Bayes Risk for the Test of Location-
The Infinite Dimensional Case*

by

Philip Cohen

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 288

June 1972

*Research was supported in part by the Office of Naval Research Contract N00014-67-A-0226-0008, Project Number NR042-216 at Purdue University.

CHAPTER I

INTRODUCTION AND FORMULATION

1.1 INTRODUCTION. Within the structure of hypothesis testing the experimenter would like to be able to judge which test one should use for a particular class of problems. Towards these ends several definitions of asymptotic relative efficiency of test procedures have been used. [The major ones in use today are Pitmann [9], Chernoff [4], and Bahadur [2].]

It would be desirable if one could unify the first two within a common framework. Besides the obvious reasons that the Bayesian framework is a natural one within which to consider such problems, Lindley [8] also pointed out that there was evidence that for a wide class of priors and loss functions the Bayes decision rules asymptotically behaved like Maximum Likelihood Estimators, which have several desirable large sample Properties.

In an attempt to see whether one could obtain a tractable measure of relative efficiency, within the Bayesian framework Rubin and Sethuraman [10] defined Bayes Risk Efficiency (BRE) to be the asymptotic relative sample sizes needed to obtain equal expected risk and considered the problem of the test of location. More specifically, they considered the question

$$H_0 , \ \theta = 0$$

vs.

$$H_1 , \ \theta \neq 0 , \quad \theta \in \Theta ,$$

where Θ is an N-dimensional parameter space. In order to state Rubin and Sethuraman's results we define a constant loss, B, for the type I error and a type II loss of

$$h(\theta/|\theta|)\,|||\theta|||^{\lambda} , \qquad \lambda > -1 .$$

Here we define $|||\theta|||^{\lambda}$ as some norm of θ and require that $h(\theta/|\theta|)$ is slowly varying.

Furthermore, we shall require a prior probability p_o that $\theta = 0$, and $(1-p_o)P(\theta)$, a prior distribution over $\{\theta:\theta \neq 0\}$. Finally, if we let $f^i(x|\theta)$ be the distribution function of the test statistic, we may state the expected risk for the i^{th} test statistic as

$$R^i = R_1^i + R_2^i , \quad i = 1, 2$$

$$= p_o B \int_{S_x} f^i(x|\theta=0)\,dx + (1-p_o) \int_{\bar{S}_x} \int_{\theta} h(\theta/|\theta|)\,|||\theta|||^{\lambda} f(x|\theta=0) P(\theta)\, d\theta dx$$

where S_x, $\bar{S}_x$ are the critical region and its compliment. Rubin and Sethuraman showed that in the finite dimensional case, for a large class of statistics, one may ignore R_1^i since

$$(1.1) \qquad R_2^i \sim (\log n) R_1^i$$

This, along with an asymptotic expression for the type II risk enabled them to invert R_2^i to obtain an expression for

$$\lim_{n\to\infty} N^{(1)}(n)/N^{(2)}(n) ,$$

the relative sample sizes needed to obtain equal expected risks. for two different tests. More specifically they showed that if, for a particular test statistic there existed an "a" such that the probability of the probability of the type I error was of the form

$$P(E_1) = \Phi(n,a)n^{-a/2} ,$$

with $\Phi_a(n,a)/\Phi(n,a) = o(\log n)$ and the asymptotic boundary of the almost-sure acceptance region in the parameter space was of the form

$$\left\{\theta: ||\theta|| > (a\ g(\theta/|\theta|)\frac{\log n}{n})^{\frac{1}{2}}\right\} ,$$

then $R_2 \sim (\log n)R_1$, and

$$(1.2) \quad R_2 \sim \left[\frac{\log n}{n}\right]^{\frac{\lambda+N}{2}} a^{\lambda+N} \int h(\theta/|\theta|)\,|||\theta|||^{\lambda} d\lambda, \quad ||\theta|| < g(\theta/|\theta|)$$

(see [9], Theorem 1). Since this expression is asymptotically invertable they obtained the result that

$$(1.3) \quad N^{(1)}(n)/N^{(2)}(n) \sim \left[\frac{\int_{\theta\in\ell_1} |||\theta|||^{\lambda} h(\theta/|\theta|)d\theta}{\int_{\theta\in\ell_2} |||\theta|||^{\lambda} h(\theta/|\theta|)d\theta}\right]^{2/(\lambda+N)} ,$$

$\ell_i = \{\theta: ||\theta||_i \le |g_i(\theta/|\theta|)|^{\frac{1}{2}}\}$, $i = 1, 2$ [see [10], Theorem 2].

