AN EMPIRICAL BAYES APPROACH TO A VARIABLES SAMPLING PLAN PROBLEM

by

James A. Craig, Jr.

Technical Report No. 98
Department of Statistics ONR Contract

February 28, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A 0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

AN EMPIRICAL BAYES APPROACH TO A VARIABLES SAMPLING PLAN PROBLEM

A Thesis Presented to the Faculty of the Graduate School

of

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Statistics

by

James A. Craig, Jr.
(B.A., The University of Texas, 1963)
(M.S., Southern Methodist University, 1968)

February 28, 1971

AN EMPIRICAL BAYES APPROACH TO A VARIABLES SAMPLING PLAN PROBLEM

A Thesis Presented to the Faculty of the Graduate School

of

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Statistics

by

James A. Craig, Jr.
(B.A., The University of Texas, 1963)
(M.S., Southern Methodist University, 1968)

February 28, 1971

ACKNOWLEDGMENTS

I wish to thank my major adviser, Dr. Richard P. Bland, for his advice and helpful comments given during the writing of this paper. In addition, I would like to thank Dr. Bill Schucany, Dr. William Tucker, and Dr. Donald B. Owen for their constructive criticism of certain aspects of this paper. In particular I would like to thank Dr. Owen for allowing me this opportunity to study by giving me a Themis research fellowship.

This paper was supported by the Office of Naval Research, Contract N00014-68-A-0515, project NR 042-260.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

DESCRIPTION OF THE PROBLEM

1.1 Introduction

Consider the following variables sampling plan problem. We are given a lot of size m + n items. Each item's quality is characterized by some continuous random variable X. If this random variable, X, for a given item is within some specification limits, say (a, b) , the item is considered acceptable. On the basis of a random sample of size n, we want to accept or reject the remaining m items. If we reject the lot, we incur a loss c_0 for each of the remaining m items that meets the specification (i.e., $a \leq X \leq b$). We incur no loss in this case when an item does not meet the specification. If we accept the lot, we incur a loss c_1 for each of the remaining m items that does not meet the specification (i.e., $X < a$, or $X > b$). We incur no loss for those items meeting the specification. The random variable, X, is known to be normally distributed with some unknown mean μ and unknown variance σ^2. The random variable, X, for any item in the lot is independent of the random variable, X, for any other item in the lot.

Furthermore, we know that μ and σ^2 are random variables with some unknown prior distribution $G(\mu,\sigma^2)$. That is, the manufacturing process for this item is a random process that chooses some μ and σ^2 at random according to a distribution $G(\mu,\sigma^2)$ and then produces m + n items according to the normal distribution with mean μ and variance σ^2.

If we knew the prior distribution $G(\mu,\sigma^2)$, we could determine a Bayesian decision rule based on the sample mean $\bar{x}$ and sample variance s^2 that would minimize the average expected loss. We show in this research that in certain cases, if one has data from past lots of size $m + n$, it is possible to estimate the Bayesian decision rule empirically. That is, we have an "empirical Bayes" decision rule, and, consequently, we have an "empirical Bayes" approach to a variables sampling plan problem. We show that as the number of past lots increase (i.e., data from past realizations of the process), the empirical Bayes decision rule converges to the Bayesian decision rule. Finally, we give an example of the theoretical results by performing a Monte Carlo simulation using a conjugate prior distribution for μ and σ^2.

1.2 General Bayesian Formulation of the Problem

We are given a lot of $m + n$ items each of which has a variable quality characteristic X. For each item the random variable X is normally distributed with mean μ and variance σ^2 independent of all other items in the lot. An item is considered acceptable if its quality characteristic is within some specification limits, i.e., $a \leq X \leq b$. We take a sample of size n and calculate $\bar{x}$ and s^2. On the basis of these sufficient statistics, we decide to accept or reject the lot. If we reject the lot, we incur a loss c_0 for each item in the remaining m items that meets the specification and no loss for each item that fails to meet the specification. If we accept the lot, we incur a loss c_1 for each item in the remaining m items that does not meet the specification and no loss for each item meeting the specification.

Since the quality characteristics, $x_1, \ldots, x_m$, of the remaining m items are independently distributed, the probability that j of the m items meets the specification is

$$\binom{m}{j} P[a \leq x \leq b \mid \mu,\sigma]^j \{1 - P[a \leq x \leq b \mid \mu,\sigma]\}^{m-j} \quad ,$$

where $P[a \leq x \leq b \mid \mu,\sigma]$ is the probability that any given item meets the specification when x is normally distributed with parameters μ and σ. Let a_0 be the action of rejecting the lot and let a_1 be the action of accepting the lot. The loss functions associated with these actions for given μ and σ are thus defined as follows:

$$L[a_0,\mu,\sigma] = c_0 \sum_{j=0}^{m} j\binom{m}{j}P[a\leq x\leq b|\mu,\sigma]^j\{1 - P[a\leq x\leq b|\mu,\sigma]\}^{m-j} \quad ,$$

and

$$L[a_1,\mu,\sigma] = c_1 \sum_{j=0}^{m} j\binom{m}{j}\{1 - P[a\leq x\leq b|\mu,\sigma]\}^j P[a\leq x\leq b|\mu,\sigma]^{m-j} \quad .$$

Examining these loss functions we see that they can be simplified as follows:

$$L[a_0,\mu,\sigma] = c_0 m\, P[a \leq x \leq b|\mu,\sigma]$$

$$L[a_1,\mu,\sigma] = c_1 m\, \{1 - P[a \leq x \leq b|\mu,\sigma]\} \quad .$$

Next we define a decision rule based on $\bar{x}$ and s^2 for taking actions a_0 and a_1. Let $t(\bar{x},s^2)$ be the probability of taking action a_1 when $\bar{x}$ and s^2 are observed, i.e., t is a randomized decision rule. The risk function corresponding to t is defined as follows:

$$R[t,\mu,\sigma] = E_{\mu,\sigma}\, \{t(\bar{x},s^2)\, L[a_1,\mu,\sigma] + [1-t(\bar{x},s^2)]\, L[a_0,\mu,\sigma]\} \quad ,$$

where $E_{\mu,\sigma}$ denotes the expectation with respect to the joint distribution $F[\bar{x},s^2|\mu,\sigma]$. The statistician's aim is to minimize $R[t,\mu,\sigma]$, but in general no t can be found that will do this uniformly for all μ and σ.

However, in our problem we further assume that μ and σ are random variables with prior distribution $G(\mu,\sigma^2)$. In this case we consider the Bayes risk of t, where the Bayes risk is defined as follows:

$$\begin{aligned} R[t,G] &= E_G\,\{R[t,\mu,\sigma]\} \\ &= E_G\,\{E_{\mu,\sigma}\{t(\bar{x},s^2)L[a_1,\mu,\sigma] + [1-t(\bar{x},s^2)]L(a_0,\mu,\sigma]\}\}, \end{aligned} \tag{1.1}$$

where E_G denotes expectation with respect to $G(\mu,\sigma^2)$. Using Fubini's theorem we re-order the integration in (1.1); and, collecting terms, we have

$$\begin{aligned} R[t,G] &= E_G\,\{L[a_0,\mu,\sigma]\} \\ &\quad - E\,\{t(\bar{x},s^2)E_{G*}\{L[a_0,\mu,\sigma]-L[a_1,\mu,\sigma]\,|\,\bar{x},s^2\}\}\ , \end{aligned}$$

where E denotes the expectation with respect to the marginal distribution $F(\bar{x},s^2)$ and $E_{G*}\{\quad|\bar{x},s^2\}$ denotes the conditional expectation of μ and σ given $\bar{x}$ and s^2. Examining the second term of this expression we see that the Bayes risk is minimized by

$$\begin{aligned} t_G(\bar{x},s^2) &= 1, \text{ for } E_{G*}[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)\ |\ \bar{x},s^2] > 0 \\ t_G(\bar{x},s^2) &= 0, \qquad\qquad \text{otherwise}\quad , \end{aligned}$$

where

$$E_{G*}[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)\ |\ \bar{x},s^2] =$$

$$\frac{\iint_\Omega [L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)]\ f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)}{\iint_\Omega f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)}\ , \tag{1.2}$$

and

$$\Omega = \{(\mu,\sigma^2)|\ -\infty < \mu < \infty,\ 0 < \sigma < \infty\}\ .$$

Examining this expression more closely we see that

$$L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma) = mc_0\ P(a\le x\le b|\mu,\sigma) - mc_1[1-P(a\le x\le b|\mu,\sigma)]$$
$$= m(c_0+c_1)\ P(a\le x\le b|\mu,\sigma) - mc_1\ .$$

Using this simplification we see that equation (1.2) becomes

$$\frac{m(c_0+c_1) \iint_\Omega P(a\le x\le b|\mu,\sigma)\ f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)}{\iint_\Omega f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)} - mc_1\ . \qquad (1.3)$$

Since $P(a \le x \le b \mid \mu,\sigma) = \int_a^b f(x|\mu,\sigma)\ dx$, we can rewrite equation (1.2) as

$$-mc_1 + \frac{m(c_0+c_1) \iint_\Omega \int_a^b f(x|\mu,\sigma)\ dx f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)}{\iint_\Omega f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)}\ . \qquad (1.4)$$

Consider the integral in the numerator of the second term of (1.4). Note that

$$\iint_\Omega \int_a^b f(x|\mu,\sigma)\ dx f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2) =$$
$$\iint_\Omega \int_a^b f(x|\mu,\sigma)\ f(\bar{x},s^2|\mu,\sigma)\ dx dG(\mu,\sigma^2). \qquad (1.5)$$

By Fubini's theorem, we can change the order of integration so that (1.5) becomes

$$\int_a^b \iint_\Omega f(x|\mu,\sigma)\ f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2) dx\ . \qquad (1.6)$$

Let $f_G(x,\bar{x},s^2) = \iint_\Omega f(x|\mu,\sigma)\ f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2)$; then we have (1.5) that can be written as

$$\int_a^b f_G(x,\bar{x},s^2)\ dx\ . \qquad (1.7)$$

Next denote the denominator in the second term of (1.4) by

$$\iint_\Omega f(\bar{x},s^2|\mu,\sigma)\ dG(\mu,\sigma^2) = f_G(\bar{x},s^2)\ . \qquad (1.8)$$

Hence, equation (1.2) can be expressed as

$$E_{G*}[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)|\bar{x}, s^2] = -mc_1 + \frac{m(c_0+c_1) \int_a^b f_G(x,\bar{x},s^2)\, dx}{f_G(\bar{x},s^2)} .$$

Therefore, the Bayes decision rule is, in general,

$$t_G(\bar{x},s^2) = 1, \text{ for } -mc_1 + \frac{m(c_0+c_1) \int_a^b f_G(x,\bar{x},s^2)\, dx}{f_G(\bar{x},s^2)} \geq 0 \qquad (1.9)$$

and $t_G(\bar{x},s^2) = 0,$ otherwise.

In the third chapter of this paper we derive the decision rule when μ and σ^2 have a conjugate prior distribution. For a definition of a conjugate prior distribution see Raiffa and Schlaifer (1961). In this case μ is distributed as a normal variable with mean α and variance σ^2/β^2, and σ^{-2} is distributed as a gamma variable with parameters γ and δ .

1.3 Empirical Bayes Decision Rules: A Survey

Robbins in his pioneering paper of 1955 established the empirical Bayes approach to statistics. Paraphrasing Robbins' 1964 paper we say that an empirical Bayes approach to statistical decision problems is sometimes applicable when the same decision problem presents itself repeatedly and independently with a fixed but unknown prior distribution of the parameter or parameters. Since 1955 and particularly since 1962 when Neyman described the empirical Bayes approach as being a "breakthrough in the theory of statistical decision-making", there have been many papers published on the subject. Before we survey the literature, we shall formulate the general empirical Bayes decision problem to expedite the discussion of the literature.

This statistical decision problem is comprised of the following:

(a) A parameter space Λ with generic element λ. λ is the "state of nature" which is unknown to us.

(b) An action space A with generic element a.

(c) A loss function $L[a,\lambda] \geq 0$ representing the loss we incur in taking action a (or making decision a) when the parameter is λ.

(d) A prior distribution G of λ on Λ. G may or may not be known to us.

(e) An observable random variable X belonging to a space $\mathfrak{X}$ on which a σ-finite measure μ is defined. When the parameter is λ, X has a specified probability density f_λ with respect to μ. In a typical problem the random variable X is a sufficient statistic for the parameter λ.

The problem is to choose a decision function, δ, defined on X and with values in A such that when we observe X we take the action $\delta(x)$ and thereby incur the loss $L[\delta(x),\lambda]$. For any δ the expected loss when λ is the parameter is

$$R[\delta\ ,\lambda] = \int_X L[\delta(x),\lambda]\ f_\lambda(x)\ d\mu(x) \quad ,$$

and hence the overall risk (or Bayes risk) when the prior distribution of λ is G is

$$R[\delta,G] = \int_\Lambda \int_X L[\delta(x),\lambda]\ f_\lambda(x)\ d\mu(x)\ dG(\lambda) \ .$$

Changing the order of integration we have that

$$R[\delta,G] = \int_X \int_\Lambda L[\delta(x),\lambda]\ dG(\lambda|x)\ f_X(x)\ d\mu(x) \ ,$$

where

$$f_X(x) = \int_\Lambda f_\lambda(x)\ dG(\lambda) \ , \text{ and}$$

$$dG(\lambda|x)f_X(x) = \frac{f_\lambda(x)dG(\lambda)}{f_X(x)} \cdot f_X(x) = f_\lambda(x)dG(\lambda) \ .$$

Hence we see that $\int_\Lambda L[\delta(x),\lambda]\ dG(\lambda|x)$ is the posterior conditional expected loss given the observation x; i.e.,

$$E\{L[\delta(X),\lambda]\ |\ X = x\} = \int_\Lambda L[\delta(x),\lambda]\ dG(\lambda|x)\ .$$

The Bayes decision rule is that decision rule δ_0 that minimizes $E\{L[\delta(X),\lambda]\ |\ X = x\}$ for each x, and hence minimizes $R[\delta,G]$.

