A BAYESIAN ANALYSIS OF A CHANGING LINEAR MODEL

BY

JAUNITA HWA-LING-TU CHIN CHOY

Bachelor of Science
National Taiwan University
Taipei, Taiwan, Republic of China
1967

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1970

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, 1977

A BAYESIAN ANALYSIS OF A CHANGING

LINEAR MODEL

Thesis approved:

Lyle Broemeling

Thesis Adviser

Don Holbert

Edward L. Smith

Ronald W. McNew

Robert D Morrison

Norman N Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

I am deeply grateful to my principal adviser, Dr. Lyle D. Broemeling, for suggesting this problem and for his guidance, concern and encouragement through this study.

Special thanks to Dr. Donald Holbert whose thesis became a guidance to this study.

I would like to thank Drs. Lyle D. Broemeling, Donald Holbert, Ronald W. McNew, Robert D. Morrison and Edward L. Smith for serving on my advisory committee and for their assistance in correcting the final copy of my thesis.

I also wish to express my thanks to Mrs. Cathy Scott for the excellent typing of this thesis.

I would also like to express my gratitude to my husband, Ed, for his encouragement and his confidence in my ability. Finally, I would like to express my gratitude to my parents, Mr. and Mrs. Chi-Lian Tu, for their encouragement and support throughout my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND STATEMENT OF THE PROBLEM

Introduction and Background

Recently increasing interest has been shown in the problem of the changing regression model for a sequence of random variables. An observed data set may be satisfied by a single regression analysis, which is normally the assumption, or it may require two or more separate regression relationships. A "switching regression problem" is one in which the observations follow a model consisting of several regression models. If a single switch occurs, this type of situation is called a two-phase regression problem. Most of the work which has been done on two-phase regression problems is with the simple linear regression case and assumes a sequence of independent random variables Y_1, Y_2, $\ldots, Y_n$ such that

$$Y_i = \alpha_1 + \beta_1 x_i + e_i, \quad i = 1,2,\ldots,m$$

and

$$Y_i = \alpha_2 + \beta_2 x_i + e_i, \quad i = m+1,\ldots,n$$

where the e_i's are i.i.d. $N(0,\sigma^2)$ and the x_i's are the values of a concomitant variable X. m is some unknown point, and when m=n there is no change and when m=2,3,...,n-2 there is one change.

Essentially, there are two problems associated with two-phase regression: i) detecting the change, i.e., is there a change occuring in a sequence of random variables?, ii) if the change does occur, estimating and making inferences about the shift point m and all the unknown regression parameters. Assuming a change does occur, Quandt (1958) estimated the switch point m and the regression parameters by a maximum likelihood technique. Hinkley (1969, 1971) under the assumption that the two-phase regression model is continuous, estimated and made inferences about the abscissa of the intersection, i.e., $\gamma = (\alpha_2-\alpha_1)/(\beta_1-\beta_2)$, of the two regression lines.

The above works are based on the classical approach where the inferences are solely based on the sample data. Sometimes prior information may exist. Bayesian approach is concerned with the combination of sample data and prior information. How can the information from two different sources be combined with each other? One way is to apply Bayes' theorem to obtain a conditional distribution which is called the posterior distribution. The posterior distribution provides the means of making all relevant inferences about a parameter or a set of parameters in which we are interested. Lindley (1965), Box and Tiao (1973), DeGroot (1970) and Zellner (1971) gave a detailed description of the Bayesian inference. Barnett (1973) described the various approaches to statistical inference and decision-making.

Based on the Bayesian approach, Holbert (1973) studied the problems of estimating the shift point m and the abscissa, γ, of the intersection of the two regression lines. Assigning a uniform proper prior to the shift point m and an improper prior to the unknown regression parameters, he derived the posterior distribution of m and γ for a number of cases. Ferriera (1975) also assigned a vague-type prior distribution to the unknown regression parameters and assigned three different prior distributions to the shift point. He obtained the marginal posterior distribution and the expected values for the shift point m and the regression parameters. Other studies related to two-phase regression problems are those of Quandt (1960, 1972), Sprent (1961), Robinson (1964), Hudson (1966) and Bacon and Watts (1971).

In studying the related decision problem of testing the presence of a switch from one regression scheme to another, Brown, Durbin and Evans (1975), employing a non-Bayesian approach, developed the tests for the constancy of a regression relationship based on the cusum and cusum of the squares of recursive residuals. Broemeling (1972) discussed a Bayesian procedure for detecting the change of distribution parameters in a sequence of random variables. He approached the problem in terms of posterior odds on 'no change'. Smith (1975) considered an informal sequential procedure to detect the change. Other studies related to this problem have been done by Quandt (1960), Bhattacharyya and Johnson (1968), Farley and Hinich (1970), Farley, Hinich and McGuire (1975), and Garbade (1977).

In this paper, the problem is generalized to the multiple linear regression case and is approached by the Bayesian method and analysed with a proper prior for all unknown parameters. It will be shown that even though $x'x$ is singular (x is the design matrix in regression analysis), one still can estimate and make inferences about the shift point and regression parameters.

The use of improper priors to represent "ignorance" has been recently criticized by Dawid, Stone and Zidek (1973), because their use can lead to logical contradictions. One of the examples that leads to a contradiction is a shifting sequence of exponential populations. Since such a contradiction cannot arise where one employs a proper prior distribution, it is important to reexamine the shift point and switching regression problems with proper prior distributions.

Another reason for using proper prior distributions is that when the shift point does occur, the posterior distribution will exist for all parameters including the shift point parameter m for all its possible values 1,2,...,n-1. With improper prior distributions, it can be shown that the posterior distribution will exist for m, but only at the mass points m = p, p+1,...,n-p, where p is the number of regression coefficients. It is unrealistic to assume that if a shift occurs only once, it occurs at only these points. Thus by using a proper prior distribution for all parameters one avoids this unrealistic assumption. Of course, one must be able to realistically formulate these priors based on the prior knowledge.

Ferreira's (1975) study emphasized the sampling properties of the point estimators of the regression coefficients in order to examine the effect of three prior distributions assigned to the switch point. His study is important in that it may convince non-Bayesians that certain Bayesian estimators have optimal sampling properties. My study is confined to switching regression problems where only the posterior distributions will be derived and from these, point and interval estimators and the highest posterior density (H.P.D.) regions providing test of hypothesis may be derived. If loss functions can realistically be assigned, then estimators and test of hypothesis can be constructed from a Bayesian decision theoretic viewpoint.

In many practical problems either the data itself will validate the assumption that there is a change in the regression relationship or or there will be reasons which make this assumption reasonable. For example, in biological systems, the threshold level of a chemical may be specific, i.e., the response of the system to the chemical is additive to the threshold level. After this level has been attained, the response stays constant or the chemical becomes toxic to the system, resulting in a decreasing response with increasing concentration. Ohki (1974) found that the top growth of cotton increased sharply with a very slight increase of manganese in the blade tissue, but after the inflection point of the nutrient calibration curve was attained the manganese content of the blade tissue increased sharply with no increase in plant growth. Pool and Borchgrevink (1964), reported on the level of the synthesis of blood factor VII (proconvertin), a coagulation factor in the blood, as a function of warfarin concentration in the liver of rats. Synthesis is inhibited when the warfarin concentration surpasses a critical level. This data set was used by Hinkley (1971) to illustrate maximum likelihood estimation of the shift point. Some other examples of this problem can be seen in the papers of Sims, Atkinson and Smitobol (1975) and Millar and Denmead (1976).

Statement of the Problem

We assume that a sequence of independent random variable Y_1, $Y_2,\ldots,Y_n$ satisfy

$$Y_i = \underset{\sim}{x}_i{}'\beta_i + e_i, \quad i = 1,2,\ldots,n$$

where $\underset{\sim}{x}_i$ is a px1 known vector of p regressor variables,

β_i is a px1 vector of regression parameters,

e_i is an error term and

$e_1, e_2, \ldots, e_n$ are i.i.d. $N(0,\sigma^2)$.

With the usual regression analysis, we assume that

$$\beta_1 = \beta_2 = \ldots = \beta_n = \underset{\sim}{\beta}_1$$

i.e., the model is

$$Y_i = \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1 + e_i, \quad i = 1,2,\ldots,n . \qquad (1.1)$$

Is this assumption valid? We need to check the consistency of this model over a set of data. It is necessary to construct a test to detect the change, i.e., with the null hypothesis H_0 of no change, vs the alternative hypothesis H_1 of one change. If H_0 is true, the model (1.1) is correct. We can then claim that the model is constant over this sequence of data and go ahead and do the usual regression analysis. If H_1 is true, we claim that there is a change point, m, and break the data set into two subsets with each subset of observations following a different regression model. This model is

$$\begin{aligned} Y_i &= \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1 + e_i, \quad i = 1,2,\ldots,m \\ &= \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_2 + e_i, \quad i = m+1,\ldots,n \end{aligned}$$

where $\underset{\sim}{\beta}_1 \neq \underset{\sim}{\beta}_2$. In this case, we need to find out where the shift occurred and make inferences about the shift point m and all unknown regression parameters.

The objective of this study is to develop Bayesian techniques to i) detect the presence of a change from one regression model to another, (ii) estimate and make inferences about the shift point and other unknown parameters in a sequence of independent random variables which change regression model at an unknown point, and (iii) estimate and make inferences about the abscissa of the intersection of two regression lines.

CHAPTER II

POSTERIOR DISTRIBUTIONS INVOLVING THE TWO-PHASE MULTIPLE REGRESSION

Basic Assumption

Suppose a sequence of normal independently distributed random variables $Y_1, Y_2, \ldots, Y_n$, follow

$$Y_i = \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1 + e_i, \quad i = 1,2,\ldots,M$$

and $$M = 1,2,\ldots,n-1$$

$$Y_i = \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_2 + e_i, \quad i = M+1,\ldots, M$$

where

$\underset{\sim}{x}_i$ is a pxl column vector of known fixed quantities on p regressors for the ith observation,

$\underset{\sim}{\beta}_1$ is a pxl column vector of regression coefficients of the first linear multiple regression model, $\underset{\sim}{\beta}_1 \in R^p$,

$\underset{\sim}{\beta}_2$ is a pxl column vector of regression coefficients of the second linear multiple regression model, $\underset{\sim}{\beta}_2 \in R^p$,

e_i's are i.i.d. $N(0,\sigma^2)$, $i = 1,2,\ldots,n$ where $\sigma^2 > 0$ and $\underset{\sim}{\beta}_1 \neq \underset{\sim}{\beta}_2$.

Thus, we assume that there is a changing regression relationship over this sequence of random variables and there is exactly one change at an unknown shift point M. We are interested in estimating the shift point M as well as any unknown regression parameters $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$ and possibly the unknown common variance σ^2 . Let $\underset{\sim}{\Theta}$ be the vector, consisting of all possible unknown regression parameters and σ^2. We assign a prior probability density function (abbreviated p.d.f.) to $\underset{\sim}{\Theta}$, denoted by $\pi(\underset{\sim}{\Theta})$, and assume that M and $\underset{\sim}{\Theta}$ are independently distributed. Throughout this paper, we assume that M has a uniform prior distribution over the space.

$$I_{n-1} = (1,2,\ldots,n-1) \ .$$

Denote $\pi_0(m)$ as the prior probability mass function (abbreviated p.m.f.) of M, then

$$\pi_0(m) = \begin{cases} \frac{1}{n-1}, & m = 1,2,\ldots,n-1 \\ 0, & \text{otherwise} \end{cases} \ .$$

Under the above assumptions, the probability density function (abbreviated p.d.f.) of $\underset{\sim}{Y} = (Y_1,Y_2,\ldots,Y_n)$ given $\underset{\sim}{x} = (x_1,x_2,\ldots,x_n)$, and m, $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$, and σ^2 is

$$f(y|m,\sigma^2,\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2) = (2\pi\sigma^2)^{-n/2}\exp\{(-\frac{1}{2\sigma^2})[\sum_{i=1}^{m}(y_i-\underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1)^2 + \sum_{i=m+1}^{n}(y_i-\underset{\sim}{x}_i{}'\underset{\sim}{\beta}_2)^2]\}$$

$$\propto \sigma^{-n}\exp(\frac{-1}{2\sigma^2})\ [\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]'[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}] \qquad (2.1)$$

where

$$\underset{\sim}{y} = \begin{pmatrix} \underset{\sim}{y}_1(m) \\ \underset{\sim}{y}_2(m) \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \end{pmatrix} \quad \text{and} \quad x(m) = \begin{pmatrix} x_1(m) & \phi \\ \phi & x_2(m) \end{pmatrix}$$

and $\underset{\sim}{y}_1(m)$, $\underset{\sim}{x}_1(m)$ and $\underset{\sim}{\beta}_1$ denote the usual observation vector, design matrix, parameter vector, respectively, for the first regression model, using the first m observations. Similarly, $y_2(m)$, $x_2(m)$ and $\underset{\sim}{\beta}_2$ correspond to the same parameters of the second regression model using the last n-m observations.

The expression of (2.1) is a function of m and $\underset{\sim}{\Theta}$ and is the likelihood function, $L(m,\underset{\sim}{\Theta})$. By the Bayes theorem, the joint posterior p.d.f. of M and $\underset{\sim}{\Theta}$ is

$$\pi(m,\underset{\sim}{\Theta}|\underset{\sim}{y}) \propto L(m,\Theta)\pi_0(m)\pi_0(\Theta)$$
$$\propto L(m,\Theta)\pi_0(\Theta) \qquad (2.2)$$

The second equation follows because $\pi(m)$ is a constant over I_{n-1} . From (2.2) we can derive the marginal p.d.f. of M and $\underset{\sim}{\Theta}$.

Many situations can give rise to the model (2.1). There are six cases to be considered:

i) σ^2 known, both regressions known;

ii) σ^2 known, one regression known, the other unknown;

iii) σ^2 known, both regressions unknown;

iv) σ^2 unknown, both regression known;

v) σ^2 unknown, one regression known, the other unknown;

vi) σ^2 unknown, both regression unknown.

We begin our study with the most general case, i.e., the case that σ^2 **is unknown,** and both regressions are unknown.

Posterior Distributions of the Unknown Parameters

The Most General Case

In this case, both regression parameter vectors $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ are unknown, and σ^2 is unknown. We need to assign a proper prior to these parameters. The joint prior distribution of $\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1{}', \underset{\sim}{\beta}_2{}')'$ and $R = 1/\sigma^2$ are assigned as follows: the conditional distribution of $\underset{\sim}{\beta}$ when $R = r(r > 0)$ is a 2p-dimensional multivariate normal distribution with mean vector $\underset{\sim}{\beta}_\mu$, and precision matrix $r\tau$ such that $\underset{\sim}{\beta}_\mu \in R^{2p}$ and

τ is a given symmetric 2px2p positive definite matrix. The marginal distribution of R is a gamma distribution with parameters a and b such that $a > 0$ and $b > 0$. If $\pi_0(\underset{\sim}{\beta}|r)$ and $\pi_0(r)$ denote the conditional p.d.f. of $\underset{\sim}{\beta}$ when R = r and the marginal p.d.f of R, respectively, then

$$\pi_0(\underset{\sim}{\beta}|r) = (2\pi)^{-(2p/2)}|r\tau|^{1/2} \exp[-\frac{r}{2}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)'\tau(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)] \qquad (2.3)$$

$$\pi_0(r) = \frac{b^a}{\Gamma(a)} r^{a-1}e^{-br} . \qquad (2.4)$$

Hence, the joint p.d.f. of $\underset{\sim}{\beta}$ and R is a normal-gamma p.d.f., which is

$$\pi_0(\underset{\sim}{\beta},r) \propto r^{a+p-1} \exp\{(-r)[b + \frac{1}{2}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)'\tau(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)]\} . \qquad (2.5)$$

From the relation (2.5), the marginal prior p.d.f. $\pi_0(\underset{\sim}{\beta})$ has the form

$$\pi_0(\underset{\sim}{\beta}) = \int_0^\infty \pi_0(\underset{\sim}{\beta},r)dr$$

$$\propto [1 + \frac{1}{2a}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)' \frac{a\tau}{b}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)]^{-(a+p)} \qquad (2.6)$$

which is the p.d.f of 2p-dimensional multivariate t distribution with 2a degrees of freedom, location parameter $\underset{\sim}{\beta}_\mu$, and precision matrix $a\tau/b$.

We assume that M is independent of $\underset{\sim}{\beta}$ and R. Therefore the joint prior distribution of M, $\underset{\sim}{\beta}$, and R is

$$\pi_0(m,\underset{\sim}{\beta}, r) = \pi_0(m) \cdot \pi_0(\underset{\sim}{\beta},r)$$

$$\propto r^{a+p-1} \exp\{(-r)[b + \frac{1}{2}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)' (\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)]\} \qquad (2.7)$$

for $m = 1,2,\ldots,n-1$, $\underset{\sim}{\beta} \in R^{2p}$ and $r > 0$.

From (2.1) the likelihood function is

$$L(m,\underset{\sim}{\beta},r) \propto r^{n/2} \exp\{-(r/2)[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]'[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]\} . \qquad (2.8)$$

By Bayes theorem, combining the likelihood function with the prior density, results in a joint posterior density of M, $\underset{\sim}{\beta}$ and R, which is

$$\pi(m,\underset{\sim}{\beta},r) \propto r^{a+p+(n/2)-1} \exp\Big\{(-r)[b + \frac{1}{2}(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)'\tau(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu) + \frac{1}{2}(\underset{\sim}{y}-x(m)\underset{\sim}{\beta})'(\underset{\sim}{y}-x(m)\underset{\sim}{\beta})]\Big\} \quad (2.9)$$

for $m = 1,2,\ldots,n-1$, $\underset{\sim}{\beta} \in R^{2p}$ and $r > 0$.

From (2.9), we can derive the following marginal posterior density for all the unknown parameters.

i) The Posterior Probability Mass Function of Shift Point M.

Integrating $\pi(m,\underset{\sim}{\beta},r)$ with respect to r and $\underset{\sim}{\beta}$, we get the marginal posterior p.m.f. for M. In order to evaluate the integral, we need to use the identity

$$(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu)'\tau(\underset{\sim}{\beta}-\underset{\sim}{\beta}_\mu) + (\underset{\sim}{y}-x(m)\underset{\sim}{\beta})'(\underset{\sim}{y}-x(m)\underset{\sim}{\beta})$$

$$= [\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]'[x(m)'x(m)+\underset{\sim}{\tau}][\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)] + \underset{\sim}{y}'\underset{\sim}{y} + \underset{\sim}{\beta}_\mu'\tau\underset{\sim}{\beta}_\mu - \underset{\sim}{\beta}^*(m)'[x(m)'x(m)+\underset{\sim}{\tau}]\underset{\sim}{\beta}^*(m) \quad (2.10)$$

where

$$\underset{\sim}{\beta}^*(m) = [x(m)'x(m)+\tau]^{-1}[\tau\underset{\sim}{\beta}_\mu+x(m)'\underset{\sim}{y}] \; . \quad (2.11)$$

Note that $[x(m)'x(m)+\tau]^{-1}$ exists even when $x(m)'x(m)$ is singular, because $x(m)'x(m)$ always is a positive semidefinite matrix and τ is a positive definite matrix.

