed Model

3.1.1. Model

Consider a set of observations Y_{ijk} classified according to genotypes and environments. The statistical model for the ijk<u>th</u> observation is represented as follows:

$$Y_{ijk} = \mu + \tau_i + \beta_i EV_j + e_{ijk} \tag{3.1}$$

$$i = 1,\ldots,n;\ j = 1,\ldots,m;\ k = 1,\ldots,r,$$

where the same properties hold as model (1.1) with restrictions $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, and $\sum_j EV_j = 0$ and with the assumption that the e_{ijk}'s are i.i.d. $N(0,\sigma^2)$, $\forall$ i,j,k.

Since $e_{ijk} \sim N(0,\sigma^2)$ then $Y_{ijk} \sim N(\mu + \tau_i + \beta_i EV_j, \sigma^2)$. For k = 1,...,r, $\underset{\sim}{Y}_k$'s are i.i.d. $N_{nm}(\mu \underset{\sim}{j}_{nm} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_m + \underset{\sim}{\beta} \otimes \underset{\sim}{EV}, \sigma^2 I_{nm})$, where $\underset{\sim}{Y}_k = (Y_{11k}, Y_{12k}, \ldots, Y_{nmk})'$; $\underset{\sim}{\tau} = (\tau_1, \tau_2, \ldots, \tau_n)'$; $\underset{\sim}{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)'$; $\underset{\sim}{EV} = (EV_1, EV_2, \ldots, EV_m)'$

and $\otimes$ is the direct product (Kronecker product). The probability density function of $\underset{\sim}{Y}_k$ is

$$f_{\underset{\sim}{Y}_k}(\underset{\sim}{y}_k;\mu,\underset{\sim}{\tau},\underset{\sim}{\beta},\underset{\sim}{EV},\sigma^2) = (2\pi)^{-\frac{nm}{2}}(\sigma^2)^{-\frac{nm}{2}}$$

$$\exp -\frac{1}{2\sigma^2}\{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ijk}-\mu-\tau_i-\beta_i EV_j)^2\}.$$

Since from (3.1) $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, and $\sum_j EV_j = 0$, then for any s and t ($s = 1,\ldots,n$; $t = 1,\ldots,m$), we have $\tau_s = -\sum_{i\neq s}\tau_i$, $\beta_s = n - \sum_{i\neq s}\beta_i$ and $EV_t = -\sum_{j\neq t} EV_j$. After reparameterization we rewrite the probability density function of $\underset{\sim}{Y}_k$ as

$$f_{\underset{\sim}{Y}_k}(\underset{\sim}{y}_k;\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\underset{\sim}{EV}^*,\sigma^2) = (2\pi)^{-\frac{nm}{2}}(\sigma^2)^{-\frac{nm}{2}}$$

$$\exp -\frac{1}{2\sigma^2}\{\sum_{i\neq s}\sum_{j\neq t}[y_{ijk}-\mu-\tau_i-\beta_i EV_j]^2 + \sum_{i\neq s}[y_{itk}-\mu-\tau_i+\beta_i(\sum_{j\neq t}EV_j)]^2$$

$$+\sum_{j\neq t}[y_{sjk}-\mu+(\sum_{i\neq s}\tau_i)-(n-\sum_{i\neq s}\beta_i)EV_j]^2+[y_{stk}-\mu+(\sum_{i\neq s}\tau_i)$$

$$+(n-\sum_{i\neq s}\beta_i)(\sum_{j\neq t}EV_j)]^2\} \qquad \forall\ k = 1,\ldots,r$$

where $\underset{\sim}{\tau}^* = (\tau_1,\ldots,\tau_{s-1},\tau_{s+1},\ldots,\tau_n)'$, $\underset{\sim}{\beta}^* = (\beta_1,\ldots,\beta_{s-1},\beta_{s+1},\ldots,\beta_n)'$, $\underset{\sim}{EV}^* = (EV_1,\ldots,EV_{t-1},EV_{t+1},\ldots,EV_m)'$.

3.1.2. Maximum Likelihood Estimators for $\underset{\sim}{\theta}$

Since $\underset{\sim}{Y}_k$'s ($k = 1,\ldots,r$) are i.i.d., then for the random sample $\underset{\sim}{y} = (\underset{\sim}{y}_1',\underset{\sim}{y}_2',\ldots,\underset{\sim}{y}_r')'$, the logarithm of the likelihood function is

$$L(\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\underset{\sim}{EV}^*,\sigma^2;\underset{\sim}{y}) = \ln f_{\underset{\sim}{Y}}(\underset{\sim}{y};\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\underset{\sim}{EV}^*,\sigma^2)$$

$$= \ln\{\prod_{k=1}^{r} f_{\underset{\sim}{Y}_k}(\underset{\sim}{y}_k;\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\underset{\sim}{EV}^*,\sigma^2)\}$$

$$= -\frac{nmr}{2}\ln(2\pi) - \frac{nmr}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{k=1}^{r}\{\sum_{i\neq s}\sum_{j\neq t}[y_{ijk}-\mu-\tau_i-\beta_i EV_j]^2$$

$$+ \sum_{i\neq s}[y_{stk} - \mu - \tau_i + \beta_i(\sum_{j\neq t} EV_j)]^2 + \sum_{j\neq t}[y_{sjk} - \mu + (\sum_{i\neq s}\tau_i)$$

$$- (n - \sum_{i\neq s}\beta_i)EV_j]^2 + [y_{stk} - \mu + (\sum_{i\neq s}\tau_i) + (n - \sum_{i\neq s}\beta_i)(\sum_{j\neq t} EV_j)]^2\} \quad (3.2)$$

Equations for obtaining $\hat{\underset{\sim}{\theta}} = (\hat{\mu}, \hat{\underset{\sim}{\tau}}^{*\prime}, \hat{\underset{\sim}{\beta}}^{*\prime}, \hat{\underset{\sim}{EV}}^{*\prime}, \hat{\sigma}^2)'$, the Maximum Likelihood estimators of $\underset{\sim}{\theta} = (\mu, \underset{\sim}{\tau}^{*\prime}, \underset{\sim}{\beta}^{*\prime}, \underset{\sim}{EV}^{*\prime}, \sigma^2)'$, come from differentiating (3.2) with respect to μ, τ_i, β_i, EV_j, σ^2 for $i = 1,\ldots,n$; $j = 1,\ldots,m$ and $i \neq s$, $j \neq t$:

$$\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2}\{y_{\cdot\cdot\cdot} - nmr\mu\}$$

$$\frac{\partial L}{\partial \tau_i} = \frac{1}{\sigma^2}\{y_{i\cdot\cdot} - y_{s\cdot\cdot} - mr\tau_i - mr(\sum_{i\neq s}\tau_i)\} \qquad \forall\, i,\ i\neq s$$

$$\frac{\partial L}{\partial \beta_i} = \frac{1}{\sigma^2}\{\sum_{j\neq t}[(y_{ij\cdot} - y_{it\cdot})EV_j] - r\beta_i[\sum_{j\neq t} EV_j^2 + (\sum_{j\neq t} EV_j)^2]$$

$$- \sum_{j\neq t}[(y_{sj\cdot} - y_{st\cdot})EV_j] + r(n - \sum_{i\neq s}\beta_i)[\sum_{j\neq t} EV_j^2 + (\sum_{j\neq t} EV_j)^2]\} \quad \forall\, i,\ i\neq s$$

$$\frac{\partial L}{\partial EV_j} = \frac{1}{\sigma^2}\{\sum_{i\neq s}[(y_{ij\cdot} - y_{it\cdot})\beta_i] - r(\sum_{i\neq s}\beta_i^2)[EV_j + (\sum_{j\neq t} EV_j)]$$

$$+ (y_{sj\cdot} - y_{st\cdot})(n - \sum_{i\neq s}\beta_i) - (n - \sum_{i\neq s}\beta_i)^2[EV_j + (\sum_{j\neq t} EV_j)]\} \quad \forall\, j,\ j\neq t$$

$$\frac{\partial L}{\partial \sigma^2} = \frac{nmr}{\sigma^2} - \frac{1}{(\sigma^2)^2}\sum_{k=1}^{r}\{\sum_{i\neq s}\sum_{j\neq t}[y_{ijk} - \mu - \tau_i - \beta_i EV_j]^2 + \sum_{i\neq s}[y_{itk} - \mu - \tau_i$$

$$+ \beta_i(\sum_{j\neq t} EV_j)]^2 + \sum_{j\neq t}[y_{sjk} - \mu + (\sum_{i\neq s}\tau_i) - (n - \sum_{i\neq s}\beta_i)EV_j]^2 + [y_{stk}$$

$$- \mu + (\sum_{j\neq s}\tau_i) + (n - \sum_{i\neq s}\beta_i)(\sum_{j\neq t} EV_j)]^2\}$$

where $y_{\cdot\cdot\cdot} = \sum_{ijk} y_{ijk}$, $y_{i\cdot\cdot} = \sum_{jk} y_{ijk}$, $y_{ij\cdot} = \sum_k y_{ijk}$.

Equating each of the above $2n + m - 1$ partial derivatives to zero, gives the Maximum Likelihood equations. The roots of the equations are

$$\hat{\mu} = \bar{y}_{\cdots} \qquad (3.3)$$

$$\hat{\tau}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots} \qquad \forall\, i,\ i \neq s \qquad (3.4)$$

$$\hat{\beta}_i = 1 + \frac{\sum_{j\neq t} [(\bar{y}_{ij\cdot} - \bar{y}_{\cdot j\cdot}) - (\bar{y}_{it\cdot} - \bar{y}_{\cdot t\cdot})]\hat{EV}_j}{[\sum_{j\neq t} \hat{EV}_j^2 + (\sum_{j\neq t} \hat{EV}_j)^2]} \qquad \forall\, i,\ i\neq s \qquad (3.5)$$

$$\hat{EV}_j = \frac{[\sum_{i\neq s} (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})\hat{\beta}_i] + [(\bar{y}_{sj\cdot} - \bar{y}_{s\cdot\cdot})(n - \sum_{i\neq s} \hat{\beta}_i)]}{[\sum_{i\neq s} \hat{\beta}_i^2 + (n - \sum_{i\neq s} \hat{\beta}_i)^2]} \qquad \forall\, j,\ j\neq t \qquad (3.6)$$

$$\hat{\sigma}^2 = \frac{1}{nmr} \sum_k \{ \sum_{i\neq s} \sum_{j\neq t} [y_{ijk} - \bar{y}_{i\cdot\cdot} - \hat{\beta}_i \hat{EV}_j]^2 + \sum_{i\neq s} [y_{itk} - \bar{y}_{i\cdot\cdot} + \hat{\beta}_i (\sum_{j\neq t} \hat{EV}_j)]^2$$
$$+ \sum_{j\neq t} [y_{sjk} - \bar{y}_{s\cdot\cdot} - (n - \sum_{i\neq s} \hat{\beta}_i)\hat{EV}_j]^2 + [y_{stk} - \bar{y}_{s\cdot\cdot} + (n - \sum_{i\neq s} \hat{\beta}_i)$$
$$(\sum_{j\neq t} \hat{EV}_j)]^2\} \qquad (3.7)$$

where $\bar{y}_{\cdots} = \frac{1}{nmr} y_{\cdots}$, $\bar{y}_{i\cdot\cdot} = \frac{1}{mr} y_{i\cdot\cdot}$, $\bar{y}_{\cdot j\cdot} = \frac{1}{nr} \sum_{ik} y_{ijk}$, and $\bar{y}_{ij\cdot} = \frac{1}{r} y_{ij\cdot}$.

From (3.5), (3.6), (3.7) we can see that none of $\hat{\beta}_i$, $\hat{EV}_j$, $\hat{\sigma}^2$ are explicit functions of y_{ijk}. Note that $\hat{\beta}_i$ is a function of $\hat{\underset{\sim}{EV}}^*$, $\hat{EV}_j$ is a function of $\hat{\underset{\sim}{\beta}}^*$, and $\hat{\sigma}^2$ is a function of both $\hat{\underset{\sim}{\beta}}^*$ and $\hat{\underset{\sim}{EV}}^*$. From (3.7) we can see that $\hat{\sigma}^2$ is strictly positive for any value of $\hat{\underset{\sim}{\beta}}^*$ and $\hat{\underset{\sim}{EV}}^*$ (except for the case $y_{ijk} - \bar{y}_{i\cdot\cdot} = \hat{\beta}_i \hat{EV}_j \quad \forall\, i,j,k$ then $\hat{\sigma}^2 = 0$). To obtain a numerical solution of $\hat{\underset{\sim}{\beta}}^*$ and $\hat{\underset{\sim}{EV}}^*$, an iterative method must be used. A computer program to obtain the numerical solution of $\hat{\underset{\sim}{\theta}}$ is attached in the Appendix A.

To show the root of the Maximum Likelihood equations is a global (or local) maximum point of the likelihood function, we need to show the Hessian matrix $H(\theta)$ which is the matrix of second partial derivatives of the logarithm of the likelihood function with respect to the parameters is negative definite when evaluated at $\hat{\underset{\sim}{\theta}}$.

$$H(\hat{\underset{\sim}{\theta}}) = \frac{\partial^2 L}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}'}\Bigg|_{\hat{\underset{\sim}{\theta}}} = -\frac{r}{\hat{\sigma}^2}\begin{bmatrix} nm & 0 & 0 & 0 & 0 \\ 0 & m(I_{n-1}+J_{n-1}) & 0 & 0 & 0 \\ 0 & 0 & \lambda_1(J_{n-1}+J_{n-1}) & A_{21} & \underset{\sim}{a}_{31} \\ 0 & 0 & A_{12} & \lambda_2(I_{m-1}+J_{m-1}) & \underset{\sim}{a}_{32} \\ 0 & 0 & \underset{\sim}{a}_{13} & \underset{\sim}{a}_{23} & \frac{nm}{\hat{\sigma}^2} \end{bmatrix}$$

where

$$\lambda_1 = [\sum_{j\neq t} \hat{EV}_j^2 + (\sum_{j\neq t} \hat{EV}_j)^2]$$

$$\lambda_2 = [\sum_{i\neq s} \hat{\beta}_i^2 + (n - \sum_{i\neq s} \hat{\beta}_i)^2]$$

$$A_{12}(q,p) = A_{21}(p,q) = 2[\hat{\beta}_p - (n - \sum_{i\neq s} \hat{\beta}_i)][\hat{EV}_q + (\sum_{j\neq t} \hat{EV}_j)] - [y_{pq\cdot} - y_{pt\cdot} - y_{sq\cdot} + y_{st\cdot}] \quad \forall\, p=1,\ldots,n-1;\ q=1,\ldots,m-1$$

$$\underset{\sim}{a}_{13}(p) = \underset{\sim}{a}_{31}(p) = \frac{1}{\hat{\sigma}^2}\{[\sum_{j\neq t} \bar{y}_{pj\cdot} - \bar{y}_{pt\cdot} - \bar{y}_{sj\cdot} + \bar{y}_{st\cdot})\hat{EV}_j] - [\hat{\beta}_p - (n - \sum_{i\neq s} \hat{\beta}_i)] [\sum_{j\neq t} EV_j^2 + (\sum_{j\neq t} EV_j)^2]\}$$

$$\underset{\sim}{a}_{23}(q) = \underset{\sim}{a}_{32}(q) = \frac{1}{\hat{\sigma}^2}[\sum_{i\neq s} (\bar{y}_{iq\cdot} - \bar{y}_{it\cdot})\hat{\beta}_i] + [(\bar{y}_{sq\cdot} - \bar{y}_{st\cdot})(n - \sum_{i\neq s} \hat{\beta}_i)] - [\sum_{i\neq s} \hat{\beta}_i^2 + (n - \sum_{i\neq s} \hat{\beta}_i)^2][\hat{EV}_q + (\sum_{j\neq t} \hat{EV}_j)]\} \quad \forall\, q=1,\ldots,m-1.$$

From (3.5) and (3.6) we have $\underset{\sim}{a}_{13} = \underset{\sim}{a}_{31}' = 0$, $\underset{\sim}{a}_{23} = \underset{\sim}{a}_{32}' = 0$, and $A_{12} = A_{21}' = [\hat{\underset{\sim}{\beta}}^* - (n - \sum_{i\neq s} \hat{\beta}_i)\underset{\sim}{j}_{n-1}][\hat{\underset{\sim}{EV}}^* + (\sum_{j\neq t} EV_j)\underset{\sim}{j}_{m-1}]'$.

For $q = 1,\dots,m-1$, we have

$$|B_q| = \left|\begin{bmatrix} \lambda_1(I_{n-1}+J_{n-1}), & [\underset{\sim}{\beta}^* - (n - \sum_{i\neq s} \hat{\beta}_i)\underset{\sim}{J}_{n-1}][\hat{\underset{\sim}{EV}}{}^*_q + (\sum_{j\neq t} EV_j)J_q]' \\ [\hat{\underset{\sim}{EV}}{}^*_q + (\sum_{j\neq t} \hat{EV}_j)J_q][\hat{\underset{\sim}{\beta}}{}^* - (n - \sum_{i\neq s} \hat{\beta}_i)\underset{\sim}{J}_{n-1}], & \lambda_2(I_q+J_q) \end{bmatrix}\right|$$

$$= (n\lambda_1^{n-1})(q+1)(\sum_{j>q} \hat{EV}_j^2)^q > 0. \tag{3.8}$$

Since λ_1, λ_2 are always positive and the determinant of $I_p + J_p$ is $p+1$, then from (3.8) we have the leading principal minors of $-H(\hat{\underset{\sim}{\theta}})$ are positive. That is, $H(\hat{\underset{\sim}{\theta}})$ is negative definite for all $\hat{\underset{\sim}{\theta}}$. Therefore, $\hat{\underset{\sim}{\theta}}$ is at least a local maximum point of the likelihood function.

$L(\underset{\sim}{\theta};\underset{\sim}{y})$ is differentiable in the space $\Omega = R^{2n+m-2} \times R^+$ $(R^+ = (0,+\infty))$, that is $L(\underset{\sim}{\theta},y)$ is differentiable in a neighborhood of the true parameter $\underset{\sim}{\theta}_0$. Then by the properties of M. L. equations, there is a root with probability 1 as $r \to \infty$, which is consistent for $\underset{\sim}{\theta}$. That is $\hat{\underset{\sim}{\theta}}$ is a consistent estimator of $\underset{\sim}{\theta}$.

