A BAYESIAN ANALYSIS OF SHIFTING SEQUENCES WITH APPLICATIONS TO TWO-PHASE REGRESSION

By

DONALD HOLBERT

Bachelor of Science
University of Oregon
Eugene, Oregon
1967

Master of Arts
Washington State University
Pullman, Washington
1969

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
May, 1973

A BAYESIAN ANALYSIS OF SHIFTING
SEQUENCES WITH APPLICATIONS
TO TWO-PHASE REGRESSION

Thesis Approved:

Lyle D. Broemeling
Thesis Adviser

J Leroy Folks

David F. Weeks

Odell L Walker

D D Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my thanks and appreciation to Professor Lyle Broemeling for suggesting the topic and for his guidance and encouragement in the preparation of this thesis.

I wish to thank Professors Leroy Folks, Odell Walker and David Weeks for serving on my advisory committee, and also to thank all of the teachers that I have had during the course of my graduate study.

Dr. Robert Morrison has taught me a great deal about what might be called "hard-core" data analysis, and I want to thank him for that.

For the friendship and help they have given me during the course of my graduate study, thanks go to my fellow students and especially to Ceal.

I would also like to thank Mrs. Mary Bonner for the first rate job she has done in typing this thesis.

There are many other people to whom I owe a vote of thanks for various contributions they have made, and since space limitations make this prohibitive I would like to thank collectively all of those who I have not mentioned individually.

Finally, for the sacrifices they made to provide me with a high school education, a sincere thank you to my mother and late father, to whom I owe much more than I can put into words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Introduction

Whenever observations are taken in an ordered sequence it can happen that the complete data set can be divided into subsets in a well defined way. Each observation may be regarded as coming from a parameterized family of distributions; those observations in a specified subset correspond to some particular value of the parameter, the parameter value changing from subset to subset.

One of the more widely known occurrences of this situation is in industry where one is interested in the quality of a product from a continuous production process. For a specific example, let us suppose that the "Surestrike Match Company" produces boxes of kitchen matches, and that the average content is advertised as fifty matches.

It seems clear enough that the company would like to know if a change occurs in the average content of a box of matches. If this average increases the company is giving away free matches, while if it decreases customers may well suspect the company of false advertising and take their business elsewhere. It also seems clear that if a change in the average content has occurred, interest would center on the time point in the sequence at which the change occurred. This knowledge may allow the company to recover the faulty product and correct the fault before distribution occurs.

Many other examples come to mind. In biological studies, the onset of a disease at some point in time may result in a reduced growth rate; the application of a treatment may inhibit the response to some stimulus; or repeated conditioning in a psychological experiment may cause a change in the proportion of correct answers given by a subject.

There is a considerable body of literature on the various problems of estimation and inference associated with parameters changing over time, and the references given in the bibliography are by no means exhaustive. Most of the recent studies have been from the classical frequency theory viewpoint. The present study is set in a Bayesian framework.

We shall now formulate the problem in mathematical terms and then review some of the recent papers on its various aspects.

We assume that we have observations on a finite sequence of random variables $X_1, \ldots, X_T$, and that for some m, $1 \leq m \leq T$, $X_1, \ldots, X_m$ are independently distributed with density $f(x; \theta_1)$ while $X_{m+1}, \ldots, X_T$ are independently distributed $f(x; \theta_2)$. We assume of course that $\theta_1 \neq \theta_2$ and that $(X_1, \ldots, X_m)$ is independent of $(X_{m+1}, \ldots, X_T)$. We refer to this point m in the sequence as the "shift point" or "switch point", since the random variables up to the m^{th} correspond to one parameter value while those after the m^{th} correspond to another parameter value. The problem may be generalized to the case where θ_1 and θ_2 are p-component vectors. The shift is then from one point p-space to another, rather than a shift along the real line.

It is clear that if $m = T$ then no shift has occurred. This enables us to identify two main problems of interest with such a sequence. The first is a detection problem. That is, has there been a parameter change in the sequence of random variables? The second problem is an estimation problem. Assuming a shift has occurred, at what point in the sequence did it occur? Along with these are the problems of estimating current and previous parameter values, and perhaps testing hypotheses about them. The present study addresses itself only to the second of these two main problems.

Review Of Some Related Literature

In a series of papers Page (1,2,3) discusses the problem of detecting a parameter change, and proposes a number of tests to this end. These are based primarily on the cumulative sums $S_r = \sum_{i=1}^{r} (X_i - \theta)$, where θ is the known initial mean. If there is no change the mean path of the cumulative sums is a horizontal path; $E(S_r) = 0$.

Quandt (4) discusses a maximum likelihood technique for estimating the switch point and the regression parameters in a two-phase regression. This is a generalization of the one parameter case discussed earlier. We assume here that the observations Y_i, $i = 1, \ldots, m$ are distributed $N(\alpha_1 + \beta_1 X_i, \sigma^2)$ while the Y_j, $j = m+1, \ldots, T$ are distributed $N(\alpha_2 + \beta_2 X_j, \sigma^2)$. Quandt's technique involves evaluating the likelihood function at each of the possible switch points. He also discusses in this article and a later paper (5) several tests of the hypothesis of no switch against the alternative of a single switch.

Sprent (6) outlines a hierarchy of possible hypotheses of interest related to two-phase regression, and suggests that the result of an initial investigation should indicate the next hypothesis to be considered. His tests, however, are based on the assumption that one knows between which two of the independent variables the switch occurs. The switch point is then the abscissa of the point at which the intersection of the two regression lines occurs, namely $\gamma = (\alpha_1 - \alpha_2)/(\beta_2 - \beta_1)$.

Chernoff and Zacks (7) study Bayesian procedures for estimating the current mean (i.e., the mean of X_T) in an observed sequence $X_1, \ldots, X_T$ of normally distributed random variables which has been subjected to occasional changes in the mean. Their estimator requires many complex computations except when the assumption of at most one change in the mean is made. A test is also given for the null hypothesis of no shift against the alternative of exactly one shift, and its power for certain alternatives is compared to that of the test proposed by Page (2). This is generalized in a later paper by Kander and Zacks (8) to the case where the distributions of the X_i's belong to the one parameter exponential family rather than the normal family in particular. The paper by Bhattacharyya and Johnson (9) derives certain optimal tests of the hypothesis just stated, their optimality criterion being the maximization of local average power.

Brown and Durbin (10) discuss methods for investigating whether a regression relationship is constant over time. Most of their techniques are graphical in nature, along the lines suggested by Tukey (11). These include plotting the residuals from a single regression fitted to the entire data set, as well as plotting the cumulative sums of residuals, in line with the cusum technique of Page (1). A further technique they discuss is that of plotting the recursive residuals $w_t = Y_t - (\hat{\alpha}_{t-1} + \hat{\beta}_{t-1} X_t)$, $t = 3, \ldots, T$, where $\hat{\alpha}_{t-1}$ and $\hat{\beta}_{t-1}$ are the least squares estimates of the regression parameters based on the first $t-1$ observations. These quantities can easily be normed in such a way that under they hypothesis of no shift they are independently distributed $N(0, \sigma^2)$. Other useful plots are the cumulative sums of the recursive residuals and the cumulative sums of squares of recursive residuals, each plotted against the time points t.

Finally we come to some of the more recent papers on shifting sequences of random variables. With regard to two-phase regression in particular, D. V. Hinkley has been a major contributor to the recent literature.

In a paper published in 1969, Hinkley (12) describes a method for finding a maximum likelihood estimate of the abscissa of the intersection point of the two regression lines, $\gamma = (\alpha_1 - \alpha_2)/(\beta_2 - \beta_1)$. This involves finding $T-3$ conditional likelihood functions, each of which has to be maximized, and then maximizing over all $T-3$ functions. This estimate is difficult to work with in that it has no explicit definition. He also proposes a likelihood ratio test for a null hypothesis of the form $H_o: \gamma = \gamma_o$. In his 1971 article, Hinkley (13) parameterizes the problem a little differently. He assumes that Y_i, $i = 1, \ldots, m$ are independently distributed $N(\theta + \beta_1(X_i - \gamma), \sigma^2)$ while the Y_j, $j = m+1, \ldots, T$ are independently distributed $N(\theta + \beta_2(X_j - \gamma), \sigma^2)$, with $X_m \leq \gamma \leq X_{m+1}$, γ being the abscissa of the intersection point of the two regression lines. He then centers interest on estimation and inference procedures related to θ and γ. Maximum likelihood estimation of θ and γ is studied under the assumption that β_2 is unknown and also under the assumption that $\beta_2 = 0$. Likelihood ratio tests for $H_0: \beta_1 = \beta_2$ and $H_0': \beta_2 = 0$ are discussed. He also derives a confidence region for γ and describes a technique for constructing a joint confidence region for θ and γ.

Hinkley's remaining papers are concerned with a change in the mean of an observed sequence of random variables. In a 1970 paper (14) he discusses maximum likelihood estimation of the shift point m in a normal sequence whose mean has been subjected to one change. He also derives the asymptotic distribution of the likelihood ratio test statistic for tests of hypotheses about m. Similar problems related to a binomial sequence are studied in a paper by Hinkley and Hinkley (15).

A Bayesian approach to the problem of estimating the shift point in an observed sequence of random variables has been given by Broemeling (16). He derives posterior densities for the shift point parameter in the case of a Bernoulli sequence, a sequence of exponentially distributed random variables, and a normal sequence with known variance. In another paper Broemeling (17) discusses Bayesian procedures for first detecting the presence or absence of a parameter change, and then making inferences after this initial decision has been made. The detection problem is handled in terms of a "posterior odds ratio" in favor of the null hypothesis, while inference procedures are as usual based on an appropriate posterior distribution.

Organization Of Thesis

We shall now describe briefly the content of this thesis in relation to the literature we have just discussed.

In Chapter II we derive a number of posterior distributions which may be used for inference about a normal sequence with unknown variance. This is an extension of the paper by Broemeling (16).

Chapters III and IV are related to two-phase regression, which may be considered as a generalization of the shifting normal sequence. In particular, Chapter III addresses itself to the problem of estimating the shift index m, while in Chapter IV we derive posterior distributions related to estimation of the abscissa of the point of intersection of the two regression lines. These chapters present an alternative to the analyses given by Quandt (4, 18), Sprent (6), Hinkley (12, 13) and others.

A brief survey of Bayesian estimation and inference techniques is presented in Chapter V, which includes examples and applications of some of our earlier results.

Chapter VI summarizes the results of this report and discusses some possibilities for future investigations.

CHAPTER II

POSTERIOR DISTRIBUTIONS RELATED TO THE NORMAL SEQUENCE

In this chapter we shall derive some posterior distributions related to the problem of estimating the time point at which a shift in the mean occurs in a finite sequence of observations on normally distributed random variables.

More specifically, we assume that we have observed a sequence $X_1, \ldots, X_n$, $n \geq 3$, of independent random variables, and that for some unknown $m\,(m = 1, \ldots, n-1)$ the distributions of the X_i's are given by:

$$X_1, \ldots, X_m \quad \text{are i.i.d.} \quad N(\phi_0, \sigma^2), \quad \text{and}$$

$$X_{m+1}, \ldots, X_n \quad \text{are i.i.d.} \quad N(\phi_1, \sigma^2),$$

with $\sigma^2 > 0$ and $\phi_0 \neq \phi_1$.

The case in which σ^2 is assumed known has been studied by Broemeling (16). We consider here only the case where σ^2 is unknown.

There are a number of subcases for the above problem which we shall consider:

(i) ϕ_0, ϕ_1, both known

(ii) Only one of ϕ_0 or ϕ_1 known

(iii) Neither ϕ_0 nor ϕ_1 known, A further consideration is whether or not the direction of the shift is known, and this will be discussed in the following sections.

While the main emphasis in this paper is on the estimation of the shift point m, posterior distributions will be derived for some of the other unknown parameters.

In order to attack the problem from a Bayesian viewpoint we shall consider m to be a discrete random variable with state space $I_{n-1} = \{1, \ldots, n-1\}$. The parameters for our problem are now m, σ^2, ϕ_0, and ϕ_1, and prior distributions will have to be assigned to those that are unknown in each of the cases considered.

In the development given in this and later chapters we shall assign independent prior distributions which may be considered appropriate for situations where prior knowledge is vague. More precisely, we shall assign in every case for m and σ^2 the prior densities:

$$\pi_0(m) = \begin{cases} 1/(n-1), & m = 1, \ldots, n-1 \\ 0 & \text{elsewhere, and} \end{cases}$$

$$\pi_0(\sigma^2) \propto \begin{cases} 1/\sigma^2, & \sigma^2 > 0 \\ 0 & \text{elsewhere .} \end{cases}$$

The prior on σ^2 is of course an improper density. It has been widely used to indicate vague prior knowledge of the variance. Its form is suggested by a number of approaches, and the reader is referred to Lindley (19) and Jeffreys (20) for a discussion of these.

We now consider in turn the three cases mentioned above.

ϕ_0 And ϕ_1 Both Known

The likelihood function for m and σ^2 is

$$L(m, \sigma^2) = (2\pi\sigma^2)^{-m/2} \exp\left\{(-1/2\sigma^2) \sum_{i=1}^{m} (X_i - \phi_0)^2\right\} (2\pi\sigma^2)^{(n-m)/2}$$

$$\cdot \exp\left\{(-1/2\sigma^2) \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right\}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} \tag{2.1}$$

where $0 < \sigma^2 < \infty$ and m belongs to I_{n-1}. In accordance with Bayes theorem the joint posterior distribution of m and σ^2 is

$$\pi_1(m, \sigma^2) \propto L(m, \sigma^2)\, \pi_0(m)\, \pi_0(\sigma^2)$$

$$\propto (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} \tag{2.2}$$

where $0 < \sigma^2 < \infty$ and m belongs to I_{n-1}.

<u>Inference About m</u>

This may be based on the marginal posterior density of m, which is given by

$$\pi_1(m) \propto \int_0^{\infty} \pi_1(m, \sigma^2)\, d\sigma^2$$

$$\propto \int_0^{\infty} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\}$$

Letting $w = 1/\sigma^2$, the integrand is seen to have the form of a gamma density, and is easily evaluated to give

$$\Gamma\left(\frac{n}{2}\right)\left[2\Big/\left(\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right)\right]^{n/2} .$$

We can thus write

$$\pi_1(m) \propto \left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]^{-n/2}, \quad m = 1, \ldots, n-1 . \tag{2.3}$$

The norming constant may be found by summing on m.