Substituting the identity (2.10) into (2.9), (2.9) can be rewritten as

$$\pi(m,\underset{\sim}{\beta},r|\underset{\sim}{y}) \propto r^{a^*+p-1} \exp\{(-r)[D(m) + \frac{1}{2}(\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m))'(x(m)'x(m) + \tau)(\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m))]\}$$

where

$$a^* = a + \frac{n}{2}$$

$$D(m) = b + \frac{1}{2}\{\underset{\sim}{y}'\underset{\sim}{y}+\underset{\sim}{\beta}_\mu{}'\tau\underset{\sim}{\beta}_\mu-\underset{\sim}{\beta}^{*\prime}(m)[x(m)'x(m)+\tau]\underset{\sim}{\beta}^*(m)\}$$

$$= b + \frac{1}{2}\{[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}^*(m)]'\underset{\sim}{y} + [\underset{\sim}{\beta}_\mu-\underset{\sim}{\beta}^*(m)]'\tau\underset{\sim}{\beta}_\mu\} \quad . \qquad (2.12)$$

Then

$$\pi(m|y) = \int_0^\infty \int_{R^{2p}}^\infty \pi(m,\underset{\sim}{\beta},r)\,d\underset{\sim}{\beta}dr$$

$$\propto \int_0^\infty r^{a^*+p-1} \exp[-D(m)r] \int_{R^{2p}}^\infty \exp\{(-r/2)$$

$$[\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]'[x'(m)x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]\}d\underset{\sim}{\beta}dr$$

$$\propto |x(m)'x(m)+\tau|^{-1/2} \int_0^\infty r^{a^*-1} \exp[-D(m)r]dr$$

$$\propto D(m)^{-a^*}|x(m)'x(m)+\tau|^{-1/2}, \qquad m = 1,2,\ldots,n-1 \qquad (2.13)$$

In going from the second line of (2.13) to the third, we use the 2p-dimensional multivariate normal density to integrate out $\underset{\sim}{\beta}$ and from the third line of (2.13) to the fourth line, we use the gamma density to integrate out r.

In order to get a more intuitive feeling of D(m), D(m) can be expanded as

$$D(m) = b + \frac{1}{2}\{[\underset{\sim}{y}-\hat{\underset{\sim}{y}}(m)]'[\underset{\sim}{y}-\hat{\underset{\sim}{y}}(m)]+[\hat{\underset{\sim}{\beta}}(m)-\underset{\sim}{\beta}_\mu]'w(m)[\hat{\underset{\sim}{\beta}}(m)-\underset{\sim}{\beta}_\mu]\},$$

where $w(m) = x(m)'x(m)[x(m)'x(m)+\tau]^{-1}\tau$ and $\hat{\underset{\sim}{y}}(m)$, $\hat{\underset{\sim}{\beta}}(m)$ are the vectors of usual least squares predicted values and least square estimators, using x(m) and $\underset{\sim}{\beta}$ as the design matrix and regression coefficient vector; i.e.,

$$\hat{\underset{\sim}{\beta}}(m) = [x(m)'x(m)]^{-1}x(m)'\underset{\sim}{y}$$

$$\hat{\underset{\sim}{y}}(m) = x(m)'\hat{\underset{\sim}{\beta}}(m) \; .$$

When $m = p, p+1, \ldots, n-p$, $[x(m)'x(m)]^{-1}$ denotes the usual inverse of $x(m)'x(m)$; whereas, when $m = 1,2,\ldots,p-1$ or $m = n-p+2,\ldots,n-1$, $[x(m)'x(m)]^{-1}$ denotes the generalized inverse of $x(m)'x(m)$ due to the singularity of $x(m)'x(m)$. Notice that $D(m)$ is invariant to the choice of the generalized inverse. Hence, $\pi(m|\underset{\sim}{y})$ is invariant to the choice of a generalized inverse.

(ii) The Posterior p.d.f. of $\underset{\sim}{\beta}$

From (2.9) we can obtain the posterior p.d.f. of $\underset{\sim}{\beta}$, which is

$$\pi(\underset{\sim}{\beta}|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \int_0^\infty \pi(m,\underset{\sim}{\beta},r)\,dr$$

$$\propto \sum_{m=1}^{n-1} \int_0^\infty r^{a^*+p-1} \exp\{(-r)[D(m)+\tfrac{1}{2}[\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]' [x(m)'x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]\}\,dr$$

$$\propto \sum_{m=1}^{n-1} \{D(m) + \tfrac{1}{2}[\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]'[x(m)'x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]\}^{-(a^*+p)}$$

$$\propto \sum_{m=1}^{n-1} D(m)^{-(a^*+p)}\{1 + \frac{1}{2a^*}[\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]p(m)[\underset{\sim}{\beta}-\underset{\sim}{\beta}^*(m)]\}^{-\frac{2a^*+2p}{2}}$$

$$\propto \sum_{m=1}^{n-1} D(m)^{-a^*}|x(m)'x(m)+\tau|^{-1/2}\, t[\underset{\sim}{\beta};2p,2a^*,\underset{\sim}{\beta}^*(m),p(m)] \qquad (2.14)$$

where

$$p(m) = (a^*/D(m))[x(m)'x(m)+\tau] \qquad (2.15)$$

and $t[\underset{\sim}{\beta};2p,2a^*,\underset{\sim}{\beta}^*(m),p(m)]$ is the p.d.f. of the 2p-dimensional multivariate t distribution of the variable vector $\underset{\sim}{\beta}$ with degrees of freedom $2a^*$, location vector $\underset{\sim}{\beta}^*(m)$, and precision matrix $p(m)$. From (2.13), we can rewrite (2.14) as

$$\pi(\underset{\sim}{\beta}|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{\beta};2p,2a^*,\underset{\sim}{\beta}^*(m),p(m)] \cdot \pi(m|\underset{\sim}{y}), \quad \underset{\sim}{\beta} \varepsilon R^{2p}$$

$$= 0, \text{ otherwise} \tag{2.16}$$

The marginal posterior distribution for any subset of the components of $\underset{\sim}{\beta}$ can be easily found because $\underset{\sim}{\beta}$ is a mixture of multivariate t distributions. Let us partition the random vector $\underset{\sim}{\beta}$, the location vector $\underset{\sim}{\beta}^*(m)$ and the precision matrix p(m) as

$$\underset{\sim}{\beta} = \begin{pmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \end{pmatrix}, \quad \underset{\sim}{\beta}^*(m) = \begin{pmatrix} \underset{\sim}{\alpha}_1^*(m) \\ \underset{\sim}{\alpha}_2^*(m) \end{pmatrix}, \quad p(m) = \begin{pmatrix} p_{11}(m) & p_{12}(m) \\ p_{21}(m) & p_{22}(m) \end{pmatrix}$$

The dimensions of $\underset{\sim}{\beta}_i$ and $\underset{\sim}{\beta}_i^*(m)$ are px1 (i=1,2) and the dimension of $p_{ij}(m)$ is pxp (i,j=1,2). Then

(iia) The Posterior p.d.f. of $\underset{\sim}{\beta}_1$ is

$$\pi(\underset{\sim}{\beta}_1|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{\beta}_1;p,2a^*,\underset{\sim}{\alpha}_1^*(m),p_1^*(m)] \cdot \pi(m|\underset{\sim}{y}) \tag{2.17}$$

where

$$p_1^*(m) = p_{11}(m) - p_{12}(m)p_{22}^{-1}(m)p_{21}(m) \tag{2.18}$$

(iib) The Posterior p.d.f. of $\underset{\sim}{\beta}_2$ is

$$(\underset{\sim}{\beta}_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{\beta}_2;p,2a^*,\underset{\sim}{\alpha}_2^*(m),p_2^*(m)] \cdot \pi(m|\underset{\sim}{y}) \tag{2.19}$$

where

$$p_2^*(m) = p_{22}(m) - p_{21}(m)p_{11}^{-1}(m)p_{12}(m) . \tag{2.20}$$

and $t[\underset{\sim}{Y};k,n,\underset{\sim}{\mu},v]$ as previously defined.

(iii) The Posterior p.d.f. of R

Let $\pi(r|\underset{\sim}{y})$ denote the marginal posterior p.d.f. of R, then

$$\pi(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \int_{R^{2p}} \pi(m,\underset{\sim}{\beta},r)\,d\underset{\sim}{\beta}$$

$$\propto \sum_{m=1}^{n-1} \int_{R^{2p}} r^{a*+p-1} \exp\{(-r)\{D(m)+\frac{1}{2}[\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]'$$

$$[x(m)'x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]\}\}d\underset{\sim}{\beta}$$

$$\propto \sum_{m=1}^{n-1} r^{a*+p-1} \exp[-D(m)r] \int_{R^{2p}} \exp\{-\frac{r}{2}[\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]'$$

$$[x(m)'x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]\}d\underset{\sim}{\beta}$$

$$\propto \sum_{m=1}^{n-1} |x(m)'x(m)+\tau|^{-1/2} r^{a*-1} \exp[-D(m)r]$$

$$\propto \sum_{m=1}^{n-1} D(m)^{-a*}|x(m)'x(m)+\tau|^{-1/2} g[r;a*,D(m)] .$$

Therefore

$$\pi(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} g[r;a*,D(m)] \cdot \pi(m|\underset{\sim}{y}), \quad r > 0$$

$$= 0, \text{ otherwise} , \tag{2.21}$$

where $g[r;a*,D(m)]$ is the p.d.f. of a gamma distribution of the variable R with parameters a* and D(m).

(iv) The Posterior p.d.f. of σ^2

Since $R = 1/\sigma^2$ is distributed as a mixture of gamma distributions, from (2.21) we can obtain the distribution of σ^2, as

$$\pi(\sigma^2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} ig[\sigma^2;a*,D(m)] \cdot \pi(m|\underset{\sim}{y}), \sigma^2 > 0$$

$$= 0, \text{ otherwise} , \tag{2.22}$$

where $ig[\sigma^2;a*,D(m)]$ is the p.d.f. of an inverse gamma distribution of the variable σ^2 with parameters a* and D(m). A random variable y has an inverse gamma distribution with parameters α and β , whose p.d.f. is

$$f(y|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{y^{\alpha+1}} e^{-\beta/y} , \quad y > 0$$

$$= 0, \text{ otherwise} .$$

Other Special Cases

Since the derivation for the most general case has been studied in great detail, it is not necessary to show the proof for other cases. If we regard the previous case as a main theorem, then we can state the other cases without proof.

Corollary 1: If the assumption given above holds, and if $\underset{\sim}{\beta}_1$ is known, $\underset{\sim}{\beta}_2$ is unknown, and σ^2 is unknown, and if a joint prior distribution to $\underset{\sim}{\beta}_2$ and $R = 1/\sigma^2$ is assigned as follows: the conditional prior distribution of $\underset{\sim}{\beta}_2$ when $R = r$ is a p-variate normal distribution with mean vector $\underset{\sim}{\beta}_2$ and precision matrix $r\tau_2$ such that $\bar{\underset{\sim}{\beta}}_2 \in R^p$ and τ_2 is a given pxp symmetric, positive definite matrix, and the marginal distribution of R is a gamma distribution with parameters a and b, such that $a > 0$, $b > 0$, then

(i) The posterior p.m.f. of M is

$$\pi(m|\underset{\sim}{y}) \propto D_2(m)^{-a^*}|x_2(m)'x_2(m)+\tau_2|^{-1/2}, \quad m = 1,2,\ldots,n-1 \qquad (2.23)$$

where

$$a^* = a + n/2$$

$$\begin{aligned} D_2(m) &= b + 1/2\{[\underset{\sim}{y}_1(m)-x_1(m)\underset{\sim}{\beta}_1]'[\underset{\sim}{y}_1(m)-x_1(m)\underset{\sim}{\beta}_1] \\ &\quad + \underset{\sim}{y}_2(m)'\underset{\sim}{y}_2(m)+\bar{\underset{\sim}{\beta}}_2'\tau_2\bar{\underset{\sim}{\beta}}_2 - \underset{\sim}{\beta}^*(m)'[x_2(m)'x_2(m)+\tau_2]\underset{\sim}{\beta}_2^*(m) \\ &= b + 1/2\{[\underset{\sim}{y}_1(m)-x_1(m)\underset{\sim}{\beta}_1]'[\underset{\sim}{y}_1(m)-x_1(m)\underset{\sim}{\beta}_1] \\ &\quad + [\underset{\sim}{y}_2(m)-x_2(m)\underset{\sim}{\beta}_2^*(m)]'y_2(m)+[\bar{\underset{\sim}{\beta}}_2-\underset{\sim}{\beta}_2^*(m)]'\tau_2\bar{\underset{\sim}{\beta}}_2\} \end{aligned} \qquad (2.24)$$

$$\underset{\sim}{\beta}_2^*(m) = [x_2(m)'x_2(m)+\tau_2]^{-1}[\tau_2\bar{\underset{\sim}{\beta}}_2+x_2(m)'\underset{\sim}{y}_2(m)] \ . \qquad (2.25)$$

(ii) The Posterior p.d.f. of $\underset{\sim}{\beta}_2$ is

$$\pi(\underset{\sim}{\beta}_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{\beta}_2;p,2a^*,\underset{\sim}{\beta}_2^*(m),p_2(m)] \cdot \pi(m|\underset{\sim}{y}) \tag{2.26}$$

where $p_2(m) = [a^*/D_2(m)](x_2(m)'x_2(m)+\tau_2)$, $t[\underset{\sim}{\beta}_2;p,2a^*,\underset{\sim}{\beta}_2^*(m),p_2(m)]$ is defined as before and $\pi(m|\underset{\sim}{y})$ is given by equation (2.23). The marginal posterior distribution of the elements of $\underset{\sim}{\beta}_2$ can be derived by (2.26).

(iii) The Posterior p.d.f. of R is

$$\pi(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} g[r;a^*,D_2(m)] \cdot \pi(m|\underset{\sim}{y}) \tag{2.27}$$

where $g[r,a^*,D_2(m)]$ is the gamma p.d.f. of the variable R with parameters a^* and $D_2(m)$, and $\pi(m|\underset{\sim}{y})$ is given by (2.23).

Corollary 2: If $\underset{\sim}{\beta}_1$ is unknown and $\underset{\sim}{\beta}_2$ is known, and σ^2 is unknown, then the results are similar to the results of Corollary 1.

Corollary 3: If the basic assumption given above holds and if σ^2 is unknown and both $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ are known, and it is assumed that $R=1/\sigma^2$ has a gamma distribution with parameters a and b, $a > 0$, $b > 0$, then

(i) The Posterior p.m.f. of M is

$$\pi(m|\underset{\sim}{y}) \propto B(m)^{-a^*}, \quad m = 1,2,\ldots,n-1 \tag{2.28}$$

where

$$a^* = a + n/2,$$

$$B(m) = b + 1/2[\underset{\sim}{y} - x(m)\underset{\sim}{\beta}]'[\underset{\sim}{y} - x(m)\underset{\sim}{\beta}] \ . \tag{2.29}$$

(ii) The Posterior p.d.f. of R is

$$\pi(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} g[r;a^*,B(m)] \cdot \pi(m|y), \tag{2.30}$$

where $g[r;a^*,B(m)]$ is previously defined and $\pi(m|\underset{\sim}{y})$ is given by (2.28).

Corollary 4: If the assumption given above holds, and if σ^2 is known, and both $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ are unknown, and $\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1{}',\underset{\sim}{\beta}_2{}')'$ has a 2p-variate normal distribution with mean vector $\underset{\sim}{\beta}_\mu$ and covariance matrix σ^2A^{-1}, such that $\underset{\sim}{\beta}_\mu' \in R^{2p}$, and A is a 2px2p symmetric, positive definite matrix, then

(i) The posterior p.m.f. of M is

$$\pi(m|\underset{\sim}{y}) \propto |x(m)'x(m)+A|^{-1/2} \exp[\frac{1}{2\sigma^2} C(m)], \quad m = 1,2,n-1 \tag{2.31}$$

where

$$C(m) = \underset{\sim}{\beta}^*(m)'[x(m)'x(m)+A]\underset{\sim}{\beta}^*(m), \tag{2.32}$$

$$\underset{\sim}{\beta}^*(m) = [x(m)'x(m)+A]^{-1}[A\underset{\sim}{\beta}_\mu+x(m)'\underset{\sim}{y}] . \tag{2.33}$$

(ii) The Posterior p.d.f. of $\underset{\sim}{\beta}$ is

$$\pi(\underset{\sim}{\beta}|y) = \sum_{m=1}^{n-1} N[\underset{\sim}{\beta};\underset{\sim}{\beta}^*(m),V(m)] \cdot \pi(m|y) \tag{2.34}$$

where

$$V(m) = \sigma^2[x(m)'x(m)+A]^{-1} \tag{2.35}$$

$\pi(m|\underset{\sim}{y})$ is given by equation (2.31) and $N[\underset{\sim}{\beta};\underset{\sim}{\beta}^*(m),V(m)]$ is the 2p-variate normal p.d.f. of the variable vector $\underset{\sim}{\beta}$ with mean vector $\underset{\sim}{\beta}^*(m)$ and covariance matrix V(m). The marginal posterior p.d.f. for any subset of the components of $\underset{\sim}{\beta}$ can be easily obtained from (2.34).