At this point the Bayes approach and the empirical Bayes approach differ. The Bayesian approach assumes that the prior distribution $G(\lambda)$ is known and hence $\delta_0(x)$ can be determined exactly. However, the empirical Bayes approach seeks to estimate $\delta_0(x)$ from past data, because $G(\lambda)$ is unknown. The empirical Bayes approach requires past data in the form of a sequence of pairs of random variables, say $(X_1, \lambda_1), \ldots, (X_\nu, \lambda_\nu)$. Each pair being independent of all other pairs, the λ_i having a common prior distribution G on Λ, and the conditional distribution of X_i given that $\lambda_i = \lambda$ being the specified probability density f_λ. When we come to make a decision about $\lambda_{\nu+1}$, we have observed $x_1, \ldots, x_{\nu+1}$ (the values $\lambda_1, \ldots, \lambda_\nu$ remaining always unknown). The object of the following literature survey is to discuss the various techniques developed to date to estimate the Bayes decision rule.

The first empirical Bayes technique for estimating the Bayes decision rule, δ_0, is applicable when the general form of the Bayes decision rule is a known function of $f_X(x)$, say $\Phi[f_X(x)]$ where

$$f_X(x) = \int_\Lambda f_\lambda(x)\ dG(\lambda)\ .$$

In this case, the empirical Bayes decision rule is defined as follows

$$\Delta_\nu\ (x|x_1, \ldots, x_\nu) = \Phi[f_\nu(x)]\ ,$$

where $f_\nu(x)$ is an estimate of the marginal density $f_X(x)$ based on the past data $(x_1, \ldots, x_\nu)$.

We see that this first empirical Bayes technique relies, heavily, on the particular density function estimation technique used; i.e., the statistical properties of this empirical Bayes technique are no better than the statistical properties of the density function estimation technique. The most desirable statistical property for the empirical Bayes rule is convergence in probability. Robbins (1964), and Samuel (1963) show that the risk for the empirical Bayes rule converges to the Bayes risk when the empirical Bayes rule converges in probability to the Bayes rule. Robbins defines the empirical Bayes rule as being asymptotically optimal in this case. If Φ is a measurable mapping, continuous on the real line, then $\Phi[f_\nu(x)]$ converges in probability to $\Phi[f_X(x)]$, if $f_\nu(x)$ converges in probability to $f_X(x)$. Hence we see that we need density function estimators that converge in probability.

Fortunately, there is considerable research on density function estimation for applications other than empirical Bayes decision rules. Parzen (1962) develops a class of density function estimators that converge in probability. Cacoullos (1966) extends Parzen's results to the multivariate case. These two authors' results, alone, guarantee us that we can find an asymptotically optimal empirical Bayes rule when the Bayes rule is a known function of the marginal density function, $f_X(x)$.

However, not all Bayes rules are a known function of $f_X(x)$. This property of the Bayes rule exists when the loss function and the density $f_\lambda(x)$ have some special and hard to define relationship. An example of this can be seen in Robbins (1955). In this paper Robbins derives the empirical Bayes estimate of the parameter in a Poisson distribution, when the loss function is squared error loss. Samuel (1963) discusses some loss functions that yield decision rules that are known functions of $f_X(x)$ when $f_\lambda(x)$ is in the exponential family. Rutherford and Krutchkoff (1969) define families of distribution functions $F(x|\lambda)$ that yield Bayes rules of this form when the loss function is squared error.

The second empirical Bayes technique for estimating the Bayes rule involves estimating the prior distribution $G(\lambda)$. This technique involves approximating $dG(\lambda)$ by a step function $dA_k(\lambda)$ where

$$dA_k(\lambda) = \frac{d\lambda}{(k-1)\ (\lambda_{j+1}-\lambda_j)} \quad , \text{ for}$$

$\lambda_j < \lambda < \lambda_{j+1}$, $j = 1, \ldots, k-1$. This yields an estimate of $F_{A_k}(x)$ defined by $F_{A_k}(x) = \int F(x|\lambda)\ dA_k(\lambda)$. The λ_j's are chosen to minimize the difference between $F_{A_k}(x)$ and $F_\nu(x)$ where $F_\nu(x)$ is the sample distribution function defined as

$$F_\nu(x) = \frac{\nu_x}{\nu}$$

where ν_x is the number of past observations of $(x_1, \ldots, x_\nu)$ less than x. For a detailed description of this technique see Maritz (1970). Once this estimate of $dG(\lambda)$ is obtained it is used to determine the Bayes rule as if it were the actual prior distribution.

The third technique involves estimating $G(\lambda)$ also, but is distinctly different from the previous method. For this technique we assume that at each realization of the process we have a sample of say n observations. That is, we have the sequence of random variables

$$(x_{11}, \ldots, x_{n1}, \lambda_1)\ , \ldots, (x_{1\nu}, \ldots, x_{n\nu}, \lambda_\nu)$$

where for each $j = 1, \ldots, \nu$ the x_{ij} are independently distributed F_{λ_j}, and the λ_j are independently distributed $G(\lambda)$. This technique assumes that a statistic $t(x_{1j}, \ldots, x_{nj})$ exists such that t converges in probability to λ_j. When n is sufficiently large, this technique uses $t(x_{1j}, \ldots, x_{nj})$ as if it were λ_j and proceeds to estimate $g(\lambda)$ where g is the probability density function associated with $G(\lambda)$. The probability density function g is estimated using the density function estimators discussed for the first technique. This estimate for $g(\lambda)$ is used in the Bayes rule as if it were the correct $g(\lambda)$. For a more detailed discussion of this technique see Lemon and Krutchkoff (1969).

Our empirical Bayes technique for estimating the Bayes decision rule in this paper is more like the first technique discussed in this survey in that we show that the Bayes decision rule (see equation 1.9) is a known function of $f_G(\bar{x},s^2)$ and $\int_a^b f_G(x,\bar{x},s^2)\,dx$. We are able to rely on previous research for estimating $f_G(\bar{x},s^2)$, but we must find and justify a method for estimating $\int_a^b f_G(x,\bar{x},s^2)\,dx$. Hence in Chapter II when we discuss our empirical Bayes decision rule, we develop and justify a method for estimating $\int_a^b f(x_1, \ldots, x_k)\,dx_1$ where $f(x_1, \ldots, x_k)$ is any uniformly continuous multivariate probability density function. We show in Chapter II that our empirical Bayes decision rule converges in probability to the Bayes rule.

This survey is in no sense comprehensive or complete, but is intended to summarize the different empirical Bayes techniques. We make no attempt here or later to discuss which technique is better in any sense. If the reader is interested in a more complete survey, see Maritz (1970).

CHAPTER II

THE EMPIRICAL BAYES DECISION RULE

2.1 Introduction

Now we seek to estimate empirically the Bayes decision rule (see equation 1.9) for accepting or rejecting the lot. We assume that in the past we have inspected ν lots taking a sample of size n from each lot - calculating $\bar{x}$ and s^2 and measuring at least one other item in each of the lots not in the samples. That is, we assume that we have had ν past realizations, (μ_i, σ_i) $i = 1, \ldots, \nu$, of the unobservable random variables (μ, σ^2) which are distributed according to the unknown joint distribution $G(\mu, \sigma^2)$. At the i^{th} realization, m + n items are produced with characteristics $X_{i1}, X_{i2}, \ldots, X_{i(m+n)}$ which are identically and independently distributed normal variables with mean μ_i and standard deviation σ_i. From each of these realizations a sample of size n is taken, $\bar{X}_i$ and S_i^2 are calculated, and one more sample is taken from the lot and its characteristic measured. Hence, we have the past data $(\bar{X}_i, S_i^2, X_{i,n+1})$ $i = 1, \ldots, \nu$.

Using this past data we want to estimate empirically the Bayes decision rule

$$t_G(\bar{x}, s^2) = 1, \quad m(c_0 + c_1) \frac{\int_a^b f_G(x, \bar{x}, s^2)\,dx}{f_G(\bar{x}, s^2)} - mc_1 \geq 0 \qquad (1.9)$$

$$t_G(\bar{x}, s^2) = 0, \qquad \text{otherwise} .$$

This is possible since the decision rule is a function of the joint marginal density function of the random variables $(\bar{X}, S^2, X)$ (see section 1.2).

Let the Bayes decision function be defined by

$$T_G(\bar{x}, s^2) = m(c_0 + c_1) \frac{\int_a^b f_G(x, \bar{x}, s^2)dx}{f_G(\bar{x}, s^2)} - mc_1 \tag{1.9a}$$

We estimate $T_G(\bar{x}, s^2)$ by $T_\nu(\bar{x}, s^2)$, which is determined by estimating $\int_a^b f_G(x, \bar{x}, s^2)dx$ and $f_G(\bar{x}, s^2)$ from the past data $(\bar{X}_i, S_i^2, X_i)$ $i = 1, \ldots, \nu$.

2.2 Estimation of $f_G(\bar{x}, s^2)$

The question now arises as to how we propose to estimate $\int_a^b f_G(x, \bar{x}, s^2)dx$ and $f_G(\bar{x}, s^2)$. First we discuss the estimation of $f_G(\bar{x}, s^2)$.

Parzen (1962) developed a class of density function estimators $f_n(x)$ of a univariate density function $f(x)$ on the basis of a random sample $X_1, \ldots, X_n$ from $f(x)$, where $f_n(x)$ is of the form

$$f_n(x) = \frac{1}{nh(n)} \sum_{j=1}^{n} K\left[\frac{x - X_j}{h(n)}\right]$$

and where

1) $\lim_{n \to \infty} h(n) = 0$

2) $\lim_{n \to \infty} nh^2(n) = \infty$

3) $\lim_{n \to \infty} nh(n) = \infty$

4) $\int_{-\infty}^{\infty} K(y)dy = 1$

5) $\lim_{|y| \to \infty} |y| \, |K(y)| = 0$

6) $\int_{-\infty}^{\infty} |K(y)| dy < \infty$, and

7) $\sup_{-\infty < y < \infty} |K(y)| < \infty$.

Parzen (1962) presented the asymptotic properties of this class of estimators. In particular, Parzen showed that $f_n(x)$ is asymptotically unbiased, consistent in quadratic mean, and converges uniformly in probability; i.e.,

$$\lim_{n \to \infty} E[f_n(x) - f(x)] = 0$$

$$\lim_{n \to \infty} E\{[f_n(x) - f(x)]^2\} = 0 \text{ , and}$$

for any $\varepsilon > 0$

$$\lim_{n \to \infty} P\{\sup_{-\infty<x<\infty} |f(x) - f_n(x)|>\varepsilon\} = 0$$

if $f(x)$ is uniformly continuous.

Cacoullos (1966) extended Parzen's results to the multivariate case. He developed a joint density function estimator, $f_n(x_1, \ldots, x_k)$, of $f(x_1, \ldots, x_k)$ of the form

$$f_n(x_1, \ldots, x_k) = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=1}^{k} \frac{1}{h_i(n)} K\left[\frac{x_i - x_{ij}}{h_i(n)}\right]$$

where $(x_{ij}, \ldots, x_{kj})$ $j = 1, \ldots, n$ are n independent observations on the k-dimensional random vector $(X_1, \ldots, X_k)$. Cacoullos' class of kernels, $K(y)$, is the same as the one developed by Parzen and the $h_i(n)$'s have the same properties as Parzen's $h(n)$. Cacoullos showed that his multivariate density function estimator was also asymptotically unbiased, consistent in quadratic mean, and converges uniformly in probability, if $f(x_1, \ldots, x_k)$ is uniformly continuous,

$$\lim_{n \to \infty} n \prod_{i=1}^{k} h_i^2(n) = \infty \text{ , and}$$

$$\lim_{n \to \infty} n \prod_{i=1}^{k} h_i(n) = \infty \text{ .}$$

Epanechnikov (1969) restricted the class of kernels considered by Parzen and Cacoullos by adding the further constraints of

1) $K(y) = K(-y)$
2) $\int_{-\infty}^{\infty} y^2 K(y)\, dy = 1$, and
3) $\int_{-\infty}^{\infty} K(y)\, y^m\, dy < \infty$ for $0 \leq m < \infty$.

In this restricted class of kernels, he was able to find a kernel that minimizes the asymptotic relative global mean square error regardless of the density function and the dimensionality, where the relative global mean square error is defined as:

$$\frac{\int \cdots \int E\{[f(x_1, \ldots, x_k) - f_n(x_1, \ldots, x_k)]^2\}\, dx_1 \ldots dx_k}{\int \cdots \int f^2(x_1, \ldots, x_k)\, dx_1, \ldots, dx_k} .$$

He further compares the asymptotic relative global mean square error of this optimum kernel to the asymptotic relative global mean square errors of other kernels in this class.

After studying these three papers, we decided to estimate $f_G(\bar{x},s^2)$ by using the following kernel,

$$K(y) = \frac{1}{\sqrt{2\pi}}\, e^{-1/2\, y^2} .$$

Hence, the estimate of $f_G(\bar{x},s^2)$ is defined as follows

$$f_\nu(\bar{x},s^2) = \frac{1}{\nu} \sum_{j=1}^{\nu} \frac{1}{2\pi h^2(\nu)} \exp\left\{-\frac{1}{2}\left[\left(\frac{\bar{x}-\bar{X}_j}{h(\nu)}\right)^2 + \left(\frac{s^2-S_j^2}{h(\nu)}\right)^2\right]\right\} \qquad (2.1)$$

where

1) $h(\nu) = (\frac{1}{\nu})^{1/7}$, and

2) ν is the number of past realizations of (μ,σ) yielding data $(\bar{X}_j, S_j^2, X_{j,n+1})$ $j=1, \ldots, \nu$; i.e., number of past lots.

We did not use the optimal kernel of Epanechnikov, because it is possible for $f_\nu(\bar{x},s^2)$ to be zero. Furthermore the ratio of the relative global mean square error for the normal density function kernel to the relative global mean square error of the optimal kernel is 1.051, asymptotically.

