BAYESIAN INFERENCE FOR THE VARIANCE COMPONENTS IN MIXED LINEAR MODELS

By

MUTHIYA RAJAGOPALAN

Bachelor of Science (Honours)
Annamalai University
Tamil Nadu, India
1960

Master of Science
Annamalai University
Tamil Nadu, India
1967

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
December, 1980

BAYESIAN INFERENCE FOR THE VARIANCE COMPONENTS IN

MIXED LINEAR MODELS

Thesis Approved:

Lyle Broemeling
Thesis Adviser

K E Case

Don Holbert

P. L. Claypool

Norman N Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my thanks and appreciation to my adviser, Dr. Lyle D. Broemeling, for suggesting this topic and for his valuable advice and encouragement.

I also wish to thank Dr. Larry Claypool, Dr. Donald Holbert, and Dr. Kenneth Case for serving on my advisory committee.

I would also like to thank my many friends who helped very much to make my stay in the United States pleasant, enjoyable and useful.

TABLE OF CONTENTS

TABLE

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The purpose of this study is to investigate the general mixed linear model using Bayesian techniques and to obtain the posterior distributions of the parameters involved in the model, especially the variance components. Such models contain both fixed and random factors and are widely used in scientific investigations especially in biological, agricultural, and engineering studies.

This problem has drawn the attention of many research workers. Their studies can be grouped in to two approaches, namely, sampling theory (or classical approach) and Bayesian.

Under classical studies there are three main approaches to the problem. They are

1. analysis of variance methods,
2. maximum likelihood methods, and
3. minimum norm methods.

Due to computational difficulties and, sometimes, due to the possession of certain undesirable properties (like a negative estimate for a variance) none of these methods is completely satisfactory in all situations. No general method is put forward to cover all situations and when alternative methods are suggested no principle has been laid down for choosing one among them as appropriate in a given problem.

Under the Bayesian approach, basically, all investigators have used Bayes' theorem after a suitable and convenient choice of a prior distribution for the parameters involved in the model. One group of workers have employed linear models, sometimes in a hierarchical order, for the prior distribution of the parameters and another group of workers have used standard probability distributions. Most of the workers in the second group have used improper priors for the parameters.

Regarding mixed linear models in multivariate cases, very little has been done under both sampling theory and Bayesian approaches.

In this study, inverse gamma distributions are used for the prior distributions of the variance components or in other words gamma distributions for the precision components and the marginal and conditional (conditional on the random factors) posterior distributions of the variance components are studied. Since the exact marginal distributions are very difficult to obtain certain approximations are considered. Using the multivariate normal as an approximation for the multivariate t, an approximation is obtained for the joint distribution of the random effects. On this basis, approximations are obtained for the distributions of the variance components. Further, closed expressions for the first two moments of the variance components are obtained using the approximate distribution of the random effects. The closeness of these approximations is numerically studied using Box's data.

The organization of this thesis is as follows. The relevant literature is reviewed in Chapter II. Chapter III describes the model and the basic assumptions associated with the model and also discusses the conditional and marginal posterior distributions necessary for solving the problem. Chapter IV deals with the approximations to the various distributions of the parameters described in Chapter III. Chapter V discusses the results of a numerical study carried out to investigate the closeness of the approximations. Chapter VI contains a brief summary of the thesis and recommendations for further research.

CHAPTER II

REVIEW OF LITERATURE

There is considerable literature on the subject of variance components. As pointed out in Chapter I these could be broadly classified under two heads, namely, sampling theory (or classical theory) methods and Bayesian methods.

An exhaustive, but brief, survey of the different methods under sampling theory is given by Searle (1978). The book by Box and Tiao (1973) provides a good account of the Bayesian methodology as applied to linear models. Kleffe (1977) gives a good survey of the different methods of estimating variance components. A review of the related literature, under the different approaches, is given below.

Analysis of Variance Methods

Until about a decade ago, methods of estimation of variance components were based on equating the sums of squares of the entries in the analysis of variance to their expected values, assuming the existence of second order moments. This was first proposed by Daniels (1939) and Winsor and Clarke (1940). Henderson (1953) has an important paper in this area where he discussed the different possibilities under three cases, methods I, II, and III. Searle (1968) gives a good description of these methods and indicates various generalizations. The properties of estimates provided by such methods have been studied by Graybill and

Wortham (1956) and Graybill and Hulquist (1961). There are no distributional assumptions involved in the application of these methods excepting that second order moments are assumed to exist. These methods yield translation invariant quadratic unbiased estimates. But negative estimates could be realized and the theoretical basis is not clear. Further, there is no single general method to cover all cases. In general, these estimates are uniformly minimum variance unbiased estimates for a large number of balanced designs. But for unbalanced designs these are not the best. Seely (1975) gives an example of an estimate of a variance component based on Analysis of Variance which is inadmissible in the class of quadratic unbiased estimates.

Maximum Likelihood Methods

Even though maximum likelihood, as a general method of estimation, was introduced by Fisher a few decades ago this method received little attention until recently. This is due to the computational difficulties involved in applying this method to the case of mixed linear models since the likelihood equation is quite complicated in such cases. Using normality assumptions on the error component, Hartley and Rao (1967) first introduced the maximum likelihood method to estimate the variance components. They also proposed a computational algorithm for solving the likelihood equation and showed that the estimates are consistent and asymptotically normally distributed under certain conditions. In general, these estimates are biased and computationally quite difficult. Harville (1977) pointed out that the likelihood equation may have multiple roots which may not lie within the parameter space at all.

Patterson and Thompson (1971) considered a variation of the maximum likelihood method which they called restricted maximum likelihood method. This method consists in assuming normality, as with maximum likelihood, but maximizing only that portion of the log likelihood which is invariant to changes in fixed effects. Corbeil and Searle (1976a, 1976b) adapted Patterson and Thompson's procedure and developed a new procedure which is applicable to mixed models also. Corbeil and Searle (1976b) obtained explicit solutions for the maximum likelihood and restricted maximum likelihood equations under normality for four common balanced variance component models. In most balanced cases the two estimates turn out to be the same.

Minimum Norm Methods

Rao (1971a, 1971b) introduced minimum norm quadratic unbiased estimates (MINQUE) and minimum variance quadratic unbiased estimates (MIVQUE). A linear function of the variance components is estimated by a quadratic function in the observations such that the estimate is translation invariant, unbiased, and a certain norm (like the variance in the case of MINQUE) is minimized. There are no distributional assumptions, but estimates could be negative. This is due to the knowledge of the parameter space not being utilized in the estimation process. Further, prior estimates of the variance components are needed, eventhough, under certain conditions, the estimates are independent of the prior estimates.

Hultquist and Atzinger (1972), assuming normality and independence, obtained minimal sufficient statistics for the parameters in a mixed effects model. They showed the existence of minimum variance unbiased estimates and how to explicitly determine them. Swallow and Searle (1978) obtained a MIVQUE of the variance components in one way random model both for balanced and unbalanced data.

Bayesian Methods.

Lindley and Smith (1972) using a hierarchical form of prior structure, derived Bayesian alternatives to the least square estimators for the parameters in a linear model. Smith (1973) studied the Bayesian approach along the same lines in some more detail and discussed the general properties of the resulting estimates. Leonard (1975) considered a Bayesian approach to the linear model with unequal variances taking a linear model for the prior structure of the means and a log linear model for the variances.

Improper priors have been extensively used in the study of variance components. Rhode (1972) used such priors to study the fixed effects model. Box and Tiao (1973) have made an extensive contribution to the Bayesian inference for the variance components of mixed linear models. They used improper prior distributions for the expected mean squares of the corresponding analysis of variance procedures. But, their methods do not explain how to handle prior information which is always available, atleast in some limited measure. Sahai (1974) obtained some formal Bayes estimates of variance components in the case of a balanced three state nested random model using suitable loss functions. All the studies, mentioned above, are restrictive in the sense that they deal with specific problems like the one way random model and numerical integration techniques are needed to normalize the posterior distributions and to evaluate the posterior moments.

Gharraf (1979) obtained conditional posterior estimates of the variance components in a general mixed linear model, fixing the random effects at their least square estimates. He used improper priors for the fixed effects and inverse gamma priors for the variance components. Broemeling and Gharraf (1979), using proper priors for all the parameters in a general mixed linear model, obtained the posterior distributions of the variance components fixing the random effects.

It is clear, from the preceding survey, that no general solution to the problem of variance components, covering all cases, currently exists. In this thesis an attempt is being made to develop a satisfactory Bayesian solution to that problem.

CHAPTER III

THE POSTERIOR ANALYSIS

The Model

The model investigated in this thesis is the general mixed linear model

$$\underset{\sim}{y} = X\underset{\sim}{\theta} + U\underset{\sim}{b} + \underset{\sim}{e} , \tag{3.1}$$

where

$\underset{\sim}{y}$ is a n-vector of observations,

X is a known $n \times p$ design matrix of full rank $(n > p)$,

$\underset{\sim}{\theta}$ is a p-vector of real unknown parameters,

U is a known $n \times m$ design matrix of rank m $(m < n-p)$,

$\underset{\sim}{b}$ is a m-vector of unobservable random variables, and

$\underset{\sim}{e}$ is a n-vector of unobservable random variables.

If the number of random factors in the model is c and the $i^{\underline{th}}$ random factor occurs at m_i levels then $\underset{\sim}{b}$ can be partitioned as

$$\underset{\sim}{b}' = (\underset{\sim}{b}_1', \underset{\sim}{b}_2', \ldots , \underset{\sim}{b}_c')$$

where

$$\underset{\sim}{b}_i' = (b_{i\,1}, b_{i\,2}, \ldots , b_{i\,m_i})$$

and $b_{i\,1}, b_{i\,2}, \ldots , b_{i\,m_i}$ are the m_i levels of the $i^{\underline{th}}$ factor,

$i = 1, 2, \ldots, c$. If $U = (U_1, U_2, \ldots, U_c)$ is the corresponding partition of U, then the above model (3.1) can be rewritten as

$$\underset{\sim}{y} = X\underset{\sim}{\theta} + \sum_{i=1}^{c} U_i \underset{\sim}{b}_i + \underset{\sim}{e}. \tag{3.2}$$

The assumptions, associated with the model, are the following:

1. $\underset{\sim}{e}$ has n-variate normal distribution with mean vector $\underset{\sim}{0}$ and dispersion matrix $\sigma^2 I_n$, σ^2 being some unknown positive number.
2. $\underset{\sim}{b}_i$ has m_i-variate normal distribution with mean vector $\underset{\sim}{0}$ and dispersion matrix $\sigma_i^2 I_{m_i}$, σ_i^2 being some unknown positive constant, $i = 1, 2, \ldots, c$.
3. $\underset{\sim}{e}, \underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_c$ are mutually independent.
4. Rank of U_i is m_i and $m_1 + m_2 + \ldots + m_c = m$.