<u>Corollary 5:</u> If the above basic assumptions hold and if σ^2 is known, $\underset{\sim}{\beta}_1$ is known, $\underset{\sim}{\beta}_2$ is unknown, and $\underset{\sim}{\beta}_2$ has a p-variate normal distribution with mean vector $\underset{\sim}{\bar{\beta}}_2$ and covariance matrix $\sigma^2 A_2^{-1}$, such that $\underset{\sim}{\bar{\beta}}_2 \varepsilon R^p$ and A_2 is a given pxp symmetric and positive definite matrix, then

(i) The Posterior p.m.f. of M is

$$\pi(m|\underset{\sim}{y}) \propto |x_2(m)'x_2(m)+A_2|^{-1/2} \exp[-\frac{1}{2\sigma^2} C_2(m)] ,$$

$$m = 1,2,\ldots,n-1 \tag{2.36}$$

where

$$C_2(m) = [x_1(m)\underset{\sim}{\beta}_1-2\underset{\sim}{y}_1(m)]'x_1(m)\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_2^*(m)'[x_2(m)'x_2(m)+A_2]\underset{\sim}{\beta}_2^*(m) \tag{2.37}$$

$$\underset{\sim}{\beta}_2^*(m) = [x_2(m)'x_2(m)+A_2]^{-1}[A_2\underset{\sim}{\bar{\beta}}_2+x_2(m)'\underset{\sim}{y}_2(m)] . \tag{2.38}$$

(ii) The Posterior p.d.f. of $\underset{\sim}{\beta}_2$ is

$$\pi(\underset{\sim}{\beta}_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} N[\underset{\sim}{\beta}_2;\underset{\sim}{\beta}_2^*(m),V_2(m)] \cdot \pi(m|y) \tag{2.39}$$

where

$$V_2(m) = \sigma^2[x_2(m)'x_2(m)+A_2]^{-1} \tag{2.40}$$

$\pi(m|\underset{\sim}{y})$ is given by (2.36) and $N[\underset{\sim}{\beta}_2,\underset{\sim}{\beta}_2^*(m),V_2(m)]$ is the p-variate normal p.d.f. of the variable vector $\underset{\sim}{\beta}_2$ with mean vector $\underset{\sim}{\beta}_2^*(m)$ and covariance matrix $V_2(m)$.

<u>Corollary 6:</u> The case for σ^2 known, $\underset{\sim}{\beta}_1$ unknown, $\underset{\sim}{\beta}_2$ known is similar to the case of Corollary 5.

<u>Corollary 7:</u> If σ^2, $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$ are all known, then the posterior p.m.f. of M is

$$\pi(m|\underset{\sim}{y}) \propto \exp\left\{-\frac{1}{2\sigma^2}[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]'[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]\right\},\ m = 1,2,\ldots,n-1. \tag{2.41}$$

Point Estimation for Parameters

In the estimation problem, we can find several different estimators corresponding to different loss functions. With a square error loss function, the estimator is the expected value of the posterior distribution. For the most general case where $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$, and σ^2 are unknown, we can find

(i) the expected values of M, $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$, R and σ^2 , which are

$$E(m|\underset{\sim}{y}) = \sum_{m=1}^{n-1} m \cdot \pi(m|\underset{\sim}{y}), \tag{2.42}$$

$$E(\underset{\sim}{\beta}_1|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \alpha_1^*(m) \cdot \pi(m|\underset{\sim}{y}), \tag{2.43}$$

$$E(\underset{\sim}{\beta}_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \alpha_2^*(m) \cdot \pi(m|\underset{\sim}{y}), \tag{2.44}$$

$$E(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} [a^*/D(m)] \cdot \pi(m|\underset{\sim}{y}), \tag{2.45}$$

and

$$E(\sigma^2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} [D(m)/(a^*-1)] \cdot \pi(m|\underset{\sim}{y}) , \tag{2.46}$$

where $\pi(m|y)$, $\alpha_1^*(m)$, $\alpha_2^*(m)$, a^*, and $D(m)$ were previously given.

(ii) The covariance matrix of the posterior distributions of m, $\underset{\sim}{\beta}_1$, $\underset{\sim}{\beta}_2$, r and σ^2 are as follows:

$$Var(m|\underset{\sim}{y}) = \sum_{m=1}^{n-1} [m-E(m|\underset{\sim}{y})]^2 \cdot \pi(m|\underset{\sim}{y}), \tag{2.47}$$

$$Cov(\underset{\sim}{\beta}_1|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \frac{a^*}{a^*-1} [p_1^*(m)]^{-1} \cdot \pi(m|\underset{\sim}{y}), \tag{2.48}$$

$$Cov(\underset{\sim}{\beta}_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \frac{a^*}{a^*-1} [p_2^*(m)]^{-1} \cdot \pi(m|\underset{\sim}{y}), \tag{2.49}$$

$$Var(r|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \frac{a^*}{(D(m))^2} \cdot \pi(m|\underset{\sim}{y}), \tag{2.50}$$

and

$$Var(\sigma^2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} \frac{(D(m))^2}{(a^*-1)^2(a^*-2)} \cdot \pi(m|\underset{\sim}{y}), \tag{2.51}$$

where $\pi(m|\underset{\sim}{y})$ is the same as (2.13).

(2.42) to (2.51) have the form $\sum_{m=1}^{n-1} h(m) \cdot \pi(m|\underset{\sim}{y})$, which can be interpreted as the expected value of h(m) under the posterior distribution of m. The value h(m) is the expected value or the variance (or covariance matrix) of the posterior distribution of those unknown parameters when it is known that the shift point is at m.

For other special cases, the estimates of the unknown parameters can be found in a similar way. Bayesian confidence intervals, the regions of highest posterior density and tests of hypothesis about the switch point and the other unknown parameters may be obtained from their posterior p.d.f., respectively. For a more detailed discussion about Bayesian inferential and decision processes, the reader is referred to DeGroot (1970), Ferguson (1967) and Zellner (1971).

Numerical Example

In this section, an example is given to illustrate the method of estimating the shift point and all the unknown regression parameters for the most general case, where both regressions and the common variance are unknown. This example is for p=2 and uses the data generated by Quandt (1958) as shown in Table I of Appendix A. This data consists of a sequence of 20 observations which is generated from the following model:

$$y_i = 2.5 + 0.7x_i + e_i, \quad i = 1,\ldots,12$$

and

$$y_i = 5 + 0.5x_i + e_i, \quad i = 13,\ldots,20 \;,$$

where e_i's are i.i.d. $N(0,1)$.

Assume that the two phase regression model is

$$y_i = \alpha_1 + \beta_1 x_i + e_i, \quad i = 1,\ldots,M$$

$$y_i = \alpha_2 + \beta_2 x_i + e_i, \quad i = M+1,\ldots,n$$

where e_i's are i.i.d. $N(0,\sigma^2)$ and M, $\underset{\sim}{\beta} = (\alpha_1,\beta_1,\alpha_2,\beta_2)'$ and σ^2 are unknown.

The first part of the example is to illustrate the effect of various prior distributions on the posterior distribution of the shift point M. The second part of the example is to make inferences about all the unknown parameters.

<u>Sensitivity of the Posterior p.m.f. of M</u>

We assume that the assumption stated previously are valid for this data set and analyze it by using two sources of prior information. The first source is a data based prior and the second source is not data based. In order to obtain the data based prior it is assumed the shift point is near 12, and we group the first 9 consecutive observations into 3 sets and group the last 6 observations into 2 sets, i.e., 3 observations in each set. Based on the 3 observations in each set the regression analysis is performed for each set. The usual least squares estimator $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}^2$ are obtained. Then from the first 3 sets, the mean and variance of $\hat{\alpha}$ and of $\hat{\beta}$ are found, and the covariance between $\hat{\alpha}$ and $\hat{\beta}$ is calculated. These values are used for obtaining the prior parameters of the first regression. A similar procedure is done for the last 2 sets and the values obtained are used for the second regression. The numerical results are shown in Table II of Appendix A. From this table, we obtain $\underset{\sim}{\beta}_{\mu}$, the mean vector of $\underset{\sim}{\beta} = [\alpha_1, \beta_1, \alpha_2, \beta_2]'$ as

$$\underset{\sim}{\beta}_{\mu} = [2.7523,\ 0.5878,\ 6.1420,\ 0.4440] \ .$$

If we assume that the coefficients of the first regression are independent of the coefficients of the second regression, then we obtain the covariance matrix of $\underset{\sim}{\beta}$, as

$$\mathrm{Cov}(\underset{\sim}{\beta}) = \begin{bmatrix} 4.2926 & -0.5704 & & \\ -0.5704 & 0.0762 & \multicolumn{2}{c}{\phi} \\ & & 4.8878 & -0.4647 \\ \multicolumn{2}{c}{\phi} & -0.4647 & 0.0442 \end{bmatrix}$$

Since R has a gamma distribution with parameters a and b, $\sigma^2 = 1/R$ has an inverse gamma distribution with parameters a and b. By the method of moments, we obtain 2 estimators for a and b. One is based on the 5 values of $\hat{\sigma}^2$ in Table II of Appendix A and the other is based on the 5 values of $r = 1/\hat{\sigma}^2$, also in the same table. Based on the $\hat{\sigma}^2$'s, the estimates are $\hat{a}_1 = 3.5032$, $\hat{b}_1 = 1.0550$ and based on r's, the estimates are $\hat{a}_2 = 0.3625$, $\hat{b}_2 = 0.0226$. Since $\underset{\sim}{\beta}$ has a 4 variate t distribution with degrees of freedom 2a, location parameter $\underset{\sim}{\beta}_\mu$ and precision matrix a/b τ, it follows from the properties of the multivariate t distribution that $Cov(\underset{\sim}{\beta}) = [b/(a-1)]\tau^{-1}$ whenever $a > 1$. In this case when $a = \hat{a}_2 = 0.3625$, $b = \hat{b}_2 = 0.0226$, $Cov(\underset{\sim}{\beta})$ will not exist, hence $\hat{a}_2$ and $\hat{b}_2$ will not be considered as the prior parameters of a and b for the purposes of this study.

Based on $a = \hat{a}_1 = 3.5032$, $b = \hat{b}_1 = 1.0550$ and $Cov(\underset{\sim}{\beta})$ stated above, we obtain $\tau = [b/(a-1)]\ Cov(\underset{\sim}{\beta})^{-1}$, which is

$$\tau = \begin{bmatrix} 17.4064 & 130.2590 & & \\ 130.2590 & 980.3119 & & \phi \\ & & 422.1041 & 4439.3427 \\ & \phi & 4439.3427 & 46698.8884 \end{bmatrix}$$

The values for a, b, $\underset{\sim}{\beta}_\mu$, and τ complete the specification of the prior normal-gamma distribution. Using these values in (2.13), the posterior p.m.f. of M is calculated and shown in Table III of the Appendix A. These results show that the p.m.f. of M at m = 12 is 0.8728 which is an extremely high probability. No doubt, the shift point is at m = 12, which is the true shift point indicated by Quandt's data.

For the second prior the values of the parameters are specified. Since the joint prior distribution of $\underset{\sim}{\beta}$ and R is a multivariate normal-gamma distribution of the type stated in (2.5);

a) $\underset{\sim}{\beta}$ has a multivariate t distribution with 2a degrees of freedom, location vector $\underset{\sim}{\beta}_{\mu}$, and precision matrix $T(\underset{\sim}{\beta}) = (a/b)\underset{\sim}{\tau}$, and it follows that $Cov(\underset{\sim}{\beta}) = [b/(a-1)]\tau^{-1}$;

b) the precision R has a gamma distribution with parameters a and b, hence $E(R) = a/b$ and $Var(R) = a/b^2$;

c) The variance σ^2 has an inverse gamma distribution with parameters a and b, hence $E(\sigma^2) = b/(a-1)$, $Var(\sigma^2) = b^2/[(a-1)^2(a-2)]$.

Two experiments are conducted in order to test the sensitivity of the probability mass function of M, (2.13) .

Experiment 1. We specify the values of $\underset{\sim}{\beta}_{\mu}$, $T(\underset{\sim}{\beta})$, E(R) and Var(R), which are assumed to be:

1) $\underset{\sim}{\beta}_{\mu} = (2.5 , 0.7, 5, 0.5)'$

2) $T(\underset{\sim}{\beta}) = \lambda I$, where I is the 4x4 identity matrix and λ takes the values 0.01, 0.1, 1, 10, 100. Therefore, all the regression coefficients are uncorrelated.

3) $E(R) = 1$.

4) Var(R) varies and takes the values 0.01, 0.1, 1, 10, 100.

Once the values of $\underset{\sim}{\beta}_{\mu}$, $T(\underset{\sim}{\beta})$, E(R) and Var(R) are specified then the values for the prior parameters $\underset{\sim}{\beta}_{\mu}$, τ , a and b are determined. The combination of values for λ and Var(R) lead to 25 different prior distributions. Based on each prior, the p.m.f. of M in (2.13) is calculated and shown in the Tables IV through VIII of Appendix A.

Experiment 2. We specify the values of $\underset{\sim}{\beta}_{\mu}$, $Cov(\underset{\sim}{\beta})$, $E(\sigma^2)$ and $Var(\sigma^2)$, which are assumed to be:

1) $\underset{\sim}{\beta}_{\mu} = (2.5, 0.7, 5, 0.5)'$ which is the same as Experiment 1.

2) $Cov(\underset{\sim}{\beta}) = vI$ where I is a 4x4 identity matrix and v takes the values 0.01, 0.1, 1, 10, 100 .

3) $E(\sigma^2) = 1$.

4) $Var(\sigma^2) = 0.01, 0.1, 1, 10, 100$.

Once the values of $\underset{\sim}{\beta}_{\mu}$, $Cov(\underset{\sim}{\beta})$, $E(\sigma^2)$ and $Var(\sigma^2)$ are specified then the values for the prior parameter $\underset{\sim}{\beta}_{\mu}$, τ, a and b are selected. Hence the combinations of v and $Var(\sigma^2)$ lead to 25 different type prior distributions. For each prior, the p.m.f. of M is calculated and shown in the Tables IX through XIII of Appendix A.

The results from the Experiment 1 show that

1) the posterior p.m.f. of M has a peak at m = 12, regardless of the values of λ and Var(R), i.e., whether $\lambda = 0.01$ or 100 and Var(R) = 0.01 or 100,

2) when λ decreases, the probability at the end points m = 1 and m = 19 increases. It is more noticable when $\lambda = 0.01$ and $\lambda = 0.1$. The reason is that when $\lambda \to 0$ (i.e., $\tau = \lambda I$ approaches singularity) and $x(m)'x(m)$ is singular at m = 1 and m = 19, $x(m)'x(m)+\tau$ approaches singularity,

3) The posterior probability at m = 12 increases with an increase in Var(R).

The results from the Experiment 2 show that

1) the posterior p.m.f. of M has a peak at m = 12, regardless of the values of v and $Var(\sigma^2)$ when v takes values between 0.01 and 100 and $Var(\sigma^2)$ takes values between 0.01 and 100. The posterior probability of m = 12 increases very little as $Var(\sigma^2)$ increases from 1 to 100.

2) When v increases, the probability at the end points m = 1 and m = 19 increases, especially at m = 1. It is more noticeable when v = 100 and v = 10. The reason is that when v increases (i.e., $\tau = v^{-1}I$

approaches singularity) and $x(m)'x(m)$ is singular when $m = 1$ and $m = 19$, $x(m)'x(m) + \tau$ approaches a singular matrix.

3) The posterior probability at $m = 12$ increases with an increase in $Var(\sigma^2)$.

From the above results we conclude that if the prior is data based or otherwise, the shift point is at $m = 12$ using Quandt's data. This conclusion is very satisfying since the true switch point is at $m = 12$.

Point and Set Estimation

In this part of the example, we are emphasizing inferences about the unknown regression parameters. We assume that the prior value for Quandt's data is as follows: $\underset{\sim}{\beta}_{\mu} = (2.5, 0.7, 5, 0.5)'$, $\tau = I_4$, $a = 3$ and $b = 2$ (i.e., $E(\underset{\sim}{\beta}) = (2.5, 0.7, 5, 0.5)'$, $Var(\underset{\sim}{\beta}) = I_4$, $E(\sigma^2) = 1$ and $Var(\sigma^2) = 1$). From these prior values the p.m.f. of M has been shown in Table XI of Appendix A and the location estimates of M are:

Mode of Posterior Distribution = 12.00

Median of Posterior Distribution = 12.00

Mean of Posterior Distribution = 11.11 .

Although the mean is at 11.11, we are willing to say that the shift point is at 12.00 because the probability at $m = 12$ is 0.6844 and the probability at $m = 11$ is 0.0495.

Inferences about the unknown regression parameters can be made either from (1) the marginal posterior distribution or (2) the conditional posterior distribution when the shift point $m = 12$. We are going to make inferences from both distributions. Although previously the marginal posterior p.d.f. was obtained for each set of unknown parameters, the conditional posterior p.d.f. was not derived. From (2.16), (2.21) and (2.22), we can easily show that

(i) the conditional posterior p.d.f. of $\underset{\sim}{\beta}$ when $M = m$ is

$$\pi(\underset{\sim}{\beta}|\underset{\sim}{y},m) = t[\underset{\sim}{\beta};2p,2a^*,\underset{\sim}{\beta}^*(m),p(m)], \qquad \underset{\sim}{\beta} \in R^{2p} \tag{2.52}$$

(ii) the conditional posterior p.d.f. of R when $M = m$ is

$$\pi(r|\underset{\sim}{y},m) = g[r;a^*,D(m)], \qquad r > 0 \tag{2.53}$$

(iii) the conditional p.d.f. of σ^2 when $M = m$ is

$$\pi(\sigma^2|\underset{\sim}{y},m) = ig[\sigma^2;a^*,D(m)], \qquad \sigma^2 > 0 . \tag{2.54}$$

The point estimates and the highest posterior density (H.P.D.) regions will be obtained for each set of parameters by employing both marginal and conditional posterior distribution. For the definition and properties of the H.P.D. region, see the paper by Box and Tiao (1965).

(i) Point estimates and H.P.D. regions for α_1:

Let $\pi(\alpha_1|\underset{\sim}{y})$ and $\pi(\alpha_1|\underset{\sim}{y},m)$ denote the marginal posterior p.d.f. of α_1 and the conditional posterior p.d.f. of α_1 when $m = 12$. In order to compare the difference in making inferences between $\pi(\alpha_1|\underset{\sim}{y})$ and $\pi(\alpha_1|\underset{\sim}{y},m)$, the point estimates and the H.P.D. regions of content 0.90, 0.95, 0.99 are calculated and presented as follows:

	$\pi(\alpha_1\|\underset{\sim}{y})$	$\pi(\alpha_1\|\underset{\sim}{y},m)$
<u>Point estimates</u>		
mean	2.36	2.29
mode	2.32	2.29
median	2.35	2.29
variance	0.2541	0.1937
<u>H.P.D. regions</u>		
90%	(1.52, 3.19)	(1.56, 3.01)
95%	(1.34, 3.42)	(1.42, 3.16)
99%	(0.86, 3.95)	(1.11, 3.46)

In order to compare the prior knowledge with the posterior information, the prior p.d.f. $\pi_0(\alpha_1)$ of α_1 and $\pi(\alpha_1|\underset{\sim}{y})$, $\pi(\alpha_1|\underset{\sim}{y},m)$ are plotted and shown in Figure 1.