2.3 Estimation of $\int_a^b f_G(x,\bar{x},s^2)\,dx$

Next we discuss the estimation of $\int_a^b f_G(x,\bar{x},s^2)dx$. First we note that with $h(\nu)$ and $(\bar{x}_j, s_j^2, x_{j,m+1})$ being defined as in 2.1 we can estimate $f_G(x,\bar{x},s^2)$ by

$$f_\nu(x,\bar{x},s^2) = \frac{1}{\nu}\sum_{j=1}^{\nu}\frac{1}{(2\pi)^{3/2}h^3(\nu)}\exp\left\{-\frac{1}{2}\left[\left(\frac{x-X_j}{h(\nu)}\right)^2+\left(\frac{\bar{x}-\bar{X}_j}{h(\nu)}\right)^2+\left(\frac{s^2-S_j^2}{h(\nu)}\right)^2\right]\right\}. \tag{2.2}$$

Intuitively we are tempted to estimate $\int_a^b f_G(x,\bar{x},s^2)dx$ by $\int_a^b f_\nu(x,\bar{x},s^2)dx$. In this section we show that this estimation procedure yields an estimator which is asymptotically unbiased, consistent in quadratic mean, and converges uniformly in probability under very general conditions. We proceed to prove these properties in the following theorems.

Theorem 1:

Let $f_\nu(x_1, \ldots, x_k)$ be an estimator of $f(x_1, \ldots, x_k)$ where $f_\nu(x_1, \ldots, x_k)$ is defined as follows:

$$f_\nu(x_1, \ldots, x_k) = \frac{1}{\nu h^k(\nu)}\sum_{j=1}^{\nu}\prod_{\ell=1}^{k} K\left[\frac{x_\ell - y_{\ell j}}{h(\nu)}\right]$$

and

a) $(y_{1j}, \ldots, y_{kj})$, $j=1, \ldots, \nu$ represent a sample of ν independent observations from the distribution with joint probability density $f(x_1, \ldots, x_k)$.

b) K(y) has the following properties:

(i) K(y) is a Borel measurable function

(ii) $\sup_{-\infty < y < \infty} |K(y)| < \infty$

(iii) $K(y) \geq 0$

(iv) $\lim_{|\underline{y}| \to \infty} |\underline{y}|^k \prod_{\ell=1}^{k} K[y_\ell] = 0,$

where $|\underline{y}| = (y_1^2 + \cdots + y_k^2)^{\frac{1}{2}}$

(v) $\int_{-\infty}^{\infty} K(y)\, dy = 1$

(vi) $K(y) = K(-y)$

c) the spreading coefficient satisfies

(i) $\lim_{\nu \to \infty} h(\nu) = 0$

(ii) $\lim_{\nu \to \infty} \nu h(\nu)^{2k} = \infty$.

Then at every continuity point of $\int_a^b f(x_1, \ldots, x_k)\, dx_1$,

$$\lim_{\nu \to \infty} E\{\int_a^b f_\nu(x_1, \ldots, x_k)\, dx_1\} = \int_a^b f(x_1, \ldots, x_k)\, dx_1 .$$

Proof:

1) First we note that, since $(y_{1j}, \ldots, y_{kj})$ $j=1, \ldots, \nu$ are all identically distributed

$E\{\int_a^b f_\nu(x_1, \ldots, x_k)\, dx_1\} =$

$$\frac{1}{h^k(\nu)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_a^b \prod_{\ell=1}^{k} K\left[\frac{x_\ell - y_\ell}{h(\nu)}\right] dx_1\, f(y_1, \ldots, y_k) \prod_{\ell=1}^{k} dy_\ell \quad ,$$

and since $f_\nu(x_1, \ldots, x_k)$ is non-negative and integrable we can apply Fubini's theorem,

$E\{\int_a^b f_\nu(x_1, \ldots, x_k)\, dx_1\} =$

$$\frac{1}{h^k(\nu)} \int_a^b \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{\ell=1}^{k} K\left[\frac{x_\ell - y_\ell}{h(\nu)}\right] f(y_1, \ldots, y_k) \prod_{\ell=1}^{k} dy_\ell\, dx_1 \quad ,$$

2) Letting $u_\ell = y_\ell - x_\ell$ $\ell = 1, \ldots, k$, we have

$$E\{\int_a^b f_\nu (x_1, \ldots, x_k)\, dx_1\} =$$

$$\frac{1}{h^k(\nu)} \int_a^b \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] f(u_1+x_1, \ldots, u_k+x_k) \prod_{\ell=1}^{k} du_\ell dx_1 .$$

3) Next we consider

$$\Delta_\nu = |\int_a^b f(x_1, \ldots, x_k)\, dx_1 - E\{\int_a^b f_\nu(x_1, \ldots, x_k)\, dx_1\}| \text{ , or}$$

$$\Delta_\nu = |\int_a^b f(x_1,\ldots,x_k)dx_1 \frac{1}{h^k(\nu)} \prod_{\ell=1}^{k} \int_{-\infty}^{\infty} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell - E\{\int_a^b f_\nu(x_1,\ldots,x_k)dx_1\}|,$$

since $\frac{1}{h(\nu)} \int_{-\infty}^{\infty} K\left[\frac{u}{h(\nu)}\right] du = 1$,

$$\Delta_\nu = |\frac{1}{h^k(\nu)} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \int_a^b [f(x_1,\ldots,x_k) - f(u_1+x_1,\ldots,u_k+x_k)]dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell| .$$

4) From this point on our proof parallels that of Cacoullos' (1966) proof of theorem 2.1. We divide the region of integration with respect to $(u_1, \ldots, u_k)$ into two regions $|\underline{u}| \leq \delta$ and $|\underline{u}| > \delta$ where

$$|\underline{u}| = \sqrt{u_1^2 + \ldots + u_k^2} .$$

Then we consider

$$\Delta_\nu = \left|\frac{1}{h^k(\nu)} \int\cdots\int_{|\underline{u}|\leq\delta} \int_a^b [f(x_1,\ldots,x_k) - f(u_1+x_1,\ldots,u_k+x_k)]dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell\right.$$

$$\left. + \frac{1}{h^k(\nu)} \int\cdots\int_{|\underline{u}|>\delta} \int_a^b [f(x_1,\ldots,x_k) - f(u_1+x_1,\ldots,u_k+x_k)]dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell\right|$$

$$\Delta_\nu \leq \left|\frac{1}{h^k(\nu)} \int\cdots\int_{|\underline{u}|\leq\delta} \int_a^b [f(x_1,\ldots,x_k) - f(u_1+x_1,\ldots,u_k+x_k)]dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell\right|$$

$$+ \left|\frac{1}{h^k(\nu)} \int\cdots\int_{|\underline{u}|>\delta} \int_a^b f(u_1+x_1,\ldots,u_k+x_k)dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell\right|$$

$$+\left|\frac{1}{h^k(\nu)} \int\cdots\int_{|\underline{u}|>\delta} \int_a^b f(x_1,\ldots,x_k)dx_1 \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right]du_\ell\right|$$

$$\Delta_\nu < \max_{|\underline{u}|\leq\delta} \left|\int_a^b[f(x_1,\ldots,x_k)-f(u_1+x_1,\ldots,u_k+x_k)]dx_1\right| \int\cdots\int_{|\underline{u}|\leq\delta} \frac{1}{h^k(\nu)} \prod_{\ell=1}^{k} K\left[\frac{u_\ell}{h(\nu)}\right] du_\ell$$

$$+ \left|\int\cdots\int_{|\underline{u}|>\delta} \frac{\int_a^b f(u_1+x_1,\ldots,u_k+x_k)dx_1}{|\underline{u}|^k} \frac{|\underline{u}|^k}{h^k(\nu)} \prod_{\ell=1}^{k} \left|K\left[\frac{u_\ell}{h(\nu)}\right]\right| du_\ell\right|$$

$$+ \left|\int_a^b f(x_1,\ldots,x_k)dx_1\right| \int\cdots\int_{|\underline{u}|>\delta} \frac{1}{h^k(\nu)} \prod_{\ell=1}^{k} \left|K\left[\frac{u_\ell}{h(\nu)}\right]\right| du_\ell \quad .$$

Let $$Z_j = \frac{u_j}{h(\nu)} \qquad j = 1, \ldots, k \ ,$$

then $$u_j = h(\nu)\, Z_j \quad .$$

If $|\underline{u}| \leq \delta$, then

$$h(\nu)\ |\underline{Z}| \leq \delta, \text{ or}$$

$$|\underline{Z}| \leq \delta/h(\nu)$$

and

$$\Delta_\nu \leq \max_{|\underline{u}|\leq\delta} \left|\int_a^b[f(x_1,\ldots,x_k)-f(x_1+u_1,\ldots,x_k+u_k)]dx_1\right| \int\cdots\int_{|\underline{Z}|\leq \frac{\delta}{h(\nu)}} \prod_{\ell=1}^{k} \left|K[Z_\ell]\right| dZ_\ell$$

$$+ \frac{1}{\delta^k} \sup_{|\underline{Z}|>\frac{\delta}{h(\nu)}} \left(\left||\underline{Z}|^k \prod_{\ell=1}^{k} K[Z_\ell]\right|\right) \int\cdots\int_{|\underline{u}|>\delta} \left|\int_a^b f(u_1+x_1,\ldots,u_k+x_k)\right| \prod_{\ell=1}^{k} du_\ell$$

$$+ \left|\int_a^b f(x_1,\ldots,x_k)dx_1\right| \int\cdots\int_{|Z|>\frac{\delta}{h(\nu)}} \prod_{\ell=1}^{k} \left|K[Z_\ell]\right| dZ_\ell \quad .$$

5) Next we note that

$$\int\cdots\int_{|\underline{Z}|\leq \frac{\delta}{h(\nu)}} \prod_{\ell=1}^{k} \left|K[Z_\ell]\right| dZ_\ell \leq \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \prod_{\ell=1}^{k} \left|K[Z_\ell]\right| dZ_\ell = 1 \ .$$

Examining the other terms of the above inequality that involve $h(\nu)$ and taking the limit of Δ_ν as $\nu \to \infty$ we see that

a) $$\lim_{\nu \to \infty} \frac{1}{\delta^k} \sup_{|\underline{Z}| > \frac{\delta}{h(\nu)}} \left| |\underline{Z}|^k \prod_{\ell=1}^{k} K[Z_\ell] \right| = 0$$

by property b(iv), and

b) $$\lim_{\nu \to \infty} \int \cdots \int_{|\underline{Z}| > \frac{\delta}{h(\nu)}} \prod_{\ell=1}^{k} \left| K[Z_\ell] \right| \; dZ_\ell = 0$$

Hence we see that

$$\lim_{\nu \to \infty} \Delta_\nu < \max_{|\underline{u}| \leq \delta} \left| \int_a^b [f(x_1,\ldots,x_k) - f(x_1+u_1,\ldots,x_k+u_k)] dx_1 \right|$$

6) Next we take the limit as $\delta \to 0$ and we see that

$$\lim_{\delta \to 0} \max_{|\underline{u}| \leq \delta} \left| \int_a^b [f(x_1,\ldots,x_k) - f(x_1+u_1,\ldots,x_k+u_k)] dx_1 \right| = 0 \quad .$$

Hence we have that

$$\lim_{\nu \to \infty} E\{\int_a^b f_\nu(x_1,\ldots,x_k) dx_1\} = \int_a^b f(x_1,\ldots,x_k) dx_1 \quad ,$$

(i.e., we have an asymptotically unbiased estimator).

Also we consider the following corollary to this theorem.

Corollary:

If $f(x_1,\ldots,x_k)$ is uniformly continuous on E_k then

$$\lim_{\nu \to \infty} \max_{(x_2,\ldots,x_k) \epsilon E_{k-1}} \left| E\{\int_a^b f_\nu(x_1,\ldots,x_k) dx_1\} - \int_a^b f(x_1,\ldots,x_k) dx_1 \right|^2 = 0 \; .$$

Proof:

1) By properties b_{ii}, b_{iv}, and b_{vi} we have that $\int_a^b f_\nu(x_1,\ldots,x_\nu) dx_1$ is uniformly continuous.

2) Since $\int_a^b f_\nu(x_1,\ldots,x_k)dx_1$ and $\int_a^b f(x_1,\ldots,x_k)dx_1$ are uniformly continuous, then

$$\max_{(x_2,\ldots,x_k)\varepsilon E_{k-1}} \left| E\{\int_a^b f\ (x_1,\ldots,x_k)dx_1\} - \int_a^b f(x_1,\ldots,x_k)dx_1 \right|^2 < \infty.$$

3) Since the point at which the maximum is obtained is a continuity point of $\int_a^b f(x_1,\ldots,x_k)dx_1$ due to $\int_a^b f(x_1,\ldots,x_k)dx_1$'s uniform continuity, we have that

$$\lim_{\nu\to\infty} \max_{(x_2,\ldots,x_k)\varepsilon E_{k-1}} \left| E\{\int_a^b f_\nu(x_1,\ldots,x_k)dx_1\} - \int_a^b f(x_1,\ldots,x_k)dx_1 \right|^2 = 0.$$

Before proving a theorem about the mean square convergence we must define some new terms and establish some new relationships. First we seek to express $f_\nu(x_1,\ldots,x_k)$ as a function, of the sample characteristic function,

$$\phi_\nu(u_1,\ldots,u_k) = \frac{1}{\nu}\sum_{j=1}^{\nu} \exp\left\{i \sum_{\ell=1}^{k} u_\ell y_{\ell j}\right\} \qquad (2.3)$$

and $(y_{1j},\ldots,y_{kj})$ $j = 1,\ldots,\nu$ is a sample of ν independent observations from a distribution with density $f(x_1,\ldots,x_k)$. Let

$$\prod_{\ell=1}^{k} J[u_\ell] = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left\{i \sum_{\ell=1}^{k} u_\ell y_\ell\right\} \prod_{\ell=1}^{k} K[y_\ell]dy_\ell, \text{ i.e.,}$$

$\prod_{\ell=1}^{k} J[u_\ell]$ is the Fourier transform of the product of the kernels. Let