The fixed effects are the elements of $\underset{\sim}{\theta}$ and the variance components are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_c^2$, and σ^2. This model covers, as particular cases, the models of almost all design problems encountered in practice.

The likelihood function of the parameters $\underset{\sim}{\theta}$, τ, $\underset{\sim}{\rho}$, and $\underset{\sim}{b}$ is given by

$$L(\underset{\sim}{\theta}, \tau, \underset{\sim}{\rho}, \underset{\sim}{b}) \propto \tau^{(n/2)} \exp\{(\underset{\sim}{y} - X\underset{\sim}{\theta} - U\underset{\sim}{b})'(\underset{\sim}{y} - X\underset{\sim}{\theta} - U\underset{\sim}{b})(-\tau/2)\}$$

$$\prod_{i=1}^{c} \{ \tau_i^{(m_i/2)} \exp(-\tau_i \underset{\sim}{b}_i{}'\underset{\sim}{b}_i / 2)\}, \tag{3.3}$$

$\underset{\sim}{\theta} \in R^p$, $\tau = 1/\sigma^2 > 0$, $\tau_i = 1/\sigma_i^2 > 0$, $\underset{\sim}{b} \in R^m$, and where

$\underset{\sim}{\rho} = (\tau_1, \tau_2, \ldots, \tau_c)$.

The prior distributions for the parameters are assigned as follows:

1. Each component of $\underset{\sim}{\theta}$ has uniform distribution in $(-\infty, \infty)$.
2. τ has gamma distribution with known parameters $\alpha > 0$, and $\beta > 0$.
3. τ_i has gamma distribution with known parameters $\alpha_i > 0$, and $\beta_i > 0$, $i = 1, 2, \ldots, c$.
4. The elements of $\underset{\sim}{\theta}$ and the precision parameters $\tau, \tau_1, \tau_2, \ldots, \tau_c$ are all mutually independent.

Thus the joint prior density function of $\underset{\sim}{\theta}$, τ, and $\underset{\sim}{\rho}$ is given by

$$\Pi_0(\underset{\sim}{\theta}, \tau, \underset{\sim}{\rho}) \propto e^{-\beta\tau} \tau^{\alpha-1} \prod_{i=1}^{c} \{e^{-\beta_i\tau_i} \tau_i^{\alpha_i-1}\}, \tag{3.4}$$

$$\underset{\sim}{\theta} \in R^p; \quad \tau, \tau_1, \tau_2, \ldots, \tau_c > 0.$$

Using (3.3), (3.4), and Bayes' theorem, the joint posterior density of $\underset{\sim}{\theta}$, τ, $\underset{\sim}{\rho}$, and $\underset{\sim}{b}$ is obtained as

$$\Pi_1(\underset{\sim}{\theta}, \tau, \underset{\sim}{\rho}, \underset{\sim}{b}) \propto \tau^{\alpha + (n/2) - 1}$$

$$\times\, e^{-(\tau/2)\{2\beta + (\underset{\sim}{y} - X\underset{\sim}{\theta} - U\underset{\sim}{b})'(\underset{\sim}{y} - X\underset{\sim}{\theta} - U\underset{\sim}{b})\}}$$

$$\times \prod_{i=1}^{c} \tau_i^{\alpha_i + (m_i/2) - 1} e^{-(\tau_i/2)(\underset{\sim}{b}_i'\underset{\sim}{b}_i + 2\beta_i)},$$

$$\underset{\sim}{\theta} \in R^p, \quad \underset{\sim}{\rho} > \underset{\sim}{0}, \quad \underset{\sim}{b} \in R^m. \tag{3.5}$$

The above joint posterior density can be rewritten as

$$\Pi_1(\underset{\sim}{\theta}, \tau, \underset{\sim}{\rho}, \underset{\sim}{b}) \propto \tau^{\{\alpha + (n/2) - 1\}}$$

$$\times \exp \{(-\tau/2) \{2\beta + \underset{\sim}{y}'R\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}$$

$$+ (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}}) + (\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}})'X'X\,(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}})\}$$

$$\times \prod_{i=1}^{c} \tau_i^{\{\alpha_i + (m_i/2) - 1\}}$$

$$\times \exp \{(-\tau_i/2)\,(2\beta_i + \underset{\sim}{b}_i'\underset{\sim}{b}_i)\} \tag{3.6}$$

where

$$\hat{\underset{\sim}{b}} = (U'R\,U)^{-}\,U'R\,\underset{\sim}{y},$$

$$\hat{\underset{\sim}{\theta}} = (X'X)^{-1}\,X'(\underset{\sim}{y} - U\,\underset{\sim}{b}),$$

$$R = I_n - X\,(X'X)^{-1}\,X', \text{ and}$$

$(U'R\,U)^{-}$ is the unique Moore-Penrose generalized inverse of $U'R\,U$.

It is relevant to remark at this stage that if the components of $\underset{\sim}{\theta}$ are apriori assumed to have independent normal distributions instead of the improper uniform distributions in $(-\infty, \infty)$ as assumed above then the resulting joint posterior distribution of $\underset{\sim}{\theta}$, τ, $\underset{\sim}{\rho}$, and $\underset{\sim}{b}$ will tend to (3.6) if the prior variances of the elements of $\underset{\sim}{\theta}$ are made to go to infinity.

The Conditional Posterior Distributions

Integrating (3.6) with respect to τ and $\underset{\sim}{\rho}$, the joint posterior density of $\underset{\sim}{\theta}$ and $\underset{\sim}{b}$ is obtained as

$$\Pi_1(\underset{\sim}{\theta}, \underset{\sim}{b}) \propto \{2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} + (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}})$$

$$+ (\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}})'X'X\,(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}})\}^{-(n + 2\alpha)/2}$$

$$\times \prod_{i=1}^{c} (2\beta_i + \underset{\sim}{b}_i{}'\underset{\sim}{b}_i)^{-(m_i + 2\alpha_i)/2} \qquad (3.7)$$

Further, integrating (3.6) with respect to $\underset{\sim}{\theta}$ the joint posterior density of τ, $\underset{\sim}{\rho}$, and $\underset{\sim}{b}$ is obtained as

$$\Pi_1(\tau, \underset{\sim}{\rho}, \underset{\sim}{b}) \propto \tau^{(\alpha + n/2 - p/2 - 1)}$$

$$\times \exp\{(-\tau/2)\,\{2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}$$

$$+ (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}})\}\}$$

$$\times \prod_{i=1}^{c} \tau_i^{(\alpha_i + m_i/2 - 1)} \exp\{(-\tau_i/2)(2\beta_i + \underset{\sim}{b}_i{}'\underset{\sim}{b}_i)\}$$

$$\underset{\sim}{\rho} > \underset{\sim}{0}, \quad \underset{\sim}{b} \in R^m . \qquad (3.8)$$

From (3.7) and (3.8) it is easily seen that

1. $\underset{\sim}{\theta}$, conditional on $\underset{\sim}{b}$, has a multivariate t distribution with mean vector $\hat{\underset{\sim}{\theta}}$ and precision matrix T where

$$T = \{\frac{n - p + 2\alpha}{2\beta + \underset{\sim}{y}'R\,\underset{\sim}{Y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} + (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}})}\}\,X'X ,$$

2. τ, conditional on $\underset{\sim}{b}$, has a gamma distribution with parameters α^* and β^* ,

3. τ_i, conditional on $\underset{\sim}{b}_i$, has a gamma distribution with parameters α_i^*, and β_i^*, $i = 1, 2, \ldots, c$, (3.9)

where

$$\alpha^* = (n - p + 2\alpha) / 2,$$

$$\alpha_i^* = (m_i + 2\alpha_i) / 2, \quad i = 1, 2, \ldots, c,$$

$$\beta^* = (2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} + Q)/2,$$

$$\beta_i^* = (2\beta_i + Q_i) / 2,$$

$$Q = (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}}), \text{ and}$$

$$Q_i = \underset{\sim}{b}_i'\underset{\sim}{b}_i, \quad i = 1, 2, \ldots c.$$

To obtain the marginal posterior density $\Pi_1(\underset{\sim}{b})$, the joint posterior density of τ, $\underset{\sim}{\rho}$, and $\underset{\sim}{b}$ in (3.8) is integrated with respect to τ and $\underset{\sim}{\rho}$. This gives

$$\Pi_1(\underset{\sim}{b}) \propto (2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} + Q)^{-(n-p+2\alpha)/2} \times \prod_{i=1}^{c} (2\beta_i + Q_i)^{-(m_i + 2\alpha_i)/2} .$$

$$\propto \{1 + (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'A\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}})\}^{-(n-p+2\alpha)/2} \times \prod_{i=1}^{c} \{(1 + \underset{\sim}{b}_i'A_i\,\underset{\sim}{b}_i)^{-(m_i + 2\alpha_i)/2}\}, \qquad (3.10)$$

where

$$A = \{ \frac{1}{2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}} \} U'R\,U ,$$

$$A_i = (1 / 2\beta_i)\ I_{m_i} \quad , \ i = 1, 2, \ldots , c.$$

Such densities are known as multiple t densities or poly-t densities and are quite difficult to handle.