3.1.3. Large Sample Distribution of the Estimators

From (3.5) we can see that $\hat{\beta}_i$ is not an explicit function of y_{ijk}. Consequently, the exact distribution of $\hat{\beta}_i$ can not be obtained. If the number of replications is large, we can employ the large sample method to derive the approximate distribution of $\hat{\beta}_i$.

$\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \dots, \underset{\sim}{Y}_r$ are i.i.d. with distribution $F_{\underset{\sim}{\theta}}$ belonging to $F = \{F_{\underset{\sim}{\theta}} ; \underset{\sim}{\theta} \in \Omega = R^{2n+m-2} \times R^+\}$. The following conditions hold:

(R1) For each $\underset{\sim}{\theta} \in \Omega$, the first derivatives of $L(\underset{\sim}{\theta};\underset{\sim}{y}_k)$ with respect to $\underset{\sim}{\theta}$ exist for all $\underset{\sim}{y}_k$.

(R2) For each $\theta_0 \in \Omega$, there exists a function $M(\underset{\sim}{y}_k)$ (possibly depending on $\underset{\sim}{\theta}_0$) such that for every $\underset{\sim}{\theta}$ belong to $N(\underset{\sim}{\theta}_0)$, a neighborhood of $\underset{\sim}{\theta}_0$, the relations

$$|L'''(\underset{\sim}{\theta};\underset{\sim}{y}_k)| < M(\underset{\sim}{y}_k)$$

hold for all $\underset{\sim}{y}_k$ and $E\{M(\underset{\sim}{y}_k)\}$ exists for $\underset{\sim}{\theta} \in N(\underset{\sim}{\theta}_0)$.

(R3) For each $\underset{\sim}{\theta} \in \Omega$, $E\{-\frac{\partial^2 L(\underset{\sim}{\theta};\underset{\sim}{y}_k)}{\partial\underset{\sim}{\theta}\,\partial\underset{\sim}{\theta}} \mid \underset{\sim}{\theta}\}$ is finite and positive defiite.

Then the asymptotic distribution of $\sqrt{r}\,(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0)$ is $N_{2n+m-1}(\underset{\sim}{0}, i^{-1}(\underset{\sim}{\theta}_0))$ where

$$i(\underset{\sim}{\theta}_0) = \frac{1}{\sigma^2}\begin{bmatrix} nm, & 0, & 0, & 0, & 0 \\ 0, & m(I_{n-1}+J_{n-1}), & 0, & 0 & 0 \\ 0, & 0, & \Sigma_{11}, & \Sigma_{12}, & 0 \\ 0, & 0, & \Sigma_{21}, & \Sigma_{22}, & 0 \\ 0, & 0, & 0, & 0, & \frac{nm}{2\sigma^2} \end{bmatrix}$$

$$\Sigma_{11} = [\sum_{j\neq t} EV_j^2 + (\sum_{j\neq t} EV_j)^2](I_{n-1} + J_{n-1})$$

$$\Sigma_{21} = \Sigma_{12}' = [E\underset{\sim}{V}^* - (\sum_{j\neq t} EV_j)\underset{\sim}{j}_{m-1}][\underset{\sim}{\beta}^* - (n - \sum_{i\neq s} \beta_i)\underset{\sim}{j}_{n-1}]'$$

$$\Sigma_{22} = [\sum_{i\neq s} \beta_i^2 + (n - \sum_{i\neq s} \beta_i)^2](I_{m-1} + J_{m-1})$$

$$\sqrt{r}\left(\begin{bmatrix}\mu\\ \hat{\tau}^*\\ \hat{\beta}^*\\ \hat{EV}^*\\ \hat{\sigma}^2\end{bmatrix} - \begin{bmatrix}\mu\\ \tau^*\\ \beta^*\\ EV^*\\ \sigma^2\end{bmatrix}\right) \overset{\mathcal{L}}{\dot{\sim}} N_{2n+m-1}\left(0,\hat{\sigma}^2\begin{bmatrix}\frac{1}{nm}, & & & & \\ & \frac{1}{m}(I_{n-1}-\frac{1}{n}J_{n-1}), & & & \\ & & V_{11}, & V_{12}, & \\ & & V_{21}, & V_{22}, & \\ & & & & \frac{2\hat{\sigma}^2}{nm}\end{bmatrix}\right) \qquad (3.9)$$

where

$$V_{11} = \frac{1}{[\sum_{j\neq t}\hat{EV}_j^2+(\sum_{j\neq t}\hat{EV}_j)^2]}\,[I_{n-1}-\frac{1}{n}J_{n-1}+\frac{1}{n}(\hat{\beta}^*-j_{n-1})(\hat{\beta}^*-j_{n-1})']$$

$$V_{21} = V_{12}' = -\frac{1}{n[\sum_{j\neq t}\hat{EV}_j^2+(\sum_{j\neq t}\hat{EV}_j)^2]}\,\hat{EV}^*(\hat{\beta}^*-j_{n-1})'$$

$$V_{22} = \frac{1}{[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]}[I_{m-1}-\frac{1}{m}J_{m-1}+\frac{[\sum_{i\neq s}(\hat{\beta}_i-1)^2+(n-\sum_{i\neq s}\hat{\beta}_i-1)^2]}{n[\sum_{j\neq t}\hat{EV}_j^2+(\sum_{j\neq t}\hat{EV}_j)^2]}EV^*\,EV^{*\prime}].$$

From the restrictions of (3.1) we have $\beta_s = n - \sum_{i\neq s}\beta_i$, $EV_t = -\sum_{j\neq t}EV_j$. That implies $\hat{\beta}_s = n - \sum_{i\neq s}\hat{\beta}_i$ and $\hat{EV}_t = \sum_{j\neq t}\hat{EV}_j$. Then from (3.9) as $r \to \infty$ we have

$$\frac{\sqrt{r}(\sum_{j=1}^{m}\hat{EV}_j^2)^{1/2}}{\sqrt{\sigma^2}[1-\frac{1}{n}+\frac{1}{n}(\hat{\beta}_s-1)^2]^{1/2}}(\hat{\beta}_s-\beta_s) \xrightarrow{\mathcal{L}} N(0,1)$$

The asymptotic distribution of $\hat{\beta}$ as $r \to \infty$ is

$$\hat{\underset{\sim}{\beta}} \overset{\mathcal{L}}{\dot{\sim}} SN(\underset{\sim}{\beta}, \frac{\hat{\sigma}^2}{r(\sum_{j=1}^{m} \hat{EV}_j^2)}[I_n - \frac{1}{n}J_n + \frac{1}{n}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)']). \tag{3.10}$$

3.1.4. Testing Hypothesis about Equality of β_i

From (3.10), the distribution of $\hat{\beta}_i - \hat{\beta}_{i'}$ is

$$\hat{\beta}_i - \hat{\beta}_{i'} \dot{\sim} N((\beta_i - \beta_{i'}), \frac{\hat{\sigma}}{r(\sum_j \hat{EV}_j^2)} [2 + \frac{1}{n}(\hat{\beta}_i - \hat{\beta}_{i'})^2]).$$

For testing the hypothesis

$$H_0: \beta_i = \beta_{i'} \quad \text{vs.} \quad H_a: \beta_i \neq \beta_{i'}.$$

The distribution of the test statistic under the null hypothesis is given as

$$S = \frac{\hat{\beta}_i - \hat{\beta}_{i'}}{\sqrt{\frac{2\hat{\sigma}^2}{r(\sum_{j=1}^{m} \hat{EV}_j^2)}}} \xrightarrow[\text{as } r\to\infty]{\mathcal{L}} N(0,1). \tag{3.11}$$

As a consequence of these tests, one could partition the β_i $(i = 1,\ldots,n)$ into several groups such that in the same group of β_i's there is no significant genotype-environment interaction. After adjusting the $\hat{\beta}_i$'s in that group, we can compare the genotype effects. Then the best genotype in that group could be selected.

3.1.5. The Sample Size r

In (3.9), we assume the sample size r is large. In applications, one needs to know what is "large." Described below are simulation results showing how large r must be to make $\hat{\underset{\sim}{\theta}}$ approximately normally distributed.

In the computer simulation, we take the number of genotypes n as 3 or 5, the number of environments m as 5 or 8, the number of replications r as 3, 10, 20, or 30, and the variance of sampling error σ^2 as 0.01 or 1. In each of the 32 cases, we generate 300 sets of nmr observations; the values of the parameters are also generated by the random number generator.

After obtaining 300 $\hat{\theta}$'s, a test of multivariate normality of these 300 $\hat{\theta}$'s should be given. But there is no suitable method available for testing multivariate normality. If $\hat{\theta}$ is distributed as multivariate normal, then $\hat{\theta}^* = (\hat{\theta} - \theta_0)(i(\theta_0))^{1/2}$ will be distributed as multivariate normal with mean zero and variance I_{2n+m-1}. Then, we test for multivariate normality of $\hat{\theta}$ by testing for univariate normality of $\hat{\theta}^*_p$ $(p = 1,\ldots, 2n+m-1)$ and testing for zero correlation between $\hat{\theta}^*_p$ and $\hat{\theta}^*_q$ $(p \neq q)$, where $\hat{\theta}^*_p$ is the component of $\hat{\theta}^*$. If any one of $\hat{\theta}^*_p$ fail to be normally distributed, then $\hat{\theta}^*$ would not be distributed as multivariate normal. That is, from the smallest O.S.L. among all the 2n+m-1 tests we can see how well the distribution of $\hat{\theta}^*$ approximates to the multivariate normal.

The simulation results show that the tests of zero correlation coefficient among $\hat{\theta}^*_p$'s are insignificant for all cases. A table of the smallest O.S.L. value among all the 2n+m-1 O.S.L.'s for testing the normality of $\hat{\theta}^*_p$ is shown as Table II; where it can be seen that for most cases, the normal distribution obtains for at least 10 replications.

We are primarily interested in the distribution of $\hat{\beta}$, and from the simulation results we know the distribution of $\hat{\beta}_i$ $(i = 1,\ldots,n-1)$ will approach to a normal distribution faster than $\hat{\sigma}^2$. That means, for fewer replications, $\hat{\beta}$ will have an approximate normal distribution than $\hat{\theta}^*$ did.

A plot of the cumulative distribution functions (CDF) of the stan-

TABLE II

SIMULATION RESULT FOR TESTING THE NORMALITY OF $\hat{\underset{\sim}{\theta}}$ IN THE FIXED MODEL

r	n	m	σ^2	O.S.L.	r	n	m	σ^2	O.S.L.
3	3	5	1	<0.01	3	5	5	1	0.079
10	3	5	1	>0.15	10	5	5	1	0.120
20	3	5	1	0.098	20	5	5	1	0.095
30	3	5	1	0.115	30	5	5	1	0.113
3	3	5	0.01	<0.01	3	5	5	0.01	0.161
10	3	5	0.01	0.082	10	5	5	0.01	0.049
20	3	5	0.01	0.108	20	5	5	0.01	0.102
30	3	5	0.01	>0.15	30	5	5	0.01	>0.15
3	3	8	1	0.073	3	5	8	1	<0.01
10	3	8	1	0.078	10	5	8	1	0.037
20	3	8	1	>0.15	20	5	8	1	0.091
30	3	8	1	0.116	30	5	8	1	>0.15
3	3	8	0.01	0.038	3	5	8	0.01	0.064
10	3	8	0.01	0.139	10	5	8	0.01	0.107
20	3	8	0.01	>0.15	20	5	8	0.01	0.089
30	3	8	0.01	>0.15	30	5	8	0.01	>0.15

dard normal and $\hat{\beta}_i^* = (\hat{\beta}_i - \beta_i)/(\mathrm{Var}(\hat{\beta}_i))^{1/2}$ provides this comparison. Two figures for the CDF of $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$ for the case $n = 3$, $m = 5$, $\sigma^2 = 1$ and $r = 3$ or 10 are given in Figure 1 and 2 of Appendix C. From Figure 1, we can see the CDF's of $\hat{\beta}_1^*$ for $r = 3$ and 10 are close to the CDF of the standard normal. The same was observed for Figure 2 and the other cases. This means, $\hat{\underset{\sim}{\beta}}$ will be approximately normally distributed for as few as 3 replications.

3.2 Mixed Model

3.2.1. Model

Consider a set of observations Y_{ijk} classified according to genotypes and environments. The statistical model for the ijkth observation is represented as follows:

$$Y_{ijk} = \mu + \tau_i + \beta_i EV_j + e_{ijk} \tag{3.12}$$

$$i = 1,\ldots,n; \quad j = 1,\ldots,m; \quad k = 1,\ldots,r.$$

Where the same properties hold as model (1.1) with restrictions $\sum_{i=1}^{n} \tau_i = 0$, $\sum_{i=1}^{n} \beta_i = n$ and with the assumption that the EV_j's are i.i.d. $N(0,\sigma_E^2)$ $\forall j$ and e_{ijk}'s are i.i.d. $N(0,\sigma_0^2)$ $\forall i,j,k$. Also we assume EV_j's are independent of e_{ijk}'s $\forall i,j,k$. The only difference from the model of the previous section is the randomness of the EV_j's.

For $j = 1,\ldots,m$, $\underset{\sim}{Y}_j$'s are i.i.d. $N_{nr}(\mu \underset{\sim}{j}_{nr} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_r, V)$, where $\underset{\sim}{Y}_j = (Y_{1j1},\ldots,Y_{1jr},\ldots,Y_{nj1},\ldots,Y_{njr})'$ and $V = \sigma_E^2(\underset{\sim}{\beta}\underset{\sim}{\beta}') \otimes J_r + \sigma_0^2 I_{nr}$. The probability density function of $\underset{\sim}{Y}_j$ $\forall j = 1,\ldots,m$ is:

$$f_{\underset{\sim}{Y}_j}(\underset{\sim}{y}_j;\mu,\underset{\sim}{\tau},\underset{\sim}{\beta},\sigma_E^2,\sigma_0^2) = (2\pi)^{-\frac{nr}{2}}|V|^{-\frac{1}{2}}$$

$$\exp -\frac{1}{2}\{[\underset{\sim}{y}_j - (\mu\underset{\sim}{j}_{nr} + \underset{\sim}{\tau}\otimes\underset{\sim}{j}_r)]'V^{-1}[\underset{\sim}{y}_j - (\mu\underset{\sim}{j}_{nr} + \underset{\sim}{\tau}\otimes\underset{\sim}{j}_r)]$$

where $|V| = (\sigma_0^2)^{nr-1}[\sigma_0^2 + r\sigma_E^2(\sum_{i=1}^{n}\beta_i^2)]$ and

$$V^{-1} = \frac{1}{\sigma_0^2}[I_{nr} - \frac{\sigma_E^2}{\sigma_0^2 + r\sigma_E^2(\sum_{i=1}^{n}\beta_i^2)}(\underset{\sim}{\beta}\otimes\underset{\sim}{j}_r)(\underset{\sim}{\beta}\otimes\underset{\sim}{j}_r)'].$$

Then

$$f_{\underset{\sim}{Y}_j}(\underset{\sim}{y}_j;\mu,\underset{\sim}{\tau},\underset{\sim}{\beta},\sigma_E^2,\sigma_0^2) = (2\pi)^{-\frac{nr}{2}}(\sigma_0^2)^{-\frac{nr-1}{2}}[\sigma_0^2 + r\sigma_E^2(\sum_{i=1}^{n}\beta_i^2)]^{-\frac{1}{2}}$$

$$\exp -\frac{1}{2\sigma_0^2}\{\sum_{ik}(y_{ijk}-\mu-\tau_i)^2 - \frac{\sigma_E^2}{\sigma_0^2+r\sigma_E^2(\sum_i\beta_i^2)}[\sum_{ij}(y_{ijk}-\mu-\tau_i)\beta_i]^2\}.$$

Since $\sum_{i=1}^{n}\tau_i = 0$ and $\sum_{i=1}^{n}\beta_i = n$ (from (3.12)), then for any s, $s = 1,2, \ldots,n$, we have $\tau_s = -\sum_{i\neq s}\tau_i$ and $\beta_s = n - \sum_{i\neq s}\beta_i$. After reparameterization we can rewrite the probability density function of $\underset{\sim}{Y}_j$ $\forall j = 1,\ldots,m$ as

$$f_{\underset{\sim}{Y}_j}(\underset{\sim}{y}_j;\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\sigma_E^2,\sigma_0^2) = (2\pi)^{\frac{nr}{2}}(\sigma_0^2)^{-\frac{nr-1}{2}}[\sigma_0^2 + r\sigma_E^2(\sum_{i\neq s}\beta_i^2 + (n -$$

$$\sum_{i\neq s}\beta_i)^2)]^{-\frac{1}{2}} \exp -\frac{1}{2\sigma_0^2}\{\sum_{i\neq s}\sum_k[y_{ijk}-\mu-\tau_i]^2 + \sum_k[y_{sjk}-\mu+(\sum_{i\neq s}\tau_i)]^2$$

$$- \frac{\sigma_E^2}{\sigma_0^2+r\sigma_E^2(\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2)}[\sum_{i\neq s}\sum_k(y_{ijk}-\mu-\tau_i)\beta_i$$

$$+ \sum_k(y_{sjk} - \mu + (\sum_{i\neq s}\tau_i))(n - \sum_{i\neq s}\beta_i)]^2\}$$

where $\underset{\sim}{\tau}^* = (\tau_1,\ldots,\tau_{s-1},\tau_{s+1},\ldots,\tau_n)'$, $\underset{\sim}{\beta}^* = (\beta_1,\ldots,\beta_{s-1},\beta_{s+1},\ldots,\beta_n)'$.