$\Phi_\nu(x_1, \ldots, x_k) =$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left\{-i \sum_{\ell=1}^{k} u_\ell x_\ell\right\} \prod_{\ell=1}^{k} J[h(\nu)u_\ell]\phi_\nu(u_1, \ldots, u_k) \prod_{\ell=1}^{k} du_\ell$$

or

$$\Phi_\nu(x_1, \ldots, x_k) =$$

$$\frac{1}{2\pi\nu} \sum_{j=1}^{\nu} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i \sum_{\ell=1}^{k} u_\ell (x_\ell - y_{\ell j})\right\} \prod_{\ell=1}^{k} J[h(\nu)u_\ell] \prod_{\ell=1}^{k} du_\ell .$$

Let $Z_\ell = h(\nu)u_\ell$ $\ell = 1, \ldots, k$, then

$$\Phi_\nu(x_1, \ldots, x_k) =$$

$$\frac{1}{2\pi\nu h^k(\nu)} \sum_{j=1}^{\nu} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i \sum_{\ell=1}^{k} Z_\ell \left(\frac{x_\ell - y_{\ell j}}{h(\nu)}\right)\right\} \prod_{\ell=1}^{k} J[Z_\ell]\, dZ_\ell .$$

Since $\frac{1}{2\pi} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i \sum_{\ell=1}^{k} Z_\ell \left(\frac{x_\ell - y_{\ell j}}{h(\nu)}\right)\right\} \prod_{\ell=1}^{k} J[Z_\ell]\, dZ_\ell$ $j = 1, \ldots, \nu$

is the inverse Fourier transform of $\prod_{\ell=1}^{k} K\left[\frac{x_\ell - y_{\ell j}}{h(\nu)}\right]$, we see that

$$\Phi_\nu(x_1, \ldots, x_k) = \frac{1}{\nu h^k(\nu)} \sum_{j=1}^{\nu} \prod_{\ell=1}^{k} K\left[\frac{x_\ell - y_{\ell j}}{h(\nu)}\right] .$$

Hence we see that

$$f_\nu(x_1, \ldots, x_k) =$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i \sum_{\ell=1}^{k} u_\ell x_\ell\right\} \prod_{\ell=1}^{k} J[h(\nu)u_\ell]\phi_\nu(u_1, \ldots, u_k) \prod_{\ell=1}^{k} du_\ell, \text{ i.e.,} \tag{2.4}$$

we have expressed $f_\nu(x_1, \ldots, x_k)$ as a function of the sample characteristic function. To simplify notation we define

$$\underline{u}'\underline{x} = \sum_{\ell=1}^{k} u_\ell x_\ell ,$$

$$\underline{u}'\underline{y}_j = \sum_{\ell=1}^{k} u_\ell y_{\ell j} \qquad j = 1, \ldots, \nu ,$$

$$J[h(\nu)\underline{u}] = \prod_{\ell=1}^{k} J[h(\nu)u_\ell] ,$$

$$\phi_\nu(\underline{u}) = \phi_\nu(u_1, \ldots, u_k) ,$$

$$d\underline{u} = \prod_{\ell=1}^{k} du_\ell ,$$

and denote $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}$ by $\underline{\int}_{-\infty}^{\infty}$.

so that equation (2.4) can be expressed as

$$f_\nu(x_1,\ldots,x_k) = \frac{1}{2\pi} \underline{\int}_{-\infty}^{\infty} e^{-i\underline{u}'\underline{x}} J[h(\nu)\underline{u}]\phi_\nu(\underline{u})d\underline{u} . \tag{2.4a}$$

Taking the expected value of $\phi_\nu(\underline{u})$ we see that

$$E[\phi_\nu(\underline{u})] = \frac{1}{\nu} \sum_{j=1}^{\nu} \underline{\int}_{-\infty}^{\infty} e^{i\underline{u}'\underline{y}_j} dF(\underline{y}_j)$$

$$= \underline{\int}_{-\infty}^{\infty} e^{i\underline{u}'\underline{y}} dF(\underline{y}) = \phi[\underline{u}]$$

where $\phi[\underline{u}]$ is the characteristic function of $F(x_1,\ldots,x_k)$.

Studying the sample characteristic function further, we derive the variance of $\phi_\nu(\underline{u})$. This will be of value when we prove the mean square convergence theorem. First we must consider $|\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2$ - the square of the modulus or absolute value of the difference between two complex variables.

$$|\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2 = \left[\phi_\nu(\underline{u})-E[\phi_\nu(\underline{u})]\right]\left[\overline{\phi_\nu(\underline{u})-E[\phi_\nu(\underline{u})]}\right] \tag{2.5}$$

where $\overline{\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]}$ is the complex conjugate of $\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]$.

Furthermore

$$\overline{\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]} = \overline{\phi_\nu(\underline{u})} - \overline{E[\phi_\nu(\underline{u})]}$$

$$= \overline{\phi_\nu(\underline{u})} - \overline{\phi(\underline{u})} .$$

Also we note that

$$\overline{\phi(\underline{u})} = \underline{\int}_{-\infty}^{\infty} e^{-i\underline{u}'\underline{y}} dF(\underline{y}) ,$$

$$E\{\overline{\phi_\nu(\underline{u})}\} = \frac{1}{\nu} \sum_{j=1}^{\nu} \underline{\int}_{-\infty}^{\infty} e^{-i\underline{u}'\underline{y}_j} dF(\underline{y}_j) = \overline{\phi(\underline{u})} ,$$

$$|\phi_\nu(\underline{u})-E[\phi_\nu(\underline{u})]|^2 = \phi_\nu(\underline{u})\overline{\phi_\nu(\underline{u})} - \phi_\nu(\underline{u})\overline{\phi(\underline{u})} - \overline{\phi_\nu(\underline{u})}\phi(\underline{u}) + \phi(\underline{u})\overline{\phi(\underline{u})}$$

and

$$\phi_\nu(\underline{u})\overline{\phi_\nu(\underline{u})} = \frac{1}{\nu^2}\sum_{j=1}^{\nu} \exp\{i[\underline{u}'\underline{y}_j - \underline{u}'\underline{y}_j]\} + \frac{2}{\nu^2}\sum_{j=1}^{\nu-1}\sum_{p=j+1}^{\nu} e^{i\underline{u}'\underline{y}_j} e^{-i\underline{u}'\underline{y}_p}$$

$$= \frac{1}{\nu} + \frac{2}{\nu^2}\sum_{j=1}^{\nu-1}\sum_{p=j+1}^{\nu} e^{i\underline{u}'\underline{y}_j}\, e^{-i\underline{u}'\underline{y}_p} \quad .$$

Taking the expected value of equation (2.5) we see that

$$E\{|\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2\} = \frac{1}{\nu} + \frac{\nu(\nu-1)}{\nu^2}\phi(\underline{u})\overline{\phi(\underline{u})} - 2\phi(\underline{u})\overline{\phi(\underline{u})} + \phi(\underline{u})\overline{\phi(\underline{u})} \ ,$$

$$= \frac{1}{\nu} - \frac{1}{\nu}\phi(\underline{u})\overline{\phi(\underline{u})} \ , \text{ or}$$

$$= \frac{1}{\nu}[1 - |\phi(\underline{u})|^2] \ . \tag{2.5a}$$

Also of later use is the following lemma.

<u>Lemma</u>: If $\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| du < \infty$, then

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|\ |\phi_\nu(\underline{u}) - \phi(\underline{u})|^2 d\underline{u} < \infty \ .$$

<u>Proof</u>:

1) $\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})] = \phi_\nu(\underline{u}) - \phi(\underline{u})$

2) $|\phi_\nu(\underline{u}) - \phi(\underline{u})| \le |\phi_\nu(\underline{u})| + |\phi(\underline{u})|$

3) Next we see that

$|\phi(\underline{u})| \le 1$

$|\phi_\nu(\underline{u})| \le \frac{1}{\nu}\sum_{j=1}^{\nu} |e^{i\underline{u}'\underline{y}_j}|$

$|\phi_\nu(\underline{u})| \le 1$.

$|\phi_\nu(\underline{u}) - \phi(\underline{u})| \le 2$

$|\phi_\nu(\underline{u}) - \phi(\underline{u})|^2 \le 4$

4) Hence we see that

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| \; |\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2 d\underline{u} \leq 4 \int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| d\underline{u} \; .$$

Therefore if $\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| \; d\underline{u} < \infty$, then

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| \; |\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2 d\underline{u} < \infty \; .$$

With these results established we proceed to prove the following theorem on mean square convergence.

Theorem 2:

If $f(x_1, \ldots, x_k)$ is uniformly continuous, $f_\nu(x_1, \ldots, x_k)$ satisfies all the conditions of theorem 1, and

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| d\underline{u} < \infty \; ,$$

we have that

$$\lim_{\nu \to \infty} \max_{(x_2,\ldots,x_k)\epsilon E_{k-1}} E\{(\int_a^b f_\nu(x_1,\ldots,x_k)dx_1 - \int_a^b f(x_1,\ldots,x_k)dx_1)^2\} = 0.$$

Proof:

1. First we note that

$$E\{|\int_a^b f_\nu(x_1,\ldots,x_k)dx_1 - \int_a^b f(x_1,\ldots,x_k)dx_1|^2\} =$$

$$E\{|\int_a^b f_\nu(x_1,\ldots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\ldots,x_k)dx_1]|^2\}$$

$$+ \{E[\int_a^b f_\nu(x_1,\ldots,x_k)dx_1] - \int_a^b f(x_1,\ldots,x_k)dx_1\}^2 .$$

2. By the corollary to theorem 1 we have that

$$\lim_{\nu\to\infty} \max_{(x_2,\ldots,x_k)\epsilon E_{k-1}} \{E[\int_a^b f_\nu(x_1,\ldots,x_k)dx_1] - \int_a^b f(x_1,\ldots,x_k)dx_1\}^2 = 0$$

3. Next we study $E\{|\int_a^b f_\nu(x_1,\dots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\dots,x_k)dx_1]|^2\}$.

Since $f_\nu(x_1,\dots,x_k)$ is non-negative and integrable we can apply Fubini's theorem.

$$\int_a^b f_\nu(x_1,\dots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\dots,x_k)dx_1]$$

$$= \int_a^b \{f_\nu(x_1,\dots,x_k) - E[f_\nu(x_1,\dots,x_k)]\}dx_1$$

$$= \frac{1}{2\pi}\int_a^b \Big\{ \int_{-\infty}^{\infty} e^{-i\underline{u}'\underline{x}} J[h(\nu)\underline{u}]\ \phi_\nu(\underline{u})d\underline{u}$$

$$- \int_{-\infty}^{\infty} e^{-i\underline{u}'\underline{x}} J[h(\nu)\underline{u}]E[\phi_\nu(\underline{u})]d\underline{u}\Big\}\ dx_1$$

$$= \frac{1}{2\pi}\int_a^b \int_{-\infty}^{\infty} e^{-i\underline{u}'\underline{x}} J[h(\nu)\underline{u}]\ [\phi_\nu(\underline{u}) - \phi(\underline{u})]d\underline{u}\ dx_1 \ .$$

4. Since $\int_a^b f_\nu(x_1,\dots,x_k)dx_1$ and $E[\int_a^b f_\nu(x_1,\dots,x_k)dx_1]$ are continuous and bounded; i.e., $F(x_1,\dots,x_k)$, we have that

$$\max_{(x_2,\dots,x_k)\varepsilon E_{k-1}} \Big| \int_a^b f_\nu(x_1,\dots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\dots,x_k)dx_1]\Big| \le \frac{|b-a|}{2\pi}\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|\ |\phi_\nu(\underline{u}) - \phi(\underline{u})|d\underline{u} \ .$$

5. Applying the Cauchy-Schwarz inequality and the lemma we see that

$$\frac{|b-a|}{2\pi}\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|\ |\phi_\nu(\underline{u}) - \phi(\underline{u})|d\underline{u} \le \tag{2.6}$$
$$\frac{b-a}{2\pi}\left\{\int_{-\infty}^{\infty}\Big[|J[h(\nu)\underline{u}]|^{\frac{1}{2}}\Big]^2 d\underline{u}\right\}^{\frac{1}{2}}\left\{\int_{-\infty}^{\infty}\Big[|J[h(\nu)\underline{u}]|^{\frac{1}{2}}\ |\phi_\nu(\underline{u})-\phi(\underline{u})|\Big]^2 d\underline{u}\right\}^{\frac{1}{2}} .$$

6. Squaring and taking the expected value of (2.6) we see that

$$E\left\{\max_{(x_2,\dots,x_k)\varepsilon E_k} \Big|\int_a^b f_\nu(x_1,\dots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\dots,x_k)dx_1]\Big|^2\right\} \le$$
$$\frac{(b-a)^2}{4\pi^2} E\left\{\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|d\underline{u} \int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|\ |\phi_\nu(\underline{u}) - \phi(\underline{u})|^2 d\underline{u}\right\} . \tag{2.7}$$

7) Examining the dominating term of (2.7) we see that

$$\frac{(b-a)^2}{4\pi^2} E\{\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|d\underline{u} \int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]| \; |\phi_\nu(\underline{u}) - \phi(\underline{u})|^2 d\underline{u}\} =$$

$$\frac{(b-a)^2}{4\pi^2} \int_{-\infty}^{\infty}|J[h(\nu)\underline{u}|d\underline{u} \int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|E\{|\phi_\nu(\underline{u}) - \phi(\underline{u})|^2\}d\underline{u}$$

Recalling equation (2.5a)

$$E\{|\phi_\nu(\underline{u}) - E[\phi_\nu(\underline{u})]|^2\} = \frac{1}{\nu}[1 - |\phi(\underline{u})|^2] \quad ,$$

we note that

$$E\{|\phi_\nu(\underline{u}) - \phi(\underline{u})|^2\} \leq \frac{1}{\nu}$$

and, therefore

$$\frac{(b-a)^2}{4\pi^2} \int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|d\underline{u} \int_{-\infty}^{\infty}|J[h(\nu)\underline{u}]|E\{|\phi_\nu(\underline{u})-\phi(\underline{u})|^2\}d\underline{u}$$
$$\leq \frac{(b-a)^2}{4\pi^2\nu} \left[\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]|d\underline{u}\right]^2 .$$

8) Letting $Z_\ell = h(\nu)u_\ell$ $\ell = 1, \ldots, k$, we have

$$h(\nu)\underline{u} = \underline{Z} ,$$

$$\underline{u} = \frac{\underline{Z}}{h(\nu)} \quad , \text{ and}$$

it follows that

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}|d\underline{u} = \frac{1}{h^k(\nu)} \int_{-\infty}^{\infty} |J[\underline{Z}]|d\underline{Z} .$$

9) Since by property c(ii) of theorem 1 we have

$$\lim_{\nu\to\infty} \nu h^{2k}(\nu) = \infty \quad \text{and therefore ,}$$

$$\lim_{\nu\to\infty} \frac{(b-a)^2}{4\pi^2\nu h^{2k}(\nu)} \left[\int_{-\infty}^{\infty} |J[\underline{Z}]|d\underline{Z}\right]^2 = 0 \quad .$$

10) Since this term dominates the left-hand side of (2.7), we have

$$\lim_{\nu \to \infty} E\left\{ \max_{(x_2,\ldots,x_k)\varepsilon E_{k-1}} \left|\int_a^b f_\nu(x_1,\ldots,x_k)dx_1 - E[\int_a^b f_\nu(x_1,\ldots,x_k)dx_1]\right|^2 \right\} = 0.$$

Therefore by the results in steps (2) and (10) of this proof we have that

$$\lim_{\nu \to \infty} E\left\{ \max_{(x_2,\ldots,x_k)\varepsilon E_{k-1}} \left|\int_a^b f_\nu(x_1,\ldots,x_k)dx_1 - \int_a^b f(x_1,\ldots,x_k)dx_1\right|^2 \right\} = 0; \text{ i.e.,}$$

we have mean square convergence or consistency in quadratic mean.