Marginal Posterior Distributions

For proper Bayesian inference on the variance components one needs their marginal distributions. The joint posterior density of τ, and $\underset{\sim}{\rho}$, namely $\Pi_1(\tau,\underset{\sim}{\rho})$, can be obtained by integrating (3.7) with respect to $\underset{\sim}{b}$. Thus

$$\Pi_1(\tau, \underset{\sim}{\rho}) \propto \int_{R^m} \tau^{\{(n - p + 2\alpha)/2 - 1\}}$$

$$\times \exp \{(-\tau/2)\{ 2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}$$

$$+ (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}})\}\}$$

$$\times \prod_{i=1}^{c} \tau_i^{\{(m_i + 2\,\alpha_i)/2 - 1\}}$$

$$\times \exp \{ (-\tau_i/2)\ (2\beta_i + \underset{\sim}{b}_i'\underset{\sim}{b}_i)\} \ d\underset{\sim}{b} \ .$$

$$\propto A \int_{R^m} \exp \{(-1/2)\ \{(\underset{\sim}{b} - \hat{\underset{\sim}{b}})\ B_1(\underset{\sim}{b} - \hat{\underset{\sim}{b}}) + \underset{\sim}{b}'B_2\,\underset{\sim}{b}\}\}\ d\underset{\sim}{b}$$

where

$$A = \tau^{\{(n - p + 2\alpha)/2 - 1\}}$$

$$\times \exp\{(-\tau/2)\,(2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}})\}$$

$$\times \prod_{i=1}^{c} \tau_i^{\{(m_i + 2\,\alpha_i)/2 - 1\}} \,.\, \exp\,(-\tau_i\beta_i) \quad ,$$

$$B_1 = \tau\,U'R\,U,$$

$$B_2 = \begin{pmatrix} \tau_1 I_{m_1} & \phi & \cdots & \phi \\ \phi & \tau_2 I_{m_2} & \cdots & \phi \\ \cdots & \cdots & \cdots & \cdots \\ \phi & \phi & \cdots & \tau_c I_{m_c} \end{pmatrix}$$

Combining the two quadratic forms in the exponent of the above integrand it follows that

$$\Pi_1(\tau, \underset{\sim}{\rho}) \propto A \int_{R^m} \exp\{(-1/2)\,(\underset{\sim}{b} - \tilde{\underset{\sim}{b}})\,B\,(\underset{\sim}{b} - \tilde{\underset{\sim}{b}})\}$$

$$\times \exp\{(-1/2)\,\hat{\underset{\sim}{b}}'B_1(B^{-1} - B_1^{-1})\,B_1\,\hat{\underset{\sim}{b}}\} \quad d\underset{\sim}{b}$$

$$\propto A \quad |B|^{-\frac{1}{2}} \exp\{(-1/2)\,\hat{\underset{\sim}{b}}\,B_1(B^{-1} - B_1^{-1})\,B_1\,\hat{\underset{\sim}{b}}\} \;, \qquad (3.11)$$

where $B = B_1 + B_2$ and $\tilde{\underset{\sim}{b}} = B^{-1}\,B_1\,\hat{\underset{\sim}{b}}$.

The above joint posterior density of τ and $\underset{\sim}{\rho}$ in (3.11) is a very complicated expression and is analytically intractable. For instance, it seems impossible to integrate it with respect to the precision components in order to obtain the marginal posterior densities of the variance components which are essential for purposes of posterior inference.

Another way of finding the marginal densities of the variance components is to combine the marginal density of $\underset{\sim}{b}$ in (3.9) with the conditional densities of the variance components in (3.8) and then removing $\underset{\sim}{b}$ by integration. But, here again, integration is a formidable problem.

CHAPTER IV

APPROXIMATIONS

The posterior distribution of $\underset{\sim}{b}$ given in (3.10) is the key factor to obtain the conditional posterior distributions in (3.9) as well as to determine the true posterior marginal distributions of the variance components. This distribution is a multiple t distribution and is very difficult to handle. One way of solving the problem is to find an approximation to the multiple t density in terms of some simple expression.

Approximation to the Posterior Distribution of $\underset{\sim}{b}$

Since a multivariate t distribution can be approximated by a multivariate normal distribution having the same first two moments as the multivariate t distribution,

$$\{1 + (1/k)\ (\underset{\sim}{x} - \underset{\sim}{\mu})'\ A\ (\underset{\sim}{x} - \underset{\sim}{\mu})\}^{-(n + k)/2}$$

can be approximated by

$$\exp\ \{(-1/2)\ ((k - 2)/k)\ (\underset{\sim}{x} - \underset{\sim}{\mu})'\ A\ (\underset{\sim}{x} - \underset{\sim}{\mu})\}$$

for any n-vector x and for any non-negative definite matrix A.

Using this approximation in (3.10) one can show that the posterior density of $\underset{\sim}{b}$ may be approximated by

$$\Pi_1(b) \;\propto\; \exp\{(-1/2)\,(\,(\underset{\sim}{b} - \underset{\sim}{\hat{b}})'A_1(\underset{\sim}{b} - \underset{\sim}{\hat{b}}) + \underset{\sim}{b}\, A_2 \underset{\sim}{b}\,)\}$$

where

$$A_1 \;=\; \{(n-p+2\alpha-m-2)\;/\;(2\beta+\underset{\sim}{y}'R\underset{\sim}{y}-\underset{\sim}{\hat{b}}U'RU\underset{\sim}{\hat{b}})\}\;U'RU\quad,$$

$$A_2 \;=\; \begin{pmatrix} ((\alpha_1-1)/\beta_1)\; I_{m_1} & \phi & \phi \\ \phi & \cdots & \phi \\ \phi & \phi & ((\alpha_c-1)/\beta_c)\; I_{m_c} \end{pmatrix}$$

Combining the two quadratic forms in the exponent it follows that

$$\Pi_1(\underset{\sim}{b}) \;\propto\; \exp\{(-1/2)\,(\underset{\sim}{b} - \underset{\sim}{b}^*)'A^*(\underset{\sim}{b} - \underset{\sim}{b}^*)\} \tag{4.1}$$

where

$$A^* = A_1 + A_2 \quad\text{and}\quad \underset{\sim}{b}^* = (A^*)^{-1}\, A_1\, \underset{\sim}{\hat{b}}\quad .$$

Thus, the posterior distribution of $\underset{\sim}{b}$ is approximately normal with mean vector $\underset{\sim}{b}^*$ and precision matrix A^*.

The mean of this posterior distribution of $\underset{\sim}{b}$, namely $\underset{\sim}{b}^*$, is a good choice to be used as the conditioning value of $\underset{\sim}{b}$ in (3.9). The results of (3.9) can now be used to make inferences about the parameters in (3.1). For instance, if conditional point estimates for the parameters are needed the mean of these conditional posterior distributions can be taken as the estimates. They are

$$\underset{\sim}{\hat{\theta}} = (X'X)^{-1}\, X'\,(\underset{\sim}{y} - U\,\underset{\sim}{b}^*),$$

$$\begin{aligned} \hat{\sigma}^2 = \;& \{1/(n + 2\alpha - p - 2)\}\,\{2\beta + \underset{\sim}{y}'R\underset{\sim}{y} - \underset{\sim}{\hat{b}}'\,U'R\,U\,\underset{\sim}{\hat{b}} \\ & + (\underset{\sim}{b}^* - \underset{\sim}{\hat{b}})\,U'R\,U\,(\underset{\sim}{b}^* - \underset{\sim}{\hat{b}})\,\}, \end{aligned} \tag{4.2}$$

$$\hat{\sigma}_i^2 = (2\beta_i + \underset{\sim}{b}_i^{*\prime}\underset{\sim}{b}_i^{*}) / (m_i + 2\alpha_i - 2) , \quad i = 1, 2, \ldots , c. \tag{4.2}$$

It is now proposed to make use of (4.1) in (3.9) to derive the posterior distribution of the variance components. For this purpose, one needs the distribution of quadratic forms in normal variables. This is derived below.

Distribution of General Non-negative Quadratic Forms in Normal Variables

Let $\underset{\sim}{z}$ be a random n-vector having multivariate normal distribution with mean vector $\underset{\sim}{\mu}$ and a positive definite dispersion matrix D, or symbolically let $\underset{\sim}{z} \sim N_n(\underset{\sim}{\mu}, D)$. Let $Q(\underset{\sim}{z}) = (\underset{\sim}{z} - \underset{\sim}{z}_0)'M(\underset{\sim}{z} - \underset{\sim}{z}_0)$ where $\underset{\sim}{z}_0$ is a known n-vector and M is a known non-negative definite matrix. It is clear that the distribution of $Q(\underset{\sim}{z})$ is the same as that of $(\underset{\sim}{z}^* - \underset{\sim}{z}_1)'M(\underset{\sim}{z}^* - \underset{\sim}{z}_1)$ where $\underset{\sim}{z}_1 = \underset{\sim}{z}_0 - \underset{\sim}{\mu}$ and $\underset{\sim}{z}^* \sim N_n(\underset{\sim}{0}, D)$.

Since D is positive definite there exists a nonsingular lower triangular matrix L such that $D = L L'$. Further, since $L'M L$ is symmetric, there exists an orthogonal matrix P such that $P'L'M L P = \Lambda$, the diagonal matrix of the eigen values of $L'M L$. Hence, using the transformation $\underset{\sim}{z}^* = L P \underset{\sim}{W}$, one can show that the distribution of $Q(\underset{\sim}{z})$ is the same as that of

$$\sum_{i=1}^{n} \lambda_i (W_i - w_i)^2$$

where

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots , \lambda_n) ,$$

$$\underset{\sim}{w}' = (w_1, w_2, \ldots , w_n) ,$$

$$\underset{\sim}{W}' = (W_1, W_2, \ldots, W_n),$$

$$\underset{\sim}{w} = P'L^{-1}\underset{\sim}{z}_1 \text{ , and } n = \# \text{ of rows of } M.$$

It is easily seen that $W_1, W_2, \ldots, W_n$ are independent standard normal variables.

If n' = rank of $M < n$, then assuming the last $n-n'$ of the λ's to be zero one can show that the distribution of $Q(\underset{\sim}{z})$ is the same as that of

$$\sum_{i=1}^{n'} \lambda_i (W_i - w_i)^2 .$$

Ruben (1960, 1962) has shown that, for a given Λ, n, and $\underset{\sim}{w}$, the distribution of $Q(\underset{\sim}{W})$ can be expressed as

$$\begin{aligned} F_n(q \mid \Lambda, \underset{\sim}{w}) &= \Pr.(Q(\underset{\sim}{z}) \leq q) \\ &= \Pr. \{ \sum_{j=1}^{n} \lambda_j (W_j - w_j)^2 \leq q \} \\ &= \sum_{j=0}^{\infty} e_j \Pr.(\chi^2_{n+2j} < q/c) \end{aligned} \tag{4.3}$$

where

$$e_0 = \exp \{ (-1/2) \sum_{j=1}^{n} w_j^2 \} \prod_{j=1}^{n} (c/\lambda_j)^{\frac{1}{2}} ,$$

$$e_r = (1/2r) \sum_{j=0}^{r-1} G_{r-j}\, e_j \quad (r \geq 1) ,$$

$$G_r = \sum_{j=1}^{n} (1 - c/\lambda_j)^r + rc \sum_{j=1}^{n} (w_j^2/\lambda_j)(1 - c/\lambda_j)^{r-1} \quad (r \geq 1)$$

and c is an arbitrary positive number.

Ruben has also shown that the series on the right of (4.1) is uniformly convergent over any finite interval of q. He also gives the bound for error in the above series at the $N^{\underline{th}}$ term. Further, it is obvious that the distribution in (4.3) can be made to be an infinite mixture of chi-square densities provided c is chosen to be less than minimum of $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, to insure each $e_i > 0$.