Another important result which they obtained was that "in all regular problems in which the Pitman efficiency is usually obtained, the BRE coincides with the Pitman efficiency." While this result was not proved rigorously, it was shown to be true for several examples which were presented.

One fairly large class of tests which was not covered by their results was the case in which the parameter space is infinite dimensional, i.e., $N = \infty$. Clearly, (1.2), (1.3) do not hold in this case. An example of such a problem might be the following:

Let $X^1, X^2, \ldots, X^n$ be the infinite dimensional elements of the power spectrum of a sample of size n taken from the output of a data channel (before clipping has occurred). Under the conditions of no noise ($\theta=\theta_o$, θ is infinite dimensional) and normal attenuation due to transmission channel characteristics, these elements have a distribution which we shall define as $f(\mathrm{X}|\theta=\theta_o)$. Furthermore, under the null hypothesis we may assume that each one of these samples is independent of the other. Suppose further that, based on previous experience we have a prior probability p_o that there is negligible noise on the channel and $(1-p_o)P(\theta)$ that there is more than a negligible amount of noise where $P(\theta)$ is a probability measure which is obtained from prior experience with the frequency of different kinds of noise which occurs on the channel. Now, if I make a type I error (guess $\theta \neq \theta_o$ when in fact $\theta = \theta_o$), then I would shut the channel down when in fact it is still profitable to operate it.

A natural form of loss for this type of error would be one which would be a constant (per unit time) whose value would be determined by such factors as capital depreciation of the channel and loss of some fraction of the customers who would transmit the data by some other means.

The loss associated with a type II error (guessing $\theta = \theta_o$ when in fact it is not) means that we would be transmitting data over a channel which was to some extent faulty. Since some data may be transmitted over such channels, a natural means of defining this loss would be $\theta' A \theta$, where A may be determined through an analysis of the frequency of different kinds of customer use and an estimate of the reliability required for each of these uses (in other words, some customers may not care that they have only received 99% of the data transmitted while others will demand a retransmission of the data).

A more general application of the infinite dimensional case is in testing of whether a sample of size n has been drawn from a population with continuous distribution function F(x). If one uses the Cramer-Von Mises Statistic W_n^2 for this test it can be shown [see [12], p. 153] that this statistic may be represented as an infinite dimensional chi-square variable which is not asymptotically normal.

The question that was raised in this thesis was whether results similar to that of (1.1) and (1.2) could be obtained for the infinite dimensional case. Because of the difficulties created by working in a non-finite dimensional space, the loss function was simplified to the case $\lambda = 0$, constant loss, and to

$$||\theta||^{\lambda} = \theta' A \theta \ ,$$

the case of quadratic loss. The case of constant loss is considered in Chapter II. The case of quadratic loss is considered in Chapter III. In both cases it was found that while the type II risk, R_2, always dominates the type I risk, R_1, the domination is not strict [see Theorem 2.1 and 3.3] and that for some cases, cited in Section 2.4, R_1 and R_2 are asymptotically proportional.

In the constant loss case an asymptotic expression for R_1, useful for moderate sample sizes ($n > 32$) was obtained.

In the more difficult quadratic loss case, a two stage asymptotic technique was required in order to obtain an expression for R_1 for the case $Tr[A] < \infty$, an upper bound for R_1 was also obtained for a more general case of quadratic loss and some examples of solution considered.

Finally the reader should be forewarned that in the introduction to both the constant loss and the quadratic loss cases, symbolic expressions for infinite dimensional distribution functions, such as $f(x|\theta)$, are used. In order to correctly evaluate these symbolic functions one must evaluate the expression for a finite dimensional parameter space and then take limits.

CHAPTER II

CONSTANT LOSS

2.1 INTRODUCTION. Let $X^1, X^2, \ldots, X^n$ be infinite dimensional independent normal variables with mean vector θ and covariance matrix I.