Inference About σ^2

This may be based on the marginal posterior distribution of σ^2, which is given by

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} \pi_1(m, \sigma^2)$$

$$\propto \sum_{m=1}^{n-1} (\sigma^2)^{-(n/2+1)} \exp\{(-1/2\sigma^2)K(m)\}, \quad 0 < \sigma^2 < \infty \tag{2.4}$$

where

$$K(m) = \left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right] .$$

If we now let $w = 1/\sigma^2$ we can write the posterior density of w as

$$\pi_1(w) \propto \sum_{m=1}^{n-1} w^{n/2-1} \exp\{-w\,K(m)/2\}, \quad 0 < w < \infty . \tag{2.5}$$

Apart from the norming constant, this is the sum of $n - 1$ gamma densities with parameters $\alpha_m = n/2$ and $\beta_m = 2/K(m)$. The posterior density of $w = 1/\sigma^2$ is thus a mixture of gamma densities. We may also make the equivalent statement that the posterior density of σ^2 is a mixture of inverted gamma densities.

An alternative distribution on which inference about σ^2 can be based is the conditional posterior distribution, $\pi_1(\sigma^2 | m)$. One could center interest on this distribution evaluated at, say, a modal value of the posterior density of m. In this case we have

$$\pi_1(\sigma^2 | m) = \pi_1(\sigma^2, m) \Big/ \pi_1(m)$$

$$\propto (\sigma^2)^{-(n/2+1)} \exp\{(-1/2\sigma^2) K(m)\} \Big/ [K(m)]^{-n/2}. \quad (2.6)$$

It is now clear that, for each fixed m, this conditional density is inverted gamma with parameters $\alpha_m = n/2$ and $\beta_m = 2/K(m)$.

Before proceeding to the next case we should perhaps remark on the mixing density which occurs in the posterior distribution of $w = 1/\sigma^2$. If we let

$$f_m(w; n/2, K(m)) = w^{n/2-1} \exp\{-w K(m)/2\}$$

then we can write the posterior density of w (see (2.5)) as

$$\pi_1(w) \propto \sum_{m=1}^{n-1} f_m(w; n/2, K(m)) . \quad (2.7)$$

The norming constant K is now given by

$$K = \int_0^\infty \sum_{m=1}^{n-1} \Gamma(n/2)[2/K(m)]^{n/2} \Big/ \Gamma(n/2)[2/K(m)]^{n/2} f_m(w; n/2, K(m))dw$$

$$= \Gamma(n/2)\, 2^{n/2} \sum_{m=1}^{n-1} [K(m)]^{-n/2} .$$

Referring to (2.7) and inserting the norming constant K we see that the m^{th} mixing constant for the posterior density of w is $K(m)^{-n/2} \Big/ \sum_{m=1}^{n-1} [K(m)]^{-n/2}$. By referring also to (2.3) we see that this is precisely the posterior density of the shift point at its m^{th} mass point. Thus the gamma densities occurring in the posterior density of $w = 1/\sigma^2$ are mixed according to the posterior density of the shift index m.

Only One Of ϕ_0 Or ϕ_1 Known

The theory for the case ϕ_0 unknown and ϕ_1 known parallels that for the case ϕ_0 known and ϕ_1 unknown, and for that reason we shall study the latter case only in this section. The results for the former case will be evident.

We now have an additional parameter, ϕ_1, whose prior density must be assigned. We shall derive the appropriate posterior densities corresponding to two different vague prior distributions. Firstly we shall assume that the direction of the shift is now known and assign the improper prior density

$$\pi_0(\phi_1) \propto \text{constant}, \quad -\infty < \phi_1 < \infty .$$

This shall be referred to as the "unconstrained" prior density. Next we shall assume that it is known that $\phi_0 < \phi_1$, and in this case we shall assign the "constrained" prior density

$$\pi_0(\phi_1) \propto \text{constant}, \quad \phi_0 < \phi_1 < \infty .$$

Unconstrained Prior, Inference About m

The joint posterior density of the three unknown parameters is easily seen to be

$$\pi_1(m,\sigma^2,\phi_1) \propto (\sigma^2)^{-(n/2+1)}\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2+\sum_{i=m+1}^{n}(X_i-\phi_1)^2\right]\right\}. \tag{2.8}$$

The marginal posterior density of m is given by

$$\pi_1(m) \propto \int_0^\infty\int_{-\infty}^{\infty}(\sigma^2)^{-(n/2+1)}\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2 + \sum_{i=m+1}^{n}(X_i-\phi_1)^2\right]\right\}d\phi_1\,d\sigma^2$$

$$\propto \int_0^\infty\int_{-\infty}^{\infty}(\sigma^2)^{-(n/2+1)}\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2+\sum_{i=m+1}^{n}(X_i-\overline{X}^n_{m+1})^2 + (n-m)(\phi_1-\overline{X}^n_{m+1})^2\right]\right\}d\phi_1\,d\sigma^2$$

where $\overline{X}^n_{m+1} = \sum_{i=m+1}^{n} X_i \Big/ (n-m)$. Making use of the Tonelli theorem (21) we can write

$$\pi_1(m) \propto \int_0^\infty(\sigma^2)^{-(n/2+1)}\exp\ (-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2+\sum_{i=m+1}^{n}(X_i-\overline{X}^n_{m+1})^2\right]$$

$$\left[\int_{-\infty}^{\infty}\exp\{(-(n-m)/2\sigma^2)(\phi_1-\overline{X}^n_{m+1})^2\}\,d\phi_1\right]d\sigma^2$$

$$\propto (n-m)^{-1/2}\int_0^\infty(\sigma^2)^{-(n+1)/2}\exp\{(-1/2\sigma^2)[K(m,\phi_0)]\}\,d\sigma^2$$

where

$$K(m,\phi_0) = \sum_{i=1}^{m}(X_i-\phi_0)^2+\sum_{i=m+1}^{n}(X_i-\overline{X}^n_{m+1})^2 .$$

Making use of the inverted gamma integral as before we obtain for the posterior density of m

$$\pi_1(m) \propto (n-m)^{-1/2}[K(m, \phi_0)]^{-(n-1)/2}, \quad m = 1, \ldots, n-1 . \qquad (2.9)$$

A comparison with (2.3) shows that, for each m, ϕ_1 has been replaced by its estimate and an additional multiplier has been introduced.

Unconstrained Prior, Inference About σ^2

This may be based on the marginal posterior distribution of σ^2, which is given by

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} \int_{-\infty}^{\infty} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]\right\} d\phi_1$$

$$\propto \sum_{m=1}^{n-1} (n-m)^{-1/2} (\sigma^2)^{-(n+1)/2} \exp\{(-1/2\sigma^2) K[m, \phi_0]\},$$

$$0 < \sigma^2 < \infty . \qquad (2.10)$$

Again it is clear that the posterior density of $w = 1/\sigma^2$ is a mixture of gamma densities with parameters $\alpha_m = (n-1)/2$ and $\beta_m = 2/K(m, \phi_0)$. As before it is a straightforward matter to show that the mixing density is precisely the marginal posterior density of m.

Unconstrained Prior, Inference About ϕ_1

Inference about ϕ_1 can be based on the marginal posterior density of ϕ_1, which is given by

$$\pi_1(\phi_1) \propto \sum_{m=1}^{n-1} \int_0^\infty (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]\right\} d\sigma^2 .$$

Making use of the inverted gamma integral we obtain

$$\pi_1(\phi_1) \propto \sum_{m=1}^{n-1}\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]^{-n/2}$$

$$\propto \sum_{m=1}^{n-1} [K(m, \phi_0)]^{-n/2} [1 + (\tau(m, \phi_0)/(n-1)) (\phi_1 - \overline{X}_{m+1}^{n})^2]^{-[(n-1)+1]/2}$$

where $\tau(m, \phi_0) = (n-1)(n-m) / K(m, \phi_0)$. Thus we can write

$$\pi_1(\phi_1) \propto \sum_{m=1}^{n-1} (n-m)^{-1/2} [K(m, \phi_0)]^{-(n-1)/2} g_m(\phi_1; n-1, \overline{X}_{m+1}^{n}, \tau(m, \phi_0)) \tag{2.11}$$

where

$$g_m(\phi_1; n-1, \mu_m, \tau_m) = \left[\tau_m^{1/2}\,\Gamma(n/2)\right] \Big/ \left[\Gamma((n-1)/2)((n-1)\pi)^{1/2}\right] \left[1 + (\tau_m/(n-1))(\phi_1 - \mu_m)^2\right]^{\frac{-((n-1)+1)}{2}}$$

is the t density with location parameter μ_m, precision parameter τ_m, and $n-1$ degrees of freedom (See (22)). We see from (2.11) that the marginal posterior density of ϕ_1 is a mixture of t densities, and it is easy to verify that the mixing density has the value

$$\left[[K(m, \phi_0)]^{-n/2} / [\tau(m, \phi_0)]^{1/2}\right] \Big/ \sum_{i=1}^{n-1} [K(i, \phi_0)]^{-n/2} [\tau(i, \phi_0)]^{1/2}$$

at its m^{th} mass point, $m = 1, \ldots, n-1$.

Constrained Prior, Inference About m

The joint posterior density of m, σ^2 and ϕ_1 is now

$$\pi_1(m,\sigma^2,\phi_1) \propto (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2 + \sum_{i=m+1}^{n}(X_i-\phi_1)^2\right]\right\} \tag{2.12}$$

where $m \in I_{n-1}$, $0 < \sigma^2 < \infty$ and $\phi_0 < \phi_1 < \infty$. The marginal of m is thus

$$\pi_1(m) \propto \int_{\phi_0}^{\infty}\int_0^{\infty} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i-\phi_0)^2 + \sum_{i=m+1}^{n}(X_i-\phi_1)^2\right]\right\} d\sigma^2\, d\phi_1 \,.$$

Integration with respect to σ^2 proceeds as before to give

$$\pi_1(m) \propto \int_{\phi_0}^{\infty}\left[\sum_{i=1}^{m}(X_i-\phi_0)^2 + \sum_{i=m+1}^{n}(X_i-\phi_1)^2\right]^{-n/2} d\phi_1$$

$$\propto \int_{\phi_0}^{\infty}\left[K(m,\phi_0) + (n-m)(\phi_1 - \overline{X}_{m+1}^{\,n})^2\right]^{-n/2} d\phi_1$$

where $K(m,\phi_0)$ and $\tau(m,\phi_0)$ are as defined in an earlier section. This expression may now be written as

$$\pi_1(m) \propto \left[K(m,\phi_0)\right]^{-n/2}\int_{\phi_0}^{\infty}\left[1 + (\tau(m,\phi_0)/(n-1))(\phi_1 - \overline{X}_{m+1}^{\,n})^2\right]^{-\frac{[(n-1)+1]}{2}} d\phi_1 \,.$$

The integral is seen to be, apart from the norming constant, the upper tail of the t density with location parameter $\overline{X}_{m+1}^{\,n}$, precision parameter $\tau(m,\phi_0)$ and $n-1$ degrees of freedom. Inserting this norming constant we may write

$$\pi_1(m) \propto [n-m]^{-1/2}[K(m,\phi_0)]^{-(n-1)/2}[1 - T_{n-1,\,\overline{X}_{m+1}^{n},\,\tau(m,\phi_0)}(\phi_0)] \tag{2.13}$$

where $T_{k,a,b}(x)$ is the cumulative distribution function of a t random variable with location a, precision b, and k degrees of freedom. Now the general t distribution referred to above may be transformed to a Student's t by a translation and change of scale. That is, if Y is such a t random variable, then $Z = b^{1/2}[Y - a]$ has a Student's t distribution with k degrees of freedom. This enables us to write the cumulative distribution function used in (2.13) in terms of the distribution function of a Student's t with $n-1$ degrees of freedom, according to the formula

$$T_{k,a,b}(x) = \psi_k\left(b^{1/2}(\phi_0 - a)\right)$$

where $\psi_k(x)$ is the distribution function of a Student's t random variable with k degrees of freedom. We can then write for the posterior density of m

$$\pi_1(m) \propto [n-m]^{-1/2}[K(m,\phi_0)]^{-(n-1)/2}\left[1 - \psi_{n-1}\left(\tau(m,\phi_0)^{1/2}(\phi_0 - \overline{X}_{m+1}^{n})\right)\right] \tag{2.14}$$

where $m=1,\ldots,n-1$. One advantage of this formula is its extreme computational ease.

Constrained Prior, Inference About σ^2

The marginal posterior density of σ^2 is given by

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} \int_{\phi_0}^{\infty} (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\,\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\,\phi_1$$

$$\propto \sum_{m=1}^{n-1} (\sigma^2)^{-(n/2+1)} \exp\{(-1/2\,\sigma^2)\, K(m, \phi_0)\}$$

$$\int_{\phi_0}^{\infty} \exp\{(-(n-m)/2\ \sigma^2)\,(\phi_1 - \overline{X}^{\,n}_{m+1})^2\}\ d\,\phi_1\ .$$

Inserting the norming constant for the normal density on the right and integrating with respect to ϕ_1 we may write

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\,\sigma^2)\, K(m, \phi_0)\right\}\left[2\pi\,\sigma^2/(n-m)\right]^{1/2}$$

$$\left[1 - N\left((\phi_0 - \overline{X}^{\,n}_{m+1})/(\sigma/\sqrt{n-m}\,)\right)\right]$$

where $N(x)$ is the cumulative distribution function of a standard normal variate. Simplifying, we obtain

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} (n-m)^{-1/2} (\sigma^2)^{-(n+1)/2} \exp\left\{(-1/2\,\sigma^2)\, K(m, \phi_0)\right\}$$

$$\left[1 - N\left((\phi_0 - \overline{X}^{\,n}_{m+1})\ \ (\sigma/\sqrt{n-m}\,)\right)\right] \qquad (2.15)$$

for $0 < \sigma^2 < \infty$.