Approximation to the Posterior Distribution of σ^2

From (3.8), τ, given $\underset{\sim}{b}$, has a gamma distribution with parameters

$$\alpha^* = (n - p + 2\alpha)/2 \text{ , and}$$

$$\beta^* = (1/2)\,(2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} + Q)$$

where

$$Q = (\underset{\sim}{b} - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b} - \hat{\underset{\sim}{b}}).$$

Because of (4.1), Q has the distribution in (4.3) with $\underset{\sim}{b}^*$, $\hat{\underset{\sim}{b}}$, A^{*-1} and U'R U in the places of $\underset{\sim}{\mu}$, $\underset{\sim}{z}_0$, D, and M respectively. s = n = rank of U'R U.

Hence, the posterior marginal density of τ is given by

$$\Pi_1(\tau) = \int_0^\infty \frac{\beta^{*\alpha^*} \exp(-\beta^*\tau)\,.\tau^{\alpha^*-1}}{\Gamma(\alpha^*)} \, .$$

$$\times \sum_{j=0}^{\infty} e_j \left\{ \frac{d}{dq} \Pr.(\chi^2_{s+2j} < q/c)\right\} dq \, .$$

After a little simplification this expression can be rewritten as

$$\sum_{j=0}^{\infty} e_j \int_0^{\infty} \frac{\{(1/2)(d+q)\}^k e^{-(1/2)(d+q)\tau} \tau^{(k-1)}}{\Gamma(k).2^{(s+2j)/2} \Gamma\{(s+2j)/2\}} e^{(-q/2c)}$$

$$\times \; (q/c)^{(s/2)+j-1} (1/c) \; dq.$$

where $k = (n - p + 2\alpha)/2$, $d = 2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}$, and c is an arbitrary positive number.

Let k be an integer. To insure this one needs to make some slight adjustments in the value of α. But the magnitude of such changes never needs to be bigger than 0.5. On this assumption, the above expression can be rewritten as

$$\Pi(\tau) \propto \sum_{j=0}^{\infty} \frac{e_j \tau^{(k-1)} \sum_{r=0}^{k} \binom{k}{r} d^{(k-r)}}{\Gamma(k)\, 2^{(s/2)+j}\, \Gamma(s/2+j)\, c\, 2^k} \,.$$

$$\int_0^{\infty} \frac{q^r e^{(-1/2)(d+q)\tau - q/(2c)} q^{(s/2)+j-1}}{c^{(s/2)+j-1}} dq \,.$$

After performing the integration and making the transformation $\tau = 1/\sigma^2$ and simplifying the resulting expression the marginal posterior density of σ^2 is seen to be

$$\Pi_1(\sigma^2) = k\, e^{(-d/2\sigma^2)} (d/2)^k (1/\sigma^2)^{(k+1)} \cdot$$

$$\sum_{j=0}^{\infty} \frac{e_j}{\Gamma(s/2 + j)\,(1 + c/\sigma^2)^{(s/2+j)}} \qquad (4.4)$$

$$\sum_{r=0}^{k} \frac{\Gamma(s/2 + r + j)}{\Gamma(r+1)\,\Gamma(k - r + 1)\,\{(d/2)(1/c + 1/\sigma^2)\}^r}$$

Approximation to the Posterior Distribution of σ_i^2

The procedure of deriving an approximation to the marginal posterior distribution of σ_i^2 is very similar to that of σ^2. From (3.7) the posterior distribution of τ_i, conditional on $\underset{\sim}{b}_i$, is a gamma with parameters

$$\alpha_i^* = (m_i + 2\alpha_i) / 2$$

$$\beta_i^* = (1/2)\,(2\beta_i + Q_i)$$

where $Q_i = \underset{\sim}{b}_i' \underset{\sim}{b}_i$. As before, because of (4.1), Q_i has the distribution in (4.3) with $\underset{\sim}{b}_i^*$, $\underset{\sim}{0}$, A_i^{*-1}, and I_{m_i} in the places of $\underset{\sim}{\mu}$, $\underset{\sim}{z}_0$, D, and M respectively. Let s_i be the rank of $I_{m_i} = m_i$.

Hence the marginal posterior distribution of the i^{th} precision component τ_i is given by

$$\Pi_1(\tau_i) = \{1/\Gamma(\alpha_i^*)\} \int_0^{\infty} \beta_i^{*\,\alpha_i^*}\, e^{-\beta_i^* \tau_i}\, \tau_i^{\alpha_i^* - 1} \times$$

$$\sum_{j=0}^{\infty} e_{ij} \left\{ \frac{d}{dq_i} \Pr. (\chi^2_{s_i+2j} < q_i/c_i) \right\} dq_i \; .$$

After a little simplification this expression can be rewritten as

$$\sum_{j=0}^{\infty} e_{ij} \int_0^{\infty} \frac{(1/2)(d_i + q_i)^{k_i} \, e^{-(1/2)(d_i + q_i)\tau_i} \, \tau_i^{(k_i - 1)}}{\Gamma(k_i) \; 2^{(s_i + 2j)/2} \; \Gamma\{(s_i + 2j)/2\}}$$

$$\times \; (q_i/c_i)^{(s_i/2 + j - 1)} \; (1/c_i) \; dq_i$$

where $k_i = (m_i + 2\alpha_i)/2$, $d_i = 2\beta_i$ and c_i is an arbitrary positive number.

As before, let k_i be assumed to be an integer. One needs to make some slight changes in the value of α_i to insure k_i to be an integer. But the magnitude of such changes never needs to be bigger than 0.5. On this assumption, the marginal posterior density of τ_i can be written as

$$\Pi_1(\tau_i) = \sum_{j=0}^{\infty} \frac{e_{ij} \, \tau_i^{k_i - 1} \sum_{r=0}^{k_i} \binom{k_i}{r} d_i^{k_i - r}}{\Gamma(k_i) \; 2^{s_i/2 + j} \; \Gamma(s_i/2 + j) \; c_i \; 2^{k_i}}$$

$$\times \int_0^{\infty} \frac{q_i^{r} \, e^{(-1/2)(d_i + q_i)\tau_i \, - \, (q_i/2c_i)} \, q_i^{s_i/2 + j - 1}}{c_i^{(s_i/2 + j - 1)}} \, dq_i$$

After performing the integration, making the transformation $\tau_i = 1/\sigma_i^2$, and simplifying the resulting expression the marginal posterior density of σ_i^2 is obtained as

$$\Pi_1(\sigma_i^2) = k_i \; e^{(-d_i/2\sigma_i^2)} \; (d_i/2)^{k_i} \; (1/\sigma_i^2)^{(k_i + 1)}$$

$$\times \sum_{j=0}^{\infty} \frac{e_{ij}}{\Gamma(s_i/2 + j) \; (1 + c_i/\sigma_i^2)^{(s_i/2 + j)}} \tag{4.6}$$

$$\times \sum_{r=0}^{k_i} \frac{\Gamma(s_i/2 + r + j)}{\Gamma(r + 1) \; \Gamma(k_i - r + 1) \; \{(d_i/2)(1/c_i + 1/\sigma_i^2)\}^r}$$

Moments of the Posterior Distribution of the Variance Components

The moments of the posterior marginal distributions of σ^2 and σ_i^2 as given by (4.4) and (4.5) are hard to obtain directly. But, they can be obtained relatively easily by noting how (4.4) and (4.5) were derived. For instance, the moments of σ^2 (especially the first two) can be obtained by noting that

1. $\sigma^2 | b$ has a gamma distribution, and
2. $\underset{\sim}{b}$ has a multivariate normal distribution

and using the formulas

1. $E(\sigma^2) = E_b\{E(\sigma^2 | b)\}$, and
2. $\mathrm{Var}(\sigma^2) = \mathrm{Var}_b \{E(\sigma^2 | b)\} + E_b \{\mathrm{Var}(\sigma^2 | b)\}$.

They are

$$\mathrm{Mean}(\sigma^2) = \{1 / (n - p + 2\alpha - 2)\} \{2\beta + \underset{\sim}{y}'R\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}} \tag{4.7}$$

$$+ \mathrm{Trace}(U'R\,U\,A^{*\,-1}) + (\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})'U'R\,U\,(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})\}$$

$$\mathrm{var}(\sigma^2) = \{2 / ((n-p+2\alpha-2)^2(n-p+2\alpha-4))\}\{2\beta + \underset{\sim}{y}'R\,\underset{\sim}{y} - \hat{\underset{\sim}{b}}'U'R\,U\,\hat{\underset{\sim}{b}}$$

$$+ \mathrm{Trace}(U'R\,U\,A^{*-1}) + (\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})'U'\,R\,U\,(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})\}^2$$

$$+ \{1 / (n-p+2\alpha-2)(n-p+2\alpha-4)\}\{\,2\,\mathrm{Trace}(U'R\,U\,A^{*-1})^2$$

$$+ 4(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})'U'R\,U\,A^{*-1}U'R\,U\,(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})\}$$

In a similar manner the first two moments of σ_i^2 are seen to be

$$\mathrm{Mean}\ (\sigma_i^2) = \{1 / (m_i+2\alpha_i-2)\}\ \{2\beta_i + \mathrm{tr}(A_i^{*-1}) + \underset{\sim}{b}_i^{*}{}'\underset{\sim}{b}_i^{*}\}$$

$$\mathrm{Var}(\sigma_i^2) = \{2 / ((m_i + 2\alpha_i - 2)^2\,(m_i + 2\alpha_i - 4)\}$$

$$\times \{2\beta_i + \mathrm{Trace}(A_i^{*-1}) + \underset{\sim}{b}_i^{*}{}'\underset{\sim}{b}_i^{*}\}^2$$

$$+ \{1 / ((m_i + 2\alpha_i - 2)(m_i + 2\alpha_i - 4)\}$$

$$\times \{2\,\mathrm{Trace}(A_i^{*-1})^2 + 4\,\underset{\sim}{b}_i^{*}{}'\,A_i^{*-1}\underset{\sim}{b}_i^{*}\}$$

In a similar manner all the higher moments of the variance components can be obtained. These expressions could be easily evaluated for any data set using the matrix procedure in SAS..