Let the null hypothesis be $\theta = 0$ and the alternative $\theta \neq 0$. Let

$$p_0 \equiv \text{prior probability that } \theta = 0$$

$$(1-p_0)\ p(\theta) \equiv \text{prior distribution over } \{\theta : \theta \neq 0\}$$

where $p(\theta) \sim N[0,\Sigma]$ and

$$\Sigma = \{\sigma_{ij}^2\} \qquad \sigma_{ij}^2 = 0 \quad i \neq j$$
$$= \sigma_i^2 \quad i = j,\ i = 1, 2, ., \infty\ ;$$

and we shall assume that

$$(2.1) \qquad \Sigma\sigma_i^4 < \infty\ .$$

The constant loss associated with the type I and type II errors are respectively B and A. In this paper the asymptotic relationships between and asymptotic expressions for R_1 and R_2, the type I and type II Bayes risks, are developed.

Definition. When we write F_n is asymptotic to f_n, $F_n \sim f_n$, we mean $F_n/f_n \to 1$. If we write F_n is order-asymptotic to f_n we mean to say that the asymptotics are not quite as good and that we only have $\log F_n/\log f_n \to 1$. Furthermore we say two results are asymptotically comparable if both are asymptotic, or order asymptotic.

The Finite Dimensional version of this problem, $\sigma_i^2 = 0 \ \forall i > N$ was "asymptotically" solved by Rubin and Sethuraman in [10]. In Section 2.2 it is shown that the results obtained in [10] for the finite dimensional case, namely that $R_1 \log n \sim R_2$, is not always true in the infinite dimensional case (the case in which $\nexists N$ such that $\sigma_i^2 = 0 \ \forall i > N$).

The theorem presented in Section 2.2 gives asymptotic expressions for R_1 and R_2 and demonstrates that in general

$$R_2 = O(R_1 \log n)$$

$$R_1 = O(R_2) \quad .$$

The asymptotic expressions for R_1 and R_2 which are obtained in this paper are as accurate as those which were obtained in [10] for the finite dimensional case.

In Sections 2.3 order-asymptotic examples are worked out for special cases of Σ, namely, $\sigma_i^2 = 1/i^{\frac{1}{2}+\delta}$, $\delta > 0$, and $\sigma_i^2 = 1/a^i$, $a > 1$.

In Section 2.4 some exact results are obtained and a numeric comparison of exact, asymptotic and order-asymptotic results are made for the case in which $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$. For this case one begins to obtain useful accuracy for sample sizes as small as 32.

Furthermore, in the cases considered, the predicted asymptotic relationship between R_1 and R_2 is also obtained for relatively small sample sizes thus eliminating the necessity of calculating both R_1 and R_2.

In Chapter 3 of this paper it is shown that the results obtained in this chapter for the case of constant loss can be used to obtain order results for the case in which one has quadratic loss for the type II error of the form $\theta' A \theta$ and a not necessarily diagonalized covariance matrix Σ. This extention motivates Sections 2.3 and 2.4 of this chapter in which the accuracy of the asymptotic risks are considered.

While the results presented in this chapter are new, similar techniques have been brought to bear in other fields. In particular the theorem stated in Section 2.2 has a parallel in Chapter 5 of the book by Hirschman and Widder [5] and in a paper by Hsu [6]. In our case, however, the results were obtained using a different technique in order to permit certain needed generalizations. Furthermore, since one of the subclasses of the kernel considered is the class known as the Polya frequency function, some of the examples and ideas used in the Exact Evaluation section of this paper, Section 3.4, were obtained from Chapter 7 of [13].

2.2 ANALYTIC RESULTS. Let $X \equiv (X_1/\sqrt{n}, X_2/\sqrt{n}, \ldots)$ be the mean vector of $X^1, \ldots, X^n$. Then $X_i \sim N[\sqrt{n}\,\theta_i, 1]$, $i = 1, \ldots$, and $X \sim N[\theta, I/n]$. We shall use $f(x|\theta)$ to denote the density function X. Furthermore since the i^{th} mean, $X_i/\sqrt{n}$, is sufficient for θ_i, we may determine the Bayes procedure by finding that set of x such that $L_1(x) = L_2(x)$ where

$$L_1(x) \equiv \frac{f(x|\theta=0)p_o B}{f(x|\theta=0)p_o + (1-p_o)\int f(x|\theta)p(\theta)d\theta}$$

and

$$L_2(x) \equiv \frac{A(1-p_o)\int f(x|\theta)p(\theta)d\theta}{p_o f(x|\theta=0) + (1-p_o)\int f(x|\theta)p(\theta)d\theta}$$

are the posterior risk when $X = x$ is observed. Since the denominator is the same for each term, the problem may be reduced to finding

$$\left\{x: \frac{f(x|\theta=0)}{\int f(x|\theta)p(\theta)d\theta} = \frac{(1-p_o)A}{p_o B} \equiv \frac{1}{\sqrt{K}}\right\}$$