(ii) Point estimates and H.P.D. regions for β_1:

Similarly, we calculate the point estimates and the H.P.D. regions for the marginal posterior p.d.f. $\pi(\beta_1|\underset{\sim}{y})$ of β_1 and the conditional posterior p.d.f. $\pi(\beta_1|\underset{\sim}{y},m)$ of β_1 when m = 12. The results are

	$\pi(\beta_1\|\underset{\sim}{y})$	$\pi(\beta_1\|\underset{\sim}{y},m)$
Point estimates		
mean	0.67	0.69
mode	0.68	0.69
median	0.69	0.69
variance	0.0069	0.0017
H.P.D. regions		
90%	(0.58, 0.77)	(0.62, 0.75)
95%	(0.55, 0.80)	(0.61, 0.77)
99%	(0.02, 0.90)	(0.58, 0.80)

The prior p.d.f. $\pi_0(\beta_1)$, marginal posterior p.d.f. $\pi(\underset{\sim}{\beta}_1|\underset{\sim}{y})$ and the conditional posterior p.d.f. $\pi(\beta_1|\underset{\sim}{y},m)$ are plotted in Figure 2.

(iii) Point estimates and H.P.D. regions for α_2

The point estimates and the H.P.D. regions for the marginal posterior p.d.f. $\pi(\alpha_2|\underset{\sim}{y})$ and the conditional posterior p.d.f. $\pi(\alpha_2|\underset{\sim}{y},m)$ are as follows:

	$\pi(\alpha_2\|\underset{\sim}{y})$	$\pi(\alpha_2\|\underset{\sim}{y},m)$
Point estimates		
mean	5.34	5.52
mode	5.45	5.52
median	5.39	5.52
variance	0.3933	0.3617
H.P.D. regions		
90%	(4.15, 6.54)	(4.50, 6.50)
95%	(3.80, 6.72)	(4.33, 6.71)
99%	(3.20, 7.12)	(3.91, 7.12)

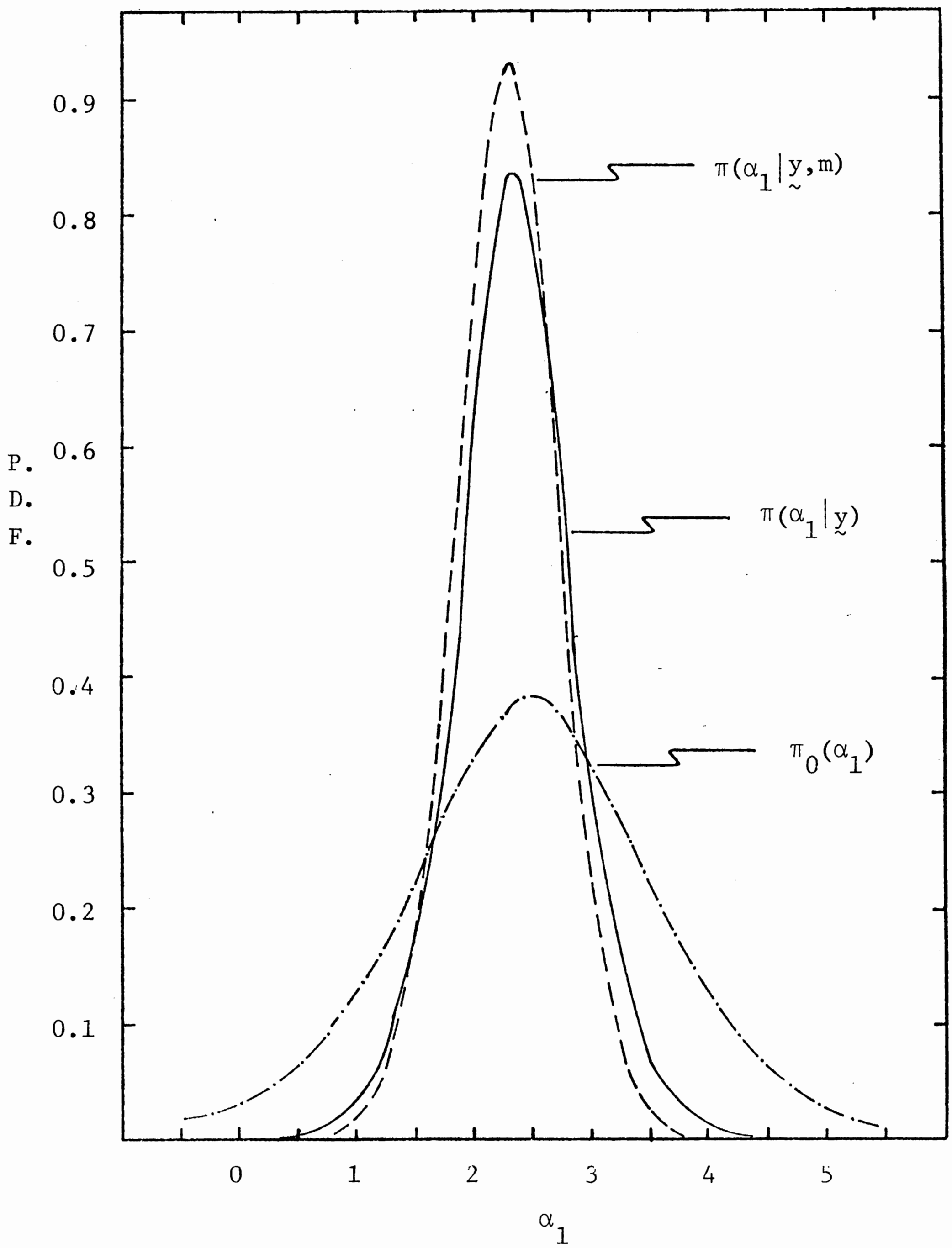

Figure 1. Prior, Marginal Posterior and Conditional Posterior P.D.F. of α_1

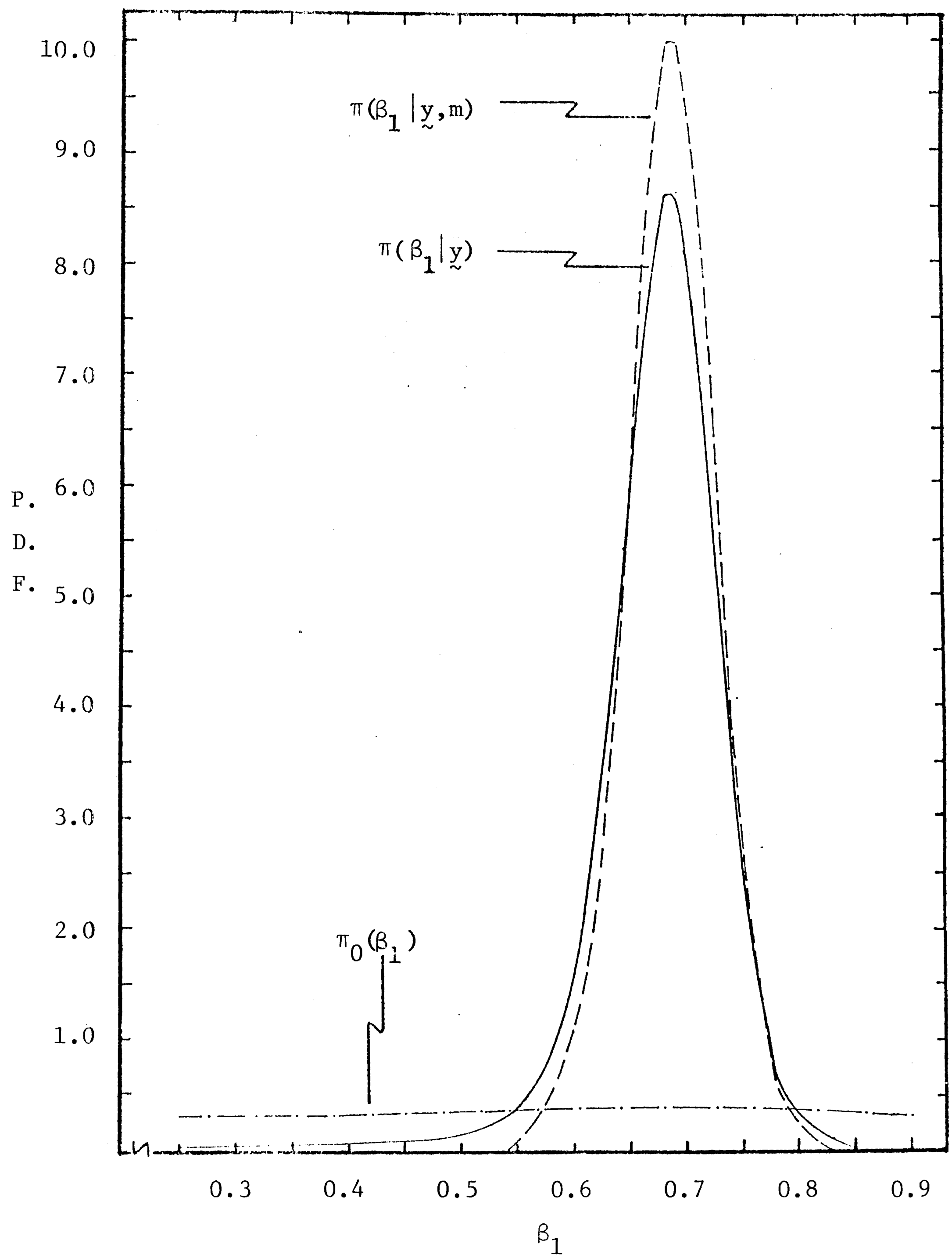

Figure 2. Prior, Marginal Posterior and Conditional Posterior P.D.F. of β_1

$\pi_0(\alpha_2)$, $\pi(\alpha_2|\tilde{y})$ and $\pi(\alpha_2|\tilde{y},m)$ are plotted and shown in Figure 3 .

(iv) Point estimates and H.P.D. regions for β_2:

The point estimates and H.P.D. regions for the marginal posterior p.d.f. of β_2 and the conditional posterior p.d.f. of β_2 when m = 12 are calculated and the results are as follows:

	$\pi(\beta_2\|\tilde{y})$	$\pi(\beta_2\|\tilde{y},m)$
Point estimates		
mean	0.52	0.51
mode	0.51	0.51
median	0.52	0.51
variance	0.0026	0.0024
H.P.D. regions		
90%	(0.43, 0.61)	(0.43, 0.59)
95%	(0.41, 0.63)	(0.41, 0.60)
99%	(0.37, 0.67)	(0.38, 0.64)

Similarly, the prior p.d.f. $\pi_0(\beta_2)$, and the posterior p.d.f. $\pi\ (\beta_2|\tilde{y})$, $\pi\ (\beta_2|\tilde{y},m)$ are plotted and shown in Figure 4.

(v) Point estimates and H.P.D. regions for R and σ^2

The estimates and H.P.D. regions for R and σ^2 are as follows:

	$\pi(r\|\tilde{y})$	$\pi(r\|\tilde{y},m)$
Point estimates		
mean	1.20	1.30
mode	1.09	1.20
median	1.17	1.27
variance	0.1126	0.1297
H.P.D. regions		
90%	(0.61, 1.78)	(0.71, 1.87)
95%	(0.54, 1.93)	(0.64, 2.02)
99%	(0.43, 2.26)	(0.51, 2.33)

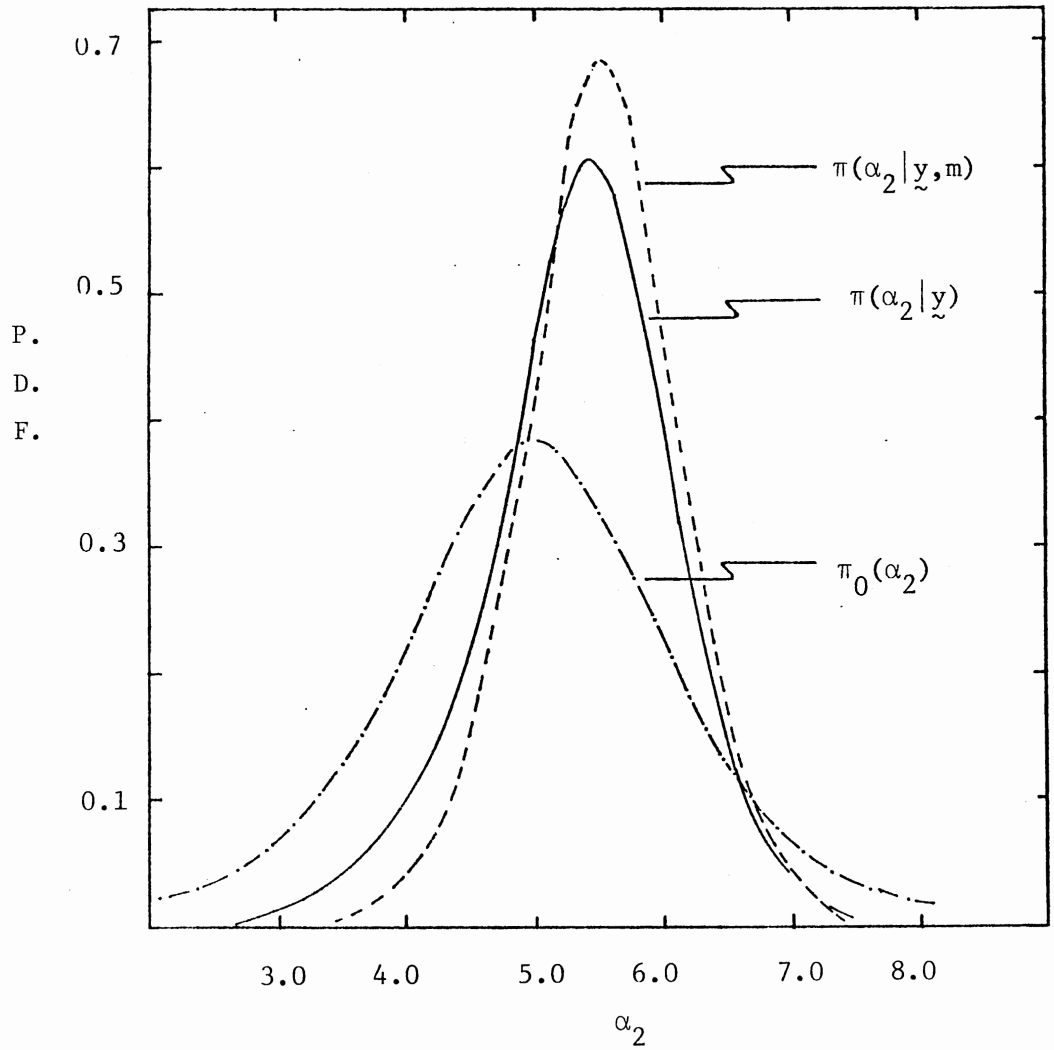

Figure 3. Prior, Marginal Posterior and Conditional Posterior P.D.F. of α_2

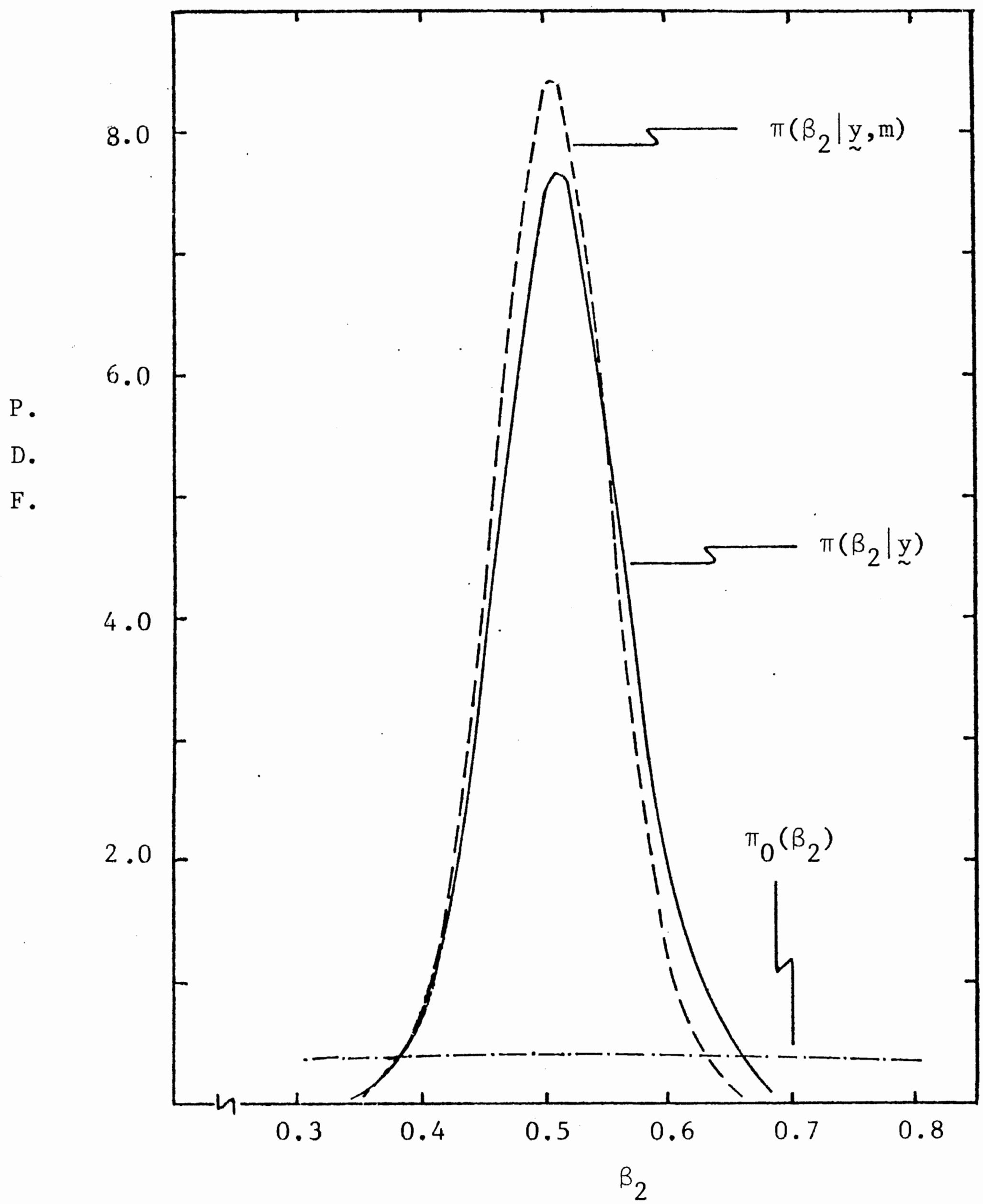

Figure 4. Prior, Marginal Posterior and Conditional Posterior P.D.F. of β_2

	$\pi(\sigma^2 \mid \underset{\sim}{y})$	$\pi(\sigma^2 \mid \underset{\sim}{y}, m)$
Point estimates		
mean	0.92	0.83
mode	0.76	0.72
median	0.86	0.80
variance	0.078	0.0633
H.P.D. regions		
90%	(0.47, 1.36)	(0.46, 1.20)
95%	(0.44, 1.55)	(0.43, 1.34)
99%	(0.38, 1.97)	(0.37, 1.66)

The prior, marginal posterior and conditional posterior p.d.f. of R and σ^2 are plotted and shown in Figure 5 and Figure 6.