Since mean square convergence implies convergence in probability we have the following corollary.

Corollary:

If $f(x_1,\ldots,x_k)$ is uniformly continuous, $f_\nu(x_1,\ldots,x_k)$ satisfies all the conditions of theorem 1, and

$$\int_{-\infty}^{\infty} |J[h(\nu)\underline{u}]d\underline{u} < \infty ,$$

$\int_a^b f_\nu(x_1, \ldots, x_k)dx_1$ converges in probability to $\int_a^b f(x_1, \ldots, x_k)dx_1$.

2.4 The Empirical Bayes Decision Rule Estimation

With the estimation of $f_G(\bar{x},s^2)$ and $\int_a^b f_G(x,\bar{x},s^2)dx$ established, we seek to estimate the Bayes decision rule as defined by equation (1.9). Our estimate of the Bayes decision rule (i.e., the empirical Bayes decision rule) is defined by

$$\begin{aligned} t_\nu(\bar{x},s^2) &= 1, \text{ for } T_\nu(\bar{x},s^2) \geq 0 \\ t_\nu(\bar{x},s^2) &= 0, \quad \text{otherwise,} \end{aligned} \tag{2.1}$$

where the empirical Bayes decision function is

$$T_\nu(\bar{x},s^2) = m(c_0+c_1)\,\frac{\int_a^b f_\nu(x,\bar{x},s^2)dx}{f_\nu(\bar{x},s^2)} - mc_1 \quad ,$$

and, $\int_a^b f_\nu(x,\bar{x},s^2)dx$ and $f_\nu(\bar{x},s^2)$ are defined in sections 2.3 and 2.2. Since $\int_a^b f_\nu(x,\bar{x},s^2)dx$ and $f_\nu(\bar{x},s^2)$ converge in probability to $\int_a^b f_G(x,\bar{x},s^2)dx$ and $f_G(\bar{x},s^2)$, respectively, we have that $T_\nu(\bar{x},s^2)$ converges in probability to $T_G(\bar{x},s^2)$ and, consequently, $t_\nu(\bar{x},s^2)$ converges in probability to $t_G(\bar{x},s^2)$. Hence we have established that our empirical Bayes decision rule converges in probability to the Bayes decision rule.

2.5 The Empirical Bayes Risk

The next question that arises involves the convergence of the empirical Bayes risk to the Bayes risk. The Bayes risk is defined by equation (1.1). The empirical Bayes risk $R[t_\nu,G]$ is as follows

$$R[t_\nu,G] = E_G[L(a_0,\mu,\sigma)] - E^*\{E_\nu[t_\nu(\bar{x},s^2)] \times E_G\left[L[a_0,\mu,\sigma] - L[a_1,\mu,\sigma]\,|\bar{x},s^2\right]\}$$

where E^* denotes expectation with respect to $(\bar{x},s^2)$ and E_ν denotes expectation with respect to $(\bar{X}_1,S^2{}_1)$, ..., $(\bar{X}_\nu,S^2{}_\nu)$ - the past data. The empirical Bayes decision rule is considered asymptotically optimal if

$$\lim_{\nu\to\infty} R[t_\nu,G] = R[t,G], \text{ i.e., if it exhibits}$$

risk convergence. This is a desirable property for an empirical Bayes rule, because it says that as the past data increases, the empirical Bayes rule becomes as good as the Bayes rule.

To prove that our empirical Bayes rule (2.1) is asymptotically optimal we rely on a theorem proved by Samuel (1963). Samuel shows that if $T_\nu(\bar{x},s^2)$ converges in probability to $E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)|\bar{x},s^2]$, then

$$\lim_{\nu \to \infty} R[t_\nu, G] = R[t, G] \ .$$

Recalling our derivation of the Bayes rule (see Chapter I eq. 1.9) we see that $T_G(\bar{x}, s^2) = E_G[L(a_0, \mu, \sigma) - L(a_1, \mu, \sigma) | \bar{x}, s^2]$. Since we established in section 2.4 that $T_\nu(\bar{x}, s^2)$ converges in probability to $T_G(\bar{x}, s^2)$, we have that our empirical Bayes rule is asymptotically optimal.

CHAPTER III

THE BAYESIAN DECISION RULE WHEN μ AND σ HAVE A CONJUGATE PRIOR DISTRIBUTION

Consider the case where μ and σ have a conjugate prior distribution; i.e., μ given σ^2 has a normal distribution with mean α and variance σ^2/β^2, while σ^{-2} has a gamma distribution with parameters $\frac{\gamma\delta}{2}$ and $\frac{\delta}{2}$. The general theory of a conjugate prior is presented in Raiffa and Schlaifer (1961). Expressed in functional form

$$dG(\mu,\sigma^2) = g(\mu,\sigma^{-2} \mid \alpha,\beta,\gamma,\delta)\ d\mu\ d\sigma^{-2} \quad ,$$

where

$$g(\mu,\sigma^{-2}|\alpha,\beta,\gamma,\delta) = \frac{\beta}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{\beta^2}{2\sigma^2}(\mu-\alpha)^2\right\} \times \frac{(\gamma\delta)^{\frac{\delta}{2}}(\sigma^{-2})^{\frac{\delta}{2}-1}}{2^{\frac{\delta}{2}}\ \Gamma(\frac{\delta}{2})} \exp\left\{-\frac{\gamma\delta\sigma^{-2}}{2}\right\}. \tag{3.1}$$

Recalling the general expression for the Bayes decision rule we see that it is

$$\begin{aligned} t_G(\bar{X},S^2) &= 1\text{, for } \int_a^b f_G(X,\bar{X},S^2)dX > \frac{c_1}{(c_0+c_1)} f_G(\bar{X},S^2) \\ t_G(\bar{X},S^2) &= 0\text{,} \qquad\qquad \text{otherwise.} \end{aligned} \tag{1.9}$$

Hence, if we are to determine the Bayes decision rule, we must first determine $f_G(X,\bar{X},S^2)$ and $f_G(\bar{X},S^2)$. Recalling equations (1.7) and (1.8) we see that

$$f_G(X,\bar{X},S^2) = \iint_\Omega f(X|\mu,\sigma)f(\bar{X},S^2|\mu,\sigma)\ dG(\mu,\sigma)$$

$$f_G(\bar{X},S^2) = \iint_\Omega f(\bar{X},S^2|\mu,\sigma)dG(\mu,\sigma) \quad .$$

When $G(\mu,\sigma^2)$ is the conjugate prior we see that

$$f_G(X,\bar{X},S^2) = \int_0^\infty\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left[\frac{X-\mu}{\sigma}\right]^2\right\} \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left[\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right]^2\right\}$$

$$\times \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}(\sigma^{-2})^{\frac{1}{2}(n-1)}}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \exp\left\{-\frac{(n-1)S^2}{2\sigma^2}\right\} \frac{\beta}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\beta^2}{2\sigma^2}(\mu-\alpha)^2\right\} \quad (3.2)$$

$$\times \frac{(\gamma\delta)^{\frac{\delta}{2}}(\sigma^{-2})^{\frac{\delta}{2}-1}}{2^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})} \exp\left\{-\frac{\gamma\delta\sigma^{-2}}{2}\right\} d\mu d\sigma^{-2} \quad .$$

Proceeding to evaluate this integral, we collect exponential terms involving X and $\bar{X}$. These are

$$\left(\frac{X-\mu}{\sigma}\right)^2 + n\left(\frac{\bar{X}-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\{X^2 - 2\mu X + \mu^2 + n\bar{X}^2 - 2n\mu\bar{X} + n\mu^2\}\text{ , or}$$

$$= \frac{1}{\sigma^2}\{X^2 + n\bar{X}^2 - 2\mu(X+n\bar{X}) + \mu^2(n+1)\} \quad .$$

Adding

$$\frac{1}{\sigma^2}\left[\frac{X^2 + 2nX\bar{X} + n^2\bar{X}^2}{n+1} - \frac{X^2 + 2nX\bar{X} + n^2\bar{X}^2}{n+1}\right] = 0 \text{ , we have}$$

that

$$\left(\frac{X-\mu}{\sigma}\right)^2 + n\left(\frac{\bar{X}-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\left[-\frac{X^2 + 2nX\bar{X} + n^2X^2}{n+1} + X^2 + n\bar{X}^2\right]$$

$$+ \frac{1}{\sigma^2}\left[\frac{X^2 + 2n\bar{X}X + n^2\bar{X}^2}{n+1} - 2\mu(X+n\bar{X}) + \mu^2(n+1)\right] \text{ ,}$$

$$= \frac{1}{\sigma^2}\left[\frac{(n+1)(X^2+n\bar{X}^2) - X^2 - 2nX\bar{X} - n^2\bar{X}^2}{n+1}\right]$$

$$+ \frac{1}{\sigma^2}\left[\left(\frac{X+n\bar{X}}{\sqrt{n+1}}\right)^2 - 2\mu\sqrt{n+1}\left(\frac{X+n\bar{X}}{\sqrt{n+1}}\right) + (\mu\sqrt{n+1})^2\right] \text{ ,}$$

$$= \frac{1}{\sigma^2}\left[\frac{nX^2 - 2nX\bar{X} + n\bar{X}^2}{n+1}\right] + \frac{1}{\sigma^2}\left[\frac{X+n\bar{X}}{\sqrt{n+1}} - \mu\sqrt{n+1}\right]^2 \text{ ,}$$

$$= \frac{n}{\sigma^2} \frac{[X-\bar{X}]^2}{n+1} + \frac{1}{\sigma^2}\left[\sqrt{n+1}\left(\frac{X+n\bar{X}}{n+1} - \mu\right)\right]^2, \text{ or}$$

finally

$$= \frac{n}{\sigma^2} \frac{(X-\bar{X})^2}{n+1} + \frac{n+1}{\sigma^2}\left[\frac{X+n\bar{X}}{n+1} - \mu\right]^2 .$$

Let

$$u = \frac{X+n\bar{X}}{n+1} \text{ , then (3.2) becomes}$$

$$f_G(X,\bar{X},S^2) = \frac{\sqrt{n}}{2\pi\sigma^2} \int_0^\infty \int_{-\infty}^\infty \exp\left\{-\frac{n+1}{2\sigma^2}(u-\mu)^2\right\} \exp\left\{-\frac{n}{2\sigma^2}\frac{(X-\bar{X})^2}{n+1}\right\}$$

$$\times \frac{[(n-1)S^2]^{\frac{n-1}{2}-1} (\sigma^{-2})^{\frac{1}{2}(n-1)}}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} \exp\left\{\frac{-(n-1)S^2}{2\sigma^2}\right\}$$

$$\times \frac{\beta}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{\beta^2}{2\sigma^2}(\mu-\alpha)^2\right\} \quad \frac{(\gamma\delta)^{\frac{\delta}{2}} (\sigma^{-2})^{\frac{\delta}{2}-1}}{2^{\frac{\delta}{2}} \Gamma(\frac{\delta}{2})} \exp -\left\{\frac{\gamma\delta\sigma^{-2}}{2}\right\} d\mu d\sigma^{-2} .$$

Collecting exponential terms involving σ^{-2} not involving μ we have

$$\exp\left\{-\frac{1}{2\sigma^2}\left[\frac{n}{n+1}(X-\bar{X})^2 + (n-1)S^2 + \gamma\delta\right]\right\} .$$

Let

$$\upsilon = \frac{n}{n+1}(X-\bar{X})^2 + (n-1)S^2 + \gamma\delta$$

then

$$f_G(X,\bar{X},S^2) = \frac{[(n-1)S^2]^{\frac{n-1}{2}-1} (\gamma\delta)^{\frac{\delta}{2}} \sqrt{n}}{2^{\frac{n+\delta+1}{2}} \Gamma(\frac{n-1}{2}) \Gamma(\frac{\delta}{2}) \pi} \int_0^\infty \int_{-\infty}^\infty \exp\left\{-\frac{n+1}{2\sigma^2}(u-\mu)^2\right\}$$

$$\times \frac{\beta}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{\beta^2}{2\sigma^2}(\mu-\alpha)^2\right\} \times (\sigma^{-2})^{\frac{n-1+\delta}{2}} \exp\left\{-\frac{\upsilon\sigma^{-2}}{2}\right\} d\mu d\sigma^{-2} .$$

Collecting terms involving μ we have

$$(n+1)(u-\mu)^2 + \beta^2(\mu-\alpha)^2 = (n+1)\ [u^2 - 2\mu u + \mu^2] + \beta^2(\mu^2 - 2\alpha\mu + \alpha^2)$$

and completing the square, we have that

$$(n+1)(u-\mu)^2 + \beta^2(\alpha-\mu)^2 = (n+1+\beta^2)\left[\mu - \frac{(n+1)u + \beta^2\alpha}{n+1+\beta^2}\right]^2 + \frac{\beta^2(n+1)\ [u-\alpha]^2}{n+1+\beta^2} \quad .$$