This, in turn, is equivalent to the problem of finding

$$\left\{x: \exp(-\tfrac{1}{2}\, n|x|^2) = \frac{1}{\sqrt{K}}\int \frac{\exp(-\frac{1}{2}\,\theta'\Sigma^{-1}\theta + n(x-\theta)'(x-\theta))d\theta}{|2\pi\Sigma|^{\frac{1}{2}}}\right\} \tag{2.2}$$

Note. that (2.2) and other subsequent expressions of this form must be treated as symbolizing limits which are taken in terms of the dimensions of the parameter space [ie $|2\pi\Sigma| = 0$]. Now if one evaluates the integral on the right hand side of the above equation and takes the log of both sides, straightforward calculations show that we will accept H_0 if X satisfies the relationship

$$\sum_{i=1}^{\infty} \left[\frac{X_i^2}{1+n\sigma_i^2} - 1\right] n\sigma_i^2 \leq U_n \quad , \tag{2.3}$$

where the constant $U_n \equiv \log\left[K \prod_i (1+n\sigma_i^2)\, e^{-n\sigma_i^2}\right]$ and $X_i \sim N[\sqrt{n}\theta_i, 1]$.

The infinite sum obtained in (2.3) can be shown to be a random variable for fixed n by a straightforward application of Kolmogorov's three-series criterion in conjunction with Chebychev's inequality. This result holds for all values of θ .

Thus the type I Bayes Risk, R_1, may be expressed in the form,

$$(2.4) \qquad R_1 = Bp_o \; P\left[\sum_{i=1}^{\infty} \left(\frac{X_i^2}{1+n\sigma_i^2} - 1 \right) n\sigma_i^2 - U_n \geq 0 \,\middle|\, \theta=0 \right].$$

In order to obtain the type II risk we must evaluate

$$(2.5) \quad R_2 = A(1-p_o) \iint_S \frac{e^{-n/2(Z-\theta)'(Z-\theta) - \frac{1}{2}\theta'\Sigma^{-1}\theta} dZd\theta}{\left|\frac{2\pi}{n}\right|^{\frac{1}{2}} \left|2\pi\Sigma\right|^{\frac{1}{2}}} \quad ,$$

where S is obtained from (2.4) and can be expressed as

$$S = \left\{ (\theta, Z) : \sum_i \left[\frac{Z_i^2}{1+n\sigma_i^2} - 1 \right] n\sigma_i^2 - U_n \geq 0, \; \theta \neq 0 \right\} ;$$

straightforward computations can be used to reduce (2.5) to

$$(2.6) \qquad R_2 = A(1-p_o) \; P\left[\sum_i (W_i^2 - 1) \; \sigma_i^2 - \frac{U_n}{n} \leq 0 \right] ,$$

where W_i are independent N[0,1] random variables. Now let

$$(2.7) \qquad Y_1 \equiv \sum_i \left(\frac{X_i^2}{1+n\sigma_i^2} - 1 \right) n\sigma_i^2$$

and

$$(2.8) \qquad Y_2 \equiv \sum_i (W_i^2 - 1) \; \sigma_i^2 \; .$$

The Laplace transforms of (2.7) and (2.8) are defined respectively as

$$\Phi_1(v-\tfrac{1}{2}) = E\left[e^{-(v-\frac{1}{2})y_1}\right] = \left[\frac{\prod_i (1+n\sigma_i^2)e^{-n\sigma_i^2}}{\prod_i (1+2vn\sigma_i^2)e^{-2vn\sigma_i^2}}\right]^{\frac{1}{2}}, \quad 0 < v \tag{2.9}$$

and

$$\Phi_2(s) \equiv E(e^{-sy_2}) = \left[\frac{1}{\prod_i (1+2s\sigma_i^2)e^{-2s\sigma_i^2}}\right]^{\frac{1}{2}}, \quad s > -\tfrac{1}{2}\ . \tag{2.10}$$

Now in order to state this chapter's main result, Theorem 2.1, some technical lemmas are needed.