Example

Now, a balanced one way random model design is considered to illustrate the preceding results. The linear model describing such a design is

$$y_{ij} = \mu + b_i + e_{ij};\quad i = 1, 2, \ldots, m;\quad j = 1, 2, \ldots, t,$$

where b_i follows a normal distribution with mean 0 and variance σ_1^2, and e_{ij} follows a normal distribution with 0 mean and variance σ^2 for all i and j. This is precisely the model defined in (3.2) with $c = 1$, $m_1 = m$, $p = 1$, and $n = mt$. Also,

$$\underset{\sim}{y} = (y_{11}, \ldots, y_{1t}, y_{21}, \ldots, y_{2t}, \ldots, y_{mt})',$$

$$X = j_n, \text{ a } n \times 1 \text{ matrix of ones},$$

$$\underset{\sim}{\theta} = (\mu),$$

$$U = \text{diagonal}(j_t, j_t, \ldots, j_t), \text{ of order } n \times m,$$

$$\underset{\sim}{b} = (b_1, b_2, \ldots, b_m)'.$$

For this particular case, it can be shown that

$$R = I_n - (1/n)J_n, \; J_n \text{ being a } n \times n \text{ matrix of ones},$$

$$U'RU = t\{I_m - (1/m)J_m\},$$

$$(U'RU)^- = (1/t)\{I_m - (1/m)J_m\},$$

$$\hat{\underset{\sim}{b}} = (\bar{y}_{1.} - \bar{y}_{..}, \bar{y}_{2.} - \bar{y}_{..}, \ldots, \bar{y}_{m.} - \bar{y}_{..})',$$

$$\hat{\underset{\sim}{b}}'U'RU\hat{\underset{\sim}{b}} = S_1, \text{ the between groups sum of squares in the AOV},$$

$$A_1 = \{(n-m+2\alpha-3)/(2\beta+S_2)\}\{t\, I_m - (t/m)\, J_m\}, \text{ where } S_2 \text{ is}$$

the within groups sum of squares in the AOV,

$$A_2 = \{(\alpha_1-1)/\beta_1\}\, I_m,$$

$$A^* = A_1 + A_2 = (at + b)\, I_m - (at/m)\, J_m, \text{ where}$$

$$a = (n - m + 2\alpha - 3)/(2\beta + S_2) \text{ and } b = (\alpha_1 - 1)/\beta_1.$$

Using the preceding results, it is straight forward to show that

$$(A^*)^{-1} = \{1/(at+b)\}I_m + \{at/(bm(at+b))\}\ J_m,$$

$$\underset{\sim}{b}^* = \{at/(at+b)\}\ (\bar{y}_{1.} - \bar{y}_{..}, \ldots, \bar{y}_{m.} - \bar{y}_{..}),$$

$$U'RU(A^{*-1}) = \{t/(at+b)\}\{I_m - (1/m)\ J_m\},$$

$$\underset{\sim}{b}^{*\prime}\underset{\sim}{b}^* = \{a^2t/(at+b)^2\}\ S_1, \text{ and}$$

$$(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}})'U'RU(\underset{\sim}{b}^* - \hat{\underset{\sim}{b}}) = \{b^2/(at+b)^2\}S_1 .$$

Hence, the conditional Bayes estimates of the variance components, as defined in (4.2), are

$$\hat{\sigma}^2 = \{1/(n + 2\alpha - 3)\}\{2\beta + S_2 + b^2S_1/(at + b)^2\}, \text{ and}$$

$$\hat{\sigma}_1^2 = \{1/(m + 2\alpha_1 - 2)\}\{\ 2\beta_1 + a^2tS_1/(at + b\ )^2\}.$$

Similarly, the Bayes estimates of the variance components based on the approximate distributions, as defined in (4.7), are the following.

$$\hat{\sigma}^2 = \{(1/(n + 2\alpha - 3)\}\{\ 2\beta + S_2 + t(m - 1)/(at + b) + b^2S_1/(at + b)^2\},$$

$$\hat{\sigma}_1^2 = \{(1/(m + 2\alpha_1 - 2)\}\{\ 2\beta_1 + (at + mb)/b(at + b) + a^2tS_1/(at + b)^2\}.$$

CHAPTER V

NUMERICAL STUDY

In this Chapter the results of a numerical study, carried out to examine the closeness of the approximate distributions of the variance components to their true distributions, is discussed. A set of data for a one-way random model, generated by Box and Tiao (1973), is used for this purpose. For the generated data the true value of the within groups variance component σ^2 is 16 and the true value of the between variance component σ_1^2 is 4.

Many different sets of values for the prior parameters α, β and α_1, β_1 were considered to examine the effects of the prior parameters on the posterior parameters and also on the closeness of the approximations. For each set of values of the prior parameters the posterior means and variances of the variance components were calculated, both for the true posterior distributions and also for their approximations. The eigen values and other constants, needed to evaluate the approximate distributions (4.4) and (4.5) and their moments, were evaluated using the MATRIX procedure in SAS (Statistical Analysis System). The true posterior distributions and their moments were evaluated using numerical integration techniques. These were done using a Fortran (Watfiv) program. To evaluate the approximate distributions (4.4) and (4.5) only 20 terms in the mixtures were taken. The contribution from the remaining terms was found to be negligible. The constant c was arbitrarily selected to be some positive number less than the smallest of the eigen values in order to make (4.4) and (4.5) mixtures. The computations were done on the IBM 370/168 computer at the Oklahoma State University.

Table I gives the different sets of values considered for the prior parameters along with the means and variances of the true and approximate posterior marginal distributions of the variance components. The graphs of the true posterior distributions of the variance components and their approximations, for some selected values of the prior parameters, are given in Figures 1-6.

The conclusions, indicated by the results of the numerical study, are summarized in the following paragraphs.

Generally, in all the cases considered, the approximations were close to the true distributions. The closeness of the approximations increase with α. Even for values of the α-parameter as low as 8 the approximations are quite close the true ones. It is interesting to note that these are true whatever the values of the β-parameter may be. That this must be true is obvious from the fact that as the value of the α-parameter increase the degrees of freedom of the multivariate t type factors in (3.9) also increase bringing these factors very close to the corresponding normal type expressions. This insures the accuracy of the approximations. But a large value of the α-parameter implies a small prior standard deviation. This is so because for an inverse gamma distribution standard deviation = β / ($(\alpha-1)\sqrt{(\alpha-2)}$) . This, in turn, implies a strongly informative prior for the variance components. Thus, when precise prior information is available the approximations will work very efficiently.

Secondly, when the value of the α-parameter is large and the prior means of the variance components are fairly close to the true values of the variance components, the posterior means are also quite close to the true values. In other words, in such situations, in addition to the true posterior distributions and their approximations being quite close to each other, the two distributions are also centered very close to the true value of the variance components. Further, in such situations the posterior variances are also small.

Thirdly, when the α-parameter is large, changes in the values of the α-parameter do not significantly affect the posterior distribution of the within variance component as compared to the between variance component. In other words, informative priors influence the between variance component relatively more than the within variance component. Therefore, unless one is quite sure of what he is doing, the value of the α-parameter for the between variance component should not be taken to be large even though it does not matter much for the within variance component. The reasons are not difficult to seek. In the first part of the multiple t density in (3.9), the degrees of freedom of the multivariate t type expression, which introduces the within variance component, is $(n - p + 2\alpha)/2$ which is always reasonably large for any data set, irrespective of the value of α. Hence, variations in the α-parameter do not affect the value of that expression very much. Whereas, in the second part which introduces the other variance components, the degrees of freedom of the multivariate t type factor is just α_i. Hence, variations in the value of the α-parameter for the between variance component significantly affect the posterior distribution of the between variance component.

Fourthly, when there is little prior information, it seems reasonable to take the value of the α-parameter to be close to 2 so that prior variance is large. It is interesting to note that, even in such cases, the approximations are good. But, whether the posterior distributions of the variance components are located at their true values or not depends on the β-parameter. From Table I, one can see that when $\alpha = 2$ and β varies from 2 to 20 the posterior mean of the within variance component goes from 13.06 to 14.82 only. Note that the true value is 16.0. One can, therefore, conclude that the posterior distribution of the within variance component is not very sensitive to changes in β. But, the case is different for the between variance component. When the α-parameter is 2 the posterior mean and variance heavily depend on the β-parameter. So a fairly good estimate of the true value is needed for the proper choice of the value of the β-parameter.

It should be remembered that the above inferences are based upon a limited study of only one data set. More extensive numerical studies need be carried out before one can formulate general rules on the closeness of the approximations and also on the choice of the prior parameters.

CHAPTER VI

SUMMARY

The main objective of this thesis is to develop a Bayesian methodology with which inferences about the variance components in general mixed linear models can be made. This objective is fulfilled by studying the marginal posterior distributions of the variance components since the joint posterior density of the variance components, which is quite easy to obtain, is a complicated function of the variance components and is analytically intractable.

The results obtained are quite general and can be used with any design, balanced or otherwise, coming under mixed linear models.

Results and Conclusions

Assuming all the parameters in the model to be apriori independent and employing gamma priors for the precision components and uniform priors for the fixed effects, it is shown that the posterior distribution of the random effects is a multiple t distribution and that of the variance components, conditional on the random effects, are independent inverse gamma distributions.

Since the exact marginal posterior distributions of the variance components is very difficult to obtain due to integration problems, certain approximations are considered. Approximating a multivariate t density by a normal multivariate density and employing distributions of quadratic forms in normal variables, where the idempotency condition is not fulfilled, approximations to the distributions of the variance components are obtained in infinite series form. Further, closed expressions for the moments of the approximate distributions of the variance components are also derived.

The results of a numerical study, carried out using Box's data for a one way random model, indicate that the approximations are good. Generally, the accuracy of the approximations increase with the α-parameter of the prior distribution of the variance components. The approximations are close even for small values of the α-parameter such as 8. The posterior distribution of the within variance component is less sensitive to variations in the β-parameter of the gamma priors as compared to that of the between variance component. When precise prior information is not available, it seems desirable to fix the α-parameter around 2. Generally, in order to fix the β-parameter, a fairly good prior estimate of the true value of the variance components is needed. But, more detailed numerical studies are needed before one can put forward these indications as general rules.

Further Work

There are many areas under mixed linear models needing further research. If moments of multiple t distributions could be obtained in a general way, then the posterior moments of the variance components could be easily obtained with out using any approximations.

The posterior analysis of the fixed effects is another area where much work still remains to be done. Again, multiple t distributions play a central role in such research work.

The results obtained in this thesis could, probably, be generalized to multivariate cases also. Any way, the approach through the conditional posterior distributions, similar to the results of Gharraf (1979), could be easily generalized to cover multivariate cases.

There are many practical situations where the random effects are correlated instead of being independent as assumed in this thesis. There are lots of possibilities in this area for further research.

Another interesting area involving variance components is estimating the heritability parameters of models used in genetic progeny trials.

A SELECTED BIBLIOGRAPHY

1. Box, G. E. P., and G. C. Tiao (1973). Bayesian Inference in Statistical Analysis. Reading, Massachusetts: Addison-Wesley.