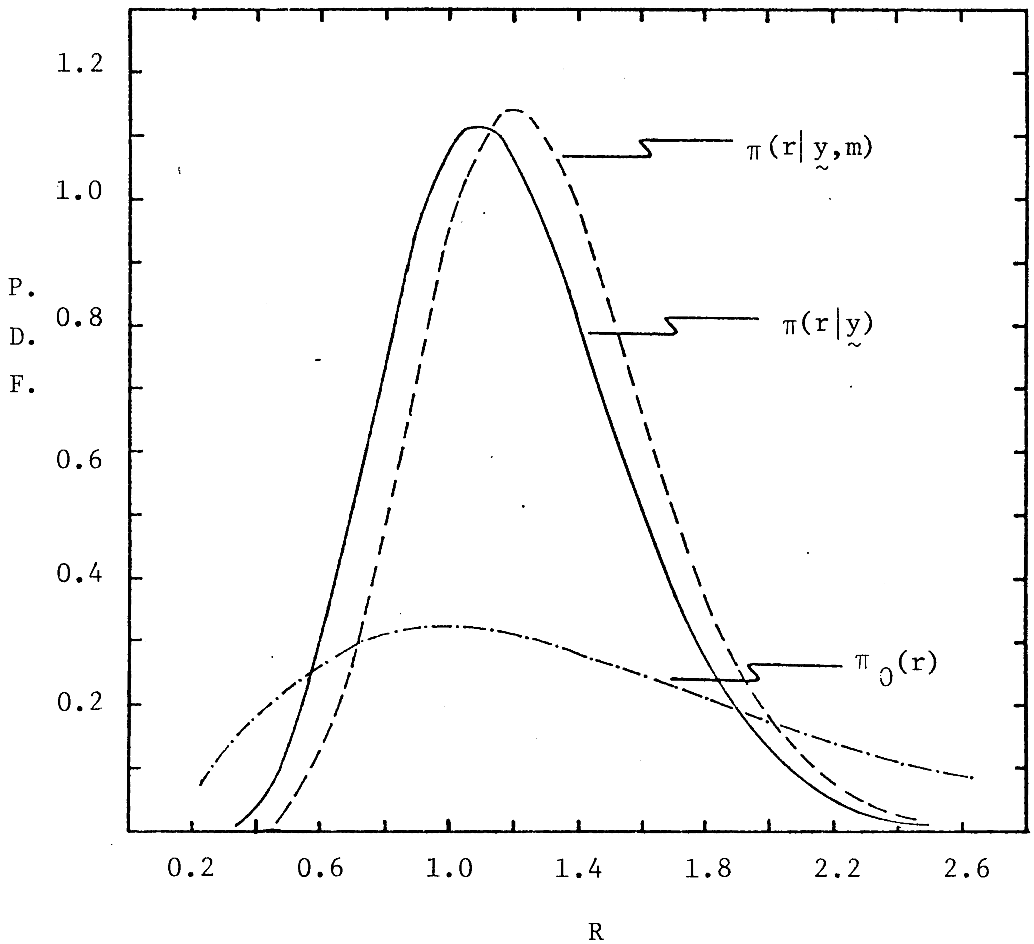

Figure 5. Prior, Marginal Posterior and Conditional Posterior P.D.F. of R

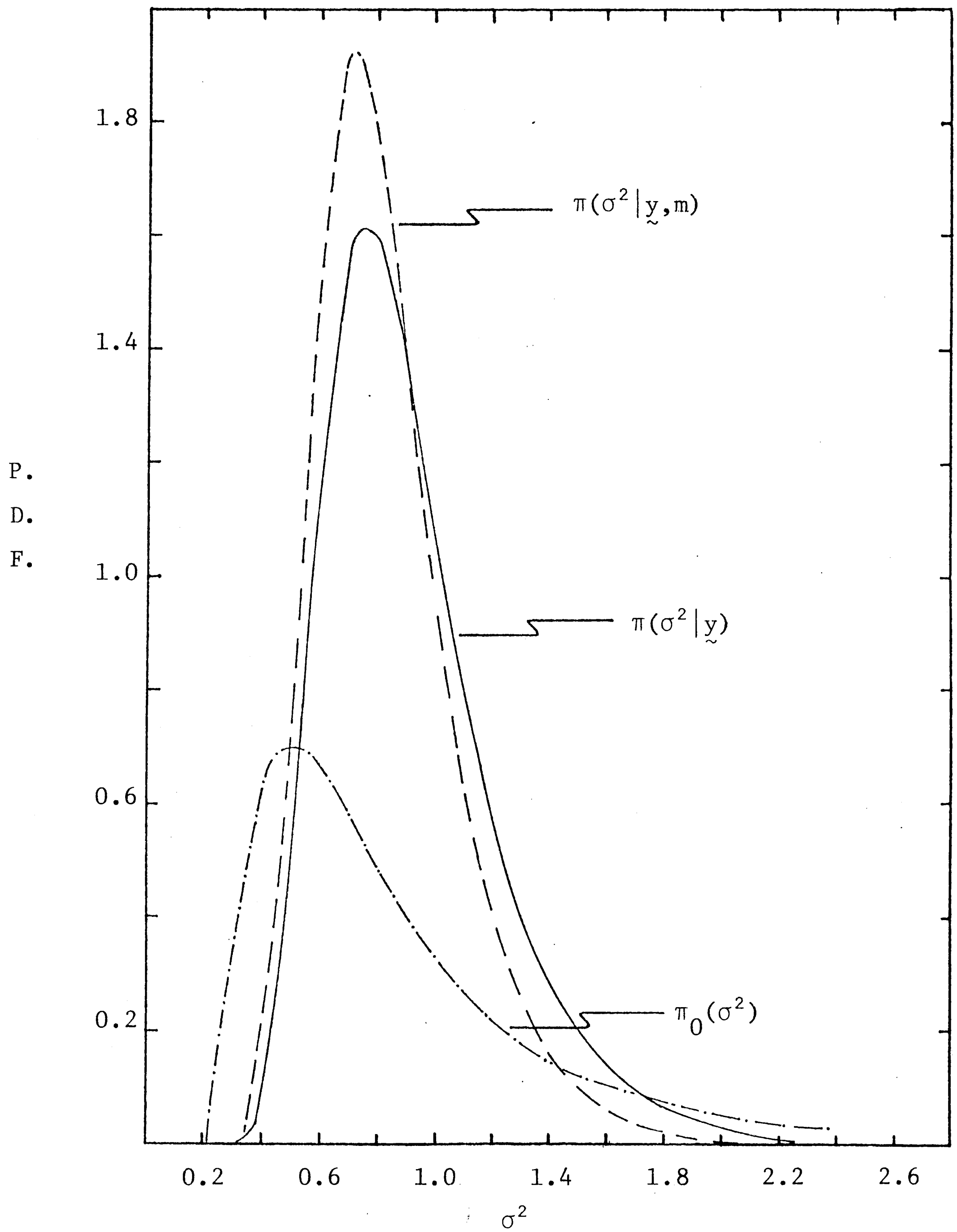

Figure 6. Prior, Marginal Posterior and Conditional Posterior P.D.F. of σ^2

CHAPTER III

INFERENCES ABOUT THE INTERSECTION OF TWO REGRESSION LINES

Suppose $y_1,\ldots,y_n$ is a sequence of random variables, such that

$$y_i = \alpha_1 + \beta_1 x_i + e_i, \quad i = 1,\ldots,M$$
$$= \alpha_2 + \beta_2 x_i + e_i, \quad i = M+1,\ldots,n \ , \tag{3.1}$$

where $e_i \sim N(0,\sigma^2)$, $i = 1,\ldots,n$ and $M = 1,\ldots,n-1$, thus, a change occurs once in this sequence of random variables. This model is a special case of the changing regression model stated in Chapter II with p=2. In Chapter II, we have derived the posterior distribution for M, $\underset{\sim}{\beta}$, R, and σ^2. In this chapter, we are interested in making inferences about the abscissa γ of the intersection point of two regression lines, therefore we need to find the posterior distribution of γ . From model (3.1), it is easy to show that $\gamma = (\alpha_2 - \alpha_1)/(\beta_1 - \beta_2)$ and is a function of the regression coefficients. Only the most general case where all parameters are unknown will be considered and a conjugate prior distribution will be employed. For other special cases, the derivation is the same and will not be discussed here.

Posterior Distribution of γ

Before we are able to find the posterior distribution of γ , we need to find the posterior distribution of $\underset{\sim}{\beta} = (\alpha_1,\beta_1,\alpha_2,\beta_2)$. When the prior distribution of $(M,\underset{\sim}{\beta},R)$ is a multivariate normal-gamma distribution, as specified by the relation (2.7), then the joint posterior p.d.f. of $\underset{\sim}{\beta} = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ is a mixture of multivariate t distributions stated in (2.14) with p=2. This means that $\underset{\sim}{\beta}$ has a posterior p.d.f. of

$$\pi(\underset{\sim}{\beta}|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{\beta};4,2a^*,\underset{\sim}{\beta}^*(m),p(m)] \cdot \pi(m|\underset{\sim}{y}) \ . \tag{3.2}$$

$\underset{\sim}{\beta}^*(m)$, $p(m)$ and $\pi(m|\underset{\sim}{y})$ are the same as (2.11), (2.12) and (2.13).

Consider the transformation

$$w_1 = (\alpha_2 - \alpha_1), \qquad \alpha_1, \alpha_2, \beta_1, \beta_2 \ \varepsilon \ R$$
$$w_2 = (\beta_1 - \beta_2)$$

which can be expressed as

$$\underset{\sim}{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \alpha_2 - \alpha_1 \\ \beta_1 - \beta_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \alpha_2 \\ \beta_2 \end{pmatrix} = T\underset{\sim}{\beta} \ . \tag{3.3}$$

In order to find the distribution of $\underset{\sim}{w}$, it is necessary to state a property of the multivariate t distribution. Suppose that a random vector $x = (x_1, \ldots, x_k)'$ has a k-dimensional multivariate t distribution with n degrees of freedom, location vector $\underset{\sim}{\mu}$, and precision matrix H and suppose A is an mxk matrix such that $AH^{-1}A'$ is nonsingular. Then the random vector $U = (U_1, \ldots, U_m)'$ defined as $U = AX$ has a m-dimensional multivariate t distribution with n degrees of freedom, location vector $A\underset{\sim}{\mu}$, and precision matrix $(AH^{-1}A')^{-1}$. From this property, the posterior p.d.f. of $\underset{\sim}{w} = (w_1, w_2)'$ is

$$\pi(\underset{\sim}{w}|\underset{\sim}{y}) = \sum_{m=1}^{n-1} t[\underset{\sim}{w};2,2a^*,\underset{\sim}{w}(m),V(m)] \cdot \pi(m|\underset{\sim}{y}) \tag{3.4}$$

where

$$\underset{\sim}{w}(m) = T\beta^*(m) \tag{3.5}$$

and

$$V(m) = [TP^{-1}(m)T']^{-1} \ . \tag{3.6}$$

Now consider the transformation

$$\gamma_1 = \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} = \frac{w_1}{w_2} \qquad \gamma_1, \gamma_2 \in R$$

$$\gamma_2 = \beta_1 - \beta_2 = w_2 \ . \tag{3.7}$$

Then the joint p.d.f. of γ_1 and γ_2 is

$$\pi(\gamma_1,\gamma_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} k(m)\{1 + (1/2a^*)[A(\gamma_1,m)\gamma_2^2 - 2B(\gamma_1,m)\gamma_2$$

$$+ C(m)]\}^{-(a^*+1)} \cdot \pi(m|\underset{\sim}{y}),$$

$$-\infty < \gamma_i < \infty, \ i = 1,2 \tag{3.8}$$

where

$$k(m) = \frac{|V(m)|^{1/2}\ \Gamma(a^*+1)}{2a^*\pi\Gamma(a^*)} \tag{3.9}$$

$$A(\gamma_1,m) = (\gamma_1,1)V(m)(\gamma_1,1)'$$

$$= \underset{\sim}{b}'(\gamma_1)V(m)\underset{\sim}{b}(\gamma_1) \ , \tag{3.10}$$

$$B(\gamma_1,m) = \underset{\sim}{b}'(\gamma_1)V(m)T\underset{\sim}{\beta}^*(m) \tag{3.11}$$

and

$$C(m) = [T\underset{\sim}{\beta}^*(m)]'V(m)[T\underset{\sim}{\beta}^*(m)] \ . \tag{3.12}$$

Before integrating (3.8) with respect to γ_2, the following algebraic manipulation needs to be done.

$$1 + (1/2a^*)[A(\gamma_1,m)\gamma_2^2 + 2B(\gamma_1,m)\gamma_2 + C(m)]$$

$$= 1 + (1/2a^*)\{A(\gamma_1,m)[\gamma_2 - B(\gamma_1,m)/A(\gamma_1,m)]^2 + C(m)$$

$$- B^2(\gamma_1,m)/A(\gamma_1,m)\}$$

$$= G(\gamma_1,m) + (A(\gamma_1,m)/2a^*)[\gamma_2 - B(\gamma_1,m)/A(\gamma_1,m)]^2 \tag{3.13}$$

where

$$G(\gamma_1,m) = 1 + (1/2a^*)[C(m) - B^2(\gamma_1,m)/A(\gamma_1,m)]. \tag{3.14}$$

Substituting (3.13) into (3.8), then

$$\pi(\gamma_1,\gamma_2|\underset{\sim}{y}) = \sum_{m=1}^{n-1} k(m)G(\gamma_1,m)^{-(a^*+1)}|\gamma_2|\{1 + [S(\gamma_1,m)/(2a^*+1)]$$

$$\cdot [\gamma_2 - Q(\gamma_1,m)]^2\}^{-[(2a^*+1)+1]/2}\pi(m|\underset{\sim}{y}) \quad (3.15)$$

where

$$Q(\gamma_1,m) = B(\gamma_1,m)/A(\gamma_1,m) \quad (3.16)$$

and

$$S(\gamma_1,m) = (2a^*+1)A(\gamma_1,m)/[2a^*G(\gamma_1,m)] \ . \quad (3.17)$$

It is shown in B.1. of Appendix B that $S(\gamma_1,m) > 0$. Integrating (3.15) with respect to γ_2, the posterior p.d.f. of γ_1 is

$$\pi(\gamma_1|\underset{\sim}{y}) = \frac{\Gamma[(2a^*+1)/2]}{(2\pi a^*)^{1/2}\Gamma(a^*)} \sum_{m=1}^{n-1} |V(m)|^{1/2} E|\gamma_0| A^{-1/2}(\gamma_1,m)$$

$$G^{-(a^*+1/2)}(\gamma_1,m) \cdot \pi(m|\underset{\sim}{y}), \quad \gamma_1 \in R^- \quad (3.18)$$

where γ_0 has a general t distribution with 2a*+1 degrees of freedom, location parameter $Q(\gamma_1,m)$, and precision parameter $S(\gamma_1,m)$. From the proof shown in B.2. of Appendix B,

$$E|\gamma_0| = \frac{(2a^*+1)^{1/2}\Gamma(a^*+1)}{a^*\sqrt{\pi}\ \Gamma[(2a^*+1)/2]} S^{-(1/2)}(\gamma_1,m)\{1 + [1/(2a^*+1)]$$

$$\cdot S(\gamma_1,m)Q^2(\gamma_1,m)\}^{-a^*} + Q(\gamma_1,m)\ \{2\psi_{(2a^*+1)}[Q(\gamma_1,m)$$

$$\cdot S^{1/2}(\gamma_1,m)] - 1\}\ , \quad (3.19)$$

where $\psi_{2a^*+1}(x)$ is the cumulative distribution function of a student t distribution with 2a*+1 degrees of freedom.

Thus, (3.18) and (3.19) complete the specification of the marginal posterior p.d.f. of γ. From the above derivation, it is easy to show that the conditional posterior p.d.f. of γ when $M = m$ is

$$\pi(\gamma|\underset{\sim}{y},m) = \frac{\Gamma[(2a^*+1)/2]}{(2\pi a^*)^{1/2}\Gamma(a^*)} |V(m)|^{1/2} E|\gamma_0| A^{-1/2}(\gamma_1,m)$$

$$\cdot G^{-[a^*+(1/2)]}, \tag{3.20}$$

where $E|\gamma_0|$ is defined as (3.19).

Although both (3.18) and (3.20) are not in an easily recognizable form, it is not difficult to compute the point estimators and interval estimators of γ with the aid of a computer. An illustration is followed by a numerical example.

Example

Data from Pool and Borchgrevink (1964) will be used and is shown in Table XIV of Appendix A. The independent variable X represents the logarithm of Warfarin concentration and the dependent variable Y is blood factor VII production. Hinkley (1971) and Holbert (1973) have used this data to illustrate the techniques which they developed. Their analyses are based on no prior information or vague type prior information, whereas our method is based on proper prior distribution. For purposes of illustration, assume that the values for the prior parameters are: $\underset{\sim}{\beta}_\mu = (0,0.2,0.95,0)$, $\tau = I_4$, $a = 2$ and $b = 0.0017$, i.e., $E\sigma^2 = 0.0017$ which is the estimate obtained by Hinkley (1971). From (2.13), the posterior p.m.f. of M is calculated and is shown in Table XIV of Appendix A. The location estimators for M are: Mode = 6.00, Mean = 6.13. From here we know that the shift index is at 6; i.e., the first 6 observations $x_1,\ldots,x_6$ follow the first regression line, whereas the remaining 9 observations $x_7,\ldots,x_{15}$ follow the second regression line. Now we are going to find the abscissa γ of the intersection of these two regression lines. When we derive the posterior and conditional p.d.f. of γ , we did not have the restriction that $x_m < \gamma < x_{m+1}$ as did Hinkley (1971). Hence γ is at the entire real line. It is seen that (3.18) and (3.20) do not yield an explicit form for the estimates. Therefore we need to use the definition of the estimation in order to be able to find the estimators. Due to the difficulty in evaluating an integral from $-\infty$ to ∞, the density given by (3.18) and (3.20) will be truncated over the interval [3.50, 6.50]. Since

$$\int_{3.5}^{6.5} \pi(\gamma|\underset{\sim}{y}) = 0.99999954$$

and

$$\int_{3.5}^{6.5} \pi(\gamma|\underset{\sim}{y},m) = 0.99999994$$

therefore the inferences based on this truncated p.d.f. will not lose any information. The marginal posterior p.d.f. $\pi(\gamma|\underset{\sim}{y})$ and the conditional posterior p.d.f. $\pi(\gamma|\underset{\sim}{y},m)$ when m=6 are plotted in Figure 7 and with the aid of the computer, the estimators are evaluated and shown in the following table.