3.2.2. Maximum Likelihood Estimator for $\underset{\sim}{\theta}$

Since $\underset{\sim}{Y}_j$'s $(j=1,\ldots,m)$ are i.i.d. then for the random sample $\underset{\sim}{y} = (\underset{\sim}{y}_1',\ldots,\underset{\sim}{y}_m')'$, the logarithm of the likelihood function is

$$L(\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\sigma_E^2,\sigma_0^2;\underset{\sim}{y}) = -\frac{nmr}{2}\ln(2\pi) - \frac{(nr-1)m}{2}\ln(\sigma_0^2)$$

$$-\frac{m}{2}\ln\{\sigma_0^2 + r\sigma_E^2[\sum_{i\neq s}\beta_i^2 + (n-\sum_{i\neq s}\beta_i)^2]\} - \frac{1}{2\sigma_0^2}\sum_j\{\sum_{i\neq s}\sum_k[y_{ijk}-\mu-\tau_i]^2$$

$$+\sum_k[y_{sjk}-\mu+(\sum_{i\neq s}\tau_i)]^2\} + \frac{\sigma_E^2}{2\sigma_0^2\{\sigma_0^2 + r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}}\sum_j\{\sum_{i\neq s}\sum_k[y_{ijk}$$

$$-\mu-\tau_i]\beta_i + \sum_k[y_{sjk}-\mu+(\sum_{i\neq s}\tau_i)](n-\sum_{i\neq s}\beta_i)\}^2 \qquad (3.13)$$

Equations for obtaining $\hat{\underset{\sim}{\theta}} = (\hat{\mu},\hat{\underset{\sim}{\tau}}^{*\prime},\hat{\underset{\sim}{\beta}}^{*\prime},\hat{\sigma}_E^2,\hat{\sigma}_0^2)'$, the Maximum Likelihood estimators of $\underset{\sim}{\theta} = (\mu,\underset{\sim}{\tau}^*,\underset{\sim}{\beta}^*,\sigma_E^2,\sigma_0^2)$, come from differentiating (3.13) with respect to μ, τ_i, β_i, σ_E^2 and σ_0^2 for $i = 1,\ldots,n$ and $i \neq s$:

$$\frac{\partial L}{\partial\mu} = \frac{1}{\sigma_0^2}\{y_{\cdots} - nmr\mu - \frac{nr\sigma_E^2}{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]}\sum_j[\sum_{i\neq s}(y_{ij\cdot}-y_{sj\cdot})\beta_i+ny_{sj\cdot}$$

$$-nr\mu-r\sum_{i\neq s}\tau_i\beta_i+r(\sum_{i\neq s}\tau_i)(n-\sum_{i\neq s}\beta_i)]\}$$

$$\frac{\partial L}{\partial\tau_i} = \frac{1}{\sigma_0^2}\{y_{i\cdot\cdot}-y_{s\cdot\cdot}-mr[\tau_i+(\sum_{i\neq s}\tau_i)] - \frac{r\sigma_E^2[\beta_i-(n-\sum_{i\neq s}\beta_i)]}{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]}\sum_j[\sum_{i\neq s}$$

$$(y_{ij\cdot}-y_{sj\cdot})\beta_i+ny_{sj\cdot}-nr\mu-r\sum_{i\neq s}\tau_i\beta_i+r(\sum_{i\neq s}\tau_i)(n-\sum_{i\neq s}\beta_i)] \qquad \forall i\neq s$$

$$\frac{\partial L}{\partial\beta_i} = -mr\sigma_E^2[\beta_i-(n-\sum_{i\neq s}\beta_i)]\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}^{-1} - \frac{r(\sigma_E^2)^2}{\sigma_0^2}$$

$$[\beta_i-(n-\sum_{i\neq s}\beta_i)]\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}^{-2}\sum_j\{\sum_{i\neq s}(y_{ij\cdot}-y_{sj\cdot})\beta_i$$

$$+ny_{sj\cdot}-nr\mu-r\sum_{i\neq s}\tau_i[\beta_i-(n-\sum_{i\neq s}\beta_i)]\}^2+\frac{\sigma_E^2}{\sigma_0^2}\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2$$

$$+(n-\sum_{i\neq s}\beta_i)^2]\}^{-1}\sum_j\{y_{ij\cdot}-y_{sj\cdot}-r(\tau_i+\sum_{i\neq s}\tau_i)\}\{\sum_{i\neq s}(y_{ij\cdot}-y_{sj\cdot})\beta_i$$

$$+ny_{sj\cdot}-nr\mu-r\sum_{i\neq s}\tau_i[\beta_i-(n-\sum_{i\neq s}\beta_i)]\}\quad \forall i\neq s$$

$$\frac{\partial L}{\partial\sigma_E^2}=-\frac{mr}{2\sigma_0^2}[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}^{-1}+\frac{1}{2\sigma_0^2}\{\sigma_0^2+r\sigma_E^2$$

$$[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}^{-2}\sum_j\{\sum_{i\neq s}(y_{ij\cdot}-r\mu-r\tau_i)\beta_i+[y_{sj\cdot}-r\mu+r(\sum_{i\neq s}\tau_i)]$$

$$(n-\sum_{i\neq s}\beta_i)\}^2$$

$$\frac{\partial L}{\partial\sigma_0^2}=-\frac{(nr-1)m}{2\sigma_0^2}-\frac{m}{2}\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2+(n-\sum_{i\neq s}\beta_i)^2]\}^{-1}+\frac{1}{2(\sigma_0^2)^2}\{\sum_{i\neq s,j,k}(y_{ijk}$$

$$-\mu+\tau_i)^2+\sum_{j,k}[y_{sjk}-\mu+(\sum_{i\neq s}\tau_i)]^2\}-\frac{\sigma_E^2}{2(\sigma_0^2)^2}\{\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2$$

$$+(n-\sum_{i\neq s}\beta_i)^2]\}^{-2}\{2\sigma_0^2+r\sigma_E^2[\sum_{i\neq s}\beta_i^2)^2]\}\sum_j\{\sum_{i\neq s}(y_{ij\cdot}-r\mu-r\tau_i)\beta_i+[y_{sj\cdot}$$

$$-r\mu+r(\sum_{i\neq s}\tau_i)](n-\sum_{i\neq s}\beta_i)\}^2.$$

Equating each of the above 2n+1 partial derivatives to zero results in the Maximum Likelihood equations and the roots of the equations are:

$$\hat{\mu}=\overline{y}_{\cdot\cdot\cdot}\qquad(3.14)$$

$$\hat{\tau}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot} \qquad \forall i = 1,\ldots,n \text{ and } i \neq s \tag{3.15}$$

$$\hat{\beta}_i = 1 + \frac{[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]\sum_j\{(\bar{y}_{ij\cdot}-\bar{y}_{\cdot j\cdot})[\sum_{i\neq s}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})\hat{\beta}_i+(\bar{y}_{sj\cdot}-\bar{y}_{s\cdot\cdot})(n-\sum_{i\neq s}\hat{\beta}_i)]\}}{\sum_j\{\sum_{i\neq s}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})\hat{\beta}_i+(\bar{y}_{sj\cdot}-\bar{y}_{s\cdot\cdot})(n-\sum_{i\neq s}\hat{\beta}_i)\}^2}$$

$$\forall i = 1,\ldots,n \text{ and } i \neq s \tag{3.16}$$

$$\hat{\sigma}_E^2 = \frac{1}{(nr-1)m[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]}\sum_j\{nr[\sum_{i\neq s}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})\hat{\beta}_i+(\bar{y}_{sj\cdot}-y_{s\cdot\cdot})$$

$$(n-\sum_{i\neq s}\beta_i)]^2-\frac{1}{r}[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]\sum_{i,k}(y_{ijk}-\bar{y}_{i\cdot\cdot})^2\} \tag{3.17}$$

$$\hat{\sigma}_0^2 = \frac{1}{(nr-1)m[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]}\sum_j\{[\sum_{i\neq s}\hat{\beta}_i^2+(n-\sum_{i\neq s}\hat{\beta}_i)^2]\sum_{i,k}(y_{ijk}-y_{i\cdot\cdot})^2$$

$$-r[\sum_{i\neq s}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})\hat{\beta}_i+(\bar{y}_{sj\cdot}-\bar{y}_{\cdot\cdot\cdot})(n-\sum_{i\neq s}\hat{\beta}_i)]^2\} \tag{3.18}$$

From (3.16), (3.17), and (3.18) we can see that none of $\hat{\beta}_i$, $\hat{\sigma}_E^2$, $\hat{\sigma}_0^2$ are explicit functions of y_{ijk}. Note that $\hat{\beta}_i$ is a function of all $\hat{\beta}_i$ and both $\hat{\sigma}_E^2$ and $\hat{\sigma}_0^2$ are also functions of all $\hat{\beta}_i$ $(i = 1,\ldots,n$ and $i \neq s)$. By using the Cauchy-Schwarz inequality, we can show that for any solution of $\hat{\beta}_i$'s, except for $\hat{\beta}_i = \lambda(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})$ $\forall i$. (If in this case, $\sigma_E^2 = 0$), $\hat{\sigma}_0^2$ is is always positive. But it is very difficult to show $\hat{\sigma}_E^2 > 0$ algebraically. To derive a numerical solution of $\hat{\underset{\sim}{\beta}}^*$, we can use an iterative method. A computer program to obtain the numerical solution is contained in the Appendix A.

By using the same methodology as in the section 3.1., we can show that the Hessian matrix $H(\underset{\sim}{\theta})$ is negative definite when evaluated at $\hat{\underset{\sim}{\theta}}$. Thus $\hat{\underset{\sim}{\theta}}$ is a local maximum point of the likelihood function.

Since $L(\underset{\sim}{\theta};\underset{\sim}{y})$ is differentiable in the space $\Omega = R^{2n-1} \times R^+ \times R^+$, that is $L(\underset{\sim}{\theta};\underset{\sim}{y})$ is differentiable in the neighborhood of the true parameter $\underset{\sim}{\theta}_0$, then by the properties of M.L. equations, there is a root with probability 1 as $m \to \infty$, which is consistent for $\underset{\sim}{\theta}$. That is, $\hat{\underset{\sim}{\theta}}$ is a consistent estimator of $\underset{\sim}{\theta}$.

3.2.3. Large Sample Distribution of the Estimators

From (3.16) we can see that $\hat{\beta}_i$ is not an explicit function of y_{ijk}, it is very hard to derive the exact distribution of $\hat{\beta}_i$. If the number of environments is large enough, we can employ the large sample method to derive the approximate distribution of $\hat{\beta}_i$.

$\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_m$ are i.i.d. with distribution $F_{\underset{\sim}{\theta}}$ belonging to $F = \{F_{\underset{\sim}{\theta}};\ \theta \in \Omega = R^{2n-1} \times R^+ \times R^+\}$ and satisfy the conditions (R1) - (R3), as shown in section 3.1. Thus the asymptotic distribution of $\sqrt{m}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0)$ is $N_{2n+1}(\underset{\sim}{0}, i^{-1}(\underset{\sim}{\theta}_0))$, where

$$i(\underset{\sim}{\theta}_0) = \begin{bmatrix} A_{n\times n} & \\ & B_{(n+1)\times(n+1)} \end{bmatrix}$$

$$A = \frac{nr}{\sigma_0^2}\begin{bmatrix} 1 - \dfrac{nr\sigma_E^2}{K_1} & , & \dfrac{r\sigma_E^2}{K_1}\underset{\sim}{\gamma}' \\ \dfrac{r\sigma_E^2}{K_1}\underset{\sim}{\gamma} & , & \dfrac{1}{n}\{I_{n-1}+J_{n-1}+\dfrac{r\ _E^2}{K_1}\underset{\sim}{\gamma}\,\underset{\sim}{\gamma}' \end{bmatrix}$$

$$B = \frac{r\sigma_E^2}{K_1}\begin{bmatrix} \frac{r\sigma_E^2\zeta}{\sigma_0^2}\{ I_{n-1}+J_{n-1} - \frac{K_2}{\zeta K_1}\underset{\sim}{\gamma}\underset{\sim}{\gamma}'\}, & \frac{r\zeta}{K_1}\underset{\sim}{\gamma}', & \frac{1}{K_1}\underset{\sim}{\gamma}' \\ \frac{r\zeta}{K_1}\underset{\sim}{\gamma}' & , \frac{r\zeta^2}{2\sigma_E^2 K_1}, & \frac{\zeta}{2\sigma_E^2 K_1} \\ \frac{1}{K_1}\underset{\sim}{\gamma}' & , \frac{\zeta}{2\sigma_E^2 K_1}, & \frac{K_1}{2r\sigma_E^2}\{ \frac{nr-1}{(\sigma_0^2)^2} + \frac{1}{K_1^2}\} \end{bmatrix}$$

$$\zeta = \sum_{i\neq s} \beta_i^2 + (n - \sum_{i\neq s} \beta_i)^2$$

$$K_1 = \sigma_0^2 + r\sigma_E^2\zeta$$

$$K_2 = \sigma_0^2 - r\sigma_E^2\zeta$$

$$\underset{\sim}{\gamma} = \underset{\sim}{\beta}^* - (n - \sum_{i\neq s} \beta_i)\underset{\sim}{j}_{n-1}$$

Since we have $i^{-1}(\hat{\underset{\sim}{\theta}}) \xrightarrow[m\to\infty]{P} i^{-1}(\underset{\sim}{\theta}_0)$, thus

$$\sqrt{m}(\begin{bmatrix} \hat{\mu} \\ \hat{\underset{\sim}{\tau}}^* \\ \hat{\underset{\sim}{\beta}}^* \\ \hat{\sigma}_E^2 \\ \hat{\sigma}_0^2 \end{bmatrix} - \begin{bmatrix} \mu \\ \underset{\sim}{\tau}^* \\ \underset{\sim}{\beta}^* \\ \sigma_E^2 \\ \sigma_0^2 \end{bmatrix}) \overset{\mathcal{L}}{\dot{\to}} N(0, \begin{bmatrix} V_{11} & & & \\ & V_{22} & V_{23} & \\ & V_{32} & V_{33} & V_{34} \\ & & V_{43} & V_{44} \end{bmatrix}) \tag{3.19}$$

where

$$V_{11} = \begin{bmatrix} \frac{\hat{\sigma}_0^2+nr\hat{\sigma}_E^2}{nr} & , & \hat{\sigma}_E^2(\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})' \\ \hat{\sigma}_E^2(\hat{\beta}^*-j_{n-1}), & & \frac{\hat{\sigma}_0^2}{r}\{I_{n-1} - \frac{1}{n} J_{n-1} + \frac{1}{n}(\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})(\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})'\} \end{bmatrix}$$

$$V_{22} = \frac{\hat{\sigma}_0^2 \hat{K}_1}{r(\hat{\sigma}_E^2)^2 \hat{\zeta}} \{I_{n-1} - \frac{1}{n} J_{n-1} + \frac{1}{n}(\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})(\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})'\}$$

$$V_{23} = V'_{32} = -\frac{2\hat{\sigma}_0^2 \hat{K}_1}{nr^2 \hat{\sigma}_E^2 \hat{\zeta}} (\hat{\underset{\sim}{\beta}}^* - \underset{\sim}{j}_{n-1})$$

$$V_{33} = \frac{2\hat{K}_1}{nr^2 \hat{\zeta}^2} [(2\hat{\zeta} - n)\hat{\sigma}_0^2 + nr\hat{\sigma}_E^2 \hat{\zeta}] + \frac{2\hat{\sigma}_0^2}{(nr-1) r^2 \hat{\zeta}^2}$$

$$V_{34} = V_{43} = -\frac{2(\hat{\sigma}_0^2)^2}{(nr-1) r\hat{\zeta}}$$

$$V_{44} = \frac{2(\hat{\sigma}_0^2)^2}{(nr-1)}$$

$$\hat{\zeta} = \sum_{i \neq s} \hat{\beta}_i^2 + (n - \sum_{i \neq s} \hat{\beta}_i)^2$$

$$\hat{K}_1 = \hat{\sigma}_0^2 + r\hat{\sigma}_E^2 \hat{\zeta}$$

From the restriction of (3.12) we have $\beta_s = n - \sum_{i \neq s} \beta_i$. That implies $\hat{\beta} = n - \sum_{i \neq s} \hat{\beta}_i$. Then from (3.19), as $m \to \infty$, we have

$$\frac{[nmr^2 (\hat{\sigma}_E^2)^2 (\sum_i \hat{\beta}_i^2)]^{1/2}}{\{\hat{\sigma}_0^2 [\hat{\sigma}_0^2 + r\hat{\sigma}_E^2 (\sum_i \hat{\beta}_i^2)][n-1+(\hat{\beta}_s - 1)^2]\}^{1/2}} (\hat{\beta}_s - \beta_s) \xrightarrow{\mathcal{L}} N(0,1)$$

The asymptotic distribution of $\hat{\underset{\sim}{\beta}}$ as $m \to \infty$ is

$$\hat{\underset{\sim}{\beta}} \overset{\mathcal{L}}{\dot{\sim}} SN_n(\underset{\sim}{\beta}, \frac{\hat{\sigma}_0^2 [\hat{\sigma}_0^2 + r\hat{\sigma}_E^2 (\sum_i \hat{\beta}_i^2)]}{mr^2 (\hat{\sigma}_E^2)(\sum_i \hat{\beta}_i^2)} [I_n - \frac{1}{n} J_n + \frac{1}{n}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{j}_n)(\hat{\underset{\sim}{\beta}} - \underset{\sim}{j}_n)']) \qquad (3.20)$$

3.2.4. Testing Hypothesis about Equality of β's

From (3.20), the distribution of $\hat{\beta}_i - \hat{\beta}_{i'}$ is

$$\hat{\beta}_i - \hat{\beta}_{i'} \;\dot{\sim}\; N\left((\beta_i - \beta_{i'}),\; \frac{\hat{\sigma}_0^2[\hat{\sigma}_0^2 + r\hat{\sigma}_E^2(\Sigma_i \hat{\beta}_i^2)][2n + (\hat{\beta}_i - \hat{\beta}_{i'})^2]}{nmr^2(\hat{\sigma}_E^2)^2(\Sigma_i \hat{\beta}_i^2)}\right).$$

For the hypothesis

$$H_0\colon \beta_i = \beta_{i'} \quad \text{vs.} \quad H_a\colon \beta_i \neq \beta_{i'}$$

the distribution of test statistics under the null hypothesis is given as

$$S = \frac{\hat{\beta}_i - \hat{\beta}_{i'}}{\sqrt{\dfrac{2\hat{\sigma}_0^2[\hat{\sigma}_0^2 + r\hat{\sigma}_E^2(\Sigma_i \hat{\beta}_i^2)]}{mr^2(\hat{\sigma}_E^2)^2(\Sigma_i \hat{\beta}_i^2)}}} \xrightarrow[m\to\infty]{L} N(0,1)$$

As a consequence of these tests, one could partition the β_i $(i = 1, \ldots, n)$ into several groups such that in the same group of β_i's there is no significant genotype-environment interaction. After adjusting the $\hat{\beta}_i$'s in that group, we can compare the genotype effects. Then the best genotype in that group could be selected.

3.2.5. The Sample Size m

To determine how large m should be to have approximately a normal distribution for $\hat{\underset{\sim}{\theta}}$, a computer simulation evaluation was carried out. The following values were chosen: the number of genotypes was 3 or 5, the number of environments was 4, 10, 15, 20, or 30, the number of replications was 2 or 10, the value of σ_E^2 was 9, 4, or 1 and the value of σ_0^2 was 4, 1, or 0.01. For each of 300 sets of observations, the smallest O.S.L. value among all the O.S.L. value for testing the normality of each individual estimator is given in the Table III.