Let $\Theta_1 = \sqrt{n+1+\beta^2} \qquad \Theta_3 = \dfrac{(n+1)u + \beta^2\alpha}{n+1+\beta^2}$, and

$\Theta_2 = \dfrac{\beta\sqrt{n+1}}{\sqrt{n+1+\beta^2}}$, so that

$$(n+1)(u-\mu)^2 + \beta^2(\mu-\alpha)^2 = \Theta_1^{\,2}[\mu-\Theta_3]^2 + \Theta_2^{\,2}[u-\alpha]^2$$

and hence

$$f_G(X,\bar{X},S^2) = \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}(\gamma\delta)^{\frac{\delta}{2}}\sqrt{n}}{2^{\frac{n+\delta+1}{2}}\,\Gamma(\frac{n-1}{2})\,\Gamma(\frac{\delta}{2})\,\pi}$$

$$\times \int_0^\infty \int_{-\infty}^\infty \frac{\beta}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left[\Theta_1^{\,2}(\mu-\Theta_3)^2 + \Theta_2^{\,2}(u-\alpha)^2\right]\right\}$$

$$\times\ (\sigma^{-2})^{\frac{n+\delta-1}{2}} \exp\left\{-\frac{\upsilon\sigma^{-2}}{2}\right\} d\mu\, d\sigma^{-2} \quad .$$

Letting $\Theta_4 = \upsilon + \Theta_2^{\,2}(u-\alpha)^2$,

the expression for (3.2) becomes

$$f_G(X,\bar{X},S^2) = \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}(\gamma\delta)^{\frac{\delta}{2}}\sqrt{n}\,\beta}{2^{\frac{n+\delta+1}{2}}\,\Gamma(\frac{n-1}{2})\,\Gamma(\frac{\delta}{2})\,\pi}$$

$$\times \int_0^\infty (\sigma^{-2})^{\frac{n+\delta+1}{2}-1} \exp\left\{-\frac{\Theta_4\sigma^{-2}}{2}\right\}\frac{1}{\sqrt{2\pi}\sigma}$$

$$\times \int_{-\infty}^\infty \exp\left\{-\frac{1}{2}\frac{\Theta_1^{\,2}}{\sigma^2}(\mu-\Theta_3)^2\right\} d\mu\, d\sigma^{-2} \quad , \text{ or}$$

$$f_G(X,\bar{X},S^2) = \frac{\sqrt{n}\ \Gamma\left(\frac{n+\delta+1}{2}\right)}{\pi\ \Gamma\left(\frac{n-1}{2}\right)\ \Gamma\left(\frac{\delta}{2}\right)} \times \frac{\beta(\gamma\delta)^{\frac{\delta}{2}}}{\Theta_1} \times \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}}{\Theta_4^{\frac{n+\delta+1}{2}}} \quad . \tag{3.3}$$

Similarly, we evaluate $f_G(\bar{X},S^2)$

$$f_G(\bar{X},S^2) = \int_0^\infty \int_{-\infty}^\infty \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{n}{2\sigma^2}\left[\frac{\bar{X}-\mu}{\sigma}\right]^2\right\}$$

$$\times \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}(\sigma^{-2})^{\frac{1}{2}(n-1)}}{2^{\frac{n-1}{2}}\ \Gamma\left(\frac{n-1}{2}\right)} \exp\left\{-\frac{(n-1)S^2}{2\sigma^2}\right\} \tag{3.4}$$

$$\times \frac{\beta}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{\beta^2}{2\sigma^2}(\mu-\alpha)^2\right\} \times \frac{(\gamma\delta)^{\frac{\delta}{2}}(\sigma^{-2})^{\frac{\delta}{2}-1}}{2^{\frac{\delta}{2}}\ \Gamma\left(\frac{\delta}{2}\right)} \exp\left\{-\frac{\gamma\delta\sigma^{-2}}{2}\right\} d\mu\, d\sigma^{-2}$$

Collecting terms involving μ, we see that

$$n(\bar{X}-\mu)^2 + \beta^2(\mu-\alpha)^2 = n\bar{X}^2 - 2n\bar{X}\mu + n\mu^2 + \beta^2\mu^2 - 2\beta^2\alpha\mu + \beta^2\alpha^2$$

and completing the square, we have

$$n(\bar{X}-\mu)^2 + \beta^2(\mu-\alpha)^2 = (n+\beta^2)\left[\frac{n\bar{X} + \beta^2\alpha}{\sqrt{\beta^2 + n}} - \mu\right]^2 + \frac{n\beta^2}{\beta^2 + n}[\bar{X} - \alpha]^2$$

so that (3.4) becomes

$$f_G(\bar{X},S^2) = \int_0^\infty \int_{-\infty}^\infty \frac{\beta\sqrt{n}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{n+\beta^2}{2\sigma^2}\left[\frac{n\bar{X}+\beta^2\alpha}{\beta^2+n} - \mu\right]^2\right\}$$

$$\times \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}(\sigma^{-2})^{\frac{1}{2}(n-1)}}{\sqrt{2\pi}\ 2^{\frac{n-1}{2}}\ \Gamma\left(\frac{n-1}{2}\right)} \times \frac{(\gamma\delta)^{\frac{\delta}{2}}}{2^{\frac{\delta}{2}}\ \Gamma\left(\frac{\delta}{2}\right)}$$

$$\times (\sigma^{-2})^{\frac{\delta}{2}-1} \exp\left\{-\frac{\sigma^{-2}}{2}\left[(n-1)S^2 + \gamma\delta + \frac{n\beta^2}{\beta^2+n}(\bar{X}-\alpha)^2\right]\right\} d\mu\, d\sigma^{-2} \quad .$$

Let $\Theta_5 = \sqrt{n + \beta^2}$,

$$\Theta_6 = \frac{n\bar{X} + \beta^2\alpha}{\beta^2 + n} ,$$

and $\Theta_7 = (n-1)S^2 + \gamma\delta + \frac{n\beta^2}{\beta^2 + n} (\bar{X}-\alpha)^2$,

then
$$f_G(\bar{X},S^2) = \sqrt{\frac{n}{\pi}} \cdot \frac{[(n-1)S^2]^{\frac{n-1}{2}-1} (\gamma\delta)^{\frac{\delta}{2}}\beta}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{\delta}{2}\right) 2^{\frac{n+\delta}{2}}}$$

$$\times \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{- \frac{\Theta_5^2}{\sigma^2} (\Theta_6 - \mu)\right\}^2$$

$$\times (\sigma^{-2})^{\frac{n+\delta}{2}-1} \exp\left\{- \frac{\Theta_7\sigma^{-2}}{2}\right\} d\mu d\sigma^{-2}$$

$$f_G(\bar{X},S^2) = \sqrt{\frac{n}{\pi}} \cdot \frac{[(n-1)S^2]^{\frac{n-1}{2}-1} (\gamma\delta)^{\frac{\delta}{2}}\beta}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{\delta}{2}\right) 2^{\frac{n+\delta}{2}}}$$

$$\times \int_0^\infty (\sigma^{-2})^{\frac{n+\delta}{2}-1} \exp\left\{- \frac{\Theta_7\sigma^{-2}}{2}\right\}$$

$$\times \int_{-\infty}^\infty \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{- \frac{\Theta_5^2}{\sigma^2} (\Theta_6 - \mu)^2\right\} d\mu d\sigma^{-2}$$

$$f_G(\bar{X},S^2) = \sqrt{\frac{n}{\pi}} \frac{[(n-1)S^2]^{\frac{n-1}{2}-1} (\gamma\delta)^{\frac{\delta}{2}} \Gamma\left(\frac{n+\delta}{2}\right)\beta}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{\delta}{2}\right) \Theta_5 \ \Theta_7^{\frac{n+\delta}{2}}} \tag{3.5}$$

Returning to the decision rule (1.9), we see that we must evaluate one more integral before we have an explicit expression for the decision rule; i.e., we must evaluate $\int_a^b f_G(X,\bar{X},S^2)dX$.

$$\int_a^b f_G(X,\bar{X},S^2)dX = \frac{n}{2\pi^2} \frac{\Gamma\left(\frac{n+\delta+1}{2}\right) \beta(\gamma\delta)^{\frac{\delta}{2}} [(n-1)S^2]^{\frac{n-1}{2}-1}}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{\delta}{2}\right) \Theta_1} \tag{3.6}$$

$$\times \int_a^b \Theta_4^{-\frac{n+\delta+1}{2}} dX ,$$

where $$\Theta_4 = \frac{n}{n+1}(X-\bar{X})^2 + (n-1)S^2 + \gamma\delta + \frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{X + n\bar{X}}{n + 1} - \alpha\right]^2 ,$$

or $$\Theta_4 = X^2\left[\frac{n}{n+1} + \frac{\beta}{(n+1+\beta^2)(n+1)}\right] - 2X\left[\frac{n\bar{X}}{n+1} - \frac{n\bar{X}\beta}{(n+1+\beta^2)(n+1)} + \frac{\alpha\beta}{n+1+\beta^2}\right]$$

$$+ \frac{n\bar{X}^2}{n+1} + (n-1)S^2 + \gamma\delta + \frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{n^2\bar{X}^2}{(n+1)^2} - \frac{2\alpha n\bar{X}}{n+1} + \alpha^2\right] .$$

Let

$$W_1 = \frac{n}{n+1} + \frac{\beta}{(n+1+\beta^2)\ (n+1)}$$

$$W_2 = -2\left[\frac{n\bar{X}}{n+1} + \frac{n\bar{X}\beta}{(n+1+\beta^2)\ (n+1)} - \frac{\alpha\beta}{n+1+\beta^2}\right]$$

$$W_3 = \frac{n\bar{X}^2}{n+1} + (n-1)S^2 + \gamma\delta + \frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{n^2\bar{X}^2}{(n+1)^2} - \frac{2\alpha n\bar{X}}{n+1} + \alpha^2\right]$$

$$\Theta_4(X) = W_1\ X^2 + W_2\ X + W_3$$

Next we seek to evaluate the integral $\int_a^b \Theta_4^{\frac{-n+\delta+1}{2}} dX$ for the following two cases. For the first case, δ is an integer and $n + \delta$ is an even integer, and we have

$$\int_a^b \Theta_4(X)^{\frac{-n+\delta+1}{2}} dX = \int_a^b \frac{dX}{\sqrt{[W_1X^2 + W_2X + W_3]^{n+\delta+1}}} .$$

Using the integral tables by Ryzhik (1965), formula 3, page 83, we have

$$\int_a^b \Theta_4(X)^{\frac{-n+\delta+1}{2}} dX = \frac{2[2W_1X + W_2]\ \Theta_4(X)^{\frac{-n+\delta+1}{2}}}{(n+\delta-1)[4W_1W_3 - W_2^2]} \times$$

$$\left\{1 + \sum_{k=1}^{\frac{1}{2}(n+\delta)-1} \frac{8^k[\frac{1}{2}(n+\delta)-1][\frac{1}{2}(n+\delta)-2]\ \cdots\ [\frac{1}{2}(n+\delta)-k]\ W_1^k\ \Theta_4^k(X)}{(n+\delta-3)\ (n+\delta-5)\ \cdots\ (n+\delta-2k-1)[4W_1W_3 - W_2^2]^k}\right\}\Bigg|_a^b .$$

For the second case, δ is not an integer and/or $n + \delta$ is not an even number. In this case, we may not be able to evaluate the integral explicitly. However, since the integral is definite, we can approximate the value of the integral by any one of several numerical integration schemes.

Since $f_G(x,\bar{x},s^2)$ is continuous and finite, either Simpson's rule or Gaussian 24 point quadrature could be used, for example. Hence, we should always be able to determine the Bayes dicision rule when we have a conjugate prior distribution.

CHAPTER IV

EVALUATION OF THE BAYES RISK FUNCTION FOR THE CASE OF A KNOWN CONJUGATE PRIOR DISTRIBUTION

Recalling the expression for the Bayes risk function, equation 1.1, we seek to determine the Bayes risk when the optimal decision rule, equation 1.9, is used, given a known conjugate prior distribution.