	$\pi(\gamma\|\underset{\sim}{y})$	$\pi(\gamma\|\underset{\sim}{y},m)$
Point estimates		
Mode	4.81	4.79
Median	4.81	4.80
Mean	4.81	4.80
Variance	0.0258	0.0227
H.P.D. regions		
90%	(4.55, 5.07)	(4.56, 5.05)
95%	(4.49, 5.13)	(4.50, 5.11)
99%	(4.37, 5.26)	(4.41, 5.23)

The above results show that the estimator for M is m=6 and the estimator for γ is at approximately 4.81. The estimators for m and γ were calculated from the posterior distributions (2.13), (3.18) and (3.20) and no restriction was placed on the value of γ when the

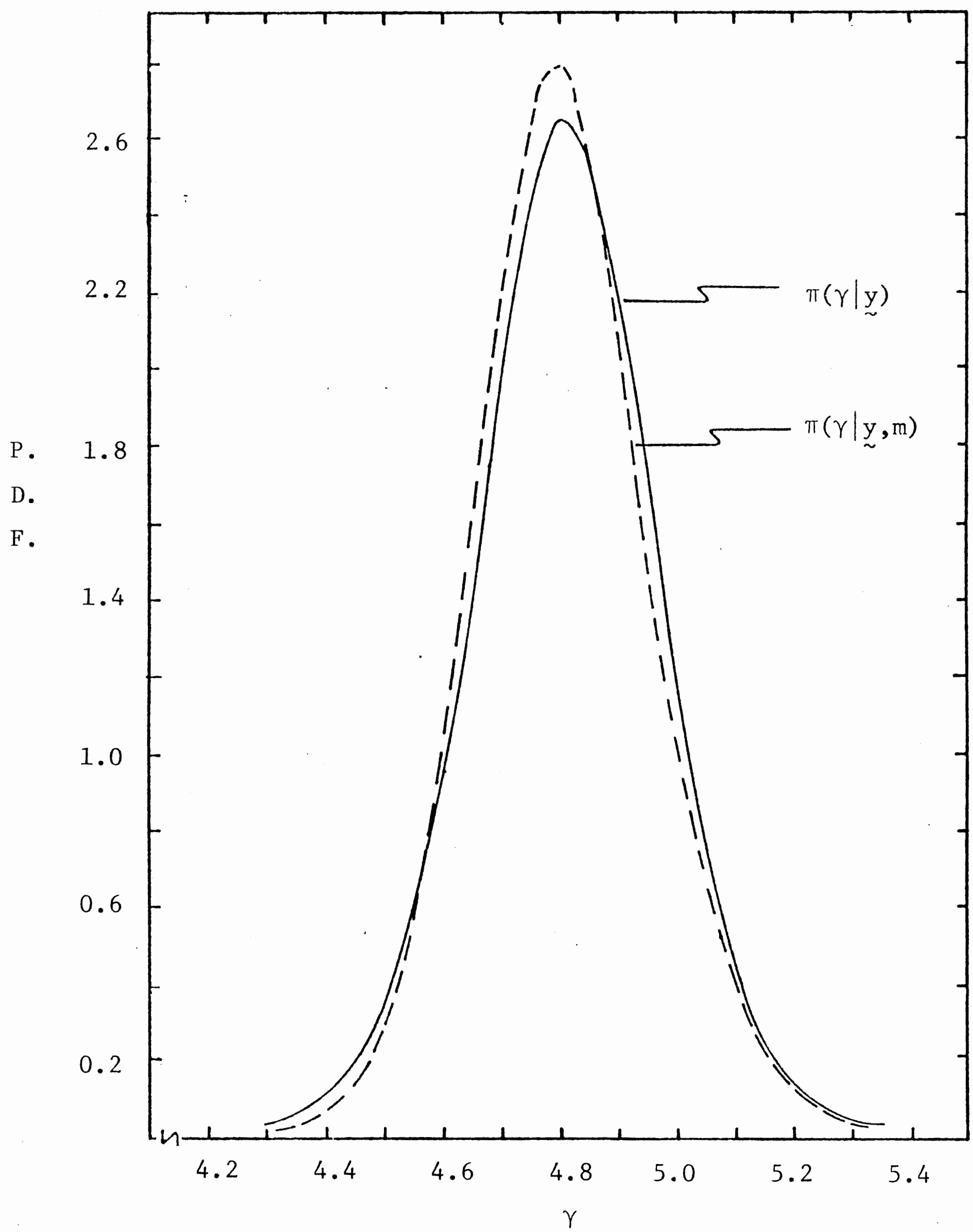

Figure 7. Marginal and Conditional Posterior P.D.F. of γ

posterior distributions of γ were computed. This method differs from that of Hinkley (1971) who restricted the value of γ between x_m and x_{m+1} in his model. Hence we can claim that our method can locate the point at which the regression model changes from one line to another when the changing model is continuous.

CHAPTER IV

THE DETECTION PROBLEM

Assume a sequence of independent, normally distributed random variables $Y_1,\ldots,Y_n$, such that

$$Y_i = \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1 + e_i, \quad i = 1,2,\ldots,\lambda$$

$$Y_i = \underset{\sim}{x}_i{}'\underset{\sim}{\beta}_2 + e_i, \quad i = \lambda+1,\ldots,n$$

where e_i's are i.i.d. $N(0,\sigma^2)$ and $\lambda = 1,\ldots,n$. When $\lambda = m$ ($m = 1,\ldots, n-1$), the first m observations are distributed $N(\underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1,\sigma^2)$ and the remaining n-m observations are distributed $N(\underset{\sim}{x}_i{}'\underset{\sim}{\beta}_2,\sigma^2)$. When $\lambda=n$, there is no change in the regression relationship in this sequence of random variables and all n observations are distributed $N(\underset{\sim}{x}_i{}'\underset{\sim}{\beta}_1,\sigma^2)$.

We need to construct a test for the null hypothesis, denoted by H_o, of no change versus the alternative hypothesis, denoted by H_1, of exactly one change, i.e.,

$$H_o: \lambda = n \qquad \text{versus} \qquad H_1: \lambda = m ,$$
$$m = 1,\ldots,n-1 .$$

We will consider only the most general case that both regressions are unknown and σ^2 is unknown. The procedure for testing the same hypothesis for the other cases can be constructed by the same technique.

Posterior Probability of 'No Change'

Since λ is unknown, we assign a prior p.m.f. for λ , which is

$$\pi_0(\lambda) = q \qquad \lambda = n$$
$$= \frac{1-q}{n-1} \qquad \lambda = m \ (m = 1,\ldots,n-1)$$

where q is a preassigned value by the researcher. The distribution of λ indicates that the prior probability of no change occurring is q and the remaining probability, uniformly distributed over the points 1,2,...,n-1, is the prior probability of exactly one change.

When $\lambda = n$, then $\underset{\sim}{\beta}_1$ and σ^2 are the unknown regression parameters. Assume that the joint p.d.f. of $(\underset{\sim}{\beta}_1,R)$ is a multivariate normal-gamma p.d.f., as stated by the following relation.

$$\pi_0(\underset{\sim}{\beta}_1,r) = \pi_0(\underset{\sim}{\beta}_1|r) \cdot \pi_0(r)$$
$$= (2\pi)^{-p/2}|r\tau_1|^{1/2} \exp[-\frac{r}{2}(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_{1\mu})'\tau_1(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_{1\mu})]$$
$$\cdot \frac{b^a}{\Gamma(a)} r^{a-1} e^{-br}$$
$$\propto r^{a+p/2-1} \exp\{(-r)[b+\frac{1}{2}(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_{1\mu})'\tau_1(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_{1\mu})]\}. \quad (4.2)$$

When $\lambda = m$, $\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1',\underset{\sim}{\beta}_2')$ and σ^2 are the unknown regression parameters and we assume that the joint p.d.f. $\pi_0(\underset{\sim}{\beta},r)$ of $\underset{\sim}{\beta}$ and R is a multivariate normal-gamma p.d.f., as specified by the relation (2.5).

The likelihood function consists of

$$L(\lambda=n,\underset{\sim}{\beta}_1,r) = (r/2\pi)^{n/2} \exp\{-\frac{r}{2}(\underset{\sim}{y}-x\underset{\sim}{\beta}_1)'(\underset{\sim}{y}-x\underset{\sim}{\beta}_1)\} \quad (4.3)$$

and

$$L(\lambda=m,\underset{\sim}{\beta},r) = (r/2\pi)^{n/2} \exp\{-\frac{r}{2}[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]'[\underset{\sim}{y}-x(m)\underset{\sim}{\beta}]\}, \quad (4.4)$$

where x is a nxp design matrix with corresponding vectors $\underset{\sim}{\beta}_1$, the px1 regression coefficient vector, and y, the nx1 observation vector.

Assume that λ is independent of the other regression parameters then by Bayes theorem

$$\pi_n(\lambda=n,\underset{\sim}{\beta}_1,r|y) \propto \pi_0(\lambda=n,\underset{\sim}{\beta}_1,r) \cdot L(\lambda=n,\underset{\sim}{\beta}_1,r) \tag{4.5}$$

$$\propto q(2\pi)^{-p/2}|\tau_1|^{1/2}\ r^{(n+p)/2+a-1}$$

$$\cdot \exp\{-r[D(n)+\frac{1}{2}(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_1*)'(x'x+\tau_1)(\underset{\sim}{\beta}_1-\underset{\sim}{\beta}_1*)]\},$$

where

$$\underset{\sim}{\beta}_1* = (x'x+\tau_1)^{-1}(\tau_1\underset{\sim}{\beta}_{1\mu}+x'\underset{\sim}{y}) \tag{4.6}$$

$$D(n) = b + \frac{1}{2}[\underset{\sim}{y}'\underset{\sim}{y} + \underset{\sim}{\beta}_{1\mu}'\tau_1\underset{\sim}{\beta}_{1\mu} - \underset{\sim}{\beta}_1*'(x'x+\tau_1)\underset{\sim}{\beta}_1*], \tag{4.7}$$

and

$$\pi_n(\lambda=m,\underset{\sim}{\beta},r|\underset{\sim}{y}) \propto \pi_0(\lambda=m,\underset{\sim}{\beta},r) \cdot L(\lambda=m,\underset{\sim}{\beta},r)$$

$$\propto (\frac{1-q}{n-1})(2\pi)^{-p}|\tau|^{1/2}\ r^{(n/2)+p+a-1}\ \exp\{(-r)\{D(m)$$

$$+ \frac{1}{2}[\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]'[x(m)'x(m)+\tau][\underset{\sim}{\beta}-\underset{\sim}{\beta}*(m)]\}\}. \tag{4.8}$$

$\underset{\sim}{\beta}*(m)$ and $D(m)$ were given by (2.11) and (2.12), respectively. From (4.5) we obtain the posterior p.m.f. of λ when $\lambda=n$

$$\pi_n(\lambda=n|\underset{\sim}{y}) = \int_{R^p}\int_0^\infty \pi_n(\lambda=n,\underset{\sim}{\beta}_1,r)\,dr\,d\underset{\sim}{\beta}_1$$

$$\propto q|\tau_1|^{1/2}D(n)^{-(n/2+a)}|x'x+\tau_1|^{-1/2} \tag{4.9}$$

which is the posterior probability of 'no change', and from (4.8) we obtain the posterior p.m.f. of λ when $\lambda = m$.

$$\pi_n(\lambda=m|\underset{\sim}{y}) = \int_{R^{2p}}\int_0^\infty \pi_n(\lambda=m,\underset{\sim}{\beta},r)\,dr\,d\underset{\sim}{\beta}$$

$$\propto (\frac{1-q}{n-1})|\tau|^{1/2}\ D(m)^{-(n/2+a)}|x'(m)x(m)+\tau|^{-1/2}$$

which is the posterior probability of a change at $\lambda=m$. (4.9) and (4.10) complete the specification of the p.m.f. of λ .

Consider a test of H_0: $\lambda=n$ versus H_1: $\lambda\neq n$, where one makes the decision from the posterior probability of no change or one uses the posterior odds ratio. The posterior odds in favor of H_0, Ω_n, is given by

$$\Omega_n = \frac{\pi_n(\lambda=n|\underset{\sim}{y})}{\pi_n(\lambda\neq n|\underset{\sim}{y})} = \frac{\pi_n(\lambda=n|\underset{\sim}{y})}{1-\pi_n(\lambda=n|\underset{\sim}{y})}$$

$$= \frac{q|\tau_1|^{1/2}D(n)^{-(a+n/2)}|x'x+\tau_1|^{-1/2}}{(1-q)|\tau|^{1/2}\sum_{m=1}^{n-1} D(m)^{-(a+n/2)}|x(m)'x(m)+\tau|^{-1/2}} \tag{4.11}$$

When $\pi_n(\lambda=n|\underset{\sim}{y}) \leq k_1$ or $\Omega_n \leq \frac{k_1}{1-k_1} = k_2$, we reject the hypothesis H_o of no change. Otherwise, we accept H_0. k_1, k_2 are pre-assigned constants specified by the researcher. Clearly, larger values of $\pi_n(\lambda=n|\underset{\sim}{y})$ indicate that H_0 is more tenable.

An Informal Sequential Procedure

Another procedure for detecting a change is the informal sequential method of Smith (1975) for testing a location parameter change in a sequence of random variables. Consecutively one takes the first t observations $y_1,y_2,\ldots,y_t$ from $y_1,y_2,\ldots,y_n$ as a set of observations, where $t=2,\ldots,n$. For each set of observations, $y_1,\ldots,y_t$, one assumes the same joint prior p.d.f. as stated in (4.2) and (2.5) for the unknown regression parameters, and assigns two types of prior distributions to λ (where $\lambda=1,\ldots,t$)

$$\pi_0(\lambda) = q, \quad \lambda = t$$

$$= \frac{1-q}{t-1}, \ \lambda = 1,\ldots,t-1 \ , \tag{4.12}$$

and

$$\pi_0(\lambda) = \frac{1}{t} \ , \quad \lambda = 1,\ldots,t \ . \tag{4.13}$$

Thus, (4.12) indicates that the prior probability of no change is q regardless of the number of observations in each set. The remaining probability is the prior probability of exactly one change and is uniformly distributed over the points $1,\ldots,t-1$. Also, (4.13) is a special case of (4.12) and indicates the prior probability of no change is $1/t$, i.e., the larger the number of observations the smaller the prior probability of no change.

For each set of t observations, $\underset{\sim}{y}_t = (y_1, y_2, \ldots, y_t)$ and using the appropriate prior, one may calculate $\pi_t(\lambda=t|\underset{\sim}{y}_t)$ and Ω_t, and plot $\pi_t(\lambda=t|\underset{\sim}{y}_t)$ or Ω_t versus t. If the plot reveals a downward trend, this indicates that a change has occurred. A numerical example will be given to illustrate this in the next section.

A Numerical Example

For Quandt's data we assume that $\underset{\sim}{\beta}_\mu = (2.5, 0.7, 5, 0.5)'$, $\tau = I_4$, $a = 1$ and $b = 1$. For four different values of the prior probability of no change, q, the posterior p.m.f. of λ, $\pi_n(\lambda|\underset{\sim}{y})$, were calculated and are shown in Table XV of Appendix A. For $q = 0.05$, $\pi_n(\lambda|\underset{\sim}{y})$ has a peak at $\lambda = 12$ and the posterior probability of no change, $\pi_n(n|\underset{\sim}{y})$, is 0.032. One would reject H_0 if he assigns $k_1 = 0.05$ and accept H_0 if he assigns $k_1 = 0.01$. When $q = 0.50$, 0.95 and 0.99, the corresponding $\pi_n(n|\underset{\sim}{y})$ is 0.3855, 0.9226, 0.9842. If $k_1 = 0.05$, one would always accept H_0 based on these high prior values of q. It is a reasonable result because (4.9) implies that the posterior probability of no change is somewhat sensitive to the value of q, the prior probability of no change. But, for all values of q the results show that the posterior probability of no change is less than the prior probability of no change and this is due to adding the information from the data. Hence, the data indicate the existence of the change.

In order to illustrate the informal sequential procedure for testing the hypothesis, Table XVI of Appendix A presents the posterior probability of no change calculated on the basis of the first t observations ($t = 2,\ldots,n$), with the prior probability stated in (4.12) and (4.13). For all prior values of $\pi_0(t)$, $\pi_t(t|\underset{\sim}{y})$ peaks at t=3 and t=11. Although when t=12, $\pi_t(t|\underset{\sim}{y})$ decreases, but the magnitude of the decrease is not as high as when t=13. For q=0.95 and 0.99, $\pi_t(t|\underset{\sim}{y})$ changes slightly with changing t, the number of observations. The result indicates that the instability of the regression relationship in this sequence of random variables, $y_1,y_2,\ldots,y_{20}$ occur after t=3 and t=11. When $q = 0.05$, the instability is most evident after $t = 12$.

CHAPTER V

SUMMARY

The main objectives of this paper are to use a proper prior for (1) detecting the presence of a change from one regression model to another; (2) estimating and making inferences about the switch point and the unknown regression parameters in a sequence of independent random variables which change regression models at an unknown point; and (3) estimating and making inferences about the abscissa of the intersection of two regression lines.

The advantage of Bayesian approach using proper prior distributions is that one can get the exact distribution for the shift point and all the unknown parameters even when the sample size, n, is less than the number, p, of regression coefficients. Although one may complain about the restriction which one places on the family of prior distributions and claim that such a restriction is perhaps unrealistic, it deserves consideration because the experimenter may have good reason for having faith in such a prior distribution. When we use a conjugate prior distribution which can represent our prior information as accurately as it can be, the mathematical operations are easier to perform. When one does not know the prior values for the prior distribution, estimates can be found by the empirical Bayes method from past observations. The example of Chapter II gave some illustrations for estimating the prior values when one has several available sets of past observations. Although we have used the present data to estimate the prior values, this was for purposes of illustration. When only one set of past observations is available, then we should (1) find the marginal distribution of the dependent variable Y and (2) estimate the prior values by using the method of moments or the maximum likelihood estimates from past observations. For more details the reader is referred to Maritz (1970).