<u>Constrained Prior, Inference About ϕ_1</u>

From (2.12) we see that the marginal posterior density of ϕ_1 is

$$\pi_1(\phi_1) \propto \sum_{m=1}^{n-1} \int_0^\infty (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\sigma^2$$

$$\propto \sum_{m=1}^{n-1} \left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]^{-n/2} .$$

Proceeding exactly as in the unconstrained case we obtain

$$\pi_1(\phi_1) \propto \sum_{m=1}^{n-1} [K(m, \phi_0)]^{-n/2} [\tau(m, \phi_0)]^{-1/2} g_m(\phi_1; n-1, \overline{X}_{m+1}^{n}, \tau(m, \phi_0)),$$
$$\phi_0 < \phi_1 < \infty \qquad (2.16)$$

where $K(m, \phi_0)$, $\tau(m, \phi_0)$ and $g(x; k, a, b)$ are as previously defined. The posterior distribution in this case is thus a mixture of truncated t distributions. In order to find the mixing density we need to compute the norming constant K. Integrating (2.16) we get

$$K = \sum_{m=1}^{n-1} [K(m, \phi_0)]^{-n/2} [\tau(m, \phi_0)]^{-1/2} \left[1 - \psi_{n-1}\left(\tau(m, \phi_0)^{1/2} (\phi_0 - \overline{X}_{m+1}^{n})\right)\right] .$$

Referring again to (2.16) we see that the value of the mixing density at its m^{th} mass point is

$$p_m = [K(m, \phi_0)]^{-n/2} \tau(m, \phi_0)^{-1/2} \Big/ K \qquad (2.17)$$

for $m=1, \ldots, n-1$. As before, $\psi_{n-1}(x)$ is the distribution function of Student's t with $n-1$ degrees of freedom.

Neither ϕ_0 Nor ϕ_1 Known

As in the previous case we shall study the present situation for two vague prior densities. Firstly we shall assume that nothing is known about the order relation between ϕ_0 and ϕ_1 and assign the improper prior density

$$\pi_0(\phi_0, \phi_1) \propto \text{constant}, \quad -\infty < \phi_i < \infty, \quad i = 0, 1 . \qquad (2.18)$$

Next we shall assume it is known that $\phi_0 < \phi_1$ and assign the prior

$$\pi_0(\phi_0, \phi_1) \propto \text{constant}, \quad -\infty < \phi_0 < \phi_1 < \infty \qquad (2.19)$$

The theory for the case in which the order relation on the ϕ_i's is reversed parallels that being presented here and will not be presented separately. As before we shall use the terms "unconstrained" prior and "constrained" prior for (2.18) and (2.19) respectively.

Unconstrained Prior, Inference About m

The joint posterior density of m, σ^2, ϕ_0 and ϕ_1 is

$$\pi_1(m, \sigma^2, \phi_0, \phi_1) \propto (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]\right\} \qquad (2.20)$$

where $m \in I_{n-1}$, $0 < \sigma^2 < \infty$, and $-\infty < \phi_i < \infty$ for $i = 0, 1$. The marginal posterior density of m is given by

$$\pi_1(m) \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{0}^{\infty} (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\,\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\,\sigma^2 \, d\,\phi_0 \, d\,\phi_1 \, .$$

Integrating first with respect to σ^2 we obtain

$$\pi_1(m) \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]^{-n/2} d\,\phi_0 \, d\,\phi_1$$

$$\propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\sum_{i=1}^{m} (X_i - \overline{X}_1^{\,m})^2 + \sum_{i=m+1}^{n} (X_i - \overline{X}_1^{\,m})^2 + m(\phi_0 - \overline{X}_1^{\,m})^2 \right.$$

$$\left. + \, (n-m)(\phi_1 - \overline{X}_{m+1}^{\,n})^2\right]^{-n/2} d\,\phi_0 \, d\,\phi_1$$

where

$$\overline{X}_1^{\,m} = \left(\sum_{i=1}^{m} X_i\right)\Big/ m \qquad \text{and} \qquad \overline{X}_{m+1}^{\,n} = \left(\sum_{i=m+1}^{n} X_i\right)\Big/ (n-m) \, .$$

Let

$$C(m) = \sum_{i=1}^{m} (X_i - \overline{X}_1^{\,m})^2 + \sum_{i=m+1}^{n} (X_i - \overline{X}_{m+1}^{\,n})^2 \, .$$

Then we may write

$$\pi_1(m) \propto [C(m)]^{-n/2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[1 + \frac{1}{n-2} (\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))' \, T(m) (\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))\right]^{\frac{-((n-2)+2)}{2}} \tag{2.21}$$

where

$$\underset{\sim}{\phi} = \begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix}, \qquad \underset{\sim}{\mu}(m) = \begin{pmatrix} \overline{X}_1^{\,m} \\ \overline{X}_{m+1}^{\,n} \end{pmatrix}, \qquad T(m) = \begin{bmatrix} \dfrac{(n-2)m}{C(m)} & 0 \\ 0 & \dfrac{(n-2)(n-m)}{C(m)} \end{bmatrix} .$$

The double integral above is, apart from the norming constant, that of a bivariate t density and is easily evaluated to give

$$[\pi \cdot \Gamma((n-2)/2)C(m)] \Big/ [\Gamma(n/2)(m(n-m))^{1/2}] .$$

Substitution in (2.21) gives

$$\pi_1(m) \propto [m(n-m)]^{-1/2} [C(m)]^{-(n-2)/2}, \quad m = 1, \ldots, n-1 . \quad (2.22)$$

A comparison of (2.3), (2.9), and (2.22) displays an interesting pattern as more parameters are assumed unknown.

Unconstrained Prior, Inference About σ^2

The marginal posterior density of σ^2 is

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\phi_0 \, d\phi_1$$

$$\propto \sum_{m=1}^{n-1} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\, C(m)\right\}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left\{(-1/2)(\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))' \Sigma^{-1}(m, \sigma^2)(\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))\right\} d\phi_0 \, d\phi_1$$

where $\underset{\sim}{\phi}$ and $\underset{\sim}{\mu}(m)$ are as before and

$$\Sigma^{-1}(m, \sigma^2) = \begin{bmatrix} m/\sigma^2 & 0 \\ 0 & (n-m)/\sigma^2 \end{bmatrix} .$$

Using the form of the bivariate normal density to evaluate the double integral we obtain for the posterior density of σ^2

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} [m(n-m)]^{-1/2} (\sigma^2)^{-n/2} \exp\{(-1/2\sigma^2) C(m)\}, \quad 0 < \sigma^2 < \infty, \tag{2.23}$$

As in the previous case considered, the posterior density of $w = 1/\sigma^2$ is a mixture of gamma densities, the mixing density being the posterior density of the shift point.

Unconstrained Prior, Inference About (ϕ_0, ϕ_1)

The joint posterior density of ϕ_0 and ϕ_1 is

$$\pi_1(\phi_0, \phi_1) \propto \sum_{m=1}^{n-1} \int_0^\infty (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\sigma^2$$

$$\propto \sum_{m=1}^{n-1} \left[C(m) + m(\phi_0 - \bar{X}_1^m)^2 + (n-m)(\phi_1 - \bar{X}_{m+1}^n)^2\right]^{-n/2} .$$

Making use of the notation introduced earlier we may write

$$\pi_1(\phi_0, \phi_1) \propto \sum_{m=1}^{n-1} [C(m)]^{-n/2}\left[1 + \frac{1}{n-2} (\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))' T(m) (\underset{\sim}{\phi} - \underset{\sim}{\mu}(m))\right]^{\frac{-((n-2)+2)}{2}}, \tag{2.24}$$

It is now clear that we have a mixture of bivariate t densities. Letting $h_m(\underset{\sim}{\phi}; n-2, \underset{\sim}{\mu}(m), T(m))$ denote the bivariate t density with $n-2$ degrees of freedom, location parameter $\underset{\sim}{\mu}(m)$, and precision matrix $T(m)$ we may write (2.24) as

$$\pi_1(\phi_0, \phi_1) \propto \sum_{m=1}^{n-1} [m(n-m)]^{-1/2} [C(m)]^{\frac{-n-2}{2}} h_m(\underset{\sim}{\phi}; n-2, \underset{\sim}{\mu}(m), T(m)),$$

$$-\infty < \phi_i < \infty, \ i = 1, 2. \quad (2.25)$$

Some straightforward algebra shows that in this case also the mixing density for the bivariate t distributions is the posterior density of the shift point m.

Constrained Prior, Inference About m

The joint posterior density for this last case to be considered is

$$\pi_1(m, \sigma^2, \phi_0, \phi_1) \propto (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]\right\} \quad (2.26)$$

where $m \in I_{n-1}$, $0 < \sigma^2 < \infty$, and $-\infty < \phi_0 < \phi_1 < \infty$. The marginal posterior density of m is thus

$$\pi_1(m) \propto \int_{-\infty}^{\infty} \int_{\phi_0}^{\infty} \int_{0}^{\infty} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]\right\} d\sigma^2 \, d\phi_1 \, d\phi_0$$

$$\propto \int_{-\infty}^{\infty} \int_{\phi_0}^{\infty} \left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \phi_1)^2\right]^{-n/2} d\phi_1 \, d\phi_0 \, .$$

Using the general t density to integrate on ϕ_1 we obtain

$$\pi_1(m) \propto \int_{-\infty}^{\infty} (n-m)^{-1/2} \left[\sum_{i=1}^{m}(X_i - \phi_0)^2 + \sum_{i=m+1}^{n}(X_i - \overline{X}_{m+1}^{\,n})^2\right]^{-(n-1)/2}$$

$$\left[1 - \psi_{n-1}(\tau(m, \phi_0)^{1/2}(\phi_0 - \overline{X}_{m+1}^{\,n}))\right] d\phi_0 \quad (2.27)$$

where $\psi_{n-1}(x)$ is the distribution function of Student's t with $n-1$ degrees of freedom and $\tau(m, \phi_0)$ is defined by

$$\tau(m, \phi_0) = (n-1)(n-m) \Big/ \left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \overline{X}_{m+1}^{\,n})^2 \right]$$

Using a well known identity it is possible to write (2.27) as

$$\pi_1(m) \propto [m(n-m)]^{1/2} [C(m)]^{\frac{-n-2}{2}} \int_{-\infty}^{\infty} \left[1 - \psi_{n-1}\left(\tau(m, \phi_0)^{1/2} (\phi_0 - \overline{X}_{m+1}^{\,n}) \right) \right]$$
$$\cdot\, g_m\left(\phi_0 ; n-2, \overline{X}_1^{\,m}, w(m) \right) d\,\phi_0 \quad (2.28)$$

where $w(m) = m(n-2)/C(m)$ and $g_m(\phi_0; k, a, b)$ is the general t density defined earlier. We can thus write (2.28) as

$$\pi_1(m) \propto [m(n-m)]^{-1/2} [C(m)]^{\frac{-n-2}{2}} E_{\phi_0} \left[1 - \psi_{n-1}\left(\tau(m, \phi_0)^{1/2} (\phi_0 - \overline{X}_{m+1}^{\,n}) \right) \right] \quad (2.29)$$

where $m = 1, \ldots, n-1$ and E_{ϕ_0} is the expectation of the indicated function of ϕ_0 taken with respect to a general t density with $n-2$ degrees of freedom, location parameter $\overline{X}_1^{\,m}$, and precision $w(m)$. Using the transformation

$$y = w(m)^{1/2} (\phi_0 - \overline{X}_1^{\,m})$$

we know that y is distributed as Student's t with $n-2$ degrees of freedom. This enables us to express the expectation in (2.29) with respect to a Student's t density. Straightforward computation shows that this expectation then becomes

$$E_y \left[1 - \psi_{n-1}(H(m, y)) \right]$$

where

$$H(m, y) = \left[(n-1)(n-2)(n-m)\Big/\Big(C(m)(y^2+n-2)\Big)\right]^{1/2}\Big[[m(n-2)\ C(m)]^{-1/2}\,y + [\bar{X}_1^m - \bar{X}_{m+1}^n]\Big] .$$

The expectation of the indicated function of y is now taken with respect to a Student's t distribution with $n-2$ degrees of freedom. We can now write the marginal posterior density of m as

$$\pi_1(m) \propto [m(n-m)]^{-1/2}\,[C(m)]^{-(n-2)/2}\,E_y\Big[1 - \Psi_{n-1}\Big(H(m, y)\Big)\Big] , \qquad m = 1, \ldots, n-1 . \tag{2.30}$$

It is clear that this formula presents considerable, though not insurmountable, computational difficulty, and that a numerical integration technique of some kind would be needed to evaluate this density for a given set of data.

Constrained Prior, Inference About σ^2

The marginal posterior density of σ^2 for this case is

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} \int_{-\infty}^{\infty} \int_{\phi_0}^{\infty} (\sigma^2)^{-(n/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\,\phi_1\, d\,\phi_0$$

$$\propto \sum_{m=1}^{n-1} (\sigma^2)^{-(n/2+1)} \int_{-\infty}^{\infty} \exp\left\{(-1/2\,\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \bar{X}_{m+1}^n)^2\right]\right\} \cdot \left(\frac{2\pi\sigma^2}{n-m}\right)^{1/2} \left[1 - N\left(\frac{\phi_0 - \bar{X}_{m+1}^n}{\sigma\big/\sqrt{n-m}}\right)\right] d\,\phi_0$$

$$\propto \sum_{n=1}^{m-1} [m(n-m)]^{1/2} (\sigma^2)^{-n/2} \exp\{(-1/2\sigma^2)\, C(m)\}\, E_{\phi_0}\left[1 - N\left(\frac{\phi_0 - \bar{X}_{m+1}^{\,n}}{\sigma/\sqrt{n-m}}\right)\right] \tag{2.31}$$

where $N(x)$ is the distribution function of a standard normal variate and the expected value of the indicated function of ϕ_0 is taken with respect to a normal density with mean $\bar{X}_1^m$ and variance σ^2/m. Letting

$$z = \left(\phi_0 - \bar{X}_1^m\right) \Big/ \left(\sigma/\sqrt{m}\right)$$

we may take the expectation with respect to a standard normal and write

$$\pi_1(\sigma^2) \propto \sum_{m=1}^{n-1} [m(n-m)]^{-1/2} \exp\{(-1/2\,\sigma^2)\, C(m)\}$$

$$\cdot E_z\left[1 - N\left(z\sqrt{\frac{n-m}{m}} + \frac{\bar{X}_1^m - \bar{X}_{m+1}^{\,n}}{\sigma/\sqrt{n-m}}\right)\right] \tag{2.32}$$

where $0 < \sigma^2 < \infty$ and the expectation of the indicated function of z is taken with respect to a standard normal variable.