From Table III, we can conclude that the number of environments should be at least 30 in order for the distribution of $\hat{\underset{\sim}{\theta}}$ to be approximately multivariate normal, regardless of the number of genotypes and replications or the value of σ_E^2 and σ_0^2.

TABLE III

SIMULATION RESULT FOR TESTING THE NORMALITY OF $\hat{\underset{\sim}{\theta}}$ IN THE MIXED MODEL

m	n	r	σ_E^2	σ_0^2	O.S.L.	m	n	r	σ_E^2	σ_0^2	O.S.L.
4	3	2	9	4	<0.01	4	3	2	4	1	<0.01
10	3	2	9	4	<0.01	10	3	2	4	1	<0.01
15	3	2	9	4	0.013	15	3	2	4	1	0.013
20	3	2	9	4	0.023	20	3	2	4	1	0.083
30	3	2	9	4	>0.15	30	3	2	4	1	>0.15
4	3	10	9	4	<0.01	4	3	2	1	0.01	<0.01
10	3	10	9	4	<0.01	10	3	2	1	0.01	<0.01
15	3	10	9	4	0.014	15	3	2	1	0.01	0.011
20	3	10	9	4	0.022	20	3	2	1	0.01	0.097
30	3	10	9	4	>0.15	30	3	2	1	0.01	0.137
15	5	2	9	4	0.027	15	5	2	4	1	0.036
20	5	2	9	4	0.076	20	5	2	4	1	0.094
30	5	2	9	4	>0.15	30	5	2	4	1	>0.15

CHAPTER IV

RESTRICTED MAXIMUM LIKELIHOOD METHOD

Several researchers have evaluated the environmental index by the mean performance of all genotypes grown in that environment. This environmental index is dependent on the phenotypic performance and the ordinary least squares estimator of a stability parameter assuming normality is a ratio of two dependent Chi-square random variables for which the density has no simple form. However, if we condition on the environmental index, then the distribution of $\hat{\underset{\sim}{\beta}}$ is easy to derive.

4.1. Model

Consider model (1.1) with restriction $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, and $\sum_j EV_j = 0$ and with the assumption that the e_{ijk}'s are i.i.d. $N(0,\sigma^2)$, $\forall i,j,k$. Then $\underset{\sim}{Y} \sim N_{nmr}(\underset{\sim}{\psi}, \sigma^2 I_{nmr})$, where $\underset{\sim}{\psi} = \mu \underset{\sim}{j}_{nmr} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_{mr} + \underset{\sim}{\beta} \otimes \underset{\sim}{EV} \otimes \underset{\sim}{j}_r$.

For $B = \frac{1}{n}\underset{\sim}{j}_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}\underset{\sim}{j}_r$ and $\hat{\underset{\sim}{EV}} = B'\underset{\sim}{Y}$, then $\hat{\underset{\sim}{EV}} \sim SN_m(\underset{\sim}{EV}, \frac{\sigma^2}{nr}(I_m - \frac{1}{m}J_m))$. The joint distribution of $\underset{\sim}{Y}$ and $\hat{\underset{\sim}{EV}}$ is

$$\begin{bmatrix} \underset{\sim}{Y} \\ \hat{\underset{\sim}{EV}} \end{bmatrix} \sim SN_{nmr+m}\left(\begin{bmatrix} \underset{\sim}{\psi} \\ \hat{\underset{\sim}{EV}} \end{bmatrix}, \ \sigma^2 \begin{bmatrix} I_{nmr}, & B \\ B' & , \frac{1}{nr}(I_m - \frac{1}{m}J_m) \end{bmatrix} \right)$$

Given $\hat{\underset{\sim}{EV}} = \underset{\sim}{\xi}$, where $\underset{\sim}{\xi}'\underset{\sim}{j}_m = 0$, then

$$\underset{\sim}{Y} | \hat{\underset{\sim}{EV}} = \underset{\sim}{\xi} \sim SN_{nmr}(\underset{\sim}{\psi} + nrB(I_m - \frac{1}{m}J_m)^{-}(\underset{\sim}{\xi} - \underset{\sim}{EV}), \ \sigma^2(I_{nmr} - A)),$$

where $A = nrB(I_m - \frac{1}{m}J_m)^- B' = \frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}J_r$ (Rao, 1966).

Since $I - A$ is singular, then a representation for the probability density function of $\underset{\sim}{Y} | \hat{E\underset{\sim}{V}} = \underset{\sim}{\xi}$ (Khatri, 1968a) is

$$f_{\underset{\sim}{Y}|\hat{E\underset{\sim}{V}}}(\underset{\sim}{y}|\underset{\sim}{\xi}) = (2\pi)^{-\frac{nmr-m+1}{1}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp - \frac{1}{2\sigma^2}\{[y - \underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^- (\underset{\sim}{\xi} - E\underset{\sim}{V})]'(I - A)^- [\underset{\sim}{y} - \underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^- (\underset{\sim}{\xi} - E\underset{\sim}{V})]\} \quad (4.1)$$

Considering those $\underset{\sim}{y}$ for which $B'\underset{\sim}{y} = \underset{\sim}{\xi}$, then the value of (4.1) is

$$f_{\underset{\sim}{Y}|\hat{E\underset{\sim}{V}}}(\underset{\sim}{y}|B'\underset{\sim}{y} = \underset{\sim}{\xi}) = (2\pi)^{-\frac{nmr-m+1}{2}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp - \frac{1}{2\sigma^2} [\underset{\sim}{y} - \underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^- B'(\underset{\sim}{y} - \underset{\sim}{\psi})]'(I - A)^- [\underset{\sim}{y} - \underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^- B' (\underset{\sim}{y} - \underset{\sim}{\psi})]\}$$

$$= (2\pi)^{-\frac{nmr-m+1}{2}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp - \frac{1}{2\sigma^2}\{[\underset{\sim}{y} - \underset{\sim}{\psi}]'(I - A)(I - A)^-(I - A)[\underset{\sim}{y} - \underset{\sim}{\psi}]\}$$

$$= (2\pi)^{-\frac{nmr-m+1}{2}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp - \frac{1}{2\sigma^2} (\underset{\sim}{y} - \underset{\sim}{\psi})'(I - A)(\underset{\sim}{y} - \underset{\sim}{\psi})\}$$

which is invariant under the choice of the generalized inverses of $(I_m - \frac{1}{m}J_m)$ and $(I - A)$.

4.2. Restricted Maximum Likelihood Estimator for $\underset{\sim}{\beta}$

Since the value of (4.1) is unique for the random sample $\underset{\sim}{y}$ for which $B'\underset{\sim}{y} = \underset{\sim}{\xi}$, the logarithm of the likelihood function is

$$L(\mu,\underset{\sim}{\tau},\underset{\sim}{\beta},\sigma^2;\underset{\sim}{y}\,|\,B'\underset{\sim}{y} = \underset{\sim}{\xi}) = -\frac{nmr-m+1}{2}\ln(2\pi) - \frac{nmr-m+1}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\{\sum_{ijk}[y_{ijk} - \overline{y}_{.j.} + \overline{y}_{...} - \mu - \tau_i - (\beta_i - 1)EV_j]^2\} - \lambda_1(\sum_i \tau_i) - \lambda_2(\sum_i \beta_i - n) \qquad (4.2)$$

After differentiating (4.2) with respect to μ, τ_i, β_i, σ^2, λ_1, and λ_2, where $i = 1,\ldots,n$, and equating each partial derivative to zero we have the Maximum Likelihood equations and the roots of the equations are:

$$\hat{\mu} = \overline{y}_{...}$$

$$\hat{\tau}_i = \overline{y}_{i..} - \overline{y}_{...} \qquad \forall i = 1,\ldots,n$$

$$\hat{\beta}_i = \frac{\sum_{j=1}^{m} \overline{y}_{ij.}(\overline{y}_{.j.} - \overline{y}_{...})}{\sum_{j=1}^{m} (\overline{y}_{.j.} - \overline{y}_{...})^2} \qquad \forall i = 1,\ldots,n \qquad (4.3)$$

$$\hat{\sigma}^2 = \frac{1}{nmr-m+1}\sum_{ijk}\{y_{ijk} - \overline{y}_i - \frac{[\sum_j \overline{y}_{ij.}(\overline{y}_{.j.} - \overline{y}_{...})](\overline{y}_{.j.} - \overline{y}_{...})}{\sum_j (\overline{y}_{.j.} - \overline{y}_{...})^2}\}^2$$

Since (4.3) maximizes (4.2) for all $\mu \in R$, $\tau_i \in R$, $\beta_i \in R$ $(i = 1,\ldots,n)$, $\sigma^2 \in R^+$ when $\sum_i \tau_i = 0$ and $\sum_i \beta_i = n$, then $\hat{\mu}$, $\hat{\tau}_i$, $\hat{\beta}_i$ and $\hat{\sigma}^2$ is the Maximum Likelihood estimator of μ, τ_i, β_i $(i = 1,\ldots,n)$, and σ^2 under the condition $B'\underset{\sim}{y} = \underset{\sim}{\xi}$.

From (4.1) and (4.3) we have

$$E(\hat{\beta}_i) = E_{\hat{\underset{\sim}{EV}}}[E_{\underset{\sim}{Y}|\underset{\sim}{EV}}(\hat{\beta}_i|\hat{\underset{\sim}{EV}})] = E_{\hat{\underset{\sim}{EV}}}[1+(\beta_i-1)\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}]$$

$$= 1 + (\beta_i - 1)E_{\hat{\underset{\sim}{EV}}}[\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}]$$

$$\mathrm{Var}(\hat{\beta}_i) = E_{\hat{\underset{\sim}{EV}}}[\mathrm{Var}_{\underset{\sim}{Y}|\hat{\underset{\sim}{EV}}}(\hat{\beta}_i|\hat{\underset{\sim}{EV}})] + \mathrm{Var}_{\hat{\underset{\sim}{EV}}}[E_{\underset{\sim}{Y}|\hat{\underset{\sim}{EV}}}(\hat{\beta}_i|\hat{\underset{\sim}{EV}})]$$

$$= \frac{(n-1)}{nr}\sigma^2 E_{\hat{\underset{\sim}{EV}}}(\frac{1}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}) + (\beta_i - 1)^2 \mathrm{Var}_{\hat{\underset{\sim}{EV}}}(\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}).$$

Since it is very difficult to obtain the exact forms for the mean and variance of the ratio of two random variables, we will obtain approximate forms of $E(\hat{\beta}_i)$ and $\mathrm{Var}(\hat{\beta}_i)$. Consider the Taylor series expansion of $\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}/\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}$ expanded about $(\underset{\sim}{EV}'\underset{\sim}{EV},\ \frac{(m-1)\sigma^2}{nr} + \underset{\sim}{EV}'\underset{\sim}{EV})$, dropping all terms of order higher than 2, we have

$$E_{\hat{\underset{\sim}{EV}}}(\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}) \approx \frac{nr\Sigma_j EV_j^2}{(m-1)\sigma^2+nr\Sigma_j EV_j^2} + \frac{2\sigma^2(nr\Sigma_j EV_j^2)^2}{[(m-1)\sigma^2+nr\Sigma_j EV_j^2]^3} .$$

Similarly, we have

$$E_{\hat{\underset{\sim}{EV}}}(\frac{1}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}) \approx \frac{nr}{(m-1)\sigma^2+nr\Sigma_j EV_j^2} + \frac{2nr\sigma^2[(m-1)\sigma^2+2nr\Sigma_j EV_j^2]}{[(m-1)\sigma^2+nr\Sigma_j EV_j^2]^3}$$

and

$$\mathrm{Var}_{\hat{\underset{\sim}{EV}}}(\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}) \approx \frac{nr\sigma^2\Sigma_j EV_j^2}{[(m-1)\sigma^2+nr\Sigma_j EV_j^2]^2} + \frac{2\sigma^2(nr\Sigma_j EV_j^2)^3}{[(m-1)\sigma^2+nr\Sigma_j EV_j^2]^4} .$$

Then

$$E(\hat{\beta}_i) \approx 1 + (\beta_i - 1)\{\frac{nr\sum_j EV_j^2}{(m-1)\sigma^2 + nr\sum_j EV_j^2} + \frac{2\sigma^2(nr\sum_j EV_j^2)^2}{[(m-1)\sigma^2 + nr\sum_j EV_j^2]^3}\} \tag{4.4}$$

and

$$\begin{aligned} Var(\hat{\beta}_i) \approx (n-1)\sigma^2\{&\frac{(m-1)(m+1)(\sigma^2)^2 + 2nr(m+1)\sigma^2\sum_j EV_j^2 + (nr\sum_j EV_j^2)^2}{[(m-1)\sigma^2 + nr\sum_j EV_j^2]^3}\} \\ &+ (\beta_i - 1)^2\sigma^2\{\frac{nr\sum_j EV_j^2}{[(m-1)\sigma^2 + nr\sum_j EV_j^2]^2} + \frac{2(nr\sum_j EV_j^2)^3}{[(m-1)\sigma^2 + nr\sum_j EV_j^2]^4}\} \end{aligned} \tag{4.5}$$

4.3. Conditional Distribution of $\hat{\underset{\sim}{\beta}}$ and Testing a Hypothesis about Equality of β's, When $B'\underset{\sim}{y} = \underset{\sim}{\xi}$

From (4.3) we can see $\hat{\beta}_i$ is a ratio of two quadratic forms. It is difficult to obtain the distribution of the ratio of two dependent random variables. Since we are interested in the comparison of β_i and $\beta_{i'}$, we can compare $\hat{\beta}_i$ and $\hat{\beta}_{i'}$ given $B'\underset{\sim}{y} = \underset{\sim}{\xi}$. The conditional distribution of $\hat{\underset{\sim}{\beta}}$ given $B'\underset{\sim}{Y} = \underset{\sim}{\xi}$ is

$$\hat{\underset{\sim}{\beta}} | B'\underset{\sim}{Y} = \underset{\sim}{\xi} \sim N_n(\underset{\sim}{j}_n + \frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}(\underset{\sim}{\beta} - \underset{\sim}{j}_n),\ \frac{\sigma^2}{r\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}(I_n - \frac{1}{n}J_n)). \tag{4.6}$$

From (4.6) we have

$$\hat{\beta}_i - \hat{\beta}_{i'} | B'\underset{\sim}{Y} = \underset{\sim}{\xi} \sim N((\beta_i - \beta_{i'})\frac{\underset{\sim}{EV}'\hat{\underset{\sim}{EV}}}{\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}},\ \frac{2\sigma^2}{r\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}}).$$

Then for testing the hypothesis

H_0: $\beta_i = \beta_{i'}$ vs. H_a: $\beta_i \neq \beta_{i'}$

the distribution of the test statistic $S = (\hat{\beta}_i - \hat{\beta}_{i'})/(\frac{2\hat{\sigma}^2}{r\hat{\underset{\sim}{EV}}'\hat{\underset{\sim}{EV}}})^{1/2}$ under the null hypothesis is t with degree of freedom nmr-m+1. The implication to the plan breeder of this hypothesis being true is that there is no genotype-environment interaction for the two genotypes being compared.

4.4. Analysis of Variance

As pointed out by Freeman and Perkins, the mistake in the A.O.V. table (see Table I) is that $\hat{\underset{\sim}{EV}}$ should be considered a random vector. In this case, the distributions of the sums of squares are complicated, unless we consider the conditional distribution of $\underset{\sim}{Y}$ given $B'\underset{\sim}{y} = \underset{\sim}{\xi}$. Without loss of generality, we may assume $r = 1$ throughout this chapter. To derive the distributions of sums of squares of each source in Table I, we need apply the theorem which is taken from Rayner and Livingston (1965, Theorem 7.2) and the results are shown as follows:

$$\text{Total S.S.} = \sum_{ij} y_{ij}^2 - \frac{(\sum_{ij} y_{ij})^2}{n\sum_j \hat{EV}_j^2} \sim \sigma^2 \chi^2_{nm-m+1} \text{ with n.c. (non-centrality)}$$

$$\frac{1}{2\sigma^2}\underset{\sim}{\psi}'(I - A)\underset{\sim}{\psi}$$

$$\text{C.F.} = \frac{(\sum_{ij} y_{ij})^2}{nm} \sim \sigma^2\chi_1^2 \text{ with n.c. } \frac{nm\mu^2}{2\sigma^2}$$

$$\text{Genotypes S.S.} = \frac{1}{m}\sum_i(\sum_j y_{ij})^2 - \text{C.F.} \sim \sigma^2\chi^2_{n-1} \text{ with n.c. } \frac{m\sum_i \tau_i^2}{2\sigma^2}$$

$$\text{Genotypes} \times \text{Environments (Linear) S.S.} = \frac{n\sum_i(\sum_j y_{ij}\hat{EV}_j)^2 - (\sum_{ij} y_{ij}\hat{EV}_j)^2}{n\sum_j \hat{EV}_j^2} \sim$$

$$\sigma^2\chi^2_{n-1} \text{ with n.c. } \frac{[\sum_i(\beta_i - 1)^2][\sum_j \hat{EV}_j^2]}{2\sigma^2}$$

$$\text{Pooled Deviation S.S.} = \sum_{i=1}^{n} \{\sum_j y_{ij}^2 - \frac{1}{m}(\sum_j y_{ij})^2 - \frac{(\sum_j y_{ij}\hat{EV}_j)^2}{\sum_j \hat{EV}_j^2}\} \sim$$

$$\sigma^2 \chi^2_{(n-1)(m-2)}.$$

These sums of squares occur in the analysis of variance presented in Table IV.