Suppose that one wishes to represent vague type prior information about unknown regression parameters. For example, for the most general case where both regressions are unknown and σ^2 is unknown. In this case one wants to find the posterior distributions for the unknown parameters in the changing regression model. If $n \geq 2p + 1$ and 2px2p matrix $x(m)'x(m)$ is nonsingular, we can let the parameter space of M be $I_M = (p, p+1, \ldots, n-p)$ and let $\tau \to 0$, $a \to -p$ and $b \to 0$ in the posterior distributions of M, $\tilde{\beta}$, R (or σ^2) and γ given by this paper. Then the same limiting posterior distribution will be obtained from an improper prior namely a joint density function of the following form:

$$\pi_0(\tilde{\beta}, R) = 1/R \quad \text{for } \tilde{\beta} \in R^{2p} \quad \text{and} \quad R > 0$$

when p=2, it is a special case of the limiting posterior distributions. These limiting posterior distributions are the same as the posterior distributions given by Holbert (1973), Ferreira (1975) and Holbert and Broemeling (1977). These workers investigated a two phase simple linear regression model by using vague type prior distributions.

In this paper we have assumed that both regression models have equal variance. This assumption is appropriate in a situation in which experiments are conducted under well controlled conditions which insure constancy of the variances of random disturbances in all experiments, whereas in some situations, this assumption is not satisfied, one may extend this study to the following two cases; (1) two regressions have unequal variance, i.e., $e_i \sim N(0, \sigma_1^2)$, $i = 1, \ldots, m$, and

$e_i \sim N(0, \sigma_2^2)$, $i = m+1,\ldots,n$, where $\sigma_1^2 \neq \sigma_2^2$, (2) all random variables $y_1,\ldots,y_n$ have unequal variances, i.e., $e_i \sim N(0,\sigma_i^2)$, $i = 1,\ldots,n$ where $\sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_n^2$. When both regressions are unknown and σ^2 is unknown, one may approach the problem by employing two multivariate normal gamma prior distribitions for the first case and a multivariate normal-Wishart prior distribution for the second case.

Most of the posterior distributions derived in this paper are mixtures of well known distributions, namely, normal, t, and gamma distribution. One can express the mean and variance of the mixture distributions in an explicit form and can easily calculate them, as shown in Chapter II. No direct formula exists for the computation of H.P.D. regions of mixtures of distributions. Thus H.P.D. regions shown in the examples of this paper were found with the aid of the computer. More investigations are needed on the properties of the mixture distributions.

In this paper, we assumed that if the change did occur, it occured once. One can extend the problem to the case when there are k changes in a sequence of random variables which are subjected to changing regression models at k unknown points. Also the problem can be extended to the multivariate case where at each time point, one observes more than one variate, say s variates, then the observation matrix is $n \times s$ instead of an $n \times 1$ observation vector.

There are two other major problems which can be studied in the area of changing regression models. One is the development of a sequential procedure to detect the change from sequential sampling. The other is the problem of predicting future observation of the sequence.

All calculations for the examples shown in this paper were done in double precision on an IBM 370/158 Computer at Oklahoma State University Computer Center.

SELECTED BIBLIOGRAPHY

Bacon, D.W. and D.G. Watts. "Estimating the Transition Between Two Intersecting Straight Lines." Biometrika, Vol. 58, (1971), 524-534.

Barnett, V. Comparative Statistical Inference. New York, New York: John Wiley and Sons, 1973.

Bhattacharyya, G.K. and R.A. Johnson. "Non Parametric Tests for Shift at an Unknown Time Point." Annals of Mathematical Statistics, Vol. 39, (1968), 1731-1734.

Box, G.E.P. and G.C. Tiao. "Multiparameter Problems From a Bayesian Point of View." Annals of Mathematical Statistics, Vol. 36, (1965), 1468-1482.

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

Broemeling, L.D. "Bayesian Procedures for Detecting a Change in a Sequence of Random Variables." Metron, Vol. XXX-N-1-4, (1972), 1-14.

Brown, R.L., J. Durbin, and J.M. Evans. "Techniques for Testing the Constancy of Regression Relationships Over Time (With Discussion)." Journal of the Royal Statistical Society, Series B, (1975), 149-92.

Dawid, A.P., M. Stone, and J.V. Zidek. "Marginalization paradoxes in Bayesian and Structural Inference (With Discussion)." Journal of the Royal Statistical Society, Series B, (1973), 189-253.

DeGroot, M.H. Optimal Statistical Decisions. New York, New York: McGraw Hill, 1970.

Farley, J.U. and M.J. Hinich. "A Test for a Shifting Slope Coefficient in a Linear Model." Journal of the American Statistical Association, Vol. 65, (1970), 1320-1329.

Farley, J.U., M.J. Hinich and T.W. McGuire. "Some Comparison of Tests for a Shift in the Slopes of a Multivariate Linear Time Series Model." J. Econometrics, Vol. 3, (1975), 297-318.

Ferguson, T.S. Mathematical Statistics - A Decision Making Approach, New York, New York: Academic Press, 1967.

Ferreira, P.E. "A Bayesian Analysis of a Switching Regression Model: Known Number of Regimes." Journal of the American Statistical Association, Vol. 70, (1975), 370-374.

Garbade, K. "Two Methods for Examining the Stability of Regression Coefficients." Journal of the American Statistical Association, Vol. 72, (1977), 54-63.

Hinkley, D.V. "Inference About the Intersection in Two-Phase Regression." Biometrika, Vol. 56, (1969), 495-504.

Hinkley, D.V. "Inference in Two-Phase Regression." Journal of the American Statistical Association, Vol. 66, (1971), 736-743.

Holbert, D. and L.D. Broemeling. "Bayesian Inference Related to Shifting Sequences and Two Phase Regression." Communications in Statistics, Vol. 6, (1977), 265-275.

Holbert, D. "A Bayesian Analysis of Shifting Sequences with Applications to Two Phase Regression. (Ph.D. thesis, Oklahoma State University, Stillwater, Oklahoma, 1973.)

Hudson, D.J. "Fitting Segmented Curves Whose Joint Points Have to be Estimated." Journal of the American Statistical Association, Vol. 61, (1966), 1097-1129.

Lindley, D.V. Introduction to Probability and Statistics From a Bayesian Viewpoint. Cambridge University Press, Cambridge, 1965.

Maritz, J.S. Empirical Bayes Methods. London and Colchester: Methuen and Co. Ltd., 1970.

Millar, B.D. and O.T. Denmead. "Water Relations of Wheat Leaves in the Field." Agronomy Journal, Vol. 68, (1976), 303-307.

Ohki, K. "Maganese Nutrition of Cotton Under Two Boron Levels. II.Critical Mn. Level." Agronomy Journal, Vol. 66, (1974), 572-575.

Pool, J. and C.F. Borchgrevink. "Comparison of Rat Liver Response to Coumarin Administered in Vivo Versus in Vitro." American Journal of Physiology, Vol. 206(1), (1964), 229-238.

Quandt, R.E. "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes." Journal of the American Statistical Association, Vol. 53, (1958), 873-880.

Quandt, R.E. "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes." Journal of the American Statistical Association, Vol. 55, (1960), 324-330.

Quandt, R.E. "A New Approach to Estimating Switching Regressions." Journal of the American Statistical Association, Vol. 67, (1972), 306-310.

Robinson, D.E. "Estimates for the Points of Intersection of Two Polynomial Regressions." Journal of the American Statistical Association, Vol. 59, (1964), 214-224.

Sims, J.L., W.D. Atkinson, and C. Smitobol. "Mo and N Effects on Growth, Yield, and Mo Composition of Burley Tobacco." Agronomy Journal, Vol. 67, (1975), 824-827.

Smith, A.F.M. "A Bayesian Approach to Inference About a Change-Point in a Sequence of Random Variables." Biometrika, Vol. 62, (1975), 407-416.

Sprent, P. "Some Hypothesis Concerning Two-Phase Regression Lines." Biometrics, Vol. 17, (1961), 634-645.

Zellner, A. An Introduction to Bayesian Inference in Econometrics. New York, New York: John Wiley and Sons, 1971.

APPENDIX A

TABLES

TABLE I

QUANDT'S DATA SET

Obs. No. (i)	1	2	3	4	5	6	7	8	9	10
x_i	4	13	5	2	6	8	1	12	17	20
y_i	3.473	11.555	5.714	5.710	6.046	7.650	3.140	10.312	13.353	17.197

Obs. No. (i)	11	12	13	14	15	16	17	18	19	20
x_i	15	11	3	14	16	10	7	19	18	9
y_i	13.036	8.264	7.612	11.802	12.551	10.296	10.014	15.472	15.65	9.871

TABLE II

LEAST SQUARES ESTIMATORS FOR EACH OF THE FIVE SETS

Set No.	Observations No. Contained in each set		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{r} = 1/\hat{\sigma}^2$
1	(1, 2, 3)		0.8009	0.8336	1.0007	0.9993
2	(4, 5, 6)		4.9266	0.2891	0.5892	1.6973
3	(7, 8, 9)		2.5294	0.6406	0.0144	69.2712
		Mean =	2.7523	0.5878		
		Var. =	4.2926	0.0762		
			$Cov(\hat{\alpha},\hat{\beta}) = -0.5704$			
4	(15,16,17)		7.7053	0.2953	0.2043	4.8960
5	(18,19,20)		4.5787	0.5925	0.2993	3.3407
		Mean =	6.1420	0.4440		
		Var. =	4.8878	0.0442		
			$Cov(\hat{\alpha},\hat{\beta}) = -0.4647$			
					$\hat{a}_1 = 3.5027$	$\hat{a}_2 = 0.3623$
					$\hat{b}_1 = 1.0550$	$\hat{b}_2 = 0.0226$

TABLE III

POSTERIOR DISTRIBUTION OF M FOR DATA BASED PRIOR

M	1	2	3	4	5	6	7	8	9	10
p.m.f	0.0002	0.0002	0.0004	0.0000	0.0002	0.0006	0.0117	0.0234	0.0279	0.0280

M	11	12	13	14	15	16	17	18	19	
p.m.f	0.0244	0.8728	0.0017	0.0022	0.0039	0.0019	0.0001	0.0001	0.0002	

TABLE IV

POSTERIOR DISTRIBUTION OF M WHEN λ=0.01,E(R)=1 AND $\underset{\sim}{\beta}_{\mu}$ = (2.5,0.7,5,0.5)'

m \ Var(R)	0.01	0.1	1	10	100
1	0.3505	0.3480	0.3410	0.3390	0.3389
2	0.0159	0.0158	0.0155	0.0154	0.0154
3	0.0132	0.0131	0.0128	0.0128	0.0127
4	0.0010	0.0013	0.0016	0.0017	0.0017
5	0.0021	0.0024	0.0027	0.0027	0.0027
6	0.0034	0.0037	0.0039	0.0039	0.0040
7	0.0186	0.0177	0.0167	0.0165	0.0164
8	0.0274	0.0254	0.0234	0.0230	0.0230
9	0.0279	0.0257	0.0235	0.0230	0.0230
10	0.0362	0.0331	0.0301	0.0295	0.0295
11	0.0347	0.0318	0.0289	0.0283	0.0283
12	0.4077	0.4173	0.4323	0.4360	0.4364
13	0.0054	0.0055	0.0054	0.0054	0.0054
14	0.0078	0.0077	0.0075	0.0074	0.0074
15	0.0169	0.0159	0.0149	0.0146	0.0146
16	0.0146	0.0139	0.0132	0.0130	0.0130
17	0.0015	0.0019	0.0022	0.0023	0.0023
18	0.0022	0.0028	0.0032	0.0033	0.0033
19	0.0129	0.0172	0.0211	0.0219	0.0220

TABLE V

POSTERIOR DISTRIBUTION OF M WHEN λ=0.1,E(R)=1 AND $\underset{\sim}{\beta}_{\mu}$ = (2.5,0.7,5,0.5)'

m \ Var(R)	0.01	0.1	1	10	100
1	0.1454	0.1458	0.1436	0.1429	0.1428
2	0.0151	0.0155	0.0155	0.0155	0.0155
3	0.0151	0.0142	0.0150	0.0150	0.0149
4	0.0013	0.0017	0.0021	0.0022	0.0022
5	0.0027	0.0032	0.0036	0.0037	0.0037
6	0.0045	0.0050	0.0053	0.0053	0.0053
7	0.0259	0.0246	0.0232	0.0229	0.0228
8	0.0381	0.0354	0.0326	0.0320	0.0320
9	0.0389	0.0358	0.0328	0.0321	0.0321
10	0.0500	0.0458	0.0417	0.0408	0.0407
11	0.0480	0.0440	0.0400	0.0392	0.0391
12	0.5459	0.5580	0.5745	0.5784	0.5789
13	0.0072	0.0073	0.0072	0.0072	0.0072
14	0.0103	0.0102	0.0099	0.0098	0.0098
15	0.0225	0.0212	0.0198	0.0196	0.0195
16	0.0189	0.0181	0.0171	0.0169	0.0169
17	0.0018	0.0023	0.0027	0.0028	0.0028
18	0.0026	0.0032	0.0038	0.0039	0.0039
19	0.0058	0.0077	0.0095	0.0099	0.0099

TABLE VI

POSTERIOR DISTRIBUTION OF M WHEN $\lambda=1$, E(R)=1 AND $\underset{\sim}{\beta}_{\mu} = (2.5,0.7,5,0.5)'$

m \ Var(R)	0.01	0.1	1	10	100
1	0.0384	0.0419	0.0444	0.0448	0.0449
2	0.0040	0.0050	0.0057	0.0059	0.0059
3	0.0071	0.0080	0.0087	0.0088	0.0088
4	0.0009	0.0014	0.0019	0.0020	0.0020
5	0.0025	0.0031	0.0037	0.0038	0.0038
6	0.0046	0.0053	0.0059	0.0060	0.0060
7	0.0351	0.0338	0.0324	0.0320	0.0320
8	0.0519	0.0489	0.0457	0.0451	0.0450
9	0.0541	0.0505	0.0469	0.0462	0.0461
10	0.0659	0.0613	0.0566	0.0557	0.0556
11	0.0629	0.0585	0.0541	0.0532	0.0531
12	0.6036	0.6120	0.6233	0.6259	0.6262
13	0.0079	0.0082	0.0084	0.0084	0.0084
14	0.0113	0.0114	0.0113	0.0113	0.0113
15	0.0250	0.0240	0.0229	0.0229	0.0226
16	0.0186	0.0183	0.0177	0.0177	0.0176
17	0.0014	0.0018	0.0023	0.0023	0.0023
18	0.0019	0.0025	0.0030	0.0030	0.0031
19	0.0029	0.0040	0.0051	0.0051	0.0053

TABLE VII

POSTERIOR DISTRIBUTION OF M WHEN $\lambda=10$, E(R)=1 AND $\underset{\sim}{\beta}_{\mu}$ = (2.5,0.7,5,0.5)'

m \ Var(R)	0.01	0.1	1	10	100
1	0.0026	0.0042	0.0060	0.0063	0.0064
2	0.0005	0.0008	0.0013	0.0014	0.0014
3	0.0015	0.0022	0.0030	0.0031	0.0032
4	0.0003	0.0005	0.0009	0.0010	0.0010
5	0.0013	0.0019	0.0026	0.0027	0.0028
6	0.0033	0.0042	0.0050	0.0052	0.0052
7	0.0431	0.0426	0.0418	0.0416	0.0416
8	0.0647	0.0624	0.0598	0.0593	0.0592
9	0.0702	0.0671	0.0639	0.0632	0.0632
10	0.0817	0.0777	0.0737	0.0728	0.0727
11	0.0776	0.0739	0.0701	0.0693	0.0692
12	0.5888	0.5944	0.6012	0.6027	0.6028
13	0.0081	0.0088	0.0093	0.0094	0.0094
14	0.0114	0.0119	0.0123	0.0124	0.0124
15	0.0247	0.0246	0.0243	0.0243	0.0243
16	0.0164	0.0169	0.0171	0.0171	0.0171
17	0.0008	0.0013	0.0017	0.0019	0.0018
18	0.0011	0.0017	0.0022	0.0024	0.0024
19	0.0018	0.0027	0.0038	0.0040	0.0040

TABLE VIII

POSTERIOR DISTRIBUTION OF M WHEN λ=100,E(R)=1 AND $\underset{\sim}{\beta}_{\mu}$ =(2.5,0.7,5,0.5)'

m \ Var(R)	0.01	0.1	1	10	100
1	0.0003	0.0006	0.0011	0.0012	0.0013
2	0.0002	0.0004	0.0007	0.0008	0.0008
3	0.0008	0.0013	0.0019	0.0020	0.0020
4	0.0002	0.0004	0.0007	0.0007	0.0007
5	0.0010	0.0015	0.0022	0.0023	0.0023
6	0.0027	0.0036	0.0045	0.0047	0.0047
7	0.0425	0.0426	0.0424	0.0424	0.0424
8	0.0644	0.0629	0.0612	0.0609	0.0608
9	0.0694	0.0673	0.0650	0.0645	0.0645
10	0.0918	0.0882	0.0846	0.0838	0.0838
11	0.0881	0.0847	0.0812	0.0805	0.0804
12	0.5822	0.5850	0.5885	0.5893	0.5894
13	0.0083	0.0092	0.0100	0.0101	0.01011
14	0.0111	0.0120	0.0127	0.0129	0.0129
15	0.0212	0.0218	0.0222	0.0222	0.0223
16	0.0134	0.0143	0.0150	0.0152	0.0152
17	0.0006	0.0010	0.0015	0.0015	0.0016
18	0.0008	0.0012	0.0018	0.0019	0.0019
19	0.0012	0.0019	0.0028	0.0030	0.0030

TABLE IX

POSTERIOR DISTRIBUTION OF $\underset{\sim}{M}$ FOR v=0.01, $E(\sigma^2)=1$ AND $\underset{\sim}{\beta}_{\mu} = (2.5,0.7,5,0.5)'$

m \ $Var(\sigma^2)$	0.01	0.1	1	10	100
1	0.0003	0.0004	0.0006	0.0006	0.0006
2	0.0002	0.0003	0.0004	0.0004	0.0004
3	0.0007	0.0010	0.0011	0.0012	0.0013
4	0.0001	0.0002	0.0003	0.0004	0.0004
5	0.0009	0.0012	0.0013	0.0014	0.0014
6	0.0026	0.0029	0.0030	0.0031	0.0031
7	0.0417	0.0387	0.0359	0.0354	0.0353
8	0.0636	0.0588	0.0542	0.0534	0.0533
9	0.0686	0.0635	0.0587	0.0578	0.0577
10	0.0911	0.0848	0.0788	0.0776	0.0775
11	0.0874	0.0813	0.0755	0.0744	0.0743
12	0.5876	0.6137	0.6387	0.6436	0.6441
13	0.0080	0.0079	0.0076	0.0075	0.0075
14	0.0108	0.0104	0.0100	0.0098	0.0098
15	0.0208	0.0195	0.0182	0.0180	0.0179
16	0.0130	0.0125	0.0118	0.0117	0.0009
17	0.0006	0.0008	0.0009	0.0009	0.0011
18	0.0007	0.0009	0.0011	0.0011	0.0018
19	0.0011	0.0015	0.0017	0.0018	0.0018