<u>Constrained Prior, Inference About (ϕ_0, ϕ_1)</u>

The joint posterior density of ϕ_0 and ϕ_1 is

$$\pi_1(\phi_0, \phi_1) \propto \sum_{m=1}^{n-1} \int_0^\infty (\sigma^2)^{-(n/2+1)}$$

$$\exp\left\{(-1/2\,\sigma^2)\left[\sum_{i=1}^{m} (X_i - \phi_0)^2 + \sum_{i=m+1}^{n} (X_i - \phi_1)^2\right]\right\} d\,\sigma^2$$

where $-\infty < \phi_0 < \phi_1 < \infty$. Evaluation of the integral and simplification of the subsequent expression proceeds exactly as in the case of the unconstrained prior density, keeping in mind the added restriction on ϕ_0 and ϕ_1. The joint posterior density of ϕ_0 and ϕ_1 may then be written as

$$\pi_1(\phi_0, \phi_1) \propto \sum_{m=1}^{n-1} [m(n-m)]^{-1/2} [C(m)]^{-(n-2)/2} h_m(\underset{\sim}{\phi}; n-2, \underset{\sim}{\mu}(m), T(m)),$$
$$-\infty < \phi_0 < \phi_1 < \infty. \tag{2.33}$$

As before, $h_m(\underset{\sim}{\phi}; n-2, \underset{\sim}{\mu}(m), T(m))$ is the bivariate t density with $n-2$ degrees of freedom, location parameter $\underset{\sim}{\mu}(m) = (\overline{X}_1^m, \overline{X}_{m+1}^n)'$, and precision matrix

$$T(m) = \begin{bmatrix} \frac{(n-2)m}{C(m)} & 0 \\ 0 & \frac{(n-2)(n-m)}{C(m)} \end{bmatrix}.$$

CHAPTER III

POSTERIOR DISTRIBUTIONS RELATED TO TWO-PHASE REGRESSION FOR THE DISCRETE CASE

As a generalization of the normal sequence studied in Chapter II we shall in this chapter derive certain posterior distributions related to making inferences about a two-phase regression.

We assume that we have observations on a sequence $Y_1, \ldots, Y_T$, $T \geq 5$, of independent random variables which follow two separate linear regression regimes. As before we shall introduce a discrete random variable m for the unknown switch point, and we shall further assume that the state space of m is the set $I_{T-2} = \{2, 3, \ldots, T-2\}$. We thereby assume that we have at least two observations on each regression. We thus have

$$Y_i, \; i = 1, \ldots, m, \text{ independently distributed } N(\alpha_1 + \beta_1 X_i, \sigma^2),$$

and

$$Y_j, \; j = m+1, \ldots, T, \text{ independently distributed } N(\alpha_2 + \beta_2 X_j, \sigma^2),$$

where $\sigma^2 > 0$, $X_1 < \ldots < X_T$ are non-stochastic regressor variables, and m is the unknown switch point.

In such a situation, interest centers on estimating the switch point m as well as any unknown regression parameters and the possibly unknown variance σ^2.

We shall study four cases in this chapter:

(i) Error variance known, first regression known

(ii) Error variance known, neither regression known

(iii) Error variance unknown, first regression known

(iv) Error variance unknown, neither regression known.

Our interest shall center primarily on the shift point m, but in some cases posterior distributions for the regression parameters also will be derived. In all cases we shall assume that prior knowledge is such that independent, diffuse prior densities for the unknown parameters will be adequate.

Error Variance Known, First Regression Known

Without loss of generality we shall assume that $\sigma^2 = 1$. α_2, β_2 and m are assumed a priori independent with prior densities

$$\pi_0(\alpha_2, \beta_2) \propto \text{constant}, \quad -\infty < \alpha_2 < \infty, \quad -\infty < \beta_2 < \infty,$$

and

$$\pi_0(m) = \begin{cases} 1/(T-3), & m = 2, \ldots, T-2 \\ 0 & \text{elsewhere} . \end{cases}$$

The likelihood function is

$$L(m, \alpha_2, \beta_2) = (2\pi)^{-T/2} \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i]^2\right]\right\}$$

resulting in a joint posterior density of

$$\pi_1(m,\alpha_2,\beta_2) \propto \exp\left\{(-1/2)\left[\sum_{i=1}^{m}[Y_i-(\alpha_1+\beta_1 X_i)]^2+\sum_{i=m+1}^{T}[Y_i-(\alpha_2+\beta_2 X_i)]^2\right]\right\} \tag{3.1}$$

for $m=2,\ldots,T-2$, $-\infty<\alpha_2<\infty$ and $-\infty<\beta_2<\infty$.

Inference About m

The marginal posterior density of m is

$$\pi_1(m) \propto \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\pi_1(m,\alpha_2,\beta_2)\,d\alpha_2\,d\beta_2, \qquad m\in I_{T-2}. \tag{3.2}$$

In order to evaluate this integral we shall use the identity

$$\sum_{i=m+1}^{T}[Y_i-(\alpha_2+\beta_2 X_i)]^2 = \sum_{i=m+1}^{T}[Y_i-\hat{Y}_i^m]^2+(\underset{\sim}{\alpha}-\underset{\sim}{\mu}^m)'\,\Sigma_m^{-1}(\underset{\sim}{\alpha}-\underset{\sim}{\mu}^m) \tag{3.3}$$

where

$$\alpha = \begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix},$$

$$\underset{\sim}{\mu}^m = \begin{pmatrix}\hat{\alpha}_2^m\\ \hat{\beta}_2^m\end{pmatrix} = \begin{pmatrix}\bar{Y}_{m+1}^T-\hat{\beta}_2^m\bar{X}_{m+1}^T\\ \sum_{i=m+1}^{T}(X_i-\bar{X}_{m+1}^T)(Y_i-\bar{Y}_{m+1}^T)\Big/\sum_{i=m+1}^{T}(X_i-\bar{X}_{m+1}^T)^2\end{pmatrix},$$

$$\Sigma_m^{-1} = \begin{bmatrix}(T-m) & \sum_{i=m+1}^{T}X_i\\ \sum_{i=m+1}^{T}X_i & \sum_{i=m+1}^{T}X_i^2\end{bmatrix},$$

$$\bar{X}_{m+1}^{T} = \left(\sum_{i=m+1}^{T} X_i \right) \Big/ \left(T - m \right), \qquad \bar{Y}_{m+1}^{T} = \left(\sum_{i=m+1}^{T} Y_i \right) \Big/ \left(T - m \right),$$

and

$$\hat{Y}_i^m = \hat{\alpha}_2^m + \hat{\beta}_2^m X_i .$$

Substituting identity (3.3) in (3.1), the posterior density (3.2) may be written

$$\pi_1(m) \propto \exp\left\{ (-1/2) \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^m]^2 \right] \right\}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{(-1/2)(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \, \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\} \, d\alpha .$$

The integral may now be easily evaluated using the bivariate normal density to give as the posterior density of m,

$$\pi_1(m) \propto \left[(T-m) \sum_{i=m+1}^{T} (X_i - \bar{X}_{m+1}^{T})^2 \right]^{-1/2}$$

$$\exp\left\{ (-1/2) \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^m]^2 \right] \right\} \qquad (3.4)$$

for $m = 2, \ldots, T-2$.

Inference About β_2

The marginal posterior density of β_2 is given by

$$\pi_1(\beta_2) \propto \sum_{m=2}^{T-2} \int_{-\infty}^{\infty} \pi_1(m, \alpha_2, \beta_2) \, d\alpha_2 , \quad -\infty < \beta_2 < \infty. \qquad (3.5)$$

To evaluate this integral we shall use the identity

$$\sum_{i=m+1}^{T}[Y_i-(\alpha_2+\beta_2 X_i)]^2 = \sum_{i=m+1}^{T}\left[(Y_i-\overline{Y}_{m+1}^{T})-\beta_2(X_i-\overline{X}_{m+1}^{T})\right]^2$$
$$+ (T-m)\left[\alpha_2-(\overline{Y}_{m+1}^{T}-\beta_2\overline{X}_{m+1}^{T})\right]^2 . \quad (3.6)$$

Substituting (3.6) in (3.1) and integrating with respect to α_2, (3.5) becomes

$$\pi_1(\beta_2) \propto \sum_{m=2}^{T-2}(T-m)^{-1/2}\exp\left\{(-1/2)\sum_{i=1}^{m}[Y_i-(\alpha_1+\beta_1 X_i)]^2\right\}$$
$$\exp\left\{(-1/2)\sum_{i=m+1}^{T}\left[(Y_i-\overline{Y}_{m+1}^{T})-\beta_2(X_i-\overline{X}_{m+1}^{T})\right]^2\right\}. \quad (3.7)$$

Substituting

$$\sum_{i=m+1}^{T}[(Y_i-\overline{Y}_1^{T})-\beta_2(X_i-\overline{X}_1^{T})]^2 = \sum_{i=m+1}^{T}(X_i-\overline{X}_{m+1}^{T})^2(\beta_2-\hat{\beta}_2^{m})^2$$
$$+\sum_{i=m+1}^{T}(Y_i-\hat{Y}_i^{m})^2$$

in (3.7) we may write the marginal posterior density of β_2 as

$$\pi_1(\beta_2) \propto \sum_{m=2}^{T-2}K(m)\, g_m(\beta_2;\hat{\beta}_2^{m},\operatorname{var}(\hat{\beta}_2^{m})), \quad -\infty<\beta_2<\infty, \quad (3.8)$$

where

$$K(m) = \left[(T-m)\sum_{i=m+1}^{T}(X_i-\overline{X}_{m+1}^{T})^2\right]^{-1/2}$$
$$\exp\left\{(-1/2)\left[\sum_{i=1}^{m}[Y_i-(\alpha_1+\beta_1 X_i)]^2+\sum_{i=m+1}^{T}[Y_i-\hat{Y}_i^{m}]^2\right]\right\},$$

$$\operatorname{var}(\hat{\beta}_2^m) = 1 \Bigg/ \sum_{i=m+1}^{T} (X_i - \bar{X}_{m+1}^{T})^2 ,$$

and $g(y\,;\,m, v)$ is the normal density with mean m and variance v. We thus obtain a mixture of normal densities, the mixing density being precisely the posterior density of the shift point m.

Inference About α_2

The derivation of the marginal posterior density of α_2 proceeds along the same lines as that of β_2 and results in the density

$$\pi_1(\alpha_2) \propto \sum_{m=2}^{T-2} K(m)\, g_m(\alpha_2\,;\,\hat{\alpha}_2^m,\, \operatorname{var}(\hat{\alpha}_2^m)) , \tag{3.9}$$

where now

$$\operatorname{var}(\hat{\alpha}_2^m) = \left(\sum_{i=m+1}^{T} X_i^2\right) \Bigg/ \left[(T-m) \sum_{i=m+1}^{T} (X_i - \bar{X}_{m+1}^{T})^2\right] .$$

Before proceeding to the next case we shall make an observation on the normal distributions involved in (3.8) and (3.9). For each m, the mean and variance of the m^{th} density in the mixture are the least squares estimate of the parameter and its variance respectively, based on the last $T - m$ data points in the sequence.

Error Variance Known, Neither
Regression Known

Assuming as before that $\sigma^2 = 1$ the likelihood function is now

$$L(m, \alpha_1, \beta_1, \alpha_2, \beta_2) = (2\pi)^{-T/2}$$
$$\exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}.$$

Assigning independent, improper uniform prior densities to the regression parameters and a discrete uniform prior to the switch parameter m we obtain the joint posterior density

$$\pi_1(m, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}. \quad (3.10)$$

for $m = 2, \ldots, T-2$, $-\infty < \alpha_i < \infty$, $i = 1, 2$ and $-\infty < \beta_i < \infty$, $i = 1, 2$.

Applying to the first m data points an identity analogous to (3.3) we may write

$$\pi_1(m, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto \exp\left\{(-1/2)\left[\sum_{i=1}^{m} (Y_i - \hat{Y}_i^{m,L})^2 + \sum_{i=m+1}^{T} (Y_i - \hat{Y}_i^{m,U})^2 + (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right]\right\} \quad (3.11)$$

where

$$\underset{\sim}{\alpha} = (\alpha_1, \beta_1, \alpha_2, \beta_2)', \qquad \underset{\sim}{\mu}^m = (\hat{\alpha}_1^m, \hat{\beta}_1^m, \hat{\alpha}_2^m, \hat{\beta}_2^m)',$$

$$\hat{Y}_i^{m,L} = \hat{\alpha}_1^m + \hat{\beta}_1^m X_i, \qquad \hat{Y}_i^{m,U} = \hat{\alpha}_2^m + \hat{\beta}_2^m X_i,$$

$$\Sigma_m^{-1} = \begin{bmatrix} m & \sum_1^m X_i & & \\ \sum_1^m X_i & \sum_1^m X_i^2 & \multicolumn{2}{c}{\emptyset} \\ & & T-m & \sum_{m+1}^T X_i \\ \multicolumn{2}{c}{\emptyset} & \sum_{m+1}^T X_i & \sum_{m+1}^T X_i^2 \end{bmatrix} .$$

Of course $\hat{\alpha}_2^m$ and $\hat{\beta}_2^m$ are as defined in the previous case and $\hat{\alpha}_1^m$ and $\hat{\beta}_1^m$ are their counterparts for the first m data points.

The regression parameters may now be integrated out quite easily using the four variate normal integral to obtain as the posterior density of m

$$\pi_1(m) \propto \left[m(T-m) \sum_{i=1}^{m} (X_i - \overline{X}_1^m)^2 \sum_{i=m+1}^{T} (X_i - \overline{X}_{m+1}^T)^2 \right]^{-1/2}$$

$$\exp\left\{ (-1/2) \left[\sum_{i=1}^{m} (Y_i - \hat{Y}_i^{m,L})^2 + \sum_{i=m+1}^{T} (Y_i - \hat{Y}_i^{m,U})^2 \right] \right\} \qquad (3.12)$$

for $m = 2, \ldots, T-2$.

A comparison of this result with (3.4) shows that the known first regression has been replaced by its estimate for each m and the weighting factor has been altered accordingly. This density will be computed for a particular set of data in Chapter V.

Error Variance Unknown, First Regression Known

The likelihood function for the present case is

$$L(m, \sigma^2, \alpha_2, \beta_2) = (2\pi\sigma^2)^{-T/2} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1+\beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2+\beta_2 X_i)]^2\right]\right\} .$$

To the new unknown parameter σ^2 we shall assign the improper prior density

$$\pi_1(\sigma^2) \propto 1/\sigma^2, \quad 0 < \sigma^2 < \infty$$

while assigning to the remaining parameters the same prior densities as before. Applying Bayes's theorem we obtain the joint posterior density

$$\pi_1(m, \sigma^2, \alpha_2, \beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\} \quad (3.13)$$

for $m = 2, \ldots, T-2$, $0 < \sigma^2 < \infty$, $-\infty < \alpha_2 < \infty$ and $-\infty < \beta_2 < \infty$.