Using the theorem, given by Khatri (1963) and Shanbhag (1966), we can prove independence among the sums of squares of mean, genotypes, genotypes × environment (linear), and pooled deviation. Then for testing the hypothesis:

$$H_0: \beta_1 = \beta_2 = \ldots = \beta_n \quad \text{vs.} \quad H_a: \text{not } H_0$$

we have the test statistic

$$S = \frac{\text{Genotypes} \times \text{Environments (Linear) S.S.}/(n-1)}{\text{Pooled Deviation S.S.}/(n-1)(m-2)}$$

and

$$S \sim F_{((n-1),(n-1)(m-2))} \quad \text{under } H_0.$$

TABLE IV

ANALYSIS OF VARIANCE WHEN ENVIRONMENTAL INDEX IS GIVEN AS MEAN PERFORMANCE OF ALL GENOTYPES GROWN IN THAT ENVIRONMENT

Source	D.F.	S.S.	E.M.S.
Total	nm-m+1	$\sum_{ij} y_{ij}^2 - \frac{(\sum_{ij} y_{ij})^2}{n\sum_j \hat{EV}_j^2}$	
Mean	1	$\frac{1}{nm}(\sum_{ij} y_{ij})^2$	$\sigma^2+nm\mu^2$
Genotype	n-1	$\frac{1}{m}\sum_i(\sum_j y_{ij})^2 - \frac{1}{nm}(\sum_{ij} y_{ij})^2$	$\sigma^2+\frac{m}{n-1}\sum_i \tau_i^2$
G × EV (Linear)	n-1	$\frac{n\sum_i(\sum_j y_{ij}\hat{EV}_j)^2 - (\sum_{ij} y_{ij}\hat{EV}_j)^2}{n\sum_j \hat{EV}_j^2}$	$\sigma^2+[\sum_i(\beta_i-1)^2][\sum_j \hat{EV}_j^2]$
Pooled Deviation	(n-1)(m-2)	$\sum_i[\sum_j y_{ij}^2 - \frac{1}{m}(\sum_j y_{ij})^2 - \frac{(\sum_j y_{ij}\hat{EV}_j)^2}{\sum_j \hat{EV}_j^2}]$	σ^2

CHAPTER V

THE ESTIMATION OF $\underset{\sim}{\beta}$ WHILE THE DESIGN IS GIVEN AS A RANDOMIZED COMPLETE BLOCK DESIGN WITHIN EACH ENVIRONMENT

In this chapter, we will consider the estimation of the stability parameters when the design is a randomized block design within each environment. Within each of the m environments, there are r blocks and n genotypes. The mathematical model for ijkth observation is represented as following:

$$Y_{ijk} = \mu + \tau_i + \beta_i EV_j + \gamma_{k(j)} + (\tau\gamma)_{ik(j)} + \varepsilon_{ijk} \tag{5.1}$$

$$i = 1,\ldots,n;\; j = 1,\ldots,m;\; k = 1,\ldots,r;$$

where

Y_{ijk} is the phenotypic performance of the ith genotype at the jth environment on the kth block;

μ is the overall mean of the population;

τ_i is the effect of the ith genotype;

β_i measures the response of the ith genotype to varying environments;

EV_j is the effect of the jth environment;

$\gamma_{k(j)}$ is the effect of the kth block within the jth environment;

$(\tau\gamma)_{ik(j)}$ is the effect of the ith genotype by the kth block interaction within the j^{th} environment;

ε_{ijk} is the random error associated with the i^{th} genotype at j^{th} environment on the k^{th} block;

and furthermore we assume $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, and $\sum_j EV_j = 0$.

In this chapter, we will obtain the estimator of $\underset{\sim}{\beta}$ for each of the following assumptions: (1) The blocks effects are random; and (2) The block effects are fixed and the genotype by block interaction effects are zero.

5.1. Estimation of $\underset{\sim}{\beta}$ for Random Block Effects

In this section, we will consider the effects of blocks to be random. That is, $\gamma_{k(j)}$'s are i.i.d. $N(0,\sigma_b^2)$ and independent of the $[(\tau\gamma)_{ik(j)} + \varepsilon_{ijk}]$'s, which are i.i.d. $N(0,\sigma_e^2)$. Then from (5.1) we have $Y_{ijk} \sim N(\mu+\tau_i + \beta_i EV_j, \sigma_b^2+\sigma_e^2)$ for all i,j, and k, and for $i \neq i'$ $Cov(Y_{ijk}, Y_{i'jk}) = \sigma_b^2$. From the above assumption, we have $\underset{\sim}{Y}_k \sim N_{nm}(\mu \underset{\sim}{j}_{nm} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_m + \underset{\sim}{\beta} \otimes \underset{\sim}{EV}, \sigma_b^2 (J_n \otimes I_m) + \sigma_e^2 I_{nm})$ where $\underset{\sim}{Y}_k = (Y_{11k}, Y_{12k}, \ldots, Y_{1mk}, \ldots, Y_{nmk})'$ and for $k \neq k'$ $Cov(\underset{\sim}{Y}_k, \underset{\sim}{Y}_{k'}) = 0$.

The probability density function of $\underset{\sim}{Y}_k$ is

$$f_{\underset{\sim}{Y}_k}(\underset{\sim}{y}_k; \mu, \underset{\sim}{\tau}', \underset{\sim}{\beta}', \underset{\sim}{EV}', \sigma_b^2, \sigma_e^2) = (2\pi)^{-\frac{nm}{2}} (\sigma_e^2)^{-\frac{(n-1)m}{2}} (n\sigma_b^2+\sigma_e^2)^{-\frac{m}{2}}$$

$$\exp -\frac{1}{2\sigma_e^2}\{ \sum_{ij} [y_{ijk}-\mu-\tau_i-\beta_i EV_j]^2 - \frac{\sigma_b^2}{n\sigma_b^2+\sigma_e^2} \sum_j [\sum_i (y_{ijk}-\mu-EV_j)]^2\}$$

Since $\underset{\sim}{j}_n' = 0$, $\underset{\sim}{j}_n'\underset{\sim}{\beta} = n$, and $\underset{\sim}{j}_m'\underset{\sim}{EV} = 0$, then for the random sample $\underset{\sim}{y} = (\underset{\sim}{y}_1', \underset{\sim}{y}_2', \ldots, \underset{\sim}{y}_r')$, the logarithm of the likelihood function of $(\mu, \underset{\sim}{\tau}', \underset{\sim}{\beta}', \underset{\sim}{EV}', \sigma_b^2, \sigma_e^2)$ is

$$L(\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV},\sigma_b^2,\sigma_e^2;\underset{\sim}{y})$$

$$= \ln \prod_{k=1}^{r} \{f_{\underset{\sim}{Y}_k}(\underset{\sim}{y}_k)\} - \lambda_1(\Sigma\tau_i) - \lambda_2(\sum_i \beta_i - n) - \lambda_3(\sum_j EV_j). \qquad (5.2)$$

Differentiating (5.2) with respect to μ, τ_i, β_i, EV_j, σ_b^2, σ_e^2, and λ_ℓ for $i=1,\ldots,n$; $j=1,\ldots,m$; $\ell=1,2,3$, and equating each of the 2n+m+6 partial derivatives to zero, we have the Maximum Likelihood equations and the roots of the equations are:

$$\hat{\mu} = \bar{y}_{\cdots}$$

$$\hat{\tau}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}$$

$$\hat{\beta}_i = 1 + \frac{\sum_j(\bar{y}_{ij\cdot} - \bar{y}_{\cdot j\cdot})\hat{EV}_j}{\sum_j \hat{EV}_j^2}$$

$$EV_j = \frac{(n\hat{\sigma}_b^2+\hat{\sigma}_e^2)\sum_i(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})\hat{\beta}_i - n\hat{\sigma}_b^2(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})}{(n\hat{\sigma}_b^2+\hat{\sigma}_e^2)(\sum_i\hat{\beta}_i^2) - n^2\hat{\sigma}_b^2}$$

$$\hat{\sigma}_b^2 = \frac{1}{(n-1)nmr}\sum_{jk}\{[\sum_i(y_{ijk} - \bar{y}_{\cdots} - \hat{EV}_j)]^2 - \sum_i(y_{ijk} - \bar{y}_{i\cdot\cdot} - \beta_i\hat{EV}_j)^2\}$$

$$\hat{\sigma}_e^2 = \frac{1}{(n-1)nmr}\sum_{jk}\{n\sum_i(y_{ijk} - \bar{y}_{i\cdot\cdot} - \hat{\beta}_i\hat{EV}_j)^2 - [\sum_i(y_{ijk} - \bar{y}_{\cdots} - \hat{EV}_j)]^2\}$$

Using the same methodology as shown in Chapter III, it can be shown the above roots maximize the likelihood function (5.2). That is, $(\hat{\mu},\hat{\underset{\sim}{\tau}}',\hat{\underset{\sim}{\beta}}',\hat{\underset{\sim}{EV}}',\hat{\sigma}_b^2,\hat{\sigma}_e^2)$ is the M.L.E. of $(\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV}',\sigma_b^2,\sigma_e^2)$. Also it can be shown that for large r $\hat{\underset{\sim}{\beta}}$ is an unbiased and consistent estimator of $\underset{\sim}{\beta}$.

Now, let us consider the Restricted Maximum Likelihood method of estimation as used in Chapter IV. The joint distribution of $\underset{\sim}{Y}$ and $\hat{\underset{\sim}{EV}}$ is

$$\begin{vmatrix} \underset{\sim}{Y} \\ \\ \hat{\underset{\sim}{EV}} \end{vmatrix} \sim SN_{nmr+m}\left(\begin{vmatrix} \underset{\sim}{\psi} \\ \\ \underset{\sim}{EV} \end{vmatrix}, \begin{vmatrix} \sigma_b^2(J_n \otimes I_{mr}) + \sigma_e^2 I_{nmr}, & (n\sigma_b^2+\sigma_e^2)B \\ \\ (n\sigma_b^2+\sigma_e^2)B' & , \frac{n\sigma_b^2+\sigma_e^2}{nr}(I_m - \frac{1}{m}J_m) \end{vmatrix} \right)$$

where $\underset{\sim}{\psi} = \mu \underset{\sim}{j}_{nmr} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_{mr} + \beta \otimes \underset{\sim}{EV} \otimes \underset{\sim}{j}_r$ and $B = \frac{1}{n}\underset{\sim}{j}_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}\underset{\sim}{j}_r$.

By Rao (1966), given $\hat{\underset{\sim}{EV}} = \underset{\sim}{\xi}$, we have $\underset{\sim}{Y}|\hat{\underset{\sim}{EV}} = \underset{\sim}{\xi}$ distributed as singular normal with mean $(I-A)\underset{\sim}{\psi} + nrB(I_m - \frac{1}{m}J_m)^{-}\underset{\sim}{\xi}$ and variance $(n\sigma_b^2 + \sigma_e^2)(I-A) - n\sigma_b^2 H$, where $A = \frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}J_r$ and $H = (I_n - \frac{1}{n}J_n) \otimes I_{mr}$. By Khatri (1968a), a representation for the probability density function of $\underset{\sim}{Y}|\hat{\underset{\sim}{EV}} = \underset{\sim}{\xi}$ is given as

$$f_{\underset{\sim}{Y}|\hat{\underset{\sim}{EV}}}(\underset{\sim}{y}|\underset{\sim}{\xi}) = (2\pi)^{-\frac{P}{2}} \left(\prod_{\ell=1}^{P} \lambda_\ell\right)^{-\frac{1}{2}} \exp - \frac{1}{2(n\sigma_b^2+\sigma_e^2)} [\underset{\sim}{y} - (I-A)\underset{\sim}{\psi} - nrB(I_n - \frac{1}{m}J_m)^{-}\underset{\sim}{\xi}]'(I - A - \frac{n\sigma_b^2}{n\sigma_0^2+\sigma_e^2}H)^{-} [\underset{\sim}{y} - (I-A)\underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^{-}\underset{\sim}{\xi}]\} \tag{5.4}$$

where P is the rank of $(I - A - \frac{n\sigma_b^2}{n\sigma_b^2+\sigma_e^2}H)$ and λ_ℓ $(\ell = 1,\ldots,P)$ is the non-zero eigenvalue of $[(n\sigma_b^2+\sigma_e^2)(I-A) - n\sigma_b^2 H]$.

For $\underset{\sim}{y}$ of which $B'\underset{\sim}{y} = \underset{\sim}{\xi}$, we have $[\underset{\sim}{y} - (I-A)\underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)^{-}\underset{\sim}{\xi}] = (I-A)(\underset{\sim}{y} - \underset{\sim}{\psi})$. One form of the generalized inverse of $[I - A - n\sigma_b^2 H/(n\sigma_b^2 + \sigma_e^2]$ can be expressed as $(I - cA + n\sigma_b^2 H/\sigma_e^2)$ for all $c \in R$. Then for any $d \in R$, all other generalized inverses can be represented as

$$(I - cA + \frac{n\sigma_b^2}{\sigma_e^2}H) - d[I - (I - cA + \frac{n\sigma_b^2}{\sigma_e^2}H)(I - A - \frac{n\sigma_b^2}{n\sigma_b^2+\sigma_e^2}H]$$

$$= I - (c-d)A + \frac{n\sigma_b^2}{\sigma_e^2} H .$$

This is the same form as $[I - cA + n\sigma_b^2 H/\sigma_e^2]$ which implies $[I - cA + n\sigma_b^2 H/\sigma_e^2]$ is the unique form for the generalized inverse of $[I - A - n\sigma_b^2 H/(n\sigma_b^2+\sigma_e^2)]$. Then the value of (5.4) for the random sample $\underset{\sim}{y}$ such that $B'\underset{\sim}{y} = \xi$ is

$$f_{\underset{\sim}{Y}|\hat{E}\underset{\sim}{V}}(\underset{\sim}{y}|\underset{\sim}{\xi}) = (2\pi)^{-\frac{nmr-m+1}{2}} (\sigma_e^2)^{-\frac{nmr-mr}{2}} (n\sigma_b^2+\sigma_e^2)^{-\frac{mr-m+1}{2}}$$

$$\exp - \frac{1}{2(n\sigma_0^2+\sigma_e^2)} \{ (\underset{\sim}{y} - \underset{\sim}{\psi})' (I - A + \frac{n\sigma_b^2}{\sigma_e^2} H)(\underset{\sim}{y} - \underset{\sim}{\psi}) \}$$

which is invariant of the choice of the generalized inverses of $(I_m - \frac{1}{m}J_m)$ and $[I_{nmr} - A - n\sigma_b^2 H/(n\sigma_b^2+\sigma_e^2)]$.

For the random sample $\underset{\sim}{y}$ such that $B'\underset{\sim}{y} = \underset{\sim}{\xi}$, the logarithm of the likelihood function of $(\mu, \underset{\sim}{\tau}', \underset{\sim}{\beta}', \sigma_b^2, \sigma_e^2)$ is

$$L(\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\sigma_b^2,\sigma_e^2;\underset{\sim}{y}|B'\underset{\sim}{y} = \underset{\sim}{\xi}) = -\frac{nmr-m+1}{2}\ln (2\pi) - \frac{nmr-mr}{2}\ln (\sigma_e^2)$$

$$- \frac{mr-m+1}{2}\ln (n\sigma_b^2+\sigma_e^2) - \frac{1}{2(n\sigma_b^2+\sigma_e^2)} \sum_{ijk} [y_{ijk} - \mu - \tau_i - \beta_i(\bar{y}_{.j.} - \bar{y}_{...})]^2$$

$$+ \frac{n\sigma_b^2}{\sigma_e^2}[y_{ijk} - \bar{y}_{.jk} - \mu - \tau_i - (\beta_i - 1)(\bar{y}_{.j.} - \bar{y}_{...})]^2\} - \lambda_1(\sum_i \tau_i)$$

$$- \lambda_2(\sum_i \beta_i - n) \tag{5.5}$$

Differentiating (5.5) with respect to μ, τ_i, β_i, σ_b^2, σ_e^2, and λ_ℓ for for $i=1,\ldots,n$, $\ell=1,2$, and equating each of the 2n+5 equations to zero, we have the Maximum Likelihood equations and their roots maximize (5.5). The solution for $\hat{\beta}_i$ which is the Restricted Maximum Likelihood estimator of β_i is given as

$$\hat{\beta}_i = \frac{\sum_j \bar{y}_{ij.}(\bar{y}_{.j.} - \bar{y}_{...})}{\sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2} \qquad \forall i = 1,..,n. \tag{5.6}$$

5.2. Estimation of $\underset{\sim}{\beta}$ for Fixed Block Effects

In this section, the effects of blocks are assumed to be fixed and $\sum_k \gamma_{k(j)} = 0$ for $j = 1,...,m$ and $(\tau\gamma)_{ik(j)} = 0$ for all i, j, and k. Also we assume ε_{ijk}'s are i.i.d. $N(0,\sigma^2)$ for all i, j, and k. Then from (5.1) we have $Y_{ijk} \sim N(\mu + \tau_i + \beta_i EV_j + \gamma_{k(j)}, \sigma^2)$ for all i, j, k. The probability density function of $\underset{\sim}{Y}$ is

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y};\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV}',\underset{\sim}{\gamma}',\sigma^2) = (2\pi)^{-\frac{nmr}{2}} (\sigma^2)^{-\frac{nmr}{2}} \exp - \frac{1}{2\sigma^2} \sum_{ijk} [y_{ijk} - \mu - \tau_i - \beta_i EV_j - \gamma_{k(j)}]^2\}$$

where $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, $\sum_j EV_j = 0$, and $\sum_k \gamma_{k(j)} = 0$ for $j = 1,...,m$.

For the random sample $\underset{\sim}{y}$, the logarithm of the likelihood function of $(\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV}',\underset{\sim}{\gamma}',\sigma^2)$ is

$$L(\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV}',\underset{\sim}{\gamma}',\sigma^2;\underset{\sim}{y}) = \ln f_Y(y;\mu,\underset{\sim}{\tau}',\underset{\sim}{\beta}',\underset{\sim}{EV}',\underset{\sim}{\gamma}',\sigma^2) - \lambda_1(\sum_i \tau_i)$$
$$- \lambda_2(\sum_i \beta_i - n) - \lambda_3(\sum_j EV_j) - \sum_j [\lambda_{j+3}(\sum_k \sigma_{k(j)})]. \tag{5.7}$$

Differentiating (5.7) with respect to μ, τ_i, β_i, EV_j, $\gamma_{k(j)}$, σ^2, and λ_ℓ for $i = 1,...,n$; $j = 1,...,m$; $k = 1,...,r$; $\ell = 1,...,m+3$ and equating each of the 2n+2m+mr+5 partial derivatives to zero, we have the Maximum Likelihood equations whose roots are:

$$\hat{\mu} = \bar{y}_{...}$$

$$\hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...} \qquad \forall i = 1,\ldots,n$$

$$\hat{\beta}_i = 1 + \frac{\sum_j (\bar{y}_{ij.} - \bar{y}_{.j.})\hat{EV}_j}{\sum_j \hat{EV}_j^2} \qquad \forall i = 1,\ldots,n \tag{5.8}$$

$$EV_j = \frac{\sum_i (\bar{y}_{ij.} - \bar{y}_{i..})\hat{\beta}_i}{\sum_i \hat{\beta}_i^2} \qquad \forall j = 1,\ldots,m$$

$$\hat{\gamma}_{k(j)} = \bar{y}_{.jk} - \bar{y}_{.j.} \qquad \forall j = 1,\ldots,m;\ k = 1,\ldots,r$$

$$\hat{\sigma}^2 = \frac{1}{nmr} \sum_{ijk} [y_{ijk} - \bar{y}_{i..} - \bar{y}_{.jk} + \bar{y}_{.j.} - \hat{\beta}_i \hat{EV}_j]^2.$$

To obtain the Restricted Maximum Likelihood estimate as before, we begin with the joint distribution of $\underset{\sim}{Y}$ and $\underset{\sim}{\hat{EV}}$:

$$\begin{bmatrix} \underset{\sim}{Y} \\ \underset{\sim}{EV} \end{bmatrix} \sim SN_{nmr+m}\left(\begin{bmatrix} \underset{\sim}{\psi} \\ EV \end{bmatrix}, \sigma^2 \begin{bmatrix} I_{nmr}, & B \\ B', & \frac{1}{nr}(I_m - \frac{1}{m}J_m) \end{bmatrix} \right)$$

where $\underset{\sim}{\psi} = \mu \underset{\sim}{j}_{nmr} + \underset{\sim}{\tau} \otimes \underset{\sim}{j}_{mr} + \underset{\sim}{\beta} \otimes \underset{\sim}{EV} \otimes \underset{\sim}{j}_r + \underset{\sim}{j}_n \otimes \underset{\sim}{\gamma}$ and $B = \frac{1}{n}\underset{\sim}{j}_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}\underset{\sim}{j}_r$.