First, let us define

$$\zeta_j(n,v) \equiv \frac{n\sigma_j^2 v}{1+2n\sigma_j^2 v}\ . \tag{2.11}$$

Furthermore let $v = v_1(n)$ be the solution to the equation

$$U_n + \sum_{j=1}^{\infty} 2n\sigma_j^2\, \zeta_j(n,v_1(n)) = 0\ , \tag{2.12}$$

where U_n is the constant defined in (2.4) and the infinite sum in (2.12) converges by (2.1).

Lemma 2.1. For each $K > 0$, where K is defined in (2.2), there exists an N_k such that for $n \geq N_k$, $v_1(n)$ exists and $0 < v_1(n) < \frac{1}{4}$.

Proof: For convenience let $L(n) \equiv U_n$ and

$$R(n,v) \equiv - \sum_{j=1}^{\infty} 2n\sigma_j^2 \zeta_j(n,v) = \sum_{j=1}^{\infty} \left[\frac{n\sigma_j^2}{1+2n\sigma_j^2 v} - n\sigma_j^2\right] .$$

Clearly, for $0 \leq n < \infty$ and $0 \leq v \leq \frac{1}{4}$, $L(n)$ and $R(n,v)$ are continuous in both n and v. Furthermore since $\log(1+a) > a/(1+(a/2))$ for all $a > 0$, we obtain that

$$0 = R(n,0) > L(n) - \log K > R(n,\tfrac{1}{4}) . \tag{2.13}$$

Thus for $K = 1$ and any $n > 0$, there exists $v_1(n) \in (0,\frac{1}{4})$ which satisfies (2.12). Now if $K \neq 1$, then we must note that since $L(n) \to -\infty$, there exists N_K^1 such that for $n > N_K^1$, we have that $R(n,0) > L(n)$. Furthermore since, for all $a > 0$,

$$\frac{d}{da}\left[\log(1+a) - a/(1+(a/2))\right] > 0,$$

it follows that $R(n,\frac{1}{4}) - L(n) \to \infty$. Hence there exists N_K^2 such that for $n > N_K^2$, $L(n) > R(n,\frac{1}{4})$. Letting $N_k = \max(N_K^1, N_K^2)$ we have the desired result.

It is not hard to show, using the techniques employed in Section 2.3, that the choice $\sigma_i^2 = (i \log^2 i)^{-\frac{1}{2}}$ (which satisfies $\Sigma\sigma_i^4 < \infty$) yields $v_1(n) \to \frac{1}{4}$. Furthermore, it is shown in Section 2.3 that in the finite dimensional case $v_1(n) \to 0$. Therefore the result $0 < v_1(n) < \frac{1}{4}$ cannot be improved.

Now in order to derive Lemma 2.3 we need the following technical result which is stated without proof.

Lemma 2.2. If for $\alpha > 0$, $U > 0$, $\log(1+U) - U/(1+\alpha U) \leq 0$ then for $0 < W < U$, $\log(1+W) - W/(1+\alpha W) \leq 0$.

This result is used to show the following.

Lemma 2.3. Without loss of generality assume $\sigma_i^2 \downarrow$. Then

$$\frac{1}{v_1(n)} = O(\log n)$$

and therefore

$$\lim_{n\to\infty} nv_1(n) = \infty .$$

Proof: Let T_n be the solution to the equation

$$\max[0,\log K] + \log(1+n\sigma_1^2/(1+2n\sigma_1^2 T_n) = 0 \quad ;$$

then $T_n \sim 1/(2 \log n)$ and $\log(1+n\sigma_1^2/(1+2n\sigma_1^2 T_n)) \leq 0$.

Now since $\sigma_i^2 \leq \sigma_j^2$ for $i > j$, we have from the preceeding lemma that for each i

$$\log(1+n\sigma_i^2) - \frac{n\sigma_i^2}{1+2n\sigma_i^2 T_n} \leq 0 .$$

Let us now subtract $n\sigma_i^2$ from the first term in the above expression and add $n\sigma_i^2$ to the second term in the above expression. If we then sum the resulting expression over i we obtain

$$L(n) - R(n,T_n) \leq 0;$$

and since $R(n,T_n)$ is monotone decreasing in T_n, we have that

$$v_1(n) \geq T_n$$

and therefore

(2.14) $$\frac{1}{v_1(n)} = O\ (\log n)\ .$$

Lemma 2.4. $v_1(n)$ is monotone decreasing in n .˙. $\exists\ v_1 \in [0,\frac{1}{4}) \ni v_1(n) \downarrow v_1$.