Integration on σ^2 can easily be performed using the inverted gamma density to obtain as the joint posterior density of m, α_2 and β_2

$$\pi_1(m, \alpha_2, \beta_2) \propto \left[\sum_{i=1}^{m} Y_i - (\alpha_1+\beta_1 X_i)^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2+\beta_2 X_i)]^2\right]^{-T/2} \quad (3.14)$$

for $m = 2, \ldots, T-2$, $-\infty < \alpha_2 < \infty$ and $-\infty < \beta_2 < \infty$. Using identity (3.3) and the same notation as used in case one we can write

$$\pi_1(m, \alpha_2, \beta_2) \propto \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^{m,U}]^2 + (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right]^{-T/2}$$

$$\propto K(m)^{-T/2} \left[1 + K(m)^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right]^{\frac{-((T-2)+2)}{2}} \quad (3.15)$$

where

$$K(m) = \sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^{m,U}]^2 .$$

We may now integrate out α_2 and β_2 using the bivariate t integral to obtain as the posterior density of m

$$\pi_1(m) \propto \left[(T-m) \sum_{i=m+1}^{T} (X_i - \overline{X}_{m+1}^{T})^2\right]^{-1/2} \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^{m,U}]^2\right]^{\frac{-(T-2)}{2}} \quad (3.16)$$

for $m = 2, \ldots, T-2$.

Error Variance Unknown, Neither Regression Known

The likelihood function for this the last case of this chapter is

$$L(m, \sigma^2, \alpha_1, \beta_1, \alpha_2, \beta_2) = (2\pi\sigma^2)^{-T/2} \exp\left\{(-1/2\,\sigma^2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}.$$

Assigning the same prior distributions as in the previous cases we obtain the joint posterior density

$$\pi_1(m, \sigma^2, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}$$

for $m = 2, \ldots, T-2$, $0 < \sigma^2 < \infty$, $-\infty\ \alpha_i < \infty$, $-\infty < \beta_i < \infty$, $i = 1, 2$.

Integration on σ^2 proceeds as in case three to give for the joint posterior density of m and the regression parameters

$$\pi_1(m, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]^{-T/2}.$$

Making use of the identities given in case two we can write, using the notation of that section,

$$\pi_1(m, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto K(m)^{-T/2}\left[1 + K(m)^{-1}(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1}(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right]^{-\frac{((T-4)+4)}{2}}$$

where the value of $K(m)$ is now given by

$$K(m) = \sum_{i=1}^{m} (Y_i - \hat{Y}_i^{m,L})^2 + \sum_{i=m+1}^{T} (Y_i - \hat{Y}_i^{m,U})^2 .$$

We may now integrate out the regression parameters using the four variate t integral to obtain as the marginal posterior density of m

$$\pi_1(m) \propto \left[m(T-m) \sum_{i=1}^{m} (X_i - \overline{X}_1^m)^2 \sum_{i=m+1}^{T} (X_i - \overline{X}_{m+1}^T)^2 \right]^{-1/2}$$

$$\left[\sum_{i=1}^{m} (Y_i - \hat{Y}_i^{m,L})^2 + \sum_{i=m+1}^{T} (Y_i - \hat{Y}_i^{m,U})^2 \right]^{-(T-4)/2} \quad (3.17)$$

for $m = 2, \ldots, T-2$.

In closing this chapter we point out the similarity between (3.16) and (3.17), and remind the reader that examples of (3.12) and (3.17) will be presented in Chapter V.

CHAPTER IV

POSTERIOR DISTRIBUTIONS RELATED TO TWO-PHASE REGRESSION FOR THE CONTINUOUS CASE

In the previous chapter our interest was centered on the index m at which the switch from one regression regime to another occurred. In some cases interest may center more on the abscissa of the point of intersection of the two regression lines. An easy calculation shows that this is given by $\gamma = (\alpha_2 - \alpha_1)/(\beta_1 - \beta_2)$.

As in chapter three there are many cases one might consider. Those discussed in this paper are

(i) Error variance known, first regression known, m known

(ii) Error variance known, first regression known, m unknown

(iii) Error variance known, neither regression known, m known

(iv) Error variance known, neither regression known, m unknown

(v)-(viii) As above, but with the error variance unknown.

The approach taken will be to find the joint posterior density of the unknown regression parameters under the assignment of diffuse prior densities, and then to obtain from this the distribution of that function of these parameters which gives the intersection. As mentioned above, the function whose distribution we require is

$$f(\alpha_1, \alpha_2, \beta_1, \beta_2) = (\alpha_2 - \alpha_1)/(\beta_1 - \beta_2) .$$

Before proceeding to study the cases in turn we remark that the cases corresponding to "second regression known" are completely analogous to our "first regression known", and for that reason are not being studied separately here.

σ^2 Known, First Regression Known, m Known

The likelihood function for this case is

$$L(\alpha_2, \beta_2) = (2\pi)^{-T/2} \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}.$$

As before we are assuming for convenience that $\sigma^2 = 1$. Combining this likelihood function with the same diffuse prior densities used in Chapter III we obtain, in accordance with Bayes's theorem

$$\pi_1(\alpha_2, \beta_2) \propto \exp\left\{(-1/2) \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right\} \tag{4.1}$$

for $-\infty < \alpha_2 < \infty$ and $-\infty < \beta_2 < \infty$. Using identity (3.3) of Chapter III we can write

$$\pi_1(\alpha_2, \beta_2) \propto \exp\{(-1/2)(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\} \tag{4.2}$$

where

$$\underset{\sim}{\alpha} = (\alpha_2, \beta_2)', \quad \underset{\sim}{\mu}^m = (\hat{\alpha}_2^m, \hat{\beta}_2^m)', \quad \Sigma_m^{-1} = \begin{bmatrix} (T-m) & \sum_{m+1}^{T} X_i \\ \sum_{m+1}^{T} X_i & \sum_{m+1}^{T} X_i^2 \end{bmatrix}.$$

Of course $\hat{\alpha}_2^m$ and $\hat{\beta}_2^m$ are the usual least squares estimates of α_2 and β_2 based on the last T-m data points of the sequence. We now transform to

$$w_1 = \alpha_2 - \alpha_1$$

$$w_2 = \beta_2 - \beta_1 \ .$$

The joint density of w_1 and w_2 is easily seen to be bivariate normal with the mean vector relocated at

$$\underset{\sim}{v}^m = (\hat{\alpha}_2^m - \alpha_1 , \hat{\beta}_2^m - \beta_2)' .$$

Finally we make the transformation to

$$\gamma_1 = -w_1/w_2$$

$$\gamma_2 = w_2$$

and obtain as the joint density of γ_1 and γ_2

$$\pi_1(\gamma_1, \gamma_2) \propto \exp\{(1/2)A(m,\gamma_1)B^2(m,\gamma_1)\} \, |\gamma_2|$$
$$\cdot \exp\{(-1/2)A(m,\gamma_1)[\gamma_2 - B(m,\gamma_1)]^2\}$$

for $-\infty < \gamma_i < \infty$, $i = 1,2$ where

$$A(m,\gamma_1) = \sum_{i=m+1}^{T} (X_i - \gamma_1)^2 ,$$

$$B(m,\gamma_1) = \sum_{i=m+1}^{T} (X_i - \gamma_1)(\hat{Y}_i^{m,U} - (\alpha_1 + \beta_1 X_i)) / A(m,\gamma_1) ,$$

$$\hat{Y}_i^{m,U} = \hat{\alpha}_2^m + \hat{\beta}_2^m X_i .$$

Integration on γ_2 now yields the posterior density of the intersection

$$\pi_1(\gamma_1) \propto A^{-1/2}(m,\gamma_1) \exp\{(1/2) A(m,\gamma_1) B^2(m,\gamma_1)\} \; E[|V|] \quad (4.3)$$

where

$$V \sim N\Big(B(m,\gamma_1),\, A^{-1}(m,\gamma_1)\Big) .$$

It may be shown that if X is a random variable normally distributed with mean μ and variance σ^2 then

$$E[|X|] = \mu[2\,\phi(\mu/\sigma) - 1] + (\pi/2\,\sigma^2)^{-1/2} \exp\{-\mu^2/2\,\sigma^2\} . \quad (4.4)$$

Using (4.4) to obtain $E[|V|]$ in (4.3) we obtain

$$\pi_1(\gamma_1) \propto A^{-1}(m,\gamma_1)\Bigg\{B(m,\gamma_1)\Big[2\,\phi\Big(B(m,\gamma_1)A^{1/2}(m,\gamma_1)\Big) - 1\Big] \exp\Big[(1/2) A(m,\gamma_1) B^2(m,\gamma_1)\Big] + \Big[\pi A(m,\gamma_1)/2\Big]^{-1/2}\Bigg\} \quad (4.5)$$

for $-\infty < \gamma_1 < \infty$.

σ^2 Known, First Regression Known, m Unknown

Proceeding as in the previous case and introducing a discrete uniform prior density on m we obtain as the joint posterior density of m, α_2 and β_2

$$\pi_1(m,\alpha_2,\beta_2) \propto \exp\Bigg\{(-1/2)\Bigg[\sum_{i=1}^{m}[Y_i - (\alpha_1+\beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2+\beta_2 X_i)]^2\Bigg]\Bigg\} \quad (4.6)$$

for $m = 2, \ldots, T-2$, $-\infty < \alpha_2 < \infty$ and $-\infty < \beta_2 < \infty$.

Using identity (3.3) we can write (4, 6) as

$$\pi_1(m, \alpha_2, \beta_2) \propto g(m) \exp\left\{(-1/2)(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1}(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right\}$$

where

$$g(m) = \exp\left\{(-1/2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}$$

and the remaining notation agrees with that introduced earlier. We again transform first to w_1 and w_2 as before and then to

$$\gamma_1 = -w_1/w_2$$

$$\gamma_2 = w_2$$

and obtain the joint density

$$\pi_1(m, \gamma_1, \gamma_2) \propto h(m)\,|\gamma_2|\,\exp\left\{(1/2)\,A(m, \gamma_1)\,B^2(m, \gamma_1)\right\}$$
$$\exp\left\{(-1/2)\,A(m, \gamma_1)(\gamma_2 - B(m, \gamma_1))^2\right\}$$

where

$$h(m) = \exp\left\{(-1/2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[\hat{Y}_i^{m,U} - (\alpha_1 + \beta_1 X_i)]^2 \right.\right.$$
$$\left.\left. + \sum_{i=m+1}^{T}[Y_i - \hat{Y}_i^{m,U}]^2\right]\right\}. \quad (4,7)$$

Integrating with respect to γ_2 and summing on m we obtain the posterior distribution of the intersection

$$\pi_1(\gamma_1) \propto \sum_{m=2}^{T-2} h(m)\,A^{-1/2}(m, \gamma_1)\left\{B(m, \gamma_1)\left[2\,\phi\left(B(m, \gamma_1)\,A^{1/2}(m, \gamma_1)\right) - 1\right]\right.$$
$$\left. \exp[(1/2)\,A(m, \gamma_1)\,B^2(m, \gamma_1)] + [\pi\,A(m, \gamma_1)/2]^{-1/2}\right\} \quad (4,8)$$

where $-\infty < \gamma_1 < \infty$. It is seen that the density in this case is a mixture of densities of the type (4.5) with the mixing density being given by equation (4.7).

σ^2 Known, Neither Regression Known, m Known

In this case the joint posterior density of the four regression parameters is easily seen to be

$$\pi_1(\alpha_1, \beta_1, \alpha_2, \beta_2) \propto \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}$$

for $-\infty < \alpha_i < \infty$ and $-\infty < \beta_i < \infty$, $i = 1, 2$.

Applying identity (3.3) and its counterpart for the first m data points we may write

$$\pi_1(\alpha_1, \beta_1, \alpha_2, \beta_2) \propto E(m) \exp\left\{(-1/2)(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1}(\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right\} \qquad (4.9)$$

where

$$E(m) = \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [Y_i - \hat{Y}_i^{m,L}]^2 + \sum_{i=m+1}^{T} [Y_i - \hat{Y}_i^{m,U}]^2\right]\right\},$$

$$\underset{\sim}{\alpha} = (\alpha_1, \beta_1, \alpha_2, \beta_2)', \qquad \underset{\sim}{\mu}^m = (\hat{\alpha}_1^m, \hat{\beta}_1^m, \hat{\alpha}_2^m, \hat{\beta}_2^m)',$$

$$\hat{Y}_i^{m,L} = \hat{\alpha}_1^m + \hat{\beta}_1^m X_i, \qquad \hat{Y}_i^{m,U} = \hat{\alpha}_2^m + \hat{\beta}_2^m X_i,$$

and

$$\Sigma_m^{-1} = \begin{bmatrix} m & \sum_1^m X_i & & \\ \sum_1^m X_i & \sum_1^m X_i^2 & \multicolumn{2}{c}{\emptyset} \\ & & (T-m) & \sum_{m+1}^T X_i \\ \multicolumn{2}{c}{\emptyset} & \sum_{m+1}^T X_i & \sum_{m+1}^T X_i^2 \end{bmatrix} .$$

$(\hat{\alpha}_1^m, \hat{\beta}_1^m)$ and $(\hat{\alpha}_2^m, \hat{\beta}_2^m)$ are the usual least squares estimates of the regression parameters based on the first m and last $T-m$ data points respectively. We now transform to

$$w_1 = \alpha_1 - \alpha_2$$

$$w_2 = \alpha_2$$

$$w_3 = \beta_1 - \beta_2$$

$$w_4 = \beta_2$$

Substitution in (4, 9) yields quadratic expressions in w_2 and w_4 in the exponent and these may be integrated out using the form of the normal density to obtain

$$\pi_1(w_1, w_3) \propto E(m) \exp\left\{(-1/2)\left[\sum_{i=1}^{m} [(w_1 - \hat{\alpha}_1^m) + (w_3 - \hat{\beta}_1^m) X_i]^2 - C^2(m, w_1, w_3) \Big/ \sum_{i=1}^{T} (X_i - \overline{X}_1^T) - B^2(m, w_1, w_3)/T\right]\right\} \qquad (4.10)$$

where

$$C(m, w_1, w_3) = \sum_{i=1}^{m} (X_i - \overline{X}_1^T)(w_1 + w_3 X_i) - \sum_{i=1}^{T} (X_i - \overline{X}_1^T)(Y_i - \overline{Y}_1^T) ,$$

$$B(m, w_1, w_3) = \sum_{i=1}^{m} (w_1 + w_3 X_i) - T\, \overline{Y}_1^T ,$$

Finally we make the transformation to

$$\gamma_1 = -w_1 / w_3$$

$$\gamma_2 = w_3$$

and obtain as the joint density of γ_1 and γ_2

$$\pi_1(\gamma_1, \gamma_2) \propto |\gamma_2| \exp\left\{(-1/2)\left[P(m, \gamma_1)\gamma_2^2 - 2\,Q(m, \gamma_1)\gamma_2 + R(m, \gamma_1)\right]\right\} \tag{4.11}$$

where

$$P(m, \gamma_1) = \sum_{i=1}^{m} (X_i - \gamma_1)^2 - \left[\sum_{i=1}^{m} (X_i - \gamma_1)\right]^2 \Big/ T - \left[\sum_{i=1}^{m} (X_i - \gamma_1)(X_i - \overline{X}_1^T)\right]^2 \Big/ \sum_{i=1}^{T} (X_i - \overline{X}_1^T)^2 ,$$

$$Q(m, \gamma_1) = \sum_{i=1}^{m} (X_i - \gamma_1)(\hat{Y}_i^{m,L} - \hat{Y}_i) ,$$

$$R(m, \gamma_1) = \sum_{i=1}^{T} (Y_i - \hat{Y}_i)^2 .$$

$\hat{Y}_i$ is the predicted Y at X_i when the regression parameters are estimated by least squares over all data points.