By Rao (1966), given $\underset{\sim}{\hat{EV}} = \underset{\sim}{\xi}$, we have $\underset{\sim}{Y} | \underset{\sim}{\hat{EV}} = \underset{\sim}{\xi}$ distributed as singular normal with mean $(I - A)\underset{\sim}{\psi} + nrB(I_m - \frac{1}{m}J_m)\underset{\sim}{\xi}$ and variance $\sigma^2(I - A)$, where $A = \frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m) \otimes \frac{1}{r}J_r$. By Khatri (1968a), a representation for the probability density function of $\underset{\sim}{Y} | \underset{\sim}{\hat{EV}} = \underset{\sim}{\xi}$ is given as

$$f_{\underset{\sim}{Y}|\underset{\sim}{\hat{EV}}}(\underset{\sim}{y}|\underset{\sim}{\xi}) = (2\pi)^{-\frac{nmr-m+1}{2}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp -\frac{1}{2\sigma^2} [\underset{\sim}{y} - (I - A)\underset{\sim}{\psi} - nrB$$

$$(I_m - \frac{1}{m}J_m)\underset{\sim}{\xi}]'(I-A)^{-}[\underset{\sim}{y} - (I-A)\underset{\sim}{\psi} - nrB(I_m - \frac{1}{m}J_m)\underset{\sim}{\xi}]\} \quad (5.9)$$

For the random sample $\underset{\sim}{y}$ such that $B'\underset{\sim}{y} = \underset{\sim}{\xi}$, the value of (5.9) is

$$(2\pi)^{-\frac{nmr-m+1}{2}} (\sigma^2)^{-\frac{nmr-m+1}{2}} \exp - \frac{1}{2\sigma^2}\{(\underset{\sim}{y} - \underset{\sim}{\psi})'(I-A)(\underset{\sim}{y} - \underset{\sim}{\psi})\},$$

which is invariant of the choice of $(I_m - \frac{1}{m}J_n)^{-}$ and $(I_{nmr} - A)^{-}$. The logarithm of the likelihood function of $(\mu, \underset{\sim}{\tau}', \underset{\sim}{\beta}', \underset{\sim}{\gamma}', \sigma^2)$ while given $B'\underset{\sim}{y} = \underset{\sim}{\xi}$ is

$$L(\mu, \underset{\sim}{\tau}', \underset{\sim}{\beta}', \underset{\sim}{\gamma}', \sigma^2; \underset{\sim}{y} | B'\underset{\sim}{y} = \underset{\sim}{\xi}) = -\frac{nmr-m+1}{2}\ln(2\pi) - \frac{nmr-m+1}{2}\ln(\sigma^2)$$

$$- \frac{1}{2\sigma^2}\sum_{ijk}\{y_{ijk} - \mu - \tau_i - \beta_i(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots}) - \gamma_{k(j)}\}^2 - \lambda_1(\sum_i \tau_i) - \lambda_2(\sum_i \beta_i - n)$$

$$- \sum_j[\lambda_{j+2}(\sum_k \gamma_{k(j)})] \quad (5.10)$$

Differentiating (5.10) with respect to μ, τ_i, β_i, $\sigma_{k(j)}$, σ^2, and λ_ℓ for $i = 1,\ldots,n$; $j = 1,\ldots,m$; $k = 1,\ldots,r$ and $\ell = 1,\ldots,m+2$, and equating each of the 2n+m+mr+4 partial derivatives to zero, we have the Maximum Likelihood equations and the roots of the equations are:

$$\hat{\mu} = \bar{y}_{\cdots}$$

$$\hat{\tau}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots} \qquad \forall i = 1,\ldots,n$$

$$\hat{\beta}_i = \frac{\sum_j y_{ij\cdot}(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})}{\sum_j(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2} \qquad \forall i = 1,\ldots,n \quad (5.11)$$

$$\hat{\gamma}_{k(j)} = \bar{y}_{\cdot jk} - \bar{y}_{\cdot j\cdot} \qquad \forall j = 1,\ldots,m;\ k = 1,\ldots,r$$

$$\hat{\sigma}^2 = \frac{1}{nmr-m+1} \sum_{ijk} \{y_{ijk}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot jk}+\bar{y}_{\cdot j\cdot}-(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdot\cdot\cdot})\frac{\sum_j \bar{y}_{ij\cdot}(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdot\cdot\cdot})}{\sum_j(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdot\cdot\cdot})^2}\}^2.$$

Since these roots maximize (5.10), then the solution for $\hat{\beta}_i$ is the Restricted Maximum Likelihood estimator of β_i.

CHAPTER VI

THE ESTIMATION OF $\underset{\sim}{\beta}$ WHILE THE NORMALITY OF THE OBSERVATIONS IS NOT ASSUMED

In the previous chapters, we obtained the estimator for the stability parameter by assuming that the random error term in model (1.1) constitute a sample from a normal population. In this chapter, we will estimate the stability parameter without normality. We will employ two estimation methods, namely, a ratio method and a generalized least squared method. The model is given as

$$Y_{ijk} = \mu + \tau_i + \beta_i EV_j + \varepsilon_{ijk} \tag{6.1}$$

$$i = 1,\ldots,n;\ j = 1,\ldots,m;\ k = 1,\ldots,r$$

with restrictions $\sum_i \tau_i = 0$, $\sum_i \beta_i = n$, and $\sum_j EV_j = 0$. And assume that ε_{ijk}'s are uncorrelated with mean zero and variance σ^2.

6.1. Ratio Estimator for $\underset{\sim}{\beta}$

By the assumption of (6.1), the expected value of Y_{ijk} is $\mu + \tau_i + \beta_i EV_j$; then $E(\overline{Y}_{ij.} - \overline{Y}_{i..}) = \beta_i EV_j$ and $E(\overline{Y}_{.j.} - \overline{Y}_{...}) = EV_j$. For any j such that $\overline{y}_{.j.} - \overline{y}_{...} \neq 0$, the suggested ratio estimator of β_i is defined as

$$\hat{\beta}_i = \frac{\overline{y}_{ij.} - \overline{y}_{i..}}{\overline{y}_{.j.} - \overline{y}_{...}} \qquad \forall i = 1,\ldots,n \tag{6.2}$$

where $\overline{y}_{.j.} - \overline{y}_{...} \neq 0$.

Assuming $EV_j \neq 0$ $\forall j$ and using the Taylor series expansion, we have the approximate forms for $E(\hat{\beta}_i)$ and $Var(\hat{\beta}_i)$ as

$$E(\hat{\beta}_i) \approx \beta_i + (\beta_i - 1)\frac{(m-1)\sigma^2}{nmrEV_j^2} \qquad \forall i,j \tag{6.3}$$

and

$$Var(\hat{\beta}_i) \approx [(\beta_i - 1)^2 + (n-1)]\frac{(m-1)\sigma^2}{nmrEV_j^2} \qquad \forall i,j. \tag{6.4}$$

The mean square errors of $\hat{\underset{\sim}{\beta}}$ for j = 1,...,m have the approximate form as:

$$MSE(\hat{\underset{\sim}{\beta}}) = \{\sum_i(\beta_i-1)^2[1+\frac{(m-1)\sigma^2}{nmrEV_j^2}] + n(n-1)\}\frac{(m-1)\sigma^2}{nmrEV_j^2} . \tag{6.5}$$

From (6.5), we can see that if we choose a j* such that $EV_{j*}^2 = \underset{j}{Max}\{EV_j^2\}$, then j* will minimize $MSE(\hat{\underset{\sim}{\beta}})$ for all j. From this we can give the ratio estimator as

$$\hat{\beta}_i = \frac{\overline{y}_{ij*.} - \overline{y}_{i..}}{\overline{y}_{.j*.} - \overline{y}_{...}} \tag{6.6}$$

where $\overline{y}_{.j*.} - \overline{y}_{...} = \underset{j}{Max}\{|\overline{y}_{.j.} - \overline{y}_{...}|\}$.

6.2. Least Squares Estimator for $\underset{\sim}{\beta}$

Without loss of generality, we may assume r = 1 throughout this chapter. From the model (6.1) we can see that it is impossible to regressor Y_{ij} on EV_j, since EV_j's are unobservable. But, since $E(B'\underset{\sim}{Y}) = \underset{\sim}{EV}$ where $B = \frac{1}{n}\underset{\sim}{j}_n \otimes (I_m - \frac{1}{m}J_m)$, we can measure $\underset{\sim}{EV}$ by $\hat{\underset{\sim}{EV}} = B'\underset{\sim}{Y}$. That is

$$\hat{\underset{\sim}{E}V} = \underset{\sim}{E}V + \underset{\sim}{e}$$

where $\underset{\sim}{e} \sim (0, \frac{\sigma^2}{n}(I_m - \frac{1}{m}J_m))$.

Now, we can consider (6.1) as the errors-in-variables model. Let's define the model as

$$\underset{\sim}{Y} = X\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}^* \tag{6.7}$$

where $X_{nm\times 2n} = [I_n \otimes \underset{\sim}{j}_m \mid I_n \otimes \hat{\underset{\sim}{E}V}]$, $\underset{\sim}{\alpha}_{2n-1} = [\underset{\sim}{\tau}^{*\prime} \mid \underset{\sim}{\beta}']'$, $\underset{\sim}{\tau}^* = \mu\underset{\sim}{j}_n + \underset{\sim}{\tau}$, and $\underset{\sim}{\varepsilon}^* = \underset{\sim}{\varepsilon} - \underset{\sim}{\beta} \otimes \underset{\sim}{e}$.

Since $\text{cov}(e_j, \varepsilon_{ij'k}) = \frac{m-1}{nm}\sigma^2 \; \forall j = j'$, $= -\frac{1}{nm}\sigma^2 \; \forall j \neq j'$, $\text{Var}(\underset{\sim}{e}) = \frac{\sigma^2}{n}(I_m - \frac{1}{m}J_m)$, and $\text{Var}(\underset{\sim}{\varepsilon}) = \sigma^2 I_{nm}$, then the variance-covariance matrix of $\underset{\sim}{\varepsilon}^*$ is $\sigma^2 V$, where $V = I_{nm} + \frac{1}{n}[(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)' - J_n] \otimes (I_m - \frac{1}{m}J_m)$.

Since V is a singular matrix, the generalized least squares estimator for $\underset{\sim}{\alpha}$ is given as

$$\hat{\underset{\sim}{\alpha}} = (X'V^-X)^{-1}X'V^-\underset{\sim}{y}$$

where V^- is a generalized inverse of V.

The general form of the generalized inverse of V can be expressed as $V^- = I_{nm} + C_1[\frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m)] + [\frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m)C_2 - [(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)' \otimes (I_m - \frac{1}{m}J_m]/(\sum_i \beta_i^2)$ for arbitrary nm×nm matrices C_1 and C_2. We have different estimate values of $\underset{\sim}{\alpha}$ by selecting different C_1 and C_2. To find a "best" estimator, we need to minimize the mean squared errors of $\hat{\underset{\sim}{\alpha}}$ with respect to C_1 and C_2. Since,

$$\begin{aligned}\min_{C_1,C_2} \text{MSE}(\hat{\underset{\sim}{\alpha}}) &= \min_{C_1,C_2} \sigma^2 \text{tr}\{(X'V^-X)^{-1}X'V^-VV^-X(X'V^-X)^{-1}\} \\ &= \sigma^2 \text{tr}\{[X'(I_{nm} - \frac{1}{\sum_i \beta_i^2}(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)' \otimes (I_m - \frac{1}{m}J_m))X]^{-1}\}.\end{aligned}$$

That is, for selecting the "best" generalized least squares estimator of $\underset{\sim}{\alpha}$, C_1 and C_2 should be chosen as zero matrices in the general form of V^-. Then, we have

$$\hat{\alpha} = \begin{bmatrix} I_n \otimes \frac{1}{m}j_m' \\ \\ I_n \otimes \frac{1}{\sum_j \hat{EV}_j^2} \hat{EV}' \end{bmatrix} y \,. \tag{6.8}$$

From (6.8) we can see that the generalized least squares estimator for α is exactly the same as the ordinary least squares estimator for α. And we have the same result in the G.L.S.E. for β and R.M.L.E. for β when given $\hat{EV} = B'y$.

To derive the approximate distribution of $\hat{\beta}$, rewrite the model (6.7) as

$$Y_m = X_m \alpha + \varepsilon_m \qquad \forall m = 2,3,\ldots$$

where $Y_m = (Y_{11},\ldots,Y_{1m},\ldots,Y_{n1},\ldots,Y_{nm})'$, $X_m = [I_n \otimes j_m \mid I_n \otimes \hat{EV}]$, $\alpha = [\mu j_n + \tau' \mid \beta']$, $\varepsilon_m \sim (0, \sigma^2 V_m)$ and $V_m = I_{nm} - \frac{1}{n}J_n \otimes (I_m - \frac{1}{m}J_m) + \frac{1}{n}[(\beta - j_n)(\beta - j_n)'] \otimes (I_m - \frac{1}{m}J_m)$.

Since V_m is symmetric and the rank of V_m is $(n-1)(m-1)$, there exists a $(n-1)(m-1)\times nm$ matrix A_m with rank $(n-1)(m-1)$ such that $A_m V_m A_m' = I_{(n-1)(m-1)}$. Then

$$A_m Y_m = A_m X_m \alpha + A_m \varepsilon_m$$

or

$$Y_m^* = X_m^* \alpha + \varepsilon_m^*$$

where

$$Y_m^* = A_m Y_m,\ X_m^* = A_m X_m,\ \text{and}\ \varepsilon_m^* = A_m \varepsilon_m \sim (0, \sigma^2 I_{(n-1)(m-1)}).$$

The least squares estimator of α is $\hat{\alpha}_m = (X_m^{*\prime} X_m^*)^{-1} X_m^{*} Y_m^*$. Since

$X^*_m(X^{*\prime}_m X^*_m)^{-1}X^{*\prime}_m = A_m X_m (X'_m V^-_m X_m)^{-1} X'_m A_m$ and

$$X_m(X'_m V^-_m X_m)^{-1}X'_m = X_m \begin{bmatrix} \frac{1}{m}I_n, & 0 \\ 0\ , & \frac{1}{\sum_j \hat{EV}_j^2}[I_n + \frac{1}{n}(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)'] \end{bmatrix} X'_m$$

$$= I_n \otimes \frac{1}{m}J_m + [I_n + \frac{1}{n}(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)'] \otimes \frac{1}{\sum_j \hat{EV}_j^2} \hat{E\underset{\sim}{V}}\, \hat{E\underset{\sim}{V}}'.$$

Then each diagonal element of $X^*_m(X^{*\prime}_m X^*_m)^{-1}X^{*\prime}_m$ has the form $\frac{1}{m}d_1 + \frac{\hat{EV}_j^2}{\sum_j \hat{EV}_j^2}d_2$ $\forall j = 1,\ldots,m$ where d_1 and d_2 are functions of n and $\underset{\sim}{\beta}$ and depend on the choice of A_m, but both are independent of m and $\hat{EV}_j$ $\forall j = 1,\ldots,m$.

Since $E\{\frac{1}{m}\sum_j \hat{EV}_j^2\} = \frac{1}{n}\sigma^2$, then as $m \to \infty$, $\frac{1}{m}\sum_j \hat{EV}_j^2 \xrightarrow{a.s.} \frac{1}{n}\sigma^2$. That implies $\sum_j \hat{EV}_j^2 \to \infty$ as $m \to \infty$, such that for $j = 1,\ldots,m$, $\frac{1}{m}d_1 + \frac{\hat{EV}_j^2}{\sum_j \hat{EV}_j^2}d_2 \longrightarrow 0$ as $m \to \infty$. Then we have the maximum diagonal element of $X^*_m(X^{*\prime}_m X^*_m)^{-1}X^{*\prime}_m$ tending 0 as $m \to \infty$, which satisfies Huber's condition (Arnold, 1980). Thus when $m \to \infty$ we have

$$\hat{\underset{\sim}{\beta}}_m \overset{\mathcal{L}}{\sim} N_n(\underset{\sim}{\beta}, \frac{\sigma^2}{\sum_j \hat{EV}_j^2}[I_n - \frac{1}{n}J_n + \frac{1}{n}(\underset{\sim}{\beta} - \underset{\sim}{j}_n)(\underset{\sim}{\beta} - \underset{\sim}{j}_n)']).$$

CHAPTER VII

A COMPARISON OF DIFFERENT ESTIMATION METHODS

In this chapter, we attempt a limited comparison of three estimates derived in the previous chapters. We will make the comparison through the MSE's of the estimators. Since the exact Mean Squared Errors are difficult to obtain, we only derive approximate forms.