Let

$$\mathfrak{X}_0 = \{\bar{X},S^2 | t_G(\bar{X},S^2) = 1\} \text{ , or equivalently}$$

$$\mathfrak{X}_0 = \{\bar{X},S^2 | \int_a^b f_G(X,\bar{X},S^2)dX \geq \frac{c_1}{c_0+c_1} f_G(\bar{X},S^2)\} \text{ ,} \tag{4.1}$$

and

$$\mathfrak{X}_1 = \{\bar{X},S^2 | t_G(\bar{X},S^2) = 0\} \text{ , or equivalently}$$

$$\mathfrak{X}_1 = \{\bar{X},S^2 | \int_a^b f_G(X,\bar{X},S^2)dX < \frac{c_1}{c_0+c_1} f_G(\bar{X},S^2)\} \text{ .}$$

The Bayes risk function becomes

$$R[t,G] = E\left\{L(a_0,\mu,\sigma) - E\left[E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma) | (\bar{X},S^2)\varepsilon\mathfrak{X}_0]\right]\right\} . \tag{4.2}$$

Since $E\left[E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma) | (\bar{X},S^2)\varepsilon\mathfrak{X}_0]\right]$ is a constant with respect to $G(\mu,\sigma^2)$, equation (4.2) becomes

$$R[t,G] = E[L(a_0,\mu,\sigma)] - E\{E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma) | (\bar{X},S^2)\varepsilon\mathfrak{X}_0]\} .$$

Examining the first term of the left hand side of equation (4.2), we see that

$$E[L(a_0,\mu,\sigma)] = \int_{-\infty}^{\infty} \int_0^{\infty} \int_a^b f(x|\mu,\sigma)dxg(\mu,\sigma|\alpha,\beta,\gamma,\delta)d\sigma^{-2}d\mu$$

$$= \int_a^b \int_{-\infty}^{\infty} \int_0^{\infty} f(x|\mu,\sigma)g(\mu,\sigma|\alpha,\beta,\gamma,\delta)d\sigma^{-2}d\mu dx$$

$$= \frac{(\gamma\delta)^{\frac{\delta}{2}}\,\Gamma\left(\frac{\delta+1}{2}\right)}{\Gamma\left(\frac{\delta}{2}\right)\sqrt{\pi(1+\beta^2}} \times \int_a^b \left[\frac{\beta^2(x-\alpha)^2}{1+\beta^2} + \gamma\delta\right]^{-\frac{\delta+1}{2}} dx \quad .$$

Considering the second term on the right hand side of equation (4.2), we have, recalling equation (1.3),

$$E\{E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)|(\bar{X},S^2)\ \varepsilon\mathfrak{X}_0]\} =$$

$$\iint_{\mathfrak{X}_0} \left\{ m(c_0+c_1) \int_{\infty}^{\infty} \int_0^{\infty} \int_a^b \frac{f(x|\mu,\sigma)dxf(\bar{X},S^2|\mu,\sigma)g(\mu,\sigma|\alpha,\beta,\gamma,\delta)d\sigma^{-2}d\mu}{\int_{\infty}^{\infty}\int_0^{\infty} f(\bar{X},S^2|\mu,\sigma)g(\mu,\sigma|\alpha,\beta,\gamma,\delta)d\sigma^{-2}d\mu} - mc_1 \right\} \times f_G(\bar{X},S^2)dS^2d\bar{X} \quad .$$

Changing the order of integration, we have

$$E\{E_G[L(a_0,\mu,\sigma) - L(a_1,\mu,\sigma)|(\bar{X},S^2)\ \varepsilon\mathfrak{X}_0]\} =$$

$$\iint_{\mathfrak{X}_0} \left\{ m(c_0+c_1) \int_a^b \int_{-\infty}^{\infty} \int_0^{\infty} \frac{f(x|\mu,\sigma)f(\bar{X},S^2|\mu,\sigma)g(\mu,\sigma|\alpha,\beta,\gamma,\delta)d\sigma^{-2}d\mu dx}{f_G(\bar{X},S^2)} - mc_1 \right\} \times f_G(\bar{X},S^2)dS^2d\bar{X}$$

$$= m(c_0+c_1) \iint_{\mathfrak{X}_0}[f_G(X,\bar{X},S^2)dx]dS^2d\bar{X} - mc_1 \iint_{\mathfrak{X}_0} f_G(\bar{X},S^2)dS^2d\bar{X} \quad . \quad (4.3)$$

To evaluate equation (4.3), we see that we must integrate over the region $\mathfrak{X}_0$. Re-examining the region $\mathfrak{X}_0$, we see that from equation (4.1)

$$\mathfrak{X}_0 = \{\bar{X},S^2 \mid \int_a^b f_G(X,\bar{X},S^2)dx \geq \frac{c_1}{c_0+c_1} f_G(\bar{X},S^2)\} \quad . \qquad (4.4)$$

Whenever the prior distribution is a conjugate prior distribution, the inequality in equation (4.4) must be considered for the two cases considered in Chapter 3. For the first case where δ is an integer and $n + \delta$ is an even integer, the inequality is

$$\sqrt{\frac{n}{2\pi^2}}\,\frac{\Gamma\left(\frac{n+\ +1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{\delta}{2}\right)}\,\frac{\beta(\gamma\delta)^{\frac{\delta}{2}}\,[(n-1)S^2]^{\frac{n-1}{2}-1}}{\Theta_1}\times\left\{\frac{2[2W_1b+W_2]\quad\Theta_4(b)^{\frac{-n+\delta-1}{2}}}{(n+\delta-1)\ [4W_1W_3-W_2{}^2]}\times\right.$$

$$\left[1+\sum_{k=1}^{\frac{1}{2}(n+\delta)-1}\frac{8^k[\frac{1}{2}(n+\delta)-1][\frac{1}{2}(n-\delta)-2]\cdots[\frac{1}{2}(n-\delta)-k]\ W_1{}^k\ \Theta_4{}^k(b)}{(n+\delta-3)(n+\delta-5)\cdots(n+\delta-2k-1)[4W_1W_3-W_2{}^2]^k}\right]-$$

$$\frac{2[2W_1a+W_2]\ \Theta_4(a)^{\frac{-n+\delta-1}{2}}}{(n+\delta-1)[4W_1W_3-W_2{}^2]}\times$$

$$\left.\left[1+\sum_{k=1}^{(n+\delta)-1}\frac{8^k[\frac{1}{2}(n+\delta)-1][\frac{1}{2}(n+\delta)-2]\cdots[\frac{1}{2}(n+\delta)-k]\ W_1{}^k\ \Theta_4{}^k(a)}{(n+\delta-3)(n+\delta-5)\cdots(n+\delta-2k-1)[4W_1W_3-W_2{}^2]^k}\right]\right\}$$

$$\geq\ \sqrt{\frac{n}{2\pi}}\ \frac{[(n-1)S^2]^{\frac{n-1}{2}-1}\quad(\gamma\delta)^{\frac{\delta}{2}}\ \Gamma\left(\frac{n-1+\delta}{2}\right)\ \beta}{\Gamma\left(\frac{n-1}{2}\right)\quad\Gamma\left(\frac{\delta}{2}\right)\qquad\Theta_7{}^{\frac{n-1+\delta}{2}}}\cdot\frac{c_1}{c_0+c_1}\quad,$$

where

$$\Theta_1 = \sqrt{n+1+\beta^2}$$

$$\Theta_4(a) = \frac{n}{n+1}(a-\bar{X})^2+(n-1)S^2+\gamma\delta+\frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{a+n\bar{X}}{n+1}-\alpha\right]^2$$

$$\Theta_4(b) = \frac{n}{n+1}(b-\bar{X})^2+(n-1)S^2+\gamma\delta+\frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{b+n\bar{X}}{n+1}-\alpha\right]^2$$

$$\Theta_5 = \sqrt{n+\beta^2}$$

$$\Theta_7 = (n-1)S^2+\gamma\delta+\frac{n\beta^2}{\beta^2+n}(\bar{X}-\alpha)^2$$

$$W_1 = \frac{n}{n+1}+\frac{\beta}{(n+1+\beta^2)(n+1)}$$

$$W_2 = -2\left[\frac{n\bar{X}}{n+1}+\frac{n\bar{X}\beta}{(n+1+\beta^2)(n+1)}-\frac{\alpha\beta}{n+1+\beta^2}\right]$$

and $$W_3 = \frac{n\bar{X}^2}{n+1} + (n-1)s^2 + \gamma\delta + \frac{\beta(n+1)}{n+1+\beta^2}\left[\frac{n^2\bar{X}^2}{(n+1)^2} - \frac{2\alpha n\bar{X}}{n+1} + \alpha^2\right] .$$

Examining this inequality for this case, we see that no simple explicit relation between $\bar{X}$ and S^2 can be derived algebraically that will allow the integral in equation (4.3) to be evaluated.

For the second case where δ is not an integer and/or $n + \delta$ is not an even integer, the integral $\int_a^b \Theta_4(x)^{\frac{-n+\delta+1}{2}} dx$ must be determined by numerical integration. Here again the inequality can not be manipulated algebraically to determine a relation between $\bar{X}$ and S^2 so that the integral in equation (4.3) can be evaluated without the aid of a computer. For this reason we do not try to compare the risk of the empirical Bayes risk to the Bayes risk in our Monte Carlo study. Also, since the region of integration is a function of the costs c_0 and c_1, any numerical results are dependent on the costs and cannot be talked about in general without extensive tabulation.

Moreover, since we have shown in Chapter II that the convergence of the empirical Bayes rule to the Bayes rule implies risk convergence, it is sufficient to demonstrate only the convergence of the empirical Bayes rule.

CHAPTER V

MONTE CARLO SIMULATION

5.1 Introduction

In this chapter we present some examples to indicate how well the empirical Bayes decision rule approximates the Bayes decision rule as the amount of past data increases. Examining the Bayes decision rule (see equation 1.9), we see that it is a function of the costs $c_0 + c_1$ and the term

$$\frac{\int_a^b f_G(x,\bar{x},s^2)dx}{f_G(\bar{x},s^2)} \quad . \tag{5.1}$$

Since the term (5.1) is the only part of the rule that is actually estimated from the data, we restrict our study to the estimation of $f_G(\bar{x},s^2)$, $\int_a^b f_G(x,\bar{x},s^2)dx$, and $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ by the estimators described in Chapter II. For each of the examples presented we let G be a conjugate prior distribution as defined in Chapter III. In the next two sections we discuss the computer program used to generate the examples. In the fourth section we present and discuss the examples. The fifth section contains the concluding remarks of this paper.

5.2 Random Number Generation

To generate examples, we must simulate the Bayesian process as described in Chapters I and II and estimate the Bayes rule by the procedure developed in Chapter II. Hence the computer program must generate at random (μ,σ^{-2}) according to the conjugate prior distribution $G(\mu,\sigma^2)$; i.e., (μ,σ^{-2}) must be jointly distributed normal-gamma with parameters $(\alpha,\beta,\frac{\gamma\delta}{2},\frac{\delta}{2})$. Then the computer must generate observations $(X_1, \ldots, X_n, X_{n+1})$ according to a normal distribution with mean μ and variance σ^2 - calculating $\bar{x}$ and s^2 from the first n observations.

The computer program accomplishes the simulation by using first a subroutine that generates random numbers from a uniform probability distribution and then using another subroutine to transform (a function of several of these uniformly distributed random variables) to a normally distributed random number with mean 0 and variance 1. The random variable σ^{-2} is generated by first generating δ independent random variables, say $Z_1, \ldots, Z_\delta$, which are normally distributed with mean 0 and variance 1. Then by letting

$$\sigma^{-2} = \frac{1}{\gamma\delta} \sum_{j=1}^{\delta} Z_i^{\,2} \quad , \tag{5.2}$$

we have generated σ^{-2} with a gamma distribution with parameters $(\frac{\gamma\delta}{2},\frac{\delta}{2})$. That is, we have that

$$f(\sigma^{-2}) = \frac{(\gamma\delta)^{\frac{\delta}{2}} (\sigma^{-2})^{\frac{\delta}{2}-1}}{2^{\frac{\delta}{2}}\ \Gamma\left(\frac{\delta}{2}\right)} \times \exp\left\{-\frac{\gamma\delta\sigma^{-2}}{2}\right\} \quad . \tag{5.3}$$

The random number u is generated by generating a normally distributed random variable u, say, with mean 0 and variance 1. Then by letting

$$\mu = \alpha + u\frac{\sigma}{\beta} \quad , \tag{5.4}$$

we have that μ is distributed normal with mean α and variance σ^2/β^2. And finally the observations $X_1, \ldots, X_n, X_{n+1}$ are generated by generating n+1 independent random numbers, $\upsilon_1, \ldots, \upsilon_{n+1}$, say, normally distributed with means 0 and variances 1. Then by letting

$$X_i = \upsilon_i\sigma + \mu \quad i = 1, \ldots, n+1 \qquad (5.5)$$

we have that the X_i $(i = 1, \ldots, n+1)$ are normally distributed with mean μ and variance σ^2.

5.3 The Computer Simulation

To study the convergence of the empirical Bayes decision rule to the Bayes decision rule for the various examples, we performed the simulation in the following manner.

We read in the parameters of the conjugate prior along with the limits (a,b), the sample size n, and the integer constants setting the limits on the number of iterations in the simulation. The simulation then procedes by performing the following steps.

Step I: Generating the Bayes Decision Rule

We generate the random variables $(\mu, \sigma, X_1,\ldots,X_n)$ by the method described in section 5.2 and calculate $\bar{x}$ and s^2. Then we calculate $f_G(\bar{x},s^2)$, $\int_a^b f_G(x,\bar{x},s^2)dx$ as they are defined in Chapter III.

Step II: Generating the Past Data and Determining the Empirical Bayes Decision Rule

We generate the random variables $(\mu, \sigma, X_1,\ldots,X_n, X_{n+1})$ by the method described in section 5.2 - calculating x_i and s^2_i from the first n x_i's. We repeat the process 20 times where each repetition is independent and different from all others. At the first, fifth, tenth, fifteenth, and twentieth repetitions we calculate $f_\nu(\bar{x},s^2)$, $\int_a^b f_\nu(x,\bar{x},s^2)dx$ and

$\int_a^b f_\nu(x,\bar{x},s^2)dx/f_\nu(\bar{x},s^2)$ where $f_\nu(\bar{x},s^2)$ and $\int_a^b f_\nu(x,\bar{x},s^2)dx$ are defined in Chapter II. That is, we estimate $f_G(\bar{x},s^2)$ and $\int_a^b f_G(x,\bar{x},s^2)dx$ (which were calculated in Step I), with one past observation, five past observations, etc. Also we calculate

$$\Delta_{1\nu} = f_G(\bar{x},s^2) - f_\nu(\bar{x},s^2)$$

$$\Delta_{2\nu} = \int_a^b f_G(x,\bar{x},s^2)dx - \int_a^b f_\nu(x,\bar{x},s^2)dx$$

and

$$\Delta_{3\nu} = \frac{\int_a^b f_G(x,\bar{x},s^2)dx}{f_G(\bar{x},s^2)} - \frac{\int_a^b f_\nu(x,\bar{x},s^2)dx}{f_\nu(\bar{x},s^2)} \; .$$

Step III:

We perform 100 iterations of Step II accumulating the sum of the $\Delta_{i\nu}$'s and the sum of the $\Delta_{i\nu}^2$'s for i = 1, 2, 3 and ν = 1, 10, 15, 20. That is, we calculate $\sum_{j=1}^{100} \Delta_{i\nu j}$ and $\sum_{j=1}^{100} \Delta_{i\nu j}^2$ for i = 1, 2, 3 and ν = 1, 10, 15, 20 where $\Delta_{i\nu j}$ is the i^{th} difference for past observations at the j^{th} iteration. At the end of the 100^{th} iteration, we calculate

$$\bar{\Delta}_{i\nu} = \sum_{j=1}^{100} \Delta_{i\nu j}/100$$

$$s^2_{\Delta i\nu} = \sum_{j=1}^{100} \Delta^2_{i\nu j}/100$$

for i = 1, 2, 3 and ν = 1, 10, 15, 20, where

1) $\bar{\Delta}_{1\nu}$ and $s^2_{\Delta 1\nu}$ are the bias and mean square error, respectively, of $f_\nu(\bar{x},s^2)$,

2) $\bar{\Delta}_{2\nu}$ and $s^2_{\Delta 2\nu}$ are the bias and mean square error, respectively, of $\int_a^b f_\nu(x,\bar{x},s^2)dx$, and

3) $\bar{\Delta}_{3\nu}$ and $s^2_{\Delta 3\nu}$ are the bias and mean square error, respectively, of $\int_a^b f_\nu(x,\bar{x},s^2)dx/f_\nu(\bar{x},s^2)$.