To integrate out γ_2 we can complete the square in the exponent and write

$$\pi_1(\gamma_1, \gamma_2) \propto |\gamma_2| \exp\left\{(-1/2)\, P(m, \gamma_1)\left[\gamma_2 - Q(m, \gamma_1)/P(m, \gamma_1)\right]^2 + R(m, \gamma_1) - Q^2(m, \gamma_1)/P(m, \gamma_1)\right\} .$$

Provided that $P(m, \gamma_1)$ is positive it is now straightforward to use the normal density form to integrate with respect to γ_2. Of course $P(m, \gamma_1)$ is a function not only of m and γ_1 but also of the X_i's from the data. The author has been unable to show that $P(m, \gamma_1)$ is always positive, but is yet to encounter a situation where this condition is not satisfied.

We shall thus assume that $P(m, \gamma_1)$ is positive for all γ_1 and proceed to obtain

$$\pi_1(\gamma_1) \propto \exp\{(-1/2)[R(m, \gamma_1) - Q^2(m, \gamma_1)]\}\, P^{-1/2}(m, \gamma_1)\, E\{|W|\}$$

where

$$W \sim N\left(Q(m, \gamma_1)/P(m, \gamma_1),\ P^{-1}(m, \gamma_1)\right) .$$

Applying formula (4.4) we can write the posterior density of the intersection as

$$\pi_1(\gamma_1) \propto P^{-1/2}(m, \gamma_1)\left\{\left[Q(m, \gamma_1)/P(m, \gamma_1)\right]\left[2\,\phi\left(Q(m, \gamma_1)\, P^{-1/2}(m, \gamma_1)\right) - 1\right] \exp\left[(-1/2)\, Q^2(m, \gamma_1)/P(m, \gamma_1)\right] + \left[\pi\, P(m, \gamma_1)/2\right]^{-1/2}\right\},$$

$$-\infty < \gamma_1 < \infty . \qquad (4.12)$$

σ^2 Known, Neither Regression Known, m Unknown

The derivation of the posterior distribution of the intersection for this case proceeds exactly as in case three, the additional step being summation on the range of the shift index m. One obtains

$$\pi_1(\gamma_1) \propto \sum_{m=2}^{T-2} P^{-1/2}(m, \gamma_1)$$
$$\left\{ \left[Q(m, \gamma_1)/P(m, \gamma_1) \right] \left[2\,\phi\!\left(Q(m, \gamma_1)\, P^{-1/2}(m, \gamma_1) \right) - 1 \right] \right.$$
$$\left. \exp\left[(1/2)\, Q^2(m, \gamma_1)/P(m, \gamma_1) \right] + \left[\pi\, P(m, \gamma_1)/2 \right]^{-1/2} \right\},$$
$$-\infty < \gamma_1 < \infty \,. \tag{4.13}$$

We remind the reader that for this case the condition discussed in case three must be satisfied for each m over the range of the summation.

σ^2 Unknown, First Regression Known, m Known

The likelihood function for the present case is

$$L(\sigma^2, \alpha_2, \beta_2) \propto (2\pi\sigma^2)^{-T/2} \exp\left\{ (-1/2\,\sigma^2) \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 \right. \right.$$
$$\left. \left. + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2 \right] \right\}.$$

We assign to σ^2 the usual improper prior density

$$\pi_0(\sigma^2) \propto 1/\sigma^2, \quad 0 < \sigma^2 < \infty \,,$$

and retain improper uniform prior densities for the regression parameters. The joint posterior density is then

$$\pi_1(\sigma^2, \alpha_2, \beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\} \quad (4.14)$$

for $0 < \sigma^2 < \infty$, $-\infty < \alpha_2 < \infty$, and $-\infty < \beta_2 < \infty$. Integrating first with respect to σ^2 we obtain

$$\pi_1(\alpha_2, \beta_2) \propto \left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]^{-T/2} . \quad (4.15)$$

We now derive from (4.15) the joint density of w_1 and w_2, where

$$w_1 = \alpha_2 - \alpha_1$$

$$w_2 = \beta_2 - \beta_1$$

and obtain

$$\pi_1(w_1, w_2) \propto \left[E(m) + (\underset{\sim}{w} - \underset{\sim}{v}^m)' \Sigma_m^{-1} (\underset{\sim}{w} - \underset{\sim}{v}^m)\right]^{-T/2}$$

where

$$\underset{\sim}{w} = (w_1, w_2)', \qquad \underset{\sim}{v}^m = (\hat{\alpha}_2^m - \alpha_1, \hat{\beta}_2^m - \beta_1)',$$

$$E(m) = \sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2 ,$$

and

$$\Sigma_m^{-1} = \begin{bmatrix} m & \sum_1^m X_i \\ \sum_1^m X_i & \sum_1^m X_i^2 \end{bmatrix} .$$

The final transformation is made to

$$\gamma_1 = -w_1/w_2$$

$$\gamma_2 = w_2 \quad .$$

This leads us to the joint density

$$\pi_1(\gamma_1, \gamma_2) \propto |\gamma_2| \left[G(m, \gamma_1) + A(m, \gamma_1)(\gamma_2 - B(m, \gamma_1))^2 \right]^{-T/2} \tag{4.17}$$

where

$$A(m, \gamma_1) = \sum_{i=m+1}^{T} (X_i - \gamma_1)^2$$

$$B(m, \gamma_1) = \left[\sum_{i=m+1}^{T} (X_i - \gamma_1)(\hat{Y}_i^{m,U} - \alpha_1 - \beta_1 X_i) \right]^2 \Bigg/ \sum_{i=m+1}^{T} (X_i - \gamma_1)^2$$

and

$$G(m, \gamma_1) = E(m) + \sum_{i=m+1}^{T} [\hat{Y}_i^{m,U} - \alpha_1 - \beta_1 X_i]^2 - A(m, \gamma_1) B^2(m, \gamma_1) .$$

We are again faced with a problem similar to that discussed in the third case considered in this chapter. Assuming that $G(m, \gamma_1)$ is positive for each γ_1 we can write (4.17) as

$$\pi_1(\gamma_1, \gamma_2) \propto G^{-T/2}(m, \gamma_1)\,|\gamma_2|$$
$$\cdot \left[1 + \frac{(T-1)\,A(m,\gamma_1)}{(T-1)\,G(m,\gamma_1)}\,(\gamma_2 - B(m,\gamma_1))^2\right]^{-[(T-1)+1]/2}$$

and make use of the general t density (22) to integrate with respect to γ_2. Although this has not been proved in general it has been checked for a number of data sets and was not violated in those cases checked. We shall therefore assume that $G(m, \gamma_1)$ is positive for each γ_1 and obtain for the marginal density of γ_1

$$\pi_1(\gamma_1) \propto G^{-T/2}(m,\gamma_1)\,[A(m,\gamma_1)/G(m,\gamma_1)]^{-1/2}\,E\{|W|\} \quad (4.18)$$

where

$$[(T-1)\,A(m,\gamma_1)/G(m,\gamma_1)]^{1/2}\,[W - B(m,\gamma_1)] \sim \text{Student's } t\ (T-1)\,.$$

Tedious but straightforward computation shows that if $t^{1/2}(X-u)$ is distributed like Student's t with n degrees of freedom, then

$$E[|X|] = 2\,t^{-1/2}\left(n^{1/2}/(n-1)\right)B^{-1}\left(n/2,\ 1/2\right)\left(1+\mu^2\,t/n\right)^{-(n-1)/2}$$
$$+\ \mu\left(2\,\psi_n(\mu\,t^{1/2}) - 1\right), \quad (4.19)$$

where $B(x,y)$ is the beta function and $\psi_n(x)$ is the cumulative distribution function of a Student's t random variable with n degrees of freedom. Applying this result to (4.18) we may write the density of the intersection as

$$\pi_1(\gamma_1) \propto G^{-(T-1)/2}(m,\gamma_1)\, A^{-1/2}(m,\gamma_1)$$
$$\Bigg\{ 2\left[A(m,\gamma_1)/G(m,\gamma_1)\right]^{-1/2}\left[(T-2)\,B\Big((T-1)/2,\, 1/2\Big)\right]^{-1}$$
$$\left[1 + A(m,\gamma_1)\,B^2(m,\gamma_1)/G(m,\gamma_1)\right]^{-(T-2)/2} \tag{4.20}$$
$$+ B(m,\gamma_1)\left[2\,\psi_{T-1}\Big(B(m,\gamma_1)\,((T-1)\,A(m,\gamma_1)/G(m,\gamma_1))^{1/2}\Big) - 1\right]\Bigg\}$$

for $-\infty < \gamma_1 < \infty$.

σ^2 Unknown, First Regression Known, m Unknown

The joint posterior density of m, σ^2, α_2 and β_2 is now

$$\pi_1(m,\sigma^2,\alpha_2,\beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\Bigg\{(-1/2\,\sigma^2)\Bigg[\sum_{i=1}^{m}\left[Y_i - (\alpha_1 + \beta_1 X_i)\right]^2$$
$$+ \sum_{i=m+1}^{T}\left[Y_i - (\alpha_2 + \beta_2 X_i)\right]^2\Bigg]\Bigg\}. \tag{4.21}$$

for $m = 2, \ldots, T-2$, $0 < \sigma^2 < \infty$, $-\infty < \alpha_2 < \infty$, and $-\infty < \beta_2 < \infty$.

Derivation of the posterior density of the intersection proceeds exactly as in the previous case with the additional step of summation over the range of the shift index m. We thus obtain a mixture of densities of the type shown in (4.20), namely

$$\pi_1(\gamma_1) \propto \sum_{m=2}^{T-2} G^{-(T-1)/2}(m,\gamma_1)\, A^{-1/2}(m,\gamma_1)$$
$$\Bigg\{ 2\left[A(m,\gamma_1)/G(m,\gamma_1)\right]^{-1/2}\left[(T-2)\,B\Big((T-1)/2,\, 1/2\Big)\right]^{-1}$$

$$\left[1 + A(m, \gamma_1) B^2(m, \gamma_1)/G(m, \gamma_1)\right]^{-(T-2)/2}$$

$$+ B(m, \gamma_1)\left[2\,\psi_{T-1}\left(B(m, \gamma_1)((T-1) A(m, \gamma_1)/G(m, \gamma_1))^{1/2}\right) - 1\right]\Bigg\} \tag{4.22}$$

for $-\infty < \gamma_1 < \infty$.

σ^2 Unknown, Neither Regression Known, m Known

The joint posterior density of the variance and the regression parameters is

$$\pi_1(\sigma^2, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}$$

for $0 < \sigma^2 < \infty$. $-\infty < \alpha_i < \infty$ and $-\infty < \beta_i < \infty$, $i = 1, 2$.

Integrating first on σ^2 we obtain

$$\pi_1(\alpha_1, \beta_1, \alpha_2, \beta_2) \propto \left[\sum_{i=1}^{m} [Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T} [Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]^{-T/2}$$

$$\propto \left[\sum_{i=1}^{m} (Y_i - \hat{Y}_i^{m,L})^2 + \sum_{i=m+1}^{T} (Y_i - \hat{Y}_i^{m,U})^2 + (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)' \Sigma_m^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{\mu}^m)\right]^{-T/2} . \tag{4.23}$$

The notation corresponds to that used in case three, and is explained immediately below equation (4.9). We next make the transformation to

$$w_1 = \alpha_1 - \alpha_2$$

$$w_2 = \alpha_2$$

$$w_3 = \beta_1 - \beta_2$$

$$w_4 = \beta_2$$

as in case three, integrate out w_2 and w_4 using the general t density shown in De Groot (22), and then obtain the joint distribution of

$$\gamma_1 = -w_1/w_3$$

and

$$\gamma_2 = w_3 \quad .$$

We obtain

$$\pi_1(\gamma_1, \gamma_2) \propto |\gamma_2| \left[P(m, \gamma_1)\gamma_2^2 - 2Q(m, \gamma_1)\gamma_2 + R(m, \gamma_1) \right]^{-(T-2)/2} \tag{4.24}$$

where

$$P(m, \gamma_1) = \sum_{i=1}^{m} (X_i - \gamma_1)^2 - \left[\sum_{i=1}^{m} (X_i - \gamma_1) \right]^2 \Big/ T$$
$$- \left[\sum_{i=1}^{m} (X_i - \gamma_1)(X_i - \overline{X}_1^T) \right]^2 \Big/ \sum_{i=1}^{T} (X_i - \overline{X}_1^T)^2$$

$$Q(m, \gamma_1) = \sum_{i=1}^{m} (X_i - \gamma_1)(\hat{Y}_i^{m,L} - \hat{Y}_i)$$

$$R(m, \gamma_1) = \sum_{i=1}^{T} (Y_i - \hat{Y}_i)^2 \quad .$$

Equation (4.24) can be written

$$\pi_1(\gamma_1, \gamma_2) \propto |\gamma_2| \left[G(m, \gamma_1) + P(m, \gamma_1)\left(\gamma_2 - G(m, \gamma_1/P(m, \gamma_1)\right)^2 \right]^{-(T-2)/2}$$

where

$$G(m, \gamma_1) = R(m, \gamma_1) - G^2(m, \gamma_1)/P(m, \gamma_1) .$$

We are again faced with the problem of the sign of $G(m,\gamma_1)$. If $G(m, \gamma_1)$ is positive we can write

$$\pi_1(\gamma_1, \gamma_2) \propto G^{-(T-2)/2}(m, \gamma_1) |\gamma_2| \left[1 + \left(P(m, \gamma_1)/G(m, \gamma_1)\right)\left(\gamma_2 - Q(m, \gamma_1)/P(m, \gamma_1)\right)^2\right]^{-(T-2)/2} \tag{4.25}$$

and integration with respect to γ_2 proceeds easily using the general t density referred to earlier. Although we have been unable to prove that $G(m, \gamma_1)$ is a positive function of m and γ_1 we shall assume that this condition is satisfied and proceed. One then obtains for the posterior density of γ_1

$$\begin{aligned}\pi_1(\gamma_1) \propto{} & P^{-1/2}(m, \gamma_1)\, G^{-(T-3)/2}(m, \gamma_1) \\ & \Big\{2\left[P(m, \gamma_1)/G(m, \gamma_1)\right]^{-1/2}\left[(T-4)\, B\left((T-3)/2,\, 1/2\right)\right]^{-1} \\ & \left[1 + Q^2(m, \gamma_1)/\left(P(m, \gamma_1)\, G(m, \gamma_1)\right)\right]^{-(T-4)/2} \\ & + \left[Q(m, \gamma_1)/P(m, \gamma_1)\right]\left[2\psi_{T-3}\left((T-3)^{1/2} Q(m,\gamma_1)/(P(m,\gamma_1)G(m,\gamma_1))^{1/2}\right)-1\right]\Big\}\end{aligned} \tag{4.26}$$

for $-\infty < \gamma_1 < \infty$.