From (3.10), when $r \to \infty$, we have the MSE of the M.L.E. for $\underset{\sim}{\beta}$ as:

$$MSE(\hat{\underset{\sim}{\beta}}_M) = E\{(\hat{\underset{\sim}{\beta}}_M - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}}_M - \underset{\sim}{\beta})\}$$

$$= \frac{\sigma^2}{r(\sum_j EV_j^2)}[n - 1 + \frac{1}{n}\sum_i(\beta_i - 1)^2]. \tag{7.1}$$

From (4.4) and (4.5), the approximate results of $E(\hat{\beta}_{i_G})$ and $Var(\hat{\beta}_{i_G})$ we have the approximate form of the MSE of R.M.L.E. for $\underset{\sim}{\beta}$ as

$$MSE(\hat{\underset{\sim}{\beta}}_G) \approx \frac{n(n-1)\sigma^2[(m^2-1)\sigma^4+2(m+1)\sigma^2\eta+\eta^2]}{[(m-1)\sigma^2+\eta]^3}$$

$$+\{\sigma^2\sum_i(\beta_i-1)^2[(m-1)^5\sigma^{10}+(m-1)^4(4m-3)\sigma^8\eta+6(m-1)^4\sigma^6\eta^2+4(m-1)^3\sigma^4\eta^3+(m+1)^2\sigma^2\eta^4$$

$$+3\eta^5]\}/[(m-1)\sigma^2+\eta]^6. \tag{7.2}$$

where $\eta = nr\sum_j EV_j^2$.

From (6.5), (7.1), and (7.2), it is difficult to decide which one is smaller. To get a comparison result, we will use the computer to simulate some special cases.

In the computer simulations, we set the number of genotypes n as 5, the number of environments m as 3 or 5, the number of replications r as 3 or 30, and the variance of sampling errors σ^2 as 1 or 100. In each case we generate two different sets of $\underset{\sim}{EV}$. One set has the sums of squares, $\underset{\sim}{EV}'\underset{\sim}{EV}$, less than 10, and the others greater than 10,000. For 100 sets of observations in each case, we will compute the average of $\sum_{\ell=1}^{100}(\hat{\underset{\sim}{\beta}}_\ell-\underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}}_\ell-\underset{\sim}{\beta})$ for three different estimation methods. The results are shown as Table V.

From Table V, we can see that Restricted Maximum Likelihood method (or Generalized Least Squares Method) always yields a small average mean squared error than Ratio method. And unless r and $\sum_j EV_j^2$ are large, R.M.L.E. is always "better" than M.L.E. If $\sum_j EV_j^2$ is small and σ^2 is large, R.E. appears to be as good as M.L.E.

TABLE V

SIMULATION RESULT FOR THE COMPARISON OF MEAN SQUARED ERRORS OF MAXIMUM LIKELIHOOD ESTIMATORS, RESTRICTED MAXIMUM LIKELIHOOD ESTIMATORS, AND RATIO ESTIMATORS OF $\underset{\sim}{\beta}$

n	m	r	σ^2	$\sum_j EV_j^2$	Average Mean Squared Errors M.L.E.	R.M.L.E.	R.E.
5	3	3	1	2.0 E+0	7.5879 E-1	7.2202 E-1	8.3090 E-1
				2.0 E+4	7.5106 E-5	7.5093 E-5	1.0029 E-4
			100	2.0 E+0	4.6195 E+2	8.8479 E+1	9.8031 E+1
				2.0 E+4	7.7773 E-3	7.7393 E-3	8.4602 E-3
		30	1	2.0 E+0	7.5351 E-2	7.4929 E-2	9.0217 E-2
				2.0 E+4	7.1882 E-6	7.1888 E-6	9.9421 E-6
			100	2.0 E+0	1.0902 E+2	9.5655 E+0	1.0772 E+1
				2.0 E+4	7.0615 E-4	7.0618 E-4	9.4806 E-4
	5	3	1	1.0 E+1	2.2645 E-1	2.1809 E-1	4.1516 E-1
				1.0 E+5	2.1157 E-5	2.1156 E-5	3.5423 E-5
			100	1.0 E+1	2.2673 E+2	6.0726 E+0	9.5449 E+0
				1.0 E+5	1.8921 E-3	1.8939 E-3	3.2559 E-3
		30	1	1.0 E+1	2.1427 E-2	2.1388 E-2	3.9562 E-2
				1.0 E+5	2.2966 E-6	2.2967 E-6	3.7314 E-6
			100	1.0 E+1	7.5892 E+0	2.0648 E+0	3.4054 E+0
				1.0 E+5	2.3004 E-4	2.3010 E-4	3.8913 E-4

CHAPTER VIII

SUMMARY

The main objectives of this thesis are to obtain estimates of stability parameters under different models and estimation methods and to derive the distributions of the estimators. We consider two different design, namely, completely randomized design within each environment and randomized blocks design within each environment. Normality is assumed in most models; however, we also consider the case that the distribution of the random error is unknown.

8.1. Results and Conclusions

In this paper, we consider four different estimation methods to estimate the stability parameters. These are Maximum Likelihood (ML), Restricted Maximum Likelihood (RML), Least Squares (LS), and a ratio (R) method. From (3.5), (3.16), (5.3), and (5.8), we can see the ML estimator for $\underset{\sim}{\beta}$ of the two different designs with two different assumptions has the same form. Also from (4.3), (5.6), and (5.11), the RML estimator for $\underset{\sim}{\beta}$ and, from (6.10), the LS estimator for $\underset{\sim}{\beta}$ are the same conditioned on $\hat{\underset{\sim}{EV}} = B'\underset{\sim}{y}$.

The simulation results for the comparison of ML, RML, and R suggests that, under normality, if the number of replications r and the variance of environments are large, we should employ the ML method to estimate the stability parameters; otherwise, we should use the RML method. If the distribution of the observations is unknown, we can employ the LS method. Since the ratio estimator yields the largest mean squared error in most cases and its distribution is very difficult to derive, its use is not recommended in spite of its ease of calculation.

8.2. Further Work

There are several areas involving genotype-environment interactions that need further research. One interesting area is that the results obtained in this thesis could, probably, be generalized to other experimental designs and to "messy data" situations. Another area is the assumption that the genotype-environment interaction term in the model is $(\beta_i - 1)EV_j$. For further research, one could consider the model as

$$Y_{ijk} = \mu + \tau_i + EV_j + \sum_{\ell=1}^{S} \alpha_{i\ell} \, EV_j^{\ell} + e_{ijk}$$

$$i = 1,\ldots,n; \; j = 1,\ldots,m; \; k = 1,\ldots,r$$

where $S \leq [mr - 1 - \frac{m}{n}]$. After obtaining the estimator for $\alpha_{i\ell}$ $\forall i,_{\ell}$, a test of $\underset{\sim}{\alpha}_{\ell} = 0$ $\forall \ell$ could be developed.

A SELECTED BIBLIOGRAPHY

Arnold, S. F. (1980). Asymptotic Validity of F Tests for the Ordinary Linear Model and Multiple Correlation Model. J.A.S.A., 75: 890-894.

Bucio Alanis, L., Perkins, J. M., and Jinks, J. L. (1969). Environmental and Genotype-Environmental Components of Variability, V. Segregating Generations. Heredity, 24:115-127.

Eberhart, S. A., and Russell, W. A. (1966). Stability Parameters for Comparing Varieties. Crop. Sci., 6:36-40.

Finlay, Y. W., and Wilkinson, G. N. (1963). The Analysis of Adaptation in a Plant-breeding Programme. Aus. J. Agric. Res., 14:742-754.

Fisher, R. A., and MacKenize, W. A. (1923). Studies in Crop Variation, II. The Manurial Response of Different Potato Varieties. J. Agric. Sci. Comb., 13:311-320.

Freeman, G. H. (1973). Statistical Methods for the Analysis of Genotype-Environment Interactions. Heredity, 31:339-354.

Freeman, G. H., and Perkins, J. M. (1971). Environmental and Genotype-Environmental Components of Variability, VIII. Relation Between Genotypes Grown in Different Environments and Measures of These Environments. Heredity, 27:15-23.

Henderson, C. R., Kempthorne, O., Searle, S. R. and von Krosig, C. M. (1959). The Estimation of Environmental and Genetic Trend from Records Subject to Culling. Biometrics, 15:192-218.

Khatri, C. G. (1963). Further Contributions to Wishartness and Independence of Second Degree Polynomials in Normal Vectors. J. Indian Stat. Assoc., 1:61-70.

Khatri, C. G. (1968). Some Results for the Singular Normal Multivariate Regression Models. Sankhyā Ser. A, 30:267-280.

Mandel, J. (1961). Non-additivity in Two-way Analysis of Variance. J.A.S.A., 56:878-888.

Mandel, J. (1971). A New Analysis of Variance Model for Non-additivity Data. Technometrics, 13:1-18.

Miezon, K., Milliken, G. A., and Liang, G. H. (1979). Using Regression Coefficient as a Stability Parameter in Plant Breeding Programs. Theor. Appl. Genet., 54:7-9.

Perkins, J. M. and Jinks, J. L. (1968b). Environmental and Genotype-Environmental Components of Variability, IV. Non-linear Interactions for Multiple Inbred Lines. Heredity, 23:525-535.

Perkins, J. M. and Jinks, J. L. (1971). Specificity of Interaction of Interaction of Genotypes with Contrasting Environments. Heredity, 26:463-474.

Rao, C. R. (1966). Generalized Inverse for Matrices and Its Applications in Mathematical Statistics. Research paper in statistics, Festschrift for J. Neyman Ed., F. N. David Wiley, N. Y.

Rayner, A. A. and Livingstone, D. (1965). On the Distribution of Quadratic Forms in Singular Normal Variables. S. African J. Agric. Sci., 8:357-370.

Rowe, P. R., and Andrew, R. A. (1964). Phenotypic Stability for a Systematic Series of Corn Genotypes. Crop. Sci., 4:563-567.

Shanbhag, D. N. (1966). On the Independence of Quadratic Form. J. Roy. Stat. Soc. (B), 28:582-583.

Yates, F. and Cochran, W. G. (1938). The Analysis of Groups of Experiments. J. Agric. Sci. Comb., 28:556-580.

APPENDICES

APPENDIX A

A SAS PROGRAM FOR THE SOLUTIONS OF THE ESTIMATORS OF THE STABILITY PARAMETERS WHEN THE DESIGN IS CRD

The following program runs using SAS (1979 version) and can only be used when the data set is balanced and the design is CRD.

Before using this program, the data set must be sorted by genotypes, environments, and replications (i.e., PROC SORT; BY GENO ENV REP;). The user should enter the number of genotypes N, environments M, and replications R into statement 11. The user also needs to specify whether the environmental effects are fixed or random. Enter MD1 into statement 55 for fixed effects or MD2 for random effects.

This program will output the estimates of the stability parameters, the estimates of the variances of the estimators, and the observed significance level for testing the hypothesis of equality of the stability parameters for two cases. The first case uses the RML method which conditions on the mean performance of all genotypes grown in each environment. The second case uses the ML method with the environmental effects specified as fixed or random by the user.

```
DATA GNXEV;
INPUT GENO ENV REP YIELD;
CARDS;
PROC SORT ; BY GENO ENV REP ;
DATA GNXEV ; SET GNXEV ; DROP GENO ENV REP ;
PROC MATRIX;
N=???; M=???; R=???;
COMMENT  N: THE NUMBER OF GENOTYPES (VARIETIES),
         M: THE NUMBER OF ENVIRONMENTS (LOCATIONS, YEARS, ETC.),
         R: THE NUMBER OF REPLICATIONS;
N1=N-1;
NM=N#M;
NR=N#R;
NR1=NR-1;
MR=M#R;
NMR=N#MR;
FETCH Y DATA=GNXEV;
YIJB=J(NM,1);
DO IJ=1 TO NM;
R11=R#(IJ-1)+1;
R12=R#IJ;
YIJB(IJ,)=(J(1,R)*Y(R11:R12,))#/R;
END;
YIB=(I(N)@J(1,M))*YIJB#/M;
YJB=(J(1,N)@I(M))*YIJB#/N;
YB=J(1,NMR)*Y#/NMR;
A1=YIJB-YIB@J(M,1);
A2=YIJB-J(N,1)@YJB;
NOTE PAGE  T H E  M O D E L;
NOTE SKIP=3 Y(I,J,K) = U + GEN(I) + BETA(I)*ENV(J) + ER(I,J,K);
NOTESKIP=3 'WHERE                                                   ';
NOTE SKIP=2 '     Y(I,J,K)  IS THE PHENOTYPIC PERFORMANCE OF THE   ';
NOTE SKIP=1 '               ITH GENOTYPE AT THE JTH ENVIRONMENT ON';
NOTE SKIP=1 '               THE KTH REPLICATE                      ';
NOTE SKIP=2 '     U         IS THE OVERALL MEANS OF THE POPULATION';
NOTE SKIP=2 '     GEN(I)    IS THE EFFECT OF THE ITH GENOTYPE      ';
NOTE SKIP=2 '     BETA(I)   IS THE STABILITY PARAMETER WHICH       ';
NOTE SKIP=1 '               MEASURES THE RESPONSE OF THE ITH       ';
NOTE SKIP=1 '               GENOTYPE TO VARYING ENVIRONMENTS       ';
NOTE SKIP=2 '     ENV(J)    IS THE EFFECT OF JTH ENVIRONMENT       ';
NOTE SKIP=2 '     ER(I,J,K) IS THE RANDOM ERROR ASSOCIATED WITH    ';
NOTE SKIP=1 '               THE ITH GENOTYPE AT THE JTH            ';
NOTE SKIP=1 '               ENVIRONMENT ON THE KTH REPLICATION AND';
NOTE SKIP=1 '               ASSUME ER(I,J,K)°N(0,V) AND ER(I,J,K) ';
NOTE SKIP=1 '               UNCORRELATED FOR ALL I,J,K.            ';
NAME1='BETA1' 'BETA2' 'BETA3' 'BETA4' 'BETA5' 'BETA6' 'BETA7' ;
```

```
LINK ???;
COMMENT MD1: GIVEN THE RESULT WHILE ASSUMING THE ENVIRONMENTAL
             EFFECTS ARE FIXED,
        MD2: GIVEN THE RESULT WHILE ASSUMING THE ENVIRONMENTAL
             EFFECTS ARE RANDOM;
STOP;
MD1:
LINK GM;
LINK FM;
RETURN;
MD2:
LINK GM;
LINK MM;
RETURN;
GM:
ENVH=YJB-YB;
ENVSS=ENVH'*ENVH;
BETAH=(I(N)@ENVH')*YIJB#/ENVSS;
DF1=NMR-M+1;
DV1=Y-(YIB@J(MR,1))-(BETAH@ENVH@J(R,1));
VH=DV1'*DV1#/DF1;
VARBH=(VH#/(R#ENVSS))#(I(N)-J(N,N)#/N);
VBD=(2#VH#/(R#ENVSS))##.5;
LINK TST;
OSL=(J(N,N)-PROBT(STAT,DF1))#2;
NOTE PAGE ESTIMATION OF THE STABILITY PARAMETER ;
NOTE SKIP=1 BY USING MAXIMUM LIKELIHOOD METHOD;
NOTE SKIP=1 WHILE THE ENVIRONMENTAL EFFECT ARE GIVEN BY THE MEAN
PERFORMANCE OF ALL THE GENOTYPES GROWN IN THAT ENVIRONMENT.;
NOTE SKIP=5 THE ESTIMATE VALUE OF BETA;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL ROWNAME=NAME1 COLNAME=NAME1;
RETURN;
FM:
BETAH=J(N,1);
```

```
DO L1=1 TO 500;
ENVH=(BETAH'@I(M))*A1#/(BETAH'*BETAH);
BETAHT=J(N,1)+(I(N)@ENVH')*A2#/(ENVH'*ENVH);
D=J(1,N)*(ABS(BETAHT-BETAH));
BETAH=BETAHT;
IF D<0.0000000001 THEN GO TO W1;
END;
W1:
DV2=Y-(YIB@J(MR,1))-(BETAH@ENVH@J(R,1));
VH=DV2'*DV2#/NMR;
VARBH=(VH#/(R#(ENVH'*ENVH)))#(I(N)-J(N,N)#/N
+(BETAH-J(N,1))*(BETAH-J(N,1))'#/N);
VBD=(2#VH#/(R#(ENVH'*ENVH)))##.5;
LINK TST;
OSL=(J(N,N)-PROBNORM(STAT))#2;
NOTE PAGE ESTIATION OF THE STABILITY PARAMETER ;
NOTE SKIP=1 BY USING MAXIMUM LIKELIHOOD METHOD;
NOTE SKIP=1 WHILE ASSUMING THAT THE ENVIRONMENTAL EFFECT ARE FIXED;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL COLNAME=NAME1 ROWNAME=NAME1;
MM:
BETAH=J(N,1);
DO L1=1 TO 500;
BETAHT=J(N,1)+(BETAH'*BETAH)#(A2'*(I(N)@(A1'*(BETAH@I(M)))'))'
#/((A1'*(BETAH@I(M)))*(A1'*(BETAH@I(M)))');
D=J(1,N)*(ABS(BETAHT-BETAH));
BETAH=BETAHT;
IF D<0.0000000001 THEN GO TO W2;
END;
W2:
BETASS=BETAH'*BETAH;
VE1=(A1'*(BETAH@I(M)))*(A1'*(BETAH@I(M)))';
VE2=(Y-(YIB@J(MR,1)))'*(Y-(YIB@J(MR,1)));
VEH=(NR#VE1-BETASS#VE2#/R)#/(NR1#M#BETASS#BETASS);
VOH=(BETASS#VE2-R#VE1)#/(NR1#M#BETASS);
VBD=(2#VOH#(VOH+R#VEH#BETASS)#/(M#R#R#VEH#VEH#BETASS))##.5;
VARBH=((VOH#(VOH+R#VEH#BETASS)#/(M#R#R#VEH#VEH#BETASS))#
(I(N)-J(N,N)#/N+(BETAH-J(N,1))*(BETAH-J(N,1))'#/N);
LINK TST;
OSL=(J(N,N)-PROBNORM(STAT))#2;
NOTE PAGE ESTIMATION OF THE STABILITY PARAMETER ;
NOTE SKIP=1 BY USING MAXIMUM LIKELIHOOD METHOD;
```

```
NOTE SKIP=1 WHILE ASSUMING THAT THE ENVIRONMENTAL EFFECT ARE RANDOM;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL COLNAME=NAME1 ROWNAME=NAME1;
RETURN;



TST:
STAT=J(N,N);
DO L2=1 TO N;
STAT(L2,)=ABS(BETAH(L2,)-BETAH')#/VBD;
END;
RETURN;
```

APPENDIX B

A SAS PROGRAM FOR THE SOLUTIONS OF THE ESTIMATORS OF THE STABILITY PARAMETERS WHEN THE DESIGN IS RCBD

The following program runs using SAS (1979 version) and can only be used when the data set is balanced and the design is RCBD.