Step IV

We perform 20 iterations of Steps I and III accumulating the sum of the mean square errors. At the end of the twentieth iteration we calculate the average of the mean square errors. That is, we have the average mean square errors in estimating $f_G(\bar{x},s^2)$, $\int_a^b f_G(x,\bar{x},s^2)dx$, and $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ for 1, 10, 15, and 20 past observations, when G is a given prior conjugate prior distribution.

5.4 Examples

In the following set of examples we present the average mean square error of the estimators of $f_G(\bar{x},s^2)$, $\int_a^b f_G(x,\bar{x},s^2)dx$, and $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ for several different conjugate prior distributions as the number of past observations increase. For our first set of examples we fix the values of all the other parameters and vary only the mean of the conjugate prior. In particular, we fix the parameters as follows

$$n = 5 \qquad \gamma = 1$$
$$\beta = 1 \qquad \delta = 2$$
$$a = -1 \qquad b = 1 \quad .$$

We then vary α from 0 to -2. This lets us study the estimation procedure as the mean of the prior varies from the center of the specification limits to a point outside the limits. Since the distributions are symmetric, we do not have to study values of α greater than zero. These examples are presented in Tables I through IV.

For our second set of examples, we consider the situation where we have one-sided specification limits; i.e., $a = -\infty$ or $b = +\infty$. Again because of the symmetry we need only to consider $a = -\infty$. This particular case is hard to justify theoretically; however, in practical situations our random variables have truncated distributions. This allows us to apply the mean value theorem for integrals which is necessary in the proof of mean square convergences (see Theorem 2, Chapter II). For this case we again fix

$$n = 5 \qquad \gamma = 1$$
$$\beta = 1 \qquad \delta = 2$$
$$b = 1,$$

and set a = -100 in the numerical integration subroutines. Setting a = -100 is a close enough approximation to $-\infty$. For this case we let α range from +2 to -2. The examples are presented in Tables V through VIII.

For each of the tables we use the same sequence of random numbers so that we can compare the tables. However this procedure has the defect that since we are only varying α we do not change the value of $f_G(\bar{x},s^2)$ or its estimate $f_\nu(\bar{x},s^2)$ at each point in the sequence of random numbers - even though the $\bar{x}$'s are shifted for different α's. Examining the expression for $f_G(\bar{x},s^2)$ (see equation 3.5) we see that $f_G(\bar{x},s^2)$ depends on α only through $(\bar{x}-\alpha)^2$. Recalling equations (5.4) and (5.5), we see that if we use the same sequence of random numbers for the cases $\alpha = \alpha'$ and $\alpha = \alpha^*$, then

$$\bar{X}_i' - \alpha_i' = \bar{X}_i^* - \alpha_i^* = \frac{\sigma}{n} \sum_{i=1}^{n} \upsilon_i - \frac{u\sigma}{n\beta} \; .$$

We can show that a similar situation exists for $f_\nu(\bar{x},s^2)$. Hence the mean square errors for the estimation of $f_G(\bar{x},s^2)$ does not change as we change α alone in the generation of our tables.

TABLE I $\alpha = 0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140128	0.000575120	0.11978479
$\nu = 10$	0.000387369	0.000237644	0.074191047
$\nu = 15$	0.000307876	0.000200854	0.069786057
$\nu = 20$	0.000262434	0.000180082	0.066457283

TABLE II $\alpha = -0.5$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140128	0.000580229	0.11432559
$\nu = 10$	0.000387369	0.000201704	0.074203577
$\nu = 15$	0.000307876	0.0001716646	0.070694492
$\nu = 20$	0.000262435	0.000154207	0.066117844

TABLE III $\alpha = -1.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140128	0.000359378	0.10312916
$\nu = 10$	0.000387689	0.000125544	0.075513902
$\nu = 15$	0.000307876	0.000104830	0.070527272
$\nu = 20$	0.000262435	0.000093696	0.064437037

TABLE IV $\alpha = -2.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140128	0.000140291	0.068035281
$\nu = 10$	0.000387369	0.000032564	0.058352461
$\nu = 15$	0.000307876	0.000024451	0.050630522
$\nu = 20$	0.000262435	0.000020658	0.047986813

TABLE V $\alpha = 2.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140127	0.000206711	0.14499654
$\nu = 10$	0.000387369	0.000483362	0.10796042
$\nu = 15$	0.000307876	0.000350282	0.098308038
$\nu = 20$	0.000262435	0.000029162	0.091224067

TABLE VI $\alpha = 1.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140127	0.000529031	0.191274494
$\nu = 10$	0.000387369	0.000161203	0.13069978
$\nu = 15$	0.000307876	0.000122711	0.11387686
$\nu = 20$	0.000262435	0.000104212	0.10458592

TABLE VII $\alpha = 0.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.001401277	0.001017486	0.18356804
$\nu = 10$	0.000387369	0.000339964	0.10138556
$\nu = 15$	0.000307876	0.000272740	0.085424967
$\nu = 20$	0.000262435	0.000236486	0.080254470

TABLE VIII $\alpha = -1.0$

Number Of Past Observations	Average Mean Square Error of Estimator for		
	$f_G(\bar{x},s^2)$	$\int_a^b f_G(x,\bar{x},s^2)dx$	$\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$
$\nu = 1$	0.00140127	0.001207624	0.15238552
$\nu = 10$	0.000387369	0.000340257	0.062985934
$\nu = 15$	0.000307876	0.0002646143	0.052881798
$\nu = 20$	0.000262435	0.000222983	0.050728625

5.5 Discussion of the Tables

Examining the numerical results presented in the tables we see that as the prior distribution changes the convergence rate for the estimates of $\int_a^b f_G(x,\bar{x},s^2)dx$ and $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ changes also. This is probably a good indication of what can happen in practice. Hence for $\nu < 20$, the empirical Bayes decision rule can be very poor for not too uncommon prior distributions, and, for certain pathological prior distributions, the empirical Bayes rule could be poor for $\nu > 20$.

However in all cases the estimator of $\int_a^b f_G(x,\bar{x},s^2)dx$ had a mean square error of relatively the same magnitude as the estimator of $f_G(\bar{x},s^2)$. In some cases (see Tables IV and V), the mean square error of the estimator of $\int_a^b f_G(x,\bar{x},s^2)dx$ is almost an order of magnitude smaller. We find this particularly encouraging as this is the first time (as far as we know) that any one has ever tried to estimate $\int_a^b f(x_1, \ldots, x_k)dx_1$. This may indicate that the slow convergence of the estimator of $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ may be due to problems in estimating $f_G(\bar{x},s^2)$.

Another possible explanation of the large mean square error encountered in estimating $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ may be due to our choice of the spreading coefficient $h(\nu)$ (see Chapter II, section 2). The results of Parzen (1962), Cacoullos (1966), and Epanechnikov (1969) indicate that we should let $h(\nu) = \nu^{-\frac{1}{6}}$ when we estimate $f_G(\bar{x},s^2)$ and let $h(\nu) = \nu^{-\frac{1}{7}}$, when we estimate $\int_a^b f_G(x,\bar{x},s^2)dx$. Whenever we tried this approach we found that we got estimates of $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ which sometimes were greater than one -- clearly an unacceptable result. Hence we chose to use $h(\nu) = \nu^{-\frac{1}{7}}$ in estimating both $\int_a^b f_G(x,\bar{x},s^2)dx$ and $f_G(\bar{x},s^2)$. This choice guarantees that $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$ will be less than one, because this causes the estimate of $\int_a^b f_G(x,\bar{x},s^2)dx$ to be a sum of terms each of which is less than or equal to its corresponding term in the estimate of $f_G(\bar{x},s^2)$. This choice may be an explanation of why the mean square error of the estimator for $f_G(\bar{x},s^2)$ was larger than the mean square error of the estimator for $\int_a^b f_G(x,\bar{x},s^2)dx$. Since the theory behind choosing the spreading coefficients was developed for density function estimation only, there may be better ways of choosing the spreading coefficient for this problem.

5.6 Conclusion

Although we prove in this paper that an empirical Bayesian approach to a variables sampling plan problem is possible, there are many unsolved problems left and possible improvements that can be made. For instance, when we estimate the Bayes decision rule $\int_a^b f_G(x,\bar{x},s^2)dx/f_G(\bar{x},s^2)$, we are in fact estimating $P[a<x<b|\bar{x},s^2]$ -- a conditional probability. To do this estimation we had to rely on the estimators of multivariate density functions. By concentrating a study on the problem of estimating conditional density functions, we may find better estimators of conditional probabilities. An unsolved problem can be found in determining how large a sample should be taken out of each lot to insure that the risk is less than some specified value. If the prior distribution is known, this would not be a problem. Another problem involves the case where, when each lot was sampled in the past, the same sample size was not drawn.

Finally we note that our results may be useful for other empirical Bayesian problems where a similar loss function is involved. That is, the loss function is a function of the probability that the observed random variable (or some known function of the observed random variable) will be within certain specified limits.

LIST OF REFERENCES

Cacoullos, T. "Estimation of a Multivariate Density." Annals of the Institute of Statistical Mathematics, 18, 174-183, 1966.

Epanechnikov, V.A. "Non-Parametric Estimation of a Multivariate Probability Density." Theory of Probability and Its Applications, Vol. 15, No. 1, 153-58, 1969.

Lemon, G. H., and R. G. Krutchkoff. "An Empirical Bayes Smoothing Technique." Biometrika, 56, No. 2, 361-65, 1969.

Maritz, J.A. Empirical Bayes Methods. London: Methuen and Co., LTD, 1970.

Neyman, J. "Two Breakthroughs in the Theory of Statistical Decision Making." Revue of the Institute of International Statistics, 30, 11-21, 1962.

Parzen, E. "On the Estimation of a Probability Density Function and Mode." The Annals of Mathematical Statistics, 33, 1065-76, 1962.

Raiffa, H., and R. Schlaifer. Applied Statistical Decision Theory. Boston: Harvard University Press, 1961.

Robbins, H. "An Empirical Bayes Approach to Statistics." Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 157-63, 1955.

Robbins, H. "The Empirical Bayes Approach to Statistical Decision Problems." The Annals of Mathematical Statistics, 35, 1-20, 1964.

Rutherford, J.R., and R. G. Krutchkoff. "Some Empirical Bayes Techniques in Point Estimation." Biometrika, 56, No. 1, 133-37, 1969.

Ryzhik, I. M., and I. D. Gradshteyn. Tables of Integrals, Series, and Products. New York: Academic Press, 1965.

Samuel, E. "An Empirical Bayes Approach to the Testing of Certain Parametric Hypothesis." The Annals of Mathematical Statistics, 34, 1370-85, 1963.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

"An empirical Bayes approach to a variables sampling plan problem."

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

James A. Craig, Jr.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 28, 1971	61	12

8a. CONTRACT OR GRANT NO. N00014-68-A-0515 b. PROJECT NO. NR 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 98 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

Consider the following variables sampling plan problem. We are given a lot of size m+n items. Each item's quality is characterized by some continuous random variable X. If this random variable, X, for a given item is within some specification limits, say (a,b), the item is considered acceptable. On the basis of a random sample of size n, we want to accept or reject the remaining m items. If we reject the lot, we incur a loss c_0 for each of the remaining m items that meets the specification (i.e., $a<X<b$). We incur no loss on this case when an item does not meet the specification. If we accept the lot, we incur a loss c_1 for each of the remaining m items that does not meet the specification (i.e., X a, or X b). We incur no loss for those items meeting the specification. The random variable, X, is known to be normally distributed with some unknown mean μ and unknown variance σ^2. The random variable, X, for any item in the lot is independent of the random variable, X, for any other item in the lot.

Furthermore, we know that μ and σ^2 are random variables with some unknown prior distribution $G(\mu,\sigma^2)$. That is, the manufacturing process for this item is a random process that chooses some μ and σ^2 at random according to a distribution $G(\mu,\sigma^2)$ and then produces m+n items according to the normal distribution with mean μ and variance σ^2.

If we knew the prior distribution $G(\mu,\sigma^2)$, we could determine a Bayesian decision rule based on the sample mean x and sample variance s^2 that would minimize the average expected loss. We show in this research that in certain cases, if one has data from past lost of size m+n, it ispossible to estimate the Bayesian ... cont'd.

DD FORM 1 NOV 65 1473 (PAGE 1)

UNCLASSIFIED

Continuation sheet.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED
3. REPORT TITLE "An empirical Bayes approach to a variables sampling plan problem"	
4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)* Technical Report	
5. AUTHOR(S) *(First name, middle initial, last name)* James A. Craig, Jr.	

6. REPORT DATE February 28, 1971	7a. TOTAL NO. OF PAGES 61	7b. NO. OF REFS 12
8a. CONTRACT OR GRANT NO. N00014-68-A-0515 b. PROJECT NO. NR 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 98 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	
10. DISTRIBUTION STATEMENT See 1st sheet.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY See 1st sheet.	

13. ABSTRACT

Continued:

decision rule empirically. That is, we have an "empirical Bayes decision rule, and, consequently, we have an "empirical Bayes" approach to a variables sampling plan problem. We show that as the number of past lots increases (i.e.,data from past realizations of the process), the empirical Bayes decision rule converges to the Bayesian decision rule. Finally, we give an example of the theoretical results by performing a Monte Carlo simulation using a conjugate prior distribution for μ and σ^2.