2. Broemeling, L. D., and M. K. Gharraf (1979). The Posterior Distribution of the Parameters of Mixed Linear Models. (Unpublished manuscript, Statistics Department, Oklahoma State University, Stillwater, Oklahoma.)

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

4. Corbeil, R. R., and S. R. Searle (1976b). A Comparison of Variance Components Estimators. Biometrics, 32, 779-791.

5. Daniels, H. E. (1939). The Estimation of Components of Variance. Journal of the Royal Statistical Society (Supp.), 6, 186-197.

6. Gharraf, M. K. (1979). A General Solution to Making Inferences about the Parameters of Mixed Linear Models. (Ph.D. dissertation, Oklahoma State University, Stillwater, Oklahoma.)

7. Graybill, F. A., and R. A. Hulquist (1961). Theorems Concerning Eisenhart's Model II. Annals of Mathematical Statistics, 32, 261-269.

8. Graybill, F. A., and A. W. Wortham (1956). A Note on Uniformly Best Unbiased Estimators for Variance Components. Journal of the American Statistical Association, 51, 266-268.

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

10. Harville, D. A. (1977). Maximum likelihood Approaches to Variance Component Estimation and to Related Problems. Journal of the American Statistical Association, 72, 320-340.

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

12. Hultquist, R., and E. M. Atzinger (1972). The Mixed Effects Model and Simultaneous Diagonalization of Symmetric Matrices. Annals of Mathematical Statistics, 43, 2024-2030.

13. Kleffe, J. (1977). Optimal Estimation of Variance Components. Sankhya B, 39, 211-244.

14. Leonard, T. (1975). A Bayesian Approach to the Linear Model with Unequal Variances. Technometrics, 17, 95-102.

15. Lindley, D. V., and A. F. M. Smith (1972). Bayes Estimates for the Linear Model. Journal of the Royal Statistical Society B, 34, 1-18.

16. Patterson, H. D., and R. Thompson (1971). Recovery of Inter-block Information when Block Sizes are Unequal. Biometrika, 58, 445-454.

17. Rao, C. R. (1971a). Estimation of Variance and Covariance Components - MINQUE Theory. Journal of Multivariate Analysis, 1, 257-275.

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

19. Rhode, C. A. (1972). Bayesian Methods in the Analysis of General Linear Model. Metrika, 18, 220-226.

20. Ruben, H. (1960). Probability Content of Regions under Spherical Normal Distributions. Annals of Mathematical Statistics, 31, 598-619.

21. Ruben, H. (1962). Probability Contents of Regions under Spherical Normal Distributions, IV: The Distribution of Homogeneous and Non-homogeneous Quadratic Function of Normal Variables. Annals of Mathematical Statistics, 33, 542-570.

22. Sahai, H. (1974). Some Formal Bayes Estimators of Variance Components in the Balanced Three State Nested Random Effects Models. Communications in Statistics, 3, 233-242.

23. Searle, S. R. (1968). Linear Models. New York: John Wiley and Sons.

24. Searle, S. R. (1978). A Summary of Recently Developed Methods of Estimating Variance Components. Technical Report, BU-338, Biometrics Unit, Cornell University, New York.

25. Seely, J. (1975). An example of an Inquadmissible Analysis of Variance Estimator for a Variance Component. Biometriks, 62, 689-690.

26. Smith, A. F. M. (1973). Bayes Estimates in One-way and Two-way Models. Biometrika, 60, 319-329.

27. Swallow, W. H., and S. R. Searle (1978). Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components. Technometrics, 20, 265-272.

28. Winsor, C. P., and G. L. Clarke (1940). A Statistical Study of Variation in the Catch of Plankton Nets. Sears Foundation Journal of Marine Research, 3, 1-34.

APPENDIX A

TABLE

TABLE I

MEAN AND VARIANCE OF THE TRUE AND THE APPROXIMATE MARGINAL POSTERIOR DISTRIBUTIONS OF THE WITHIN AND BETWEEN VARIANCE COMPONENTS FOR BOX'S DATA

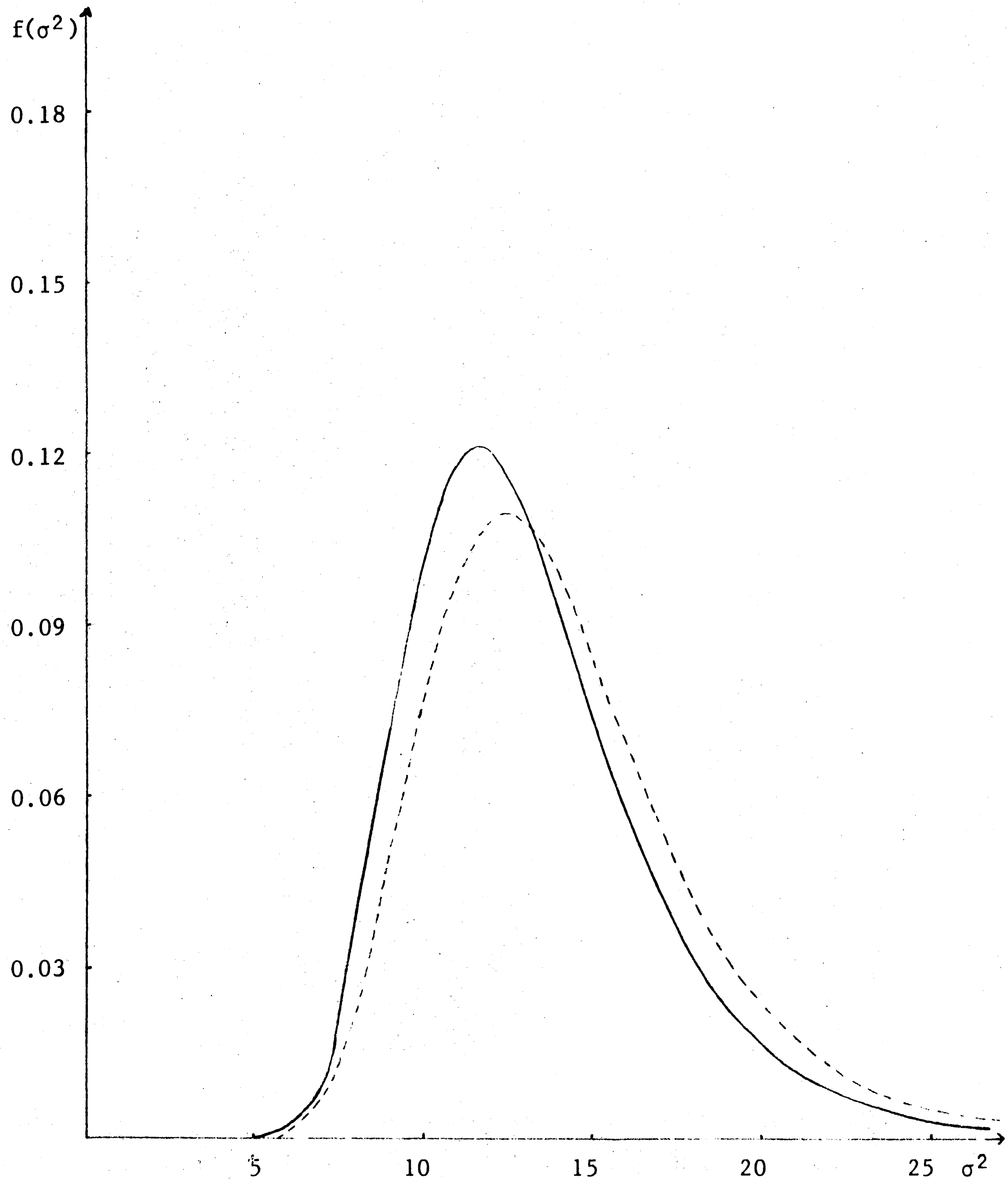
Prior Parameters				Within Variance Component				Between Variance Component			
				Mean		Variance		Mean		Variance	
α	β	α_1	β_1	True	App	True	App	True	App	True	App
2	2	2	2	12.96	13.06	11.72	12.24	1.22	1.62	0.84	1.07
2	5	2	5	13.14	13.56	11.15	13.67	2.40	3.01	1.97	2.66
2	8	2	8	12.49	13.81	12.84	14.22	3.29	3.69	2.63	4.02
2	20	2	20	14.59	14.82	16.16	17.02	7.44	8.97	16.87	24.40
2	8	5	5	13.35	13.39	12.30	12.68	1.11	1.17	0.30	0.30
3	3	3	3	12.21	12.28	9.92	10.13	1.17	1.33	0.56	0.60
3	5	3	5	12.32	12.50	9.63	10.73	1.79	2.05	1.07	1.23
3	10	3	10	12.75	12.97	10.96	11.67	3.07	3.36	2.12	2.61
3	20	3	20	13.56	13.86	12.92	13.58	5.77	6.50	9.37	11.90
3	50	3	50	15.80	16.18	17.86	19.08	12.07	12.23	24.22	26.73
3	3	5	5	12.21	12.26	9.91	10.07	1.12	1.16	0.30	0.30
4	5	4	5	11.59	11.69	8.24	8.74	1.38	1.50	0.56	0.56
4	10	4	10	11.96	12.08	9.23	9.40	2.54	2.77	1.72	2.00
4	20	4	20	12.68	12.85	10.52	10.77	4.64	5.03	5.15	6.08
4	50	4	50	14.76	14.85	14.53	14.81	10.25	10.70	18.29	20.76
5	5	5	5	10.95	11.00	7.23	7.24	1.11	1.17	0.30	0.30
5	10	5	10	11.26	11.34	7.67	7.79	2.07	2.20	0.95	1.03

TABLE I (Continued)

Prior Parameters				Within Variance Component				Between Variance Component			
				Mean		Variance		Mean		Variance	
α	β	α_1	β_1	True	App	True	App	True	App	True	App
5	20	5	20	11.91	12.04	8.70	8.89	3.85	4.08	2.98	3.35
5	50	5	50	13.85	13.98	11.98	12.15	8.76	9.13	13.86	14.37
8	8	8	8	9.48	9.53	4.98	4.65	1.08	1.09	0.16	0.16
10	10	20	10	8.79	8.81	3.55	3.57	0.52	0.52	0.01	0.01
10	40	20	50	10.10	10.14	4.77	4.82	2.51	2.52	0.32	0.32
10	100	20	100	12.86	12.94	7.55	7.87	4.90	4.89	1.10	1.25
10	200	20	200	17.56	17.31	14.62	14.90	9.70	9.74	4.65	4.71
20	100	50	100	8.82	8.84	2.49	2.50	2.01	2.02	0.08	0.06
20	400	50	200	18.14	18.02	10.58	11.26	4.00	4.00	0.43	0.33
32	500	20	80	15.49	15.51	5.52	5.55	3.98	3.96	0.80	0.79