TABLE X

POSTERIOR DISTRIBUTION OF M FOR v=0.1, $E(\sigma^2)=1$ AND $\underset{\sim}{\beta}_\mu = (2.5,0.7,5,0.5)'$

m \ $Var(\sigma^2)$	0.01	0.1	1	10	100
1	0.0025	0.0031	0.0036	0.0037	0.0037
2	0.0004	0.0006	0.0007	0.0008	0.0008
3	0.0015	0.0017	0.0019	0.0019	0.0019
4	0.0003	0.0004	0.0005	0.0005	0.0005
5	0.0012	0.0015	0.0016	0.0017	0.0017
6	0.0031	0.0033	0.0034	0.0034	0.0034
7	0.0424	0.0387	0.0352	0.0346	0.0345
8	0.0639	0.0581	0.0529	0.0518	0.0517
9	0.0695	0.0634	0.0578	0.0567	0.0566
10	0.0810	0.0744	0.0681	0.0669	0.0668
11	0.0770	0.0706	0.0646	0.0635	0.0633
12	0.5945	0.6248	0.6541	0.6598	0.6604
13	0.0079	0.0075	0.0071	0.0070	0.0070
14	0.0111	0.0104	0.0097	0.0095	0.0095
15	0.0242	0.0222	0.0203	0.0199	0.0199
16	0.0160	0.0149	0.0137	0.0135	0.0135
17	0.0008	0.0010	0.0011	0.0011	0.0011
18	0.0011	0.0013	0.0014	0.0143	0.0014
19	0.0017	0.0021	0.0023	0.0024	0.0024

TABLE XI

POSTERIOR DISTRIBUTION OF M FOR $v=1, E(\sigma^2)=1$
AND $\underset{\sim}{\beta}_{\mu} = (2.5, 0.75, 0.5)'$

m \ $Var(\sigma^2)$	0.01	0.1	1	10	100
1	0.0372	0.0351	0.0325	0.0319	0.0319
2	0.0039	0.0040	0.0039	0.0039	0.0039
3	0.0068	0.0066	0.0062	0.0061	0.0061
4	0.0009	0.0011	0.0011	0.0011	0.0011
5	0.0023	0.0025	0.0025	0.0024	0.0024
6	0.0044	0.0043	0.0041	0.0041	0.0041
7	0.0344	0.0305	0.0269	0.0262	0.0261
8	0.0512	0.0453	0.0398	0.0388	0.0387
9	0.0534	0.0474	0.0418	0.0408	0.0406
10	0.0653	0.0584	0.0519	0.0506	0.0505
11	0.0624	0.0557	0.0495	0.0483	0.0481
12	0.6104	0.6478	0.6844	0.6915	0.6923
13	0.0077	0.0071	0.0064	0.0063	0.0062
14	0.0110	0.0100	0.0089	0.0087	0.0086
15	0.0245	0.0217	0.0191	0.0186	0.0186
16	0.0182	0.0162	0.0144	0.0140	0.0140
17	0.0013	0.0014	0.0015	0.0015	0.0015
18	0.0018	0.0019	0.0020	0.0019	0.0019
19	0.0027	0.0031	0.0032	0.0032	0.0032

TABLE XII

POSTERIOR DISTRIBUTION OF M FOR v=10, $E(\sigma^2)=1$ AND $\underset{\sim}{\beta}_{\mu} = (2.5,0.7,5,0.5)'$

m \ Var(σ^2)	0.01	0.1	1	10	100
1	0.1419	0.1274	0.1128	0.1099	0.1096
2	0.0147	0.0134	0.0119	0.0116	0.0116
3	0.0147	0.0132	0.0117	0.0114	0.0114
4	0.0013	0.0014	0.0014	0.0014	0.0014
5	0.0026	0.0026	0.0025	0.0024	0.0024
6	0.0043	0.0041	0.0038	0.0037	0.0037
7	0.0254	0.0223	0.0194	0.0189	0.0188
8	0.0377	0.0330	0.0286	0.0278	0.0277
9	0.0385	0.0339	0.0295	0.0286	0.0285
10	0.0497	0.0441	0.0387	0.0377	0.0377
11	0.0477	0.0423	0.0371	0.0361	0.0360
12	0.5541	0.6008	0.6475	0.6566	0.6576
13	0.0070	0.0063	0.0056	0.0055	0.0055
14	0.0101	0.0090	0.0079	0.0077	0.0077
15	0.0221	0.0194	0.0169	0.0164	0.0163
16	0.0185	0.0163	0.0142	0.0138	0.0137
17	0.0017	0.0019	0.0019	0.0019	0.0018
18	0.0025	0.0026	0.0025	0.0025	0.0025
19	0.0055	0.0060	0.0061	0.0061	0.0061

TABLE XIII

POSTERIOR DISTRIBUTION OF M FOR $v=100, E(\sigma^2)=1$ AND $\underset{\sim}{\beta}_\mu = (2.5,0.7,5,0.5)'$

m \ Var(σ^2)	0.01	0.1	1	10	100
1	0.3442	0.3152	0.2846	0.2784	0.2777
2	0.0156	0.0143	0.0129	0.0126	0.0126
3	0.0130	0.0119	0.0107	0.0105	0.0105
4	0.0010	0.0011	0.0011	0.0011	0.0011
5	0.0020	0.0020	0.0020	0.0020	0.0019
6	0.0033	0.0032	0.0030	0.0030	0.0030
7	0.0184	0.0165	0.0147	0.0144	0.0143
8	0.0273	0.0244	0.0217	0.0212	0.0211
9	0.0278	0.0250	0.0223	0.0217	0.0217
10	0.0362	0.0329	0.0296	0.0289	0.0288
11	0.0347	0.0315	0.0283	0.0277	0.0276
12	0.4164	0.4645	0.5143	0.5256	0.5267
13	0.0053	0.0049	0.0045	0.0044	0.0044
14	0.0077	0.0070	0.0063	0.0062	0.0061
15	0.0167	0.0150	0.0133	0.0130	0.0130
16	0.0144	0.0130	0.0116	0.0113	0.0113
17	0.0015	0.0015	0.0016	0.0016	0.0016
18	0.0022	0.0023	0.0023	0.0023	0.0023
19	0.0124	0.0138	0.0144	0.0144	0.0144

TABLE XIV

DATA FROM POOL AND BORCHGREVINK (1964)
AND THE POSTERIOR DISTRIBUTION OF M

Obs. No. (i)	x_i	y_i	M	$\pi(m \mid \underset{\sim}{y})$
1	2.00000	0.370483	1	0.00000
2	2.52288	0.537970	2	0.00000
3	3.00000	0.607684	3	0.00001
4	3.52288	0.723323	4	0.00053
5	4.00000	0.761856	5	0.19744
6	4.52288	0.892063	6	0.48151
7	5.00000	0.956707	7	0.31535
8	5.52288	0.940349	8	0.00513
9	6.00000	0.898609	9	0.00002
10	6.52288	0.953850	10	0.00000
11	7.00000	0.990834	11	0.00000
12	7.52288	0.890291	12	0.00000
13	8.00000	0.990779	13	0.00000
14	8.52288	1.050865	14	0.00000
15	9.00000	0.982785		

TABLE XV

POSTERIOR PROBABILITY MASS FUNCTION OF λ

λ	Prior Probability of no change, q.			
	0.05	0.50	0.95	0.99
1	0.0430	0.0273	0.0034	0.0007
2	0.0056	0.0035	0.0004	0.0001
3	0.0084	0.0053	0.0007	0.0001
4	0.0018	0.0011	0.0001	0.0000
5	0.0036	0.0023	0.0003	0.0001
6	0.0057	0.0036	0.0005	0.0001
7	0.0313	0.0199	0.0025	0.0005
8	0.0443	0.0281	0.0035	0.0007
9	0.0454	0.0288	0.0036	0.0007
10	0.0548	0.0348	0.0044	0.0009
11	0.0523	0.0332	0.0042	0.0009
12	0.6034	0.3830	0.0482	0.0099
13	0.0081	0.0052	0.0006	0.0001
14	0.0110	0.0070	0.0009	0.0001
15	0.0222	0.0141	0.0018	0.0004
16	0.0172	0.0109	0.0014	0.0003
17	0.0022	0.0014	0.0002	0.0000
18	0.0029	0.0018	0.0002	0.0000
19	0.0049	0.0031	0.0004	0.0001
20	0.0320	0.3855	0.9226	0.9842

TABLE XVI

POSTERIOR PROBABILITIES OF 'NO CHANGE' FOR AN INFORMAL SEQUENTIAL PROCEDURE

No. of Obs.	Prior Probability of No Change				
	q				
t	0.05	0.50	0.95	0.99	1/t
1	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.0484	0.4914	0.9483	0.9900	0.4914
3	0.1815	0.8082	0.9877	0.9978	0.6781
4	0.0206	0.2858	0.8838	0.9754	0.1177
5	0.0420	0.4546	0.9406	0.9880	0.1725
6	0.0712	0.5931	0.9651	0.9931	0.2257
7	0.2263	0.8475	0.9906	0.9982	0.4809
8	0.3411	0.9077	0.9947	0.9990	0.5842
9	0.3830	0.9218	0.9956	0.9991	0.5959
10	0.4830	0.9467	0.9970	0.9994	0.6636
11	0.5233	0.9542	0.9975	0.9995	0.6759
12	0.4682	0.9436	0.9969	0.9994	0.6033
13	0.0255	0.3321	0.9043	0.9801	0.3979
14	0.1002	0.6790	0.9757	0.9952	0.1400
15	0.1062	0.6930	0.9772	0.9955	0.1388
16	0.0790	0.6198	0.9687	0.9938	0.0980
17	0.0214	0.2938	0.8877	0.9763	0.0253
18	0.0348	0.4062	0.9286	0.9855	0.0387
19	0.0478	0.4883	0.9477	0.9895	0.0503
20	0.0320	0.3855	0.9226	0.9842	0.0320

APPENDIX B

THEOREMS

THEOREMS

B.1. Prove that $S(\gamma_1,m) > 0$.

<u>Proof</u>.

From Eq. (3.17)

$$S(\gamma_1,m) = (2a^* + 1)A(\gamma_1,m)/[2a^*G(\gamma_1,m)]$$

where

(1) $a^* = a + n/2 > 0$, $2a^* > 0$, and $2a^* + 1 > 0$,

since $a > 0$ and $n > 0$

then $a^* > 0$, $2a^* > 0$, and $2a^* + 1 > 0$;

(2) $A(\gamma_1,m) = \underset{\sim}{b}'(\gamma_1)V(m)\underset{\sim}{b}(\gamma_1) > 0$,

since $V(m)$ is p.d., from the definition of p.d.,

then $A(\gamma_1,m) > 0$ for γ_1 ,

and

(3) $G(\gamma_1,m) = 1 + (1/2a^*)[C(m) - B^2(\gamma_1,m)/A(\gamma_1 m)] > 0$.

Eq. (3.4) and (3.8) show that

$A(\gamma_1,m)\gamma_2^{\,2} - 2B(\gamma_1,m)\gamma_2 + C(m) \geq 0$.

If we divide both sides by $A(\gamma_1,m)$ (where $A(\gamma_1,m) > 0$ from (2)), we get

$$\gamma_2^2 - \frac{2B(\gamma_1,m)}{A(\gamma_1,m)}\gamma_2 + \frac{C(m)}{A(\gamma_1,m)} \geq 0$$

$$[\gamma_2 - \frac{B(\gamma_1,m)}{A(\gamma_1,m)}]^2 + \frac{A(\gamma_1,m)C(m) - B^2(\gamma_1,m)}{A^2(\gamma_1,m)} \geq 0 .$$

Since $[\gamma_2 - \dfrac{B(\gamma_1,m)}{A(\gamma_1,m)}]^2 \geq 0$, then

$$\frac{A(\gamma_1,m)C(m) - B^2(\gamma_1,m)}{A^2(\gamma_1,m)} \geq 0$$

$$C(m) - B^2(\gamma_1,m)/A(\gamma_1,m) \geq 0 \ .$$

Therefore $G(\gamma_1,m) > 0$.

From (1), (2) and (3), $S(\gamma_1,m) > 0$ and the proof is complete.

B.2. Suppose x is distributed as a general t distribution with n degrees of freedom, location parameter μ and precision τ ($n > 0$, $-\infty < \mu < \infty$, $\tau > 0$), then

$$E|x| = \frac{2n^{1/2}\ \Gamma[(n+1)/2]}{(n-1)(\tau\pi)^{1/2}\Gamma(n/2)}\ (1+\tau\mu^2/n)^{-(n-1)/2}$$

$$+\ \mu[2\psi_n(\mu\tau^{1/2}) - 1]$$

where $\psi_n(x)$ is the cumulative distribution function of a student t distribution with n degrees of freedom.

Proof.

The p.d.f. of x is

$$f(x|n,\mu,\tau) = k[1 + \frac{\tau}{n}(x-\mu)^2]^{-(n+1)/2} \tag{B2.1}$$

where

$$k = \frac{\tau^{1/2}\ \Gamma[(n+1)/2]}{(n\pi)^{1/2}\Gamma(n/2)} \ .$$

The expectation of x is defined by

$$E|x| = \int_{-\infty}^{\infty} |x|\ f(x|n,\mu,\tau)dx$$

$$= \int_0^{\infty} x\ f(x|n,\mu,\tau)dx + \int_{-\infty}^{0} (-x)f(x|n,\mu,\tau)dx \ . \tag{B2.2}$$

Now, let us evaluate the first term of the right hand side of (B2.2).

$$\int_0^{\infty} x\, f(x|n,\mu,\tau)\, dx = \int_0^{\infty} (x - \mu) f(x|n,\mu,\tau) dx + \mu \int_0^{\infty} f(x|n,\mu,\tau) dx$$

$$= k \int_0^{\infty} (x-\mu)\left[1 + \frac{\tau}{n}(x-\mu)^2\right]^{-(n+1)/2} dx + k\mu \int_0^{\infty} \left[1 + \frac{\tau}{n}(x-\mu)^2\right]^{-\frac{n+1}{2}} dx. \qquad \text{(B2.3)}$$

$$\text{Let}\quad y = \tau(x-\mu)^2 \qquad\qquad z = \tau^{1/2}(x-\mu)$$

$$dy = 2\tau(x-\mu)dx \qquad\qquad dz = \tau^{1/2} dx$$

$$dx = 1/(2\tau(x-\mu))\, dy \qquad\qquad dx = \tau^{-1/2} dz \quad . \qquad \text{(B2.4)}$$

Then (B2.3) becomes

$$\frac{k}{2\tau} \int_{\tau\mu^2}^{\infty} \left[1 + \frac{1}{n} y\right]^{-(n+1)/2} dy + \frac{k\mu}{\tau^{1/2}} \int_{-\mu\tau^{1/2}}^{\infty} \left[1 + \frac{1}{n} z^2\right]^{-(n+1)/2} dz$$

$$= \frac{nk}{\tau(n-1)} \left[1 + \frac{1}{n}\tau\mu^2\right]^{-(n-1)/2} + \mu\psi_n(\mu\tau^{1/2}) \qquad \text{(B2.5)}$$

where $\psi_n(x)$ is the cumulative distribution function of a student t distribution with n degrees of freedom.

Similarly, we substitute the same transformation as (B2.4) to the second term of the right hand side of integral (B2.2), then we obtain

$$\int_{-\infty}^{0} (-x) f(x|n,\mu,\tau) dx$$

$$= -\int_{-\infty}^{0} (x-\mu) f(x|n,\mu,\tau) dx - \mu \int_{-\infty}^{0} f(x|n,\mu,\tau) dx$$

$$= -k \int_{-\infty}^{0} (x-\mu)\left[1 + \frac{\tau}{n}(x-\mu)^2\right]^{-(n+1)/2} dx$$

$$- k\mu \int_{-\infty}^{0} \left[1 + \frac{\tau}{n}(x-\mu)^2\right]^{-(n+1)/2} dx$$

$$= -\frac{k}{2\tau}\int_{\infty}^{\tau\mu^2}[1+\frac{1}{n}y]^{-(n+1)/2}dy - k\mu\tau^{-1/2}\int_{-\infty}^{-\mu\tau^{1/2}}$$

$$[1+\frac{1}{n}z^2]^{-(n+1)/2}\,dz$$

$$= \frac{nk}{\tau(n-1)}\;[1+\frac{1}{n}\tau\mu^2]^{-(n-1)/2} - \mu[1-\psi_n(\mu\tau^{1/2})]\;. \qquad (B2.6)$$

Substituting (B2.5) and (B2.6) to (B2.2), (B2.2) becomes

$$E|x| = \frac{2nk}{\tau(n-1)}\;[1+\frac{1}{n}\tau\mu^2]^{-(n-1)/2} + \mu[2\psi_n(\mu\tau^{1/2}) - 1]$$

$$= \frac{2n^{1/2}\Gamma[(n+1)/2]}{(n-1)(\tau\pi)^{1/2}\Gamma(n/2)}\;[1+\frac{1}{n}\mu\tau^2]^{-(n-1)/2}$$

$$+ \mu[2\psi_n(\mu\tau^{1/2}) - 1]\;,$$

and this completes the proof.

VITA

Jaunita Hwa-ling-Tu Chin Choy

Candidate for the Degree of

Doctor of Philosophy

Thesis: A BAYESIAN ANALYSIS OF A CHANGING LINEAR MODEL

Major Field: Statistics

Biographical:

Personal Data: Born in Hunan province, Republic of China, October 15, 1944, the daughter of Chi-Lian and Lee-ing Tu.

Education: Attended elementary school and graduated from Taipei Girls' First Provincial High School, Taipei, Taiwan, China, in 1963; received the Bachelor of Science degree from the National Taiwan University in June, 1967; received Master of Science degree in Agronomy from the Oklahoma State University in May, 1970; completed requirements for the degree of Doctor of Philosophy at Oklahoma State University in July, 1977.

Professional Experience: High school tutor, Taipei, February, 1964 to May, 1967; graduate research assistant, Department of Agronomy, Oklahoma State University, September, 1967 to May, 1970; Laboratory Technician II, Department of Animal Sciences, Oklahoma State University; graduate research assistant, Department of Statistics, Oklahoma State University, 1975.

Organizations: American Statistical Association, Mu Sigma Rho, Pi Mu Epsilon.