σ^2 Unknown, Neither Regression Known, m Unknown

The joint posterior density of the unknown parameters in this the final case is

$$\pi_1(m, \sigma^2, \alpha_1, \beta_1, \alpha_2, \beta_2) \propto (\sigma^2)^{-(T/2+1)} \exp\left\{(-1/2\sigma^2)\left[\sum_{i=1}^{m}[Y_i - (\alpha_1 + \beta_1 X_i)]^2 + \sum_{i=m+1}^{T}[Y_i - (\alpha_2 + \beta_2 X_i)]^2\right]\right\}. \tag{4.27}$$

The posterior density of the intersection is a mixture of densities of the same type as (4.26). Its derivation proceeds as in the previous case with the additional step being summation on the shift index m. The density is

$$\pi_1(\gamma_1) \propto \sum_{m=2}^{T-2} P^{-1/2}(m, \gamma_1)\, G^{-(T-3)/2}(m, \gamma_1) \left\{2\left[P(m,\gamma_1)/G(m,\gamma_1)\right]^{-1/2}\left[(T-4)\,B\left((T-3)/2, 1/2\right)\right]^{-1} \left[1 + Q^2(m,\gamma_1)/\left(P(m,\gamma_1)\,G(m,\gamma_1)\right)\right]^{-(T-4)/2} + \left[Q(m,\gamma_1)/P(m,\gamma_1)\right] \left[2\,\psi_{T-3}\left((T-3)^{1/2} Q(m,\gamma_1)/(P(m,\gamma_1)\,G(m,\gamma_1))^{1/2}\right) - 1\right]\right\} \tag{4.28}$$

for $-\infty < \gamma_1 < \infty$.

An illustration of densities (4.22) and (4.28) will be presented in the following chapter.

CHAPTER V

INFERENCE PROCEDURES AND SOME EXAMPLES

In Chapters II, III and IV we determined the posterior distributions, under a variety of assumptions, for an unknown parameter, say θ. The parameter θ was a scalar in most cases, although sometimes θ was a vector as we saw in Chapter II with the posterior distribution of (ϕ_0, ϕ_1), equation (2.25).

We shall in this chapter describe some ways in which posterior distributions are used to make inference, and illustrate some of these techniques on posterior distributions selected from Chapters II, III and IV.

All of the computations necessary to present the examples in this chapter were done with programs written by the author and run on the IBM - 360, Mod 65 computer at the Oklahoma State University Computer Center.

Plot Of The Posterior Density

In cases where the dimension of the parameter vector is at most two, a plot of the posterior density (in the univariate case) or contours of constant posterior density (in the bivariate case) can be useful.

Example 5.1

Hinkley (13) used some data from an article by Pool and Borchgrevink (23) to illustrate the techniques which he suggested. In order to compare our techniques with his, we shall use the same set of data, which is given in Table I below. The independent variable X represents the logarithm of warfarin concentration and the dependent variable Y is blood factor VII production. $\pi_k(m)$ and $\pi_u(m)$ are explained after the table.

TABLE I

DATA AND RESULTING POSTERIOR DENSITIES

m	X_m	Y_m	$\pi_k(m)$	$\pi_u(m)$
1	2.00000	0.370483	—	—
2	2.52288	0.537970	0.000057	0.002121
3	3.00000	0.607684	0.005115	0.011831
4	3.52288	0.723323	0.031579	0.034446
5	4.00000	0.761856	0.297597	0.281910
6	4.52288	0.892063	0.276329	0.266278
7	5.00000	0.956707	0.351680	0.365501
8	5.52288	0.940349	0.037518	0.055949
9	6.00000	0.898609	0.000117	0.001142
10	6.52288	0.953850	0.000006	0.000376
11	7.00000	0.990834	0.000002	0.000322
12	7.52288	0.890291	0.000000	0.000049
13	8.00000	0.990779	0.000000	0.000075
14	8.52288	1.050865	—	—
15	9.00000	0.982785	—	—

A plot of the data given in Table I suggests that two-phase regression may be an appropriate analysis.

We have computed the posterior density of the shift index m for two cases, namely variance known (equation (3.12)) and variance unknown (equation (3.17)). Both cases assume that neither regression is known. Our results for equation (3.12) are denoted in Table I above by $\pi_k(m)$, while those for equation (3.17) are denoted by $\pi_u(m)$. In computing $\pi_k(m)$ we have taken as the "known" variance the estimate obtained by Hinkley (13), namely $\hat{\sigma}^2 = .00166$.

The posterior density $\pi_u(m)$ is plotted in Figure 1 below.

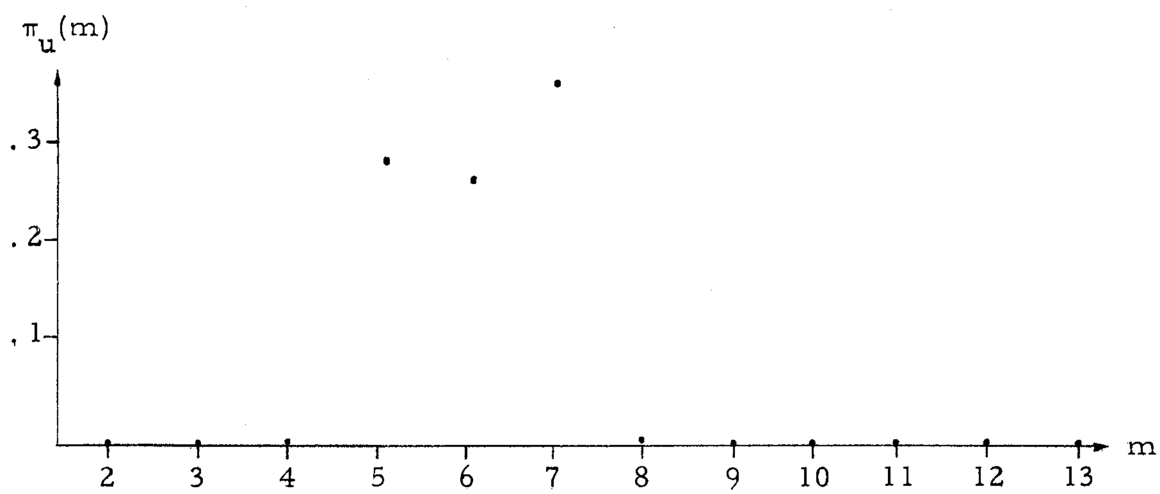

Figure 1. Plot of Posterior Density $\pi_u(m)$

The density $\pi_k(m)$ is not plotted since it very nearly coincides with $\pi_u(m)$.

There is strong evidence, using either posterior density, that the switch in regimes occurs at $m = 5$, $m = 6$, or $m = 7$. We remark in passing that Hinkley's 99% confidence interval on the intersection of the regression lines is (3.641, 5.441), which contains X_5, X_6, and X_7.

Point Estimates For Parameters

As a point estimate of a parameter θ one may use the mean, median or mode of the posterior distribution. Each estimate can be "justified" in at least one way, depending in some cases on the particular loss function assigned. These estimates are discussed in many modern inference texts and the reader is referred to Ferguson (24) for a survey of this subject. We shall present here the numerical values of the three point estimates for our data of Example 5.1:

	$\pi_k(m)$	$\pi_u(m)$
Mode of Posterior Distribution :	7.00000	7.00000
Median of Posterior Distribution :	6.00000	6.00000
Mean of Posterior Distribution :	6.05077	6.04998

The variance of the posterior distribution is sometimes given along with measures of central tendency to help describe the distribution. For our example,

$$\text{Var}_k(m) = 0.97123$$

and

$$\text{Var}_u(m) = 1.09669 .$$

Regions Of Highest Posterior Density

Let $\pi(\theta; y)$ denote the posterior density function for a parameter θ. A region R in the parameter space of θ is called a Highest Posterior Density (H. P. D.) region of content $(1 - \alpha)$ if

(i) $Pr(\theta \in R; y) = 1 - \alpha$, and

(ii) For $\theta_1 \in R$ and $\theta_2 \notin R$, $\pi(\theta_1; y) \geq \pi(\theta_2; y)$.

It is immediately clear that if θ is a discrete random variable then an H. P. D. region will not exist for all values of α. In fact, if θ has, say, n distinct mass points of positive density, then H. P. D. regions exist for at most n distinct values of α. For a more complete discussion of H. P. D. regions the reader is referred to a paper by Box and Tiao (25).

For the data of Example 5.1, the set

$$R = \{5, 6, 7\}$$

is an H. P. D. region for m, based on $\pi_k(m)$, of content 0.925606, while its content based on $\pi_u(m)$ is 0.913689.

H. P. D. regions for symmetrical, unimodal distributions are easily obtained by numerical integration. A further illustration will be given later in this chapter in Example 5.2.

Hypothesis Testing

Suppose we have observed data y from an experiment and wish to judge whether or not the data support some specified hypothesis, say $H: \theta \in S_0$, about a parameter θ on which the assumed distribution of the observation vector depends, Of course we must have S_0 a subset of S, the entire parameter space of θ. We shall discuss briefly three possible approaches to this problem.

First Approach

If one wishes to accept or reject H at some predetermined significance level, say α, then

(i) Construct a $(1-\alpha)\cdot 100\%$ H.P.D. region R for θ.

(ii) Reject H if and only if $S_0 \cap R$ is empty.

This implies that the posterior probability of H is at most α when H is rejected.

Second Approach

Compute the posterior odds ratio

$$r_1 = \int_{S_0} \pi_1(\theta; y)\, d\theta \bigg/ \int_{S-S_0} \pi_1(\theta; y)\, d\theta$$

or alternatively just the posterior probability of S_0, namely

$$r_1^* = \int_{S_0} \pi_1(\theta; y)\, d\theta \ .$$

r_1^* is normed in the sense that $0 \leq r_1^* \leq 1$. Clearly large values of either r_1 or r_1^* lend credence to H. This procedure seems to be innappropriate if $\pi_1(\theta; y)$ is continuous and S_0 is a single point of S. A possible alternative in such a situation is given in the third approach.

<u>Third Approach</u>

This is based on the ratio of ordinates of the posterior density. The ratio is defined by

$$r_2 = \sup_{\theta \epsilon S_0} \pi_1(\theta; y) \Big/ \sup_{\theta \epsilon S} \pi_1(\theta; y) .$$

Here again $0 \leq r_2 \leq 1$ and values of r_2 near one lend credence to H. This approach appears to overcome the difficulty posed by r_1 and r_1^* when H is a simple hypothesis of the form $H : \theta = \theta_0$.

For the data of our Example 5.1, we compute r_1, r_1^*, and r_2, based on $\pi_u(m)$, for a test of the hypothesis $H : m \epsilon S_0$, where $S_0 = \{5, 6, 7\}$. Here the parameter space is $S = \{2, \ldots, 13\}$. We have

$$r_1 = 10.586$$

$$r_1^* = 0.914$$

$$r_2 = 1.000$$

Other Techniques

There is a variety of other techniques available for applying the posterior density function to particular statistical problems.

Where θ is a vector, marginal posterior distributions of certain parameters may be of interest. In fact, the net result in most sections of this report is a marginal posterior distribution of some kind. Another tool used in certain applications is the predictive density, used to make inferences about data which is to be observed at some future time. Corresponding to H.P.D. regions one may define "prediction regions", or regions of "highest predictive density".

For a more complete survey of uses and applications of the posterior probability density function, the reader is referred to texts by Zellner (26), Lindley (19), LaValle (27) and De Groot (22).

An Ad Hoc Technique

All of the distributions derived in Chapter IV are not based on any prior constraints on the point of intersection of the two population regression lines.

In many cases it may be known that the intersection, namely $\gamma = (\alpha_1 - \alpha_2)/(\beta_2 - \beta_1)$, is constrained such that $X_m \leq \gamma \leq X_{m+1}$, where m is the shift index. If this is too restrictive, then $X_1 \leq \gamma \leq X_T$ would not seem unreasonable for most practical purposes. That is, we require that the regression lines intersect at some point over the observed range of the independent variable X.

When this latter assumption can be made, this author suggests truncating the appropriate distribution over the interval $[X_1, X_T]$ rather than working with it over the entire real line.

We shall now illustrate this technique on some of the distributions derived in Chapter IV, using the data of Example 5.1. Hinkley (13) gave confidence intervals for the cases "second regression known" (with slope equal to zero) and "neither regression known". We shall follow his lead in Examples 5.2 and 5.3 respectively.

Example 5.2 (Second Regression Known)

We assume here that $\beta_2 = 0$ and that $\alpha_2 = 0.961674$. This value of α_2 is the average of the last nine Y_i's in the sequence, since the median of our density $\pi_u(m)$ was six and the mean was near six. It also agrees to four decimal places with the estimate from Hinkley's paper. In Figure 2 below we show a plot of the posterior density (4.22) for the case "second regression known". The density has been truncated on the interval $[X_1, X_T]$. Since the ordinates of the density are relatively small outside of the interval $[3, 6]$, the density is shown only over that interval.