Proof: Holding $v_1(n)$ fixed and differentiating $L(n) - P(n,v_1(n))$ with respect to n we obtain

$$\frac{\partial}{\partial n}\ [L(n) - R(n,v_1(n))] = 2nv_1 \sum_{j=1}^{\infty} \left(\frac{\sigma_i^2}{1+n\sigma_i^2}\right)^2 > 0\ ;$$

thus, for any n

$$L(n+1) - R(n+1,v_1(n)) > 0\ ,$$

and since $R(n+1,v_1(n))$ is monotone decreasing in $v_1(n)$ we have that $v_1(n)$ is monotone decreasing in n.

Now let us define $k_n(v)$ by

(2.15) $$k_n(v) \equiv \log\ \Phi_1(v-\tfrac{1}{2}) + (v-\tfrac{1}{2})\ U_n\ .$$

Then its second derivative $k_n''(v)$ is

(2.16) $$k_n''(v) = -2\sum_i \left[\frac{n\sigma_i^2}{1+2vn\sigma_i^2}\right] < 0\ .$$

Theorem 2.1 If $\sum_{i=1}^{\infty} \sigma_i^4 < \infty$, then as $n \to \infty$

(2.17)
$$R_1 \sim \frac{2\ Bp_o\ \exp(k_n(v_1))}{\sqrt{-2\pi k_n''(v_1)}\ (1-2v_1)} \cdot C_N \equiv d_n$$

(2.18)
$$R_2 \sim \frac{Bp_o\ \exp(k_n(v_1))}{v_1\sqrt{-2\pi k_n''(v_1)}} \cdot C_N ,$$

and therefore

(2.19)
$$R_2 \sim \frac{(1-2v_1)\ R_1}{2\ v_1} ,$$

where v_1 is a real valued function of n and satisfies (2.12).

The constant, C_N, that appears in (2.17) and (2.18) is equal to 1 for the Infinite Dimensional Case, $N = \infty$.

In the Finite Dimensional Case, $N < \infty$, we have that

(2.20)
$$C_N = \frac{\sqrt{\pi}\ N^{(N-1)/2}\ e^{-(1+N/2)}}{\Gamma(N/2)\ 2^{(N/2-1)}} .$$

Note that $\lim_{n\to\infty} C_N = 1$.

Proof of (2.17): Since $\Phi_1(s)$ is analytic to the right of $s = -\frac{1}{2}$ for all n, we may use the upper tail bilateral Laplace inversion formula ([7],p.242) to rewrite (2.4) as

$$(2.21) \qquad R_1 = \frac{Bp_o}{2\pi i} \int_{s_1-\infty}^{s_1+i\infty} \frac{\Phi_1(s)}{s} e^{sU_n} ds = \frac{Bp_o}{2\pi i} \int_{s_1-i\infty}^{s_1+i\infty} \frac{e^{k_n(s+\frac{1}{2})}}{s} ds$$

where $\frac{1}{2} < s_1 \equiv v_1 - \frac{1}{2} < 0$ (see Lemma 2.1).

It is clear that $\Phi_1(s)$ and its inverse exists if one recognizes that $[\Phi_1(s)]^2$ is just the bilateral Laplace transform of a distribution function of the Polya type ([7] p.333).

Now in order to prove that R_1 in (2.21) is asymptotic to (2.17) we shall apply a slight variation of Laplace's asymptotic technique (see [13] p.277 and [6]).