APPENDIX B

FIGURES

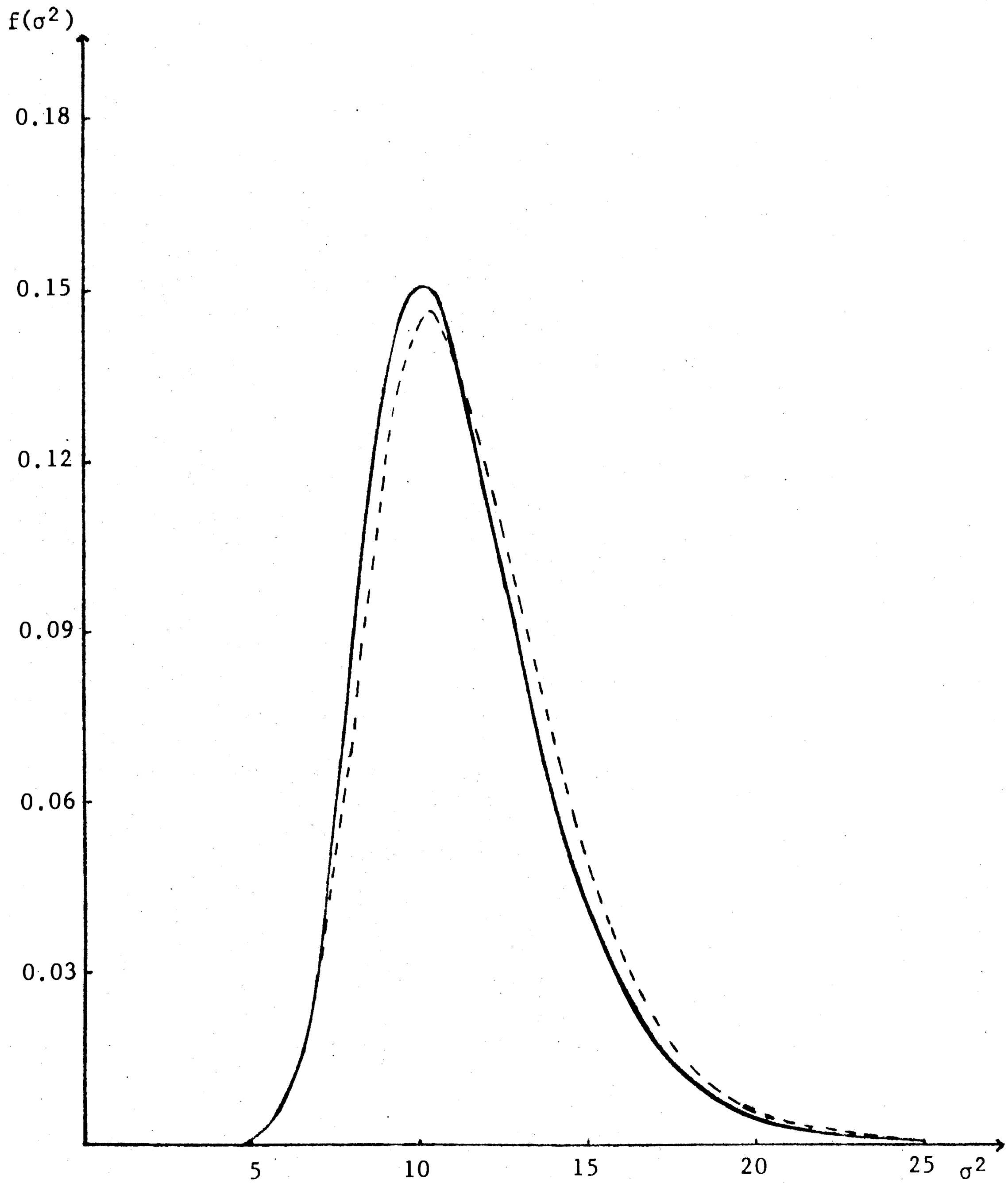


$\alpha = 2,\ \beta = 5,\ \alpha_1 = 2,\ \beta_1 = 5.$

----- Approximation

——— True Distribution

Figure 1. Marginal Posterior Density of the Within Variance Component

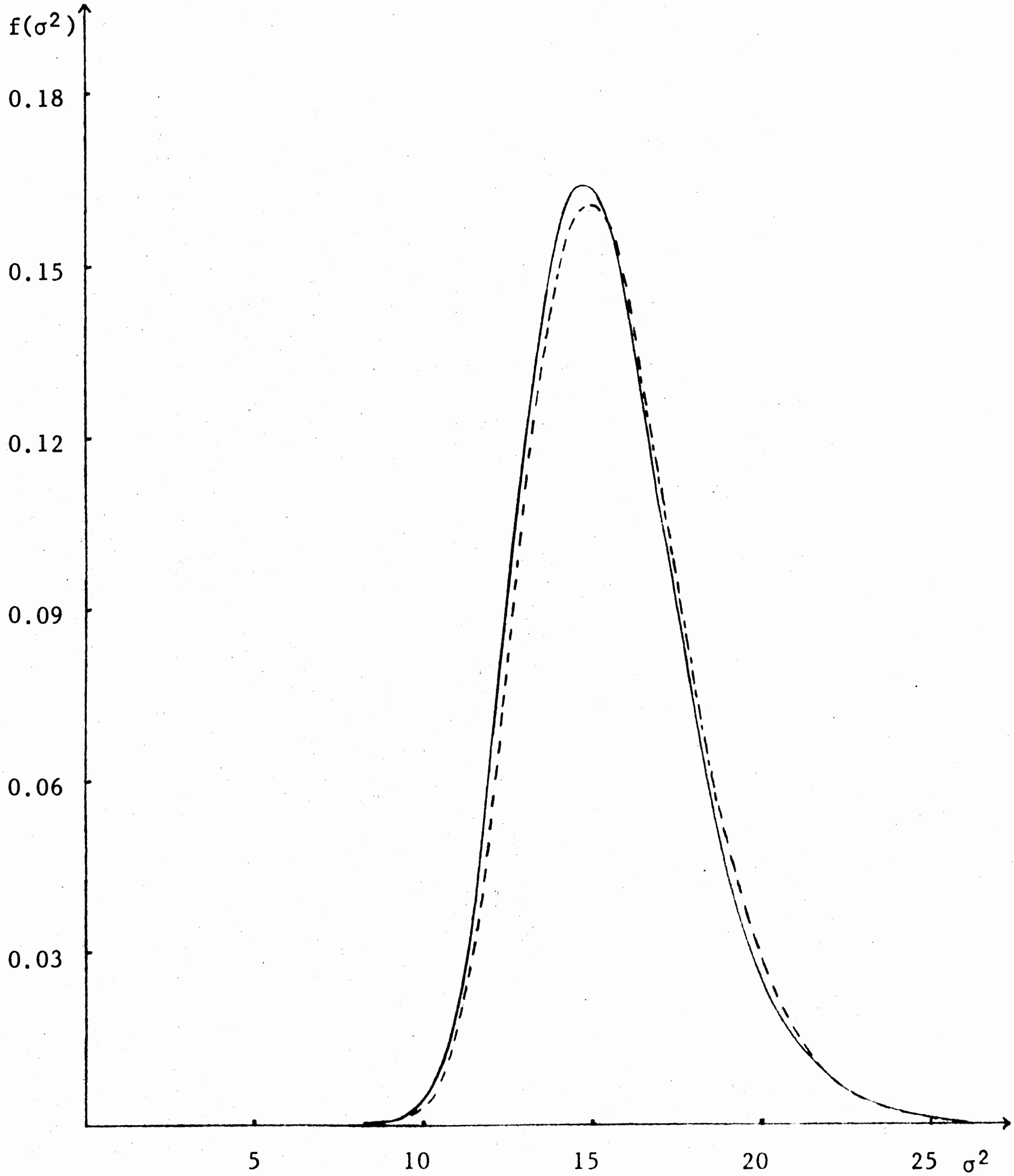


$\alpha = 5,\ \beta = 10,\ \alpha_1 = 5,\ \beta_1 = 10$

----- Approximation

——— True Distribution

Figure 2. Marginal Posterior Density of the Within Variance Component



α = 32, β = 500, α_1 = 20, β_1 = 80

----- Approximation

——— True Distribution

Figure 3. Marginal Posterior Density of the Within Variance Component

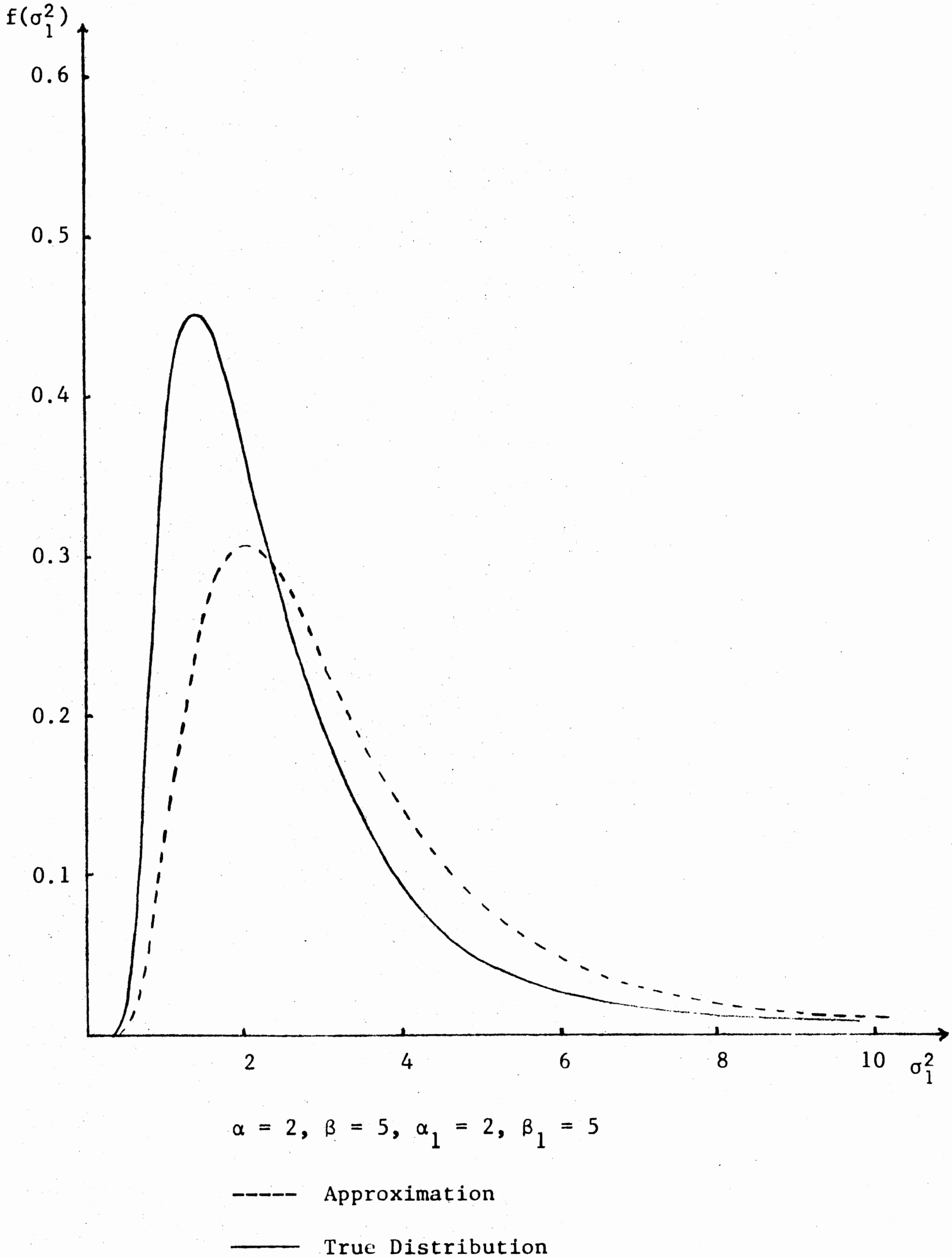
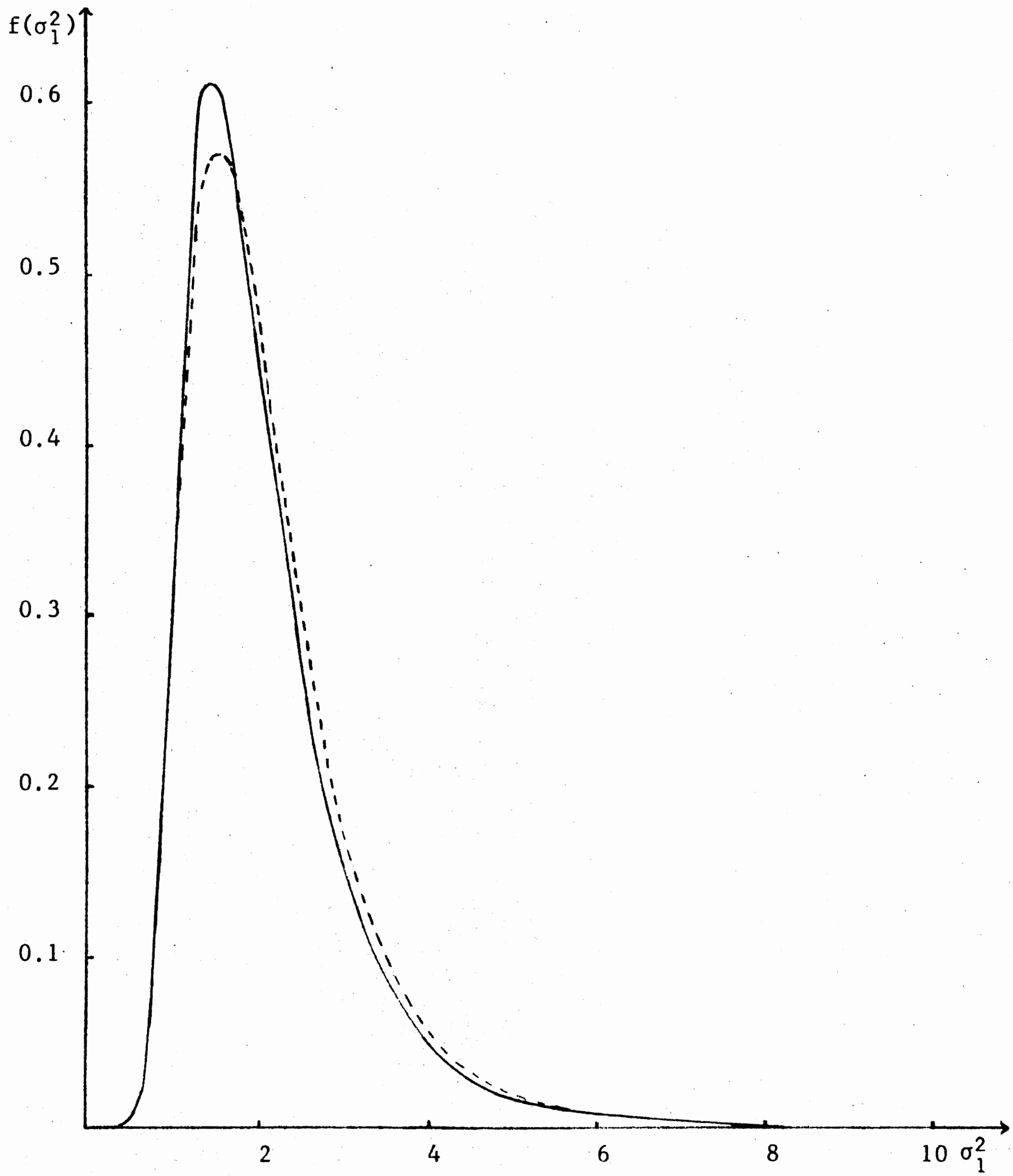


Figure 4. Marginal Posterior Density of the Between Variance Component

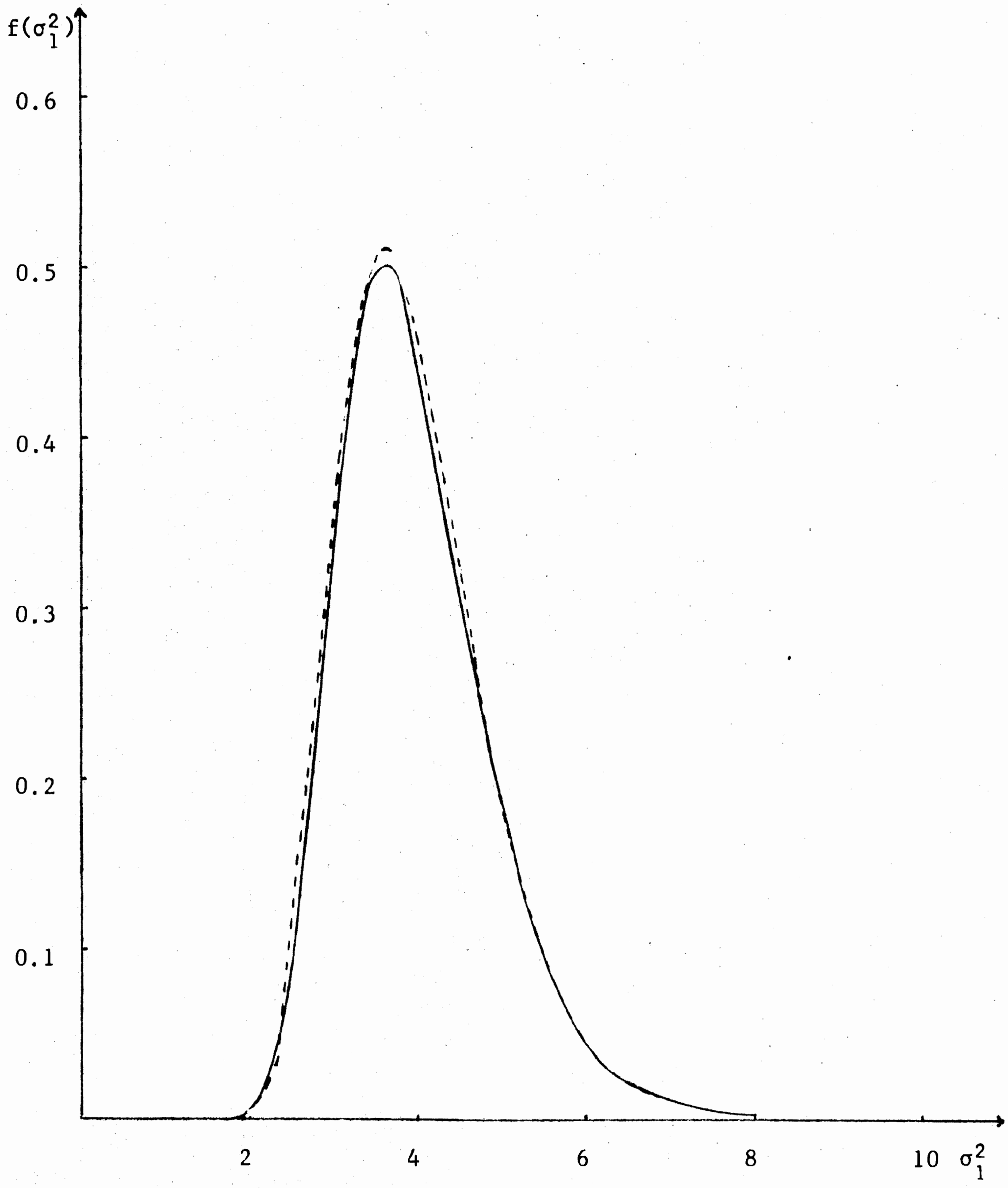


$\alpha = 5,\ \beta = 10,\ \alpha_1 = 5,\ \beta_1 = 10$

----- Approximation

_____ True Distribution

Figure 5. Marginal Posterior Distribution of the Between Variance Component



$\alpha = 32, \beta = 500, \alpha_1 = 20, \beta_1 = 80$

----- Approximation

——— True Distribution

Figure 6. Marginal Posterior Density of the Between Variance Component

VITA [2]

Muthiya Rajagopalan

Candidate for the Degree of

Doctor of Philosophy

Thesis: BAYESIAN INFERENCE FOR THE VARIANCE COMPONENTS IN MIXED LINEAR MODELS

Major Field: Statistics

Biographical:

Personal Data: Born in Mayuram, Tamil Nadu, Republic of India, January 5, 1939, the second son of Mr. and Mrs. Muthiya Pillai.

Education: Graduated from Municipal High School, Mayuram, India, in 1954; received Bachelor of Science (Honours) degree in Statistics from Annamalai University, India, in 1960; received Master of Science (by research) degree in Statistics from Annamalai University, India, in May, 1967; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in December, 1980.

Professional Experience: Lecturer, P. S. G. Arts College, Coimbatore, India, 1961; Lecturer in Statistics, Annamalai University, India, 1962-1977; Graduate teaching associate, Oklahoma State University, 1977-1980.

Professional Organizations: American Statistical Association, Mu Sigma Rho.