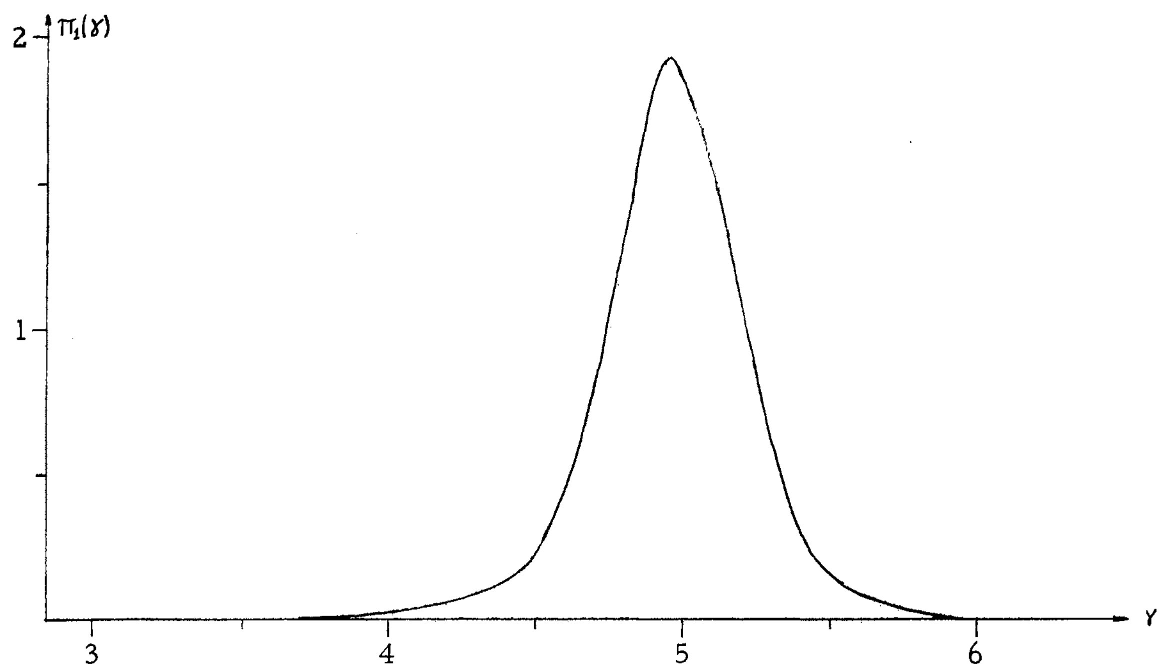

Figure 2. Posterior Density (4.22) : Second Regression Known

From this density we calculated the mode as well as H. P. D. regions of content 0.90, 0.95 and 0.99. For comparative purposes we now present the results from our posterior density as well as those from Hinkley's paper:

	M. L. E./Mode	90% Interval	95% Interval	99% Interval
Hinkley	4.88	(4.55, 5.29)	(4.45, 5.39)	(4.25, 5.66)
Density (4.22)	4.89	(4.54, 5.31)	(4.46, 5.43)	(4.28, 5.79)

Example 5.3 (Neither Regression Known)

For the same data set as we have used in the previous examples we plot in Figure 3 below the posterior density (4.28) truncated on $[X_1, X_T]$. Again it is plotted only over the interval $[3, 6]$.

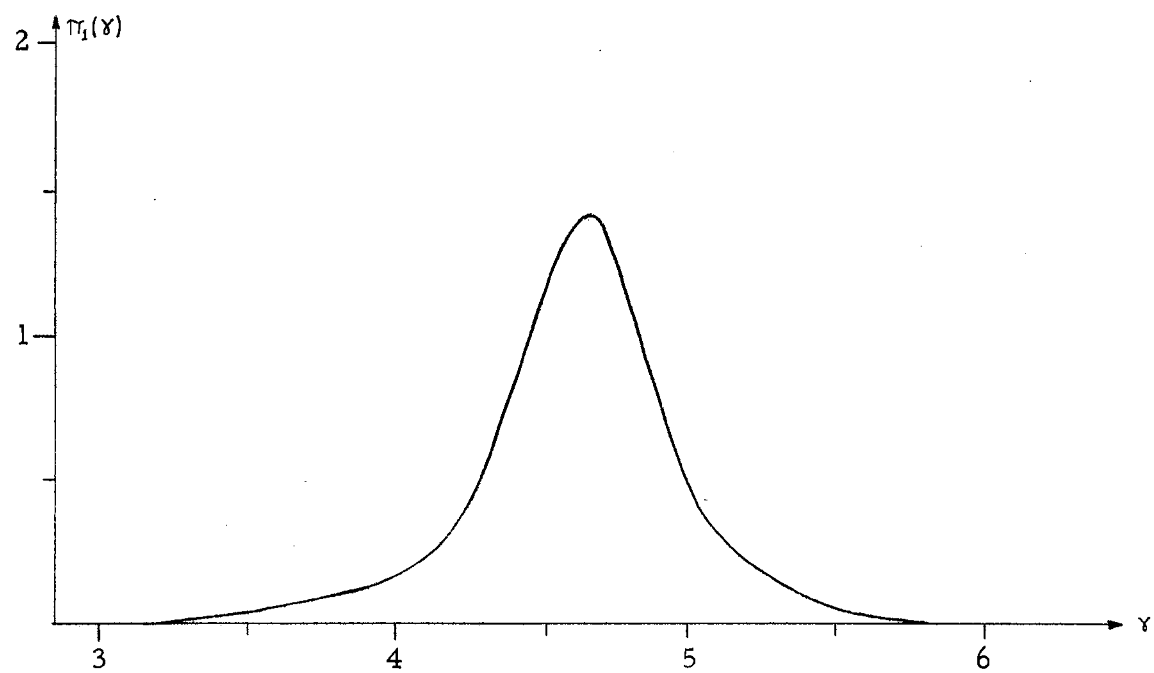

Figure 3. Posterior Density (4.28): Neither Regression Known

A comparison of our H. P. D. regions from density (4.28) with the confidence intervals given by Hinkley is again presented:

	M. L. E./Mode	90% Interval	95% Interval	99% Interval
Hinkley	4.65	(4.16, 5.10)	(4.09, 5.20)	(3.64, 5.44)
Density (4.28)	4.60	(4.04, 5.17)	(3.85, 5.33)	(3.32, 5.78)

Some General Comments

The author feels that the H. P. D. regions obtained from densities (4.22) and (4.28) compare quite favorably with the confidence intervals given by Hinkley. In making such a comparison, one must keep in mind that Hinkley's confidence intervals are based on the restriction that $X_m \leq \gamma < X_{m+1}$, whereas the H. P. D. regions derived from our posterior distribution are subject to no such restrictions.

A further consideration is that the confidence intervals derived by Hinkley are based on the asymptotic chi-square distributions of the likelihood ratio statistics, with the result that the approximation for small to moderate sample sizes is somewhat questionable. On the other hand, the posterior distributions given in this report are exact, allowing computation of H. P. D. regions to any desired degree of accuracy. We must caution, though, that we have not established whether or not the posterior distributions derived in Chapter IV are unimodal. A bimodal distribution may lead to an H. P. D. region which is not a single interval. Hinkley (13) also points out that the confidence interval algorithm suggested by him does not necessarily lead to a single interval.

With regard to computational considerations, the procedures given here present little problem. The discrete distributions given in Chapter III for the shift point obviously present no difficulty, while computation of the distributions of the intersection given in Chapter IV require only referral to the standard normal distribution function (variance known) and Student's t distribution function (variance unknown).

We close this chapter with a comment on the incompleteness of the present study. A complete analysis of a two-phase regression situation would of course include estimates not only of the shift point and intersection, but also of the regression parameters and the error variance. We have directed our attention almost exclusively to the switch point and the intersection.

CHAPTER VI

SUMMARY AND POSSIBLE EXTENSIONS

Summary Of The Study

This research was undertaken with the intention of making some contribution to the general problem of estimating the time point at which a parameter change occurs in an observed sequence of random variables. No attempt has been made to study the related problem of detecting whether or not a parameter shift has occurred in the sequence.

A Bayesian approach was employed for each of the cases studied, and vague type prior densities were assigned resulting in posterior distributions appropriate to situations where prior knowledge is imprecise. We remark here that even if prior knowledge does not fit this description, it may still be informative to look at an analysis under the assignment of diffuse prior densities.

In Chapter II we directed attention to the special case of a normal sequence with unknown variance, and derived posterior densities corresponding to a variety of assumptions on the parameters of the problem.

Chapters III and IV focused on the more general setting known as two-phase regression. In particular, in Chapter III we derived posterior densities for the shift index itself, while in Chapter IV we studied the problem of estimating the abscissa of the point of intersection of the two regression lines.

In Chapter V we surveyed some of the possible uses to which the posterior distributions may be put, and gave examples based on a set of data used by Hinkley (13) to illustrate his solution to the same general problem.

Some Possible Extensions

Little attempt has been made in this study to investigate the properties of the posterior distributions derived in Chapters II, III and IV. One of the reasons for this is the apparent algebraic complexity of the form of the distributions. This point is well illustrated by equation (4.28), for example. A considerable amount of numerical work remains to be done with regard to means, modes and variances of the distributions.

Again, we have not directed any attention to the problem of estimating the error variance in the two-phase regression setting. Its posterior distribution could, hopefully, be computed, and the estimators resulting from it could be compared to those given by Hinkley (13). The same is true for the regression parameters of the two regressions under study.

Another obvious and natural extension would be to the multivariate case, where at each time point one obtains a vector of observations rather than a single observation. In this case, of course, the ϕ_i's of Chapter II would be replaced by a vector of means, while σ^2 would be replaced by a variance-covariance or dispersion matrix.

Finally we point out that the study by Hinkley (13) assumes that γ, the abscissa of the intersection point of the two regression lines, is constrained by $X_m \leq \gamma < X_{m+1}$, where m is the (usually unknown) switch point of the sequence. The author has investigated this situation in the Bayesian framework for the first case, namely where m, σ^2 and the first regression are all known. The resulting posterior distribution for γ is simply our distribuion (4. 5) truncated on the interval $[X_m, X_{m+1})$. One could proceed to study some of the more complex cases under this added restriction to see how the resulting posterior distributions compare to those arrived at in this dissertation.

A SELECTED BIBLIOGRAPHY

(1) Page, E. S. "Continuous Inspection Schemes." Biometrika, Vol. 41, 1954, 100-115.

(2) Page, E. S. "A Test For A Change In A Parameter Occurring At An Unknown Time Point." Biometrika, Vol. 42, 1955, 523-527.

(3) Page, E. S. "On Problems In Which A Change In A Parameter Occurs At An Unknown Time Point." Biometrika, Vol. 44, 1957, 248-252.

(4) 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.

(5) Quandt, R. E. "Tests Of The Hypothesis That A Linear Regression Obeys Two Separate Regimes." Journal of the American Statistical Association, Vol. 55, 1960, 324-330.

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

(7) Chernoff, H. and Zacks, S. "Estimating The Current Mean Of A Normal Distribution Which Is Subjected To Changes In Time." Annals of Mathematical Statistics, Vol. 35, 1966, 999-1018.

(8) Kander, Z. and Zacks, S. "Test Procedures For Possible Changes In Parameters Of Statistical Distributions Occuring At Unknown Time Points." Annals of Mathematical Statistics, Vol. 37, 1966, 1196-1210.

(9) Bhattacharyya, G. K. and Johnson, R. A. "Non Parametric Tests For Shift At An Unknown Time Point." Annals of Mathematical Statistics, Vol. 39, 1968, 1731-1743.

(10) Brown, R. L. and Durbin, J. "Methods Of Investigating Whether A Regression Relationship Is Constant Over Time." Selected Statistical Papers I, Amsterdam: Mathematisch Centrum, European Meeting, 1968, 37-45.

(11) Tukey, J. "The Future Of Data Analysis." Annals of Mathematical Statistics, Vol. 33, 1962, 1-67.

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

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

(14) Hinkley, D. V. "Inference About The Change Point In A Sequence Of Random Variables." Biometrika, Vol. 57, 1970, 1-17.

(15) Hinkley, D. V. and Hinkley, E. A. "Inference About The Change Point In A Sequence Of Binomial Random Variables." Biometrika, Vol. 57, 477-488.

(16) Broemeling, L. D. "Some Bayesian Inferences About A Changing Sequence Of Random Variables." Submitted to Utilitas Mathematica, 1972.

(17) Broemeling, L. D. "Bayesian Procedures For Detecting A Change In A Sequence Of Random Variables." Submitted to Metron, 1972.

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

(19) Lindley, D. V. Introduction To Probability And Statistics From A Bayesian Viewpoint, Cambridge University Press, Cambridge, 1965.

(20) Jeffreys, H. Theory Of Probability, Clarendon Press, Oxford, 1961.

(21) Royden, H. L. Real Analysis, The Macmillan Company, New York, 1968.

(22) De Groot, M. H. Optimal Statistical Decisions, McGraw-Hill Book Company, New York, 1970.

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

(24) Ferguson, Thomas S. Mathematical Statistics - A Decision Theoretic Approach, Academic Press, New York, 1967.

(25) Box, G. E. P. and Tiao, G. C. "Multi-parameter Problems From A Bayesian Point Of View." Annals of Mathematical Statistics, Vol. 36, 1965, 1468-1482.

(26) Zellner, A. An Introduction To Bayesian Inference In Econometrics, John Wiley and Sons, New York, 1971.

(27) LaValle, I. H. An Introduction To Probability, Decision And Inference, Hold, Rinehart, and Winston, New York, 1970.

VITA

Donald Holbert

Candidate for the Degree of

Doctor of Philosophy

Thesis: A BAYESIAN ANALYSIS OF SHIFTING SEQUENCES WITH APPLICATIONS TO TWO-PHASE REGRESSION

Major Field: Statistics

Biographical:

Personal Data: Born in Tea Gardens, New South Wales, Australia, August 27, 1941, the son of Mr. and Mrs. Frederick George Holbert.

Education: Received secondary school education at Maitland High School, Maitland, N. S. W., Australia; graduated from Newcastle Teachers' College, Newcastle, N. S. W., Australia, in 1960; received Bachelor of Science degree in mathematics from the University of Oregon, Eugene, Oregon, in May, 1967; received Master of Arts degree in mathematics from Washington State University, Pullman, Washington, in May, 1969; completed requirements for the Degree of Doctor of Philosophy at Oklahoma State University in May, 1973.

Professional Experience: Mathematics Instructor at Cronulla High School, Cronulla, N. S. W., Australia, 1961-1962; Mathematics Instructor at Farrer Memorial Agricultural High School, Tamworth, N. S. W., Australia, 1963-1965; graduate teaching assistant at Washington State University, Pullman, Washington, 1967-1969; graduate research assistant at Oklahoma State University, Stillwater, Oklahoma, 1969-1973; member of the American Statistical Association and the Biometric Society.