Before using this program, the data set must be sorted by genotypes, environments, blocks, and replications (i.e., PROC SORT; BY GENO ENV BLK REP;). The user should enter the number of genotypes N, environments M, block R, and replications S into statement 7. The user also needs to specify whether the block effects are fixed or random. Entering MD1 into statement 61 for fixed effects and MD2 for random effects.

This program will output the estimates of the stability parameters, the estimates of variances of the estimators, and the observed significance level for testing the hypothesis of equality of the stability parameters for two cases. The first case is the result using the RML method and the second case is the result using the ML method.

```
DATA GNXEV;
INPUT GENO ENV BLK REP YIELD;
CARDS;
PROC SORT ; BY GENO ENV BLK REP ;
DATA GNXEV ; SET GNXEV ; DROP GENO ENV BLK REP ;
PROC MATRIX;
N=???; M=???; R=???; S=???;
COMMENT  N: THE NUMBER OF GENOTYPES (VARIETIES),
.        M: THE NUMBER OF ENVIRONMENTS (LOCATIONS, YEARS, ETC.),
.        R: THE NUMBER OF BLOCKS,
.        S: THE NUMBER OF REPLICATIONS;
N1=N-1;
NM=N#M;
NR=N#R;
NR1=NR-1;
MR=M#R;
NMR=N#MR;
NMRS=NMR#S;
MRS=MR#S;
RS=R#S;
FETCH YS DATA=GNXEV;
Y=J(NMR,1);
DO IJK=1 TO NMR;
S11=S#(IJK-1)+1;
S12=S#IJK;
Y(IJK,)=(J(1,S)*YS(S11:S12,))#/S;
END;
YIJB=J(NM,1);
DO IJ=1 TO NM;
R11=R#(IJ-1)+1;
R12=R#IJ;
YIJB(IJ,)=(J(1,R)*Y(R11:R12,))#/R;
END;
YIB=(I(N)@J(1,M))*YIJB#/M;
YJB=(J(1,N)@I(M))*YIJB#/N;
YJKB=(J(1,N)@I(MR))*Y#/N;
YB=J(1,NMR)*Y#/NMR;
A1=YIJB-YIB@J(M,1);
A2=YIJB-J(N,1)@YJB;
NOTE PAGE  T H E  M O D E L;
NOTE SKIP=3 Y(I,J,K,L)=U+GEN(I)+BETA(I)ENV(J)+BLK(K(J))+ER(I,J,K,L);
NOTE SKIP=3 'WHERE                                                    ';
NOTE SKIP=2 '   Y(I,J,K,L)  IS THE PHENOTYPIC PERFORMANCE OF THE  ';
NOTE SKIP=1 '               ITH GENOTYPE AT THE JTH ENVIROFMENT ON';
NOTE SKIP=1 '               THE KTH BLOCK ON THE KTH REPLICATE    ';
NOTE SKIP=2 '   U           IS THE OVERALL MEANS OF THE POPULATION';
NOTE SKIP=2 '   GEN(I)      IS THE EFFECT OF THE ITH GENOTYPE     ';
NOTE SKIP=2 '   BETA(I)     IS THE STABILITY PARAMETER WHICH      ';
NOTE SKIP=1 '               MEASURES THE RESPONSE OF THE ITH      ';
NOTE SKIP=1 '               GENOTYPE TO VARYING ENVIRONMENTS      ';
NOTE SKIP=2 '   ENV(J)      IS THE EFFECT OF JTH ENVIRONMENT      ';
NOTE SKIP=2 '   BLOCK(K(J)) IS THE EFFECT OF KTH BLOCK WITHIN THE ';
NOTE SKIP=1 '               JTH ENVIRONMENT                       ';
NOTE SKIP=2 '   ER(I,J,K,L) IS THE RANDOM ERROR ASSOCIATED WITH   ';
```

```
NOTE SKIP=1 '                 THE ITH GENOTYPE OF THE JTH            ';
NOTE SKIP=1 '                 ENVIRONMENT IN THE KTH BLOCK ON STH    ';
NOTE SKIP=1 '                 REPLICATION.                           ';
NOTE SKIP=1 '                 ASSUME ER(I,J,K,L)°N(0,V) AND FOR ALL ';
NOTE SKIP=1 '                 I,J,K,L ER(I,J,K,L) ARE UNCORRELATED  ';
NAME1='BETA1' 'BETA2' 'BETA3' 'BETA4' 'BETA5' 'BETA6' 'BETA7';
LINK ???;
COMMENT MD1: GIVEN THE RESULT WHILE ASSUMING THE BLOCK
.            EFFECTS ARE FIXED,
.       MD2: GIVEN THE RESULT WHILE ASSUMING THE BLOCK
.            EFFECTS ARE RANDOM;
STOP;

MD1:
ENVH=YJB-YB;
ENVSS=ENVH'*ENVH;
BETAH=(I(N)@ENVH')*YIJB#/ENVSS;
DF1=NMRS-S#(M-1);
DV1=YS-(YIB@J(MRS,1))-(BETAH@ENVH@J(RS,1))-(J(N,1)@(YJKB-(YJB@
J(R,1)))@J(S,1));
VH=DV1'*DV1#/DF1;
VARBH=(VH#/(R#ENVSS))#(I(N)-J(N,N)#/N);
VBD=(2#VH#/(R#ENVSS))##.5;
LINK TST;
OSL=(J(N,N)-PROBT(STAT,DF1))#2;
NOTE PAGE ESTIMATION OF THE STABILITY PARAMETER FOR FIXED BLOCK
EFFECTS BY USING RESTRICTED MAXIMUM LIKELIHOOD METHOD;
NOTE SKIP=1 WHILE THE ENVIRONMENTAL EFFECT ARE GIVEN BY THE MEAN
PERFORMANCE OF ALL THE GENOTYPES GROWN IN THAT ENVIRONMENT.;
NOTE SKIP=5 THE ESTIMATE VALUE OF BETA;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL ROWNAME=NAME1 COLNAME=NAME1;
BETAH=J(N,1);
DO L1=1 TO 500;
ENVH=(BETAH'@I(M))*A1#/(BETAH'*BETAH);
BETAHT=J(N,1)+(I(N)@ENVH')*A2#/(ENVH'*ENVH);
D=J(1,N)*(ABS(BETAHT-BETAH));
BETAH=BETAHT;
IF D<0.0000000001 THEN GO TO W1;
END;
W1:
DV2=YS-(YIB@J(MRS,1))-(J(N,1)@(YJKB-(YJB@J(R,1))@J(S,1)-(BETAH@
ENVH@J(RS,1));
VH=DV2'*DV2#/NMRS;
VARBH=(VH#/(R#(ENVH'*ENVH))#(I(N)-J(N,N)#/N
+(BETAH-J(N,1))*(BETAH-J(N,1))'#/N);
VBD=(2#VH#/(R#(ENVH'*ENVH)))##.5;
LINK TST;
OSL=(J(N,N)-PROBNORM(STAT))#2;
```

```
NOTE PAGE ESTIATION OF THE STABILITY PARAMETER FOR FIXED BLOCK
EFFECTS BY USING MAXIMUM LIKELIHOOD METHOD;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL COLNAME=NAME1 ROWNAME=NAME1;
RETURN;


MD2:
ENVH=YJB-YB;
ENVSS=ENVH'*ENVH;
BETAH=(I(N)@ENVH')*YIJB#/ENVSS;
DF1=NMRS-MRS;
DV1=YS-(YIB@J(MRS,1))-((BETAH-J(N,1))@ENVH@J(RS,1))-(J(N,1)@
YJKB@J(S,1));
VH=DV1'*DV1#/DF1;
VARBH=(VH#/R#ENVSS)#(I(N)-J(N,N)#/N);
VBD=(2#VH#/(R#ENVSS))##.5;
LINK TST;
OSL=(J(N,N)-PROBT(STAT,DF1))#2;
NOTE PAGE ESTIMATION OF THE STABILITY PARAMETER FOR RANDOM BLOCK
EFFECTS BY USING RESTRICTED MAXIMUM LIKELIHOOD METHOD;
NOTE SKIP=1 WHILE THE ENVIRONMENTAL EFFECT ARE GIVEN BY THE MEAN
PERFORMANCE OF ALL THE GENOTYPES GROWN IN THAT ENVIRONMENT.;
NOTE SKIP=5 THE ESTIMATE VALUE OF BETA;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL ROWNAME=NAME1 COLNAME=NAME1;
DO L=1 TO 500;
BETAHT=J(N,1)+(I(N)@ENVH')*A2#/(ENVH'*ENVH);
V1=(YS-J(NMRS,1,YB)-(I(N)@ENVH@J(RS,1)))'*(J(N,1)@I(MRS));
V2=YS-(YIB@J(MRS,1))-(BETAH@ENVH@J(RS,1));
VB=((V1'*V1)-(V2'*V2))#/(N1#NMRS);
VE=(N#(V2'*V2)-(V1'*V1))#/(N1#NMRS);
IF VE LT 0 THEN GO TO W5;
IF VB LT 0 THEN GO TO W5;
ENVHT=(N#VB+VE)#((BETAH'@I(M))*A1-N#VB#(YJB-YB))#/
((N#VB+VE)#(BETAH'*BETAH)-N#N#VB);
D=J(1,M)*(ABS(ENVHT-ENVH));
ENVH=ENVHT;
IF D<0.0000000001 THEN GO TO W4;
END;
VARBH=VE#(I(N)-J(N,N)#/N+(N#VB+VE)#(BETAH-J(N,1))'*(BETAH
-J(N,1))#/(N#N1#VB+N#VE))#/(ENVH'*ENVH);
VBD=(2#VE#/(R#S#(ENVH'*ENVH)))##.5;
LINK TST;
```

```
OSL=(J(N,N)-PROBNORM(STAT))#2;
NOTE PAGE ESTIMATION OF THE STABILITY PARAMETER FOR RANDOM BLOCK
EFFECTS BY USING MAXIMUM LIKELIHOOD METHOD;
BETAH=BETAH';
PRINT BETAH COLNAME=NAME1;
NOTE SKIP=2 THE ESTIMATE VARIANCE OF THE ESTIMATE BETA;
PRINT VARBH COLNAME=NAME1 ROWNAME=NAME1;
NOTE PAGE COMPARISON OF BETA(I) AND BETA(I");
NOTE SKIP=3 THE OBSERVE SIGNIFICANCE LEVEL ;
PRINT OSL COLNAME=NAME1 ROWNAME=NAME1;
W5:
RETURN;
TST:
STAT=J(N,N);
DO L2=1 TO N;
STAT(L2,)=ABS(BETAH(L2,)-BETAH')#/VBD;
END;
RETURN;
```

APPENDIX C

FIGURES

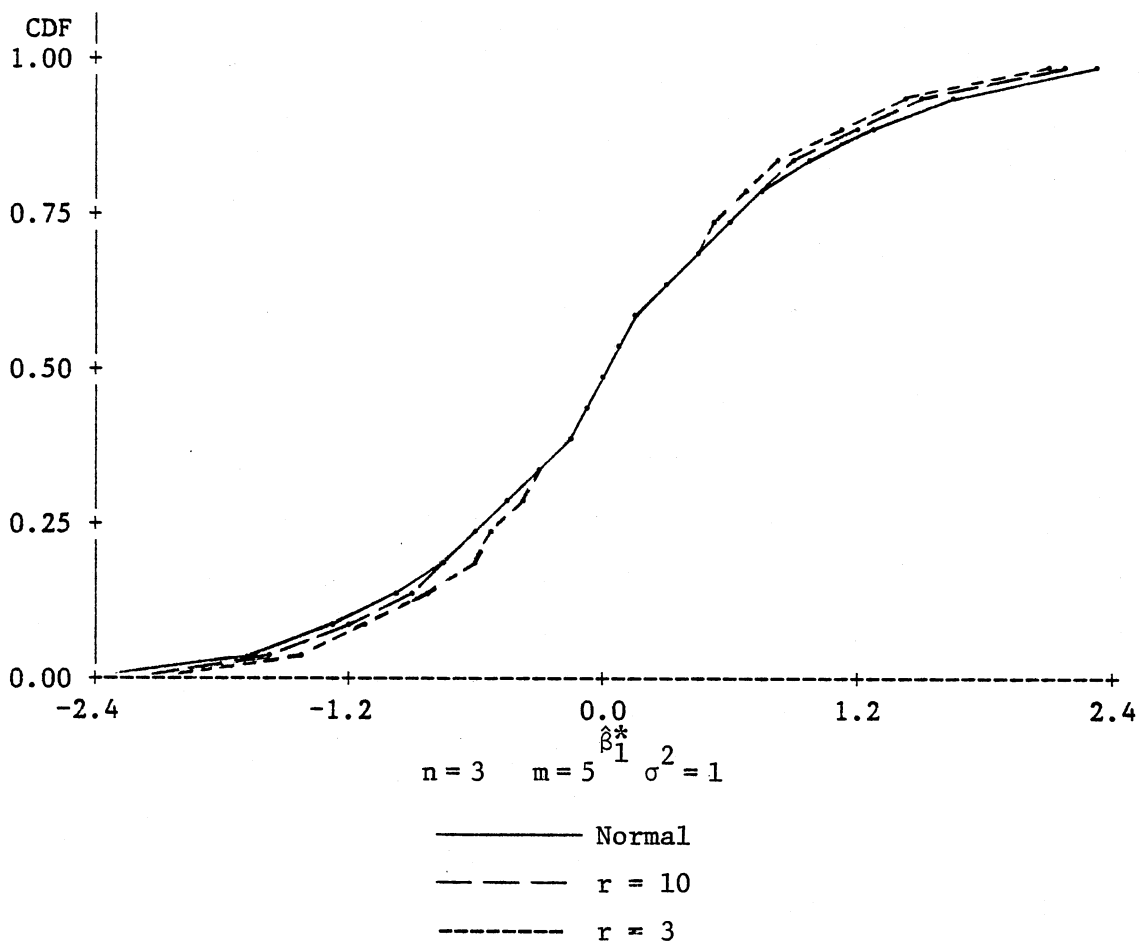

Figure 1. The Cumulative Distribution Function of $\hat{\beta}_1^*$ When the Numbers of Replications are 3 and 10

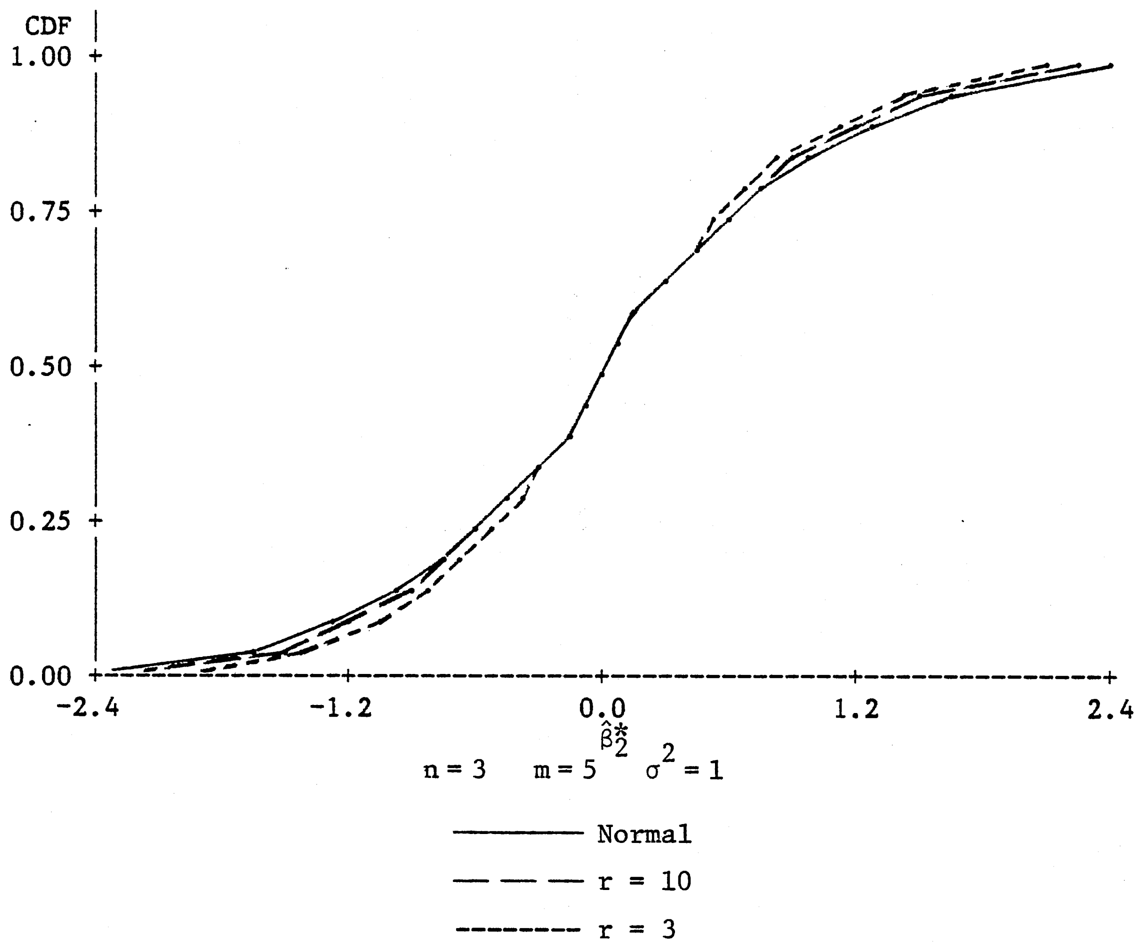

Figure 2. The Cumulative Distribution Function of $\hat{\beta}_2^*$ When the Numbers of Replications are 3 and 10

VITA

Chung-Hsien Sung

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE ESTIMATION OF STABILITY PARAMETERS

Major Field: Statistics

Biographical:

Personal Data: Born in Ilan, Taiwan, Republic of China, September 6, 1950, the first son of Mr. and Mrs. Yane-Keng Sung.

Education: Graduated from Kaohsiung High School, Kaohsiung, Taiwan, in 1969; received Bachelor of Science degree in Mathematical Statistics from Tamkang College, Taipei, Taiwan, in 1974; received Master of Arts degree in Mathematics from Northeast Missouri State University, Kirksville, Missouri, in 1979; completed requirements for the Doctor of Philosophy degree at Oklahoma State University, July, 1983.

Professional Experience: Associate Actuary, Cathay Life Insurance Co., Taipei, Taiwan, 1977-1978; Graduate Teaching Associate, Northeast Missouri State University, Kirksville, Missouri, 1978-1979; Graduate Teaching Associate, Oklahoma State University, 1979-1983.