Let us first make the transformation

$$s = s_1 + \frac{iv_1 v}{2\sqrt{\sum_j \zeta_j^2}} = s_1\left(1 + \frac{iv}{\ell(n)}\right)$$

where

$$(2.22) \qquad \ell(n) \equiv 2\frac{s_1\sqrt{\Sigma\, \zeta_j^2}}{v_1} = \sqrt{2}\, s_1\sqrt{-k_n''(v_1)}$$

$\zeta_j \equiv \zeta_j(n,v_1)$ and $\zeta_j(n,v)$ is defined in (2.11)

and $k_n''(v_1)$ is defined in (2.15). Equation (2.21) may be rewritten as

$$(2.23) \qquad R_1 = \frac{Bp_o}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(k_n[s_1(1 + iv/\ell(n)) + \frac{1}{2}])}{\ell(n) + iv} dv .$$

Now the object of this proof is to show that $R_1/d_n \to 1$, or that

(2.24) $$\int_{-\infty}^{\infty} K_n(v)\, dv \to 1 ,$$

where d_n is defined in (2.17) and

(2.25) $$K_n(v) \equiv \frac{Bp_o}{2\pi} \frac{e^{k_n(s_1[1 + iv/\ell(n)] + \frac{1}{2})}}{(\ell(n) + iv) \cdot d_n} .$$

Now at this point we must break the proof up into two cases.

(2.26) Case I: $\nexists\, N \ni \sigma_i^2 = 0 \quad \forall i > N$ - The infinite dimensional case.

(2.27) Case II: $\exists\, N \ni \sigma_i^2 = 0 \quad \forall i > N$ - The finite dimensional case.

Case I will be proved using the Dominated Convergence Theorem. Case II will be proved by a simpler argument. We begin by proving Case I.

Case I - The infinite dimensional case. From the definition (2.24) and (2.25)

$$I_m\left[\int_{-\infty}^{\infty} K_n(v)\right] = 0, \quad n = 1,2\ldots ,$$

we can obtain the result in (2.24) from the Dominated Convergence Theorem if we can show that for fixed $v \in (-\infty,\infty)$

(2.28) $$\mathrm{Re}[K_n(v)] \to \frac{e^{-v^2/4}}{2\sqrt{\pi}}$$

and

(2.29) $$|K_n(v)| \leq 1/[1 + v^2 c_1]^{c_2} ,$$

where c_1 and c_2 are non-negative constants. Now in order to prove (2.28) we first note that $k_n(s_1(1+iv/\ell(n)) + \frac{1}{2})$ may be rewritten as

$$k_n(s_1(1+iv/\ell(n)) + \tfrac{1}{2}) =$$

$$k_n(v_1) - \tfrac{1}{2}\log\Big[\prod_k \Big[1+\frac{\zeta_k iv}{\sqrt{\Sigma\zeta_j^2}}\Big]\exp\Big[\frac{-\zeta_k iv}{\sqrt{\Sigma\zeta_j^2}}\Big]\Big]$$

$$+ \frac{iv_1 v}{2\sqrt{\Sigma\zeta_j^2}}\Big[U_n + \sum_j 2\zeta_j\ n\sigma_j^2\Big];$$

and using the definition of v_1 in (2.12) we have that the last term in the preceeding equation is zero,so that

$$k_n(s_1(1+iv/\ell(n)) + \tfrac{1}{2}) =$$

(2.30)

$$k_n(v_1) - \tfrac{1}{2}\log\Big[\prod_k \Big[1+\frac{\zeta_k iv}{\sqrt{\Sigma\zeta_j^2}}\Big]\exp\Big[\frac{-\zeta_k iv}{\sqrt{\Sigma\zeta_j^2}}\Big]\Big].$$

Now in order to prove (2.28) we need only show that for fixed v the last term in (2.30) converges to $(\frac{1}{2}\sqrt{\pi})\exp(-v^2/4)$. To do this we let

$$W_j(n) = \frac{-i\zeta_j v}{\sqrt{\sum_k \zeta_k^2}} \quad . \tag{2.31}$$

Then for fixed v

$$\sum_j W_j^2(n) = -v^2.$$

Now for the case under consideration, Case I, defined in (2.26), we have the condition that, as $n \to \infty$,

$$\sum_i \zeta_i^2 \to \infty \ .$$

Furthermore since $|\zeta_j| < \frac{1}{2}$ for all j, we have the condition that, for fixed v, $W_j(n)$ convergent to zero uniformly, in j as $n \to \infty$. These conditions are sufficient to prove the following.

<u>Lemma</u> <u>2.5</u>

$$\prod_{j=1}^{\infty} (1 - W_j(n)) \ e^{W_j(n)} \to e^{v^2/2}$$

<u>Proof</u>. The uniform convergence condition implies that given $\epsilon > 0$. $\exists$ N_ϵ such that for $n > N_\epsilon$ $\max_j |W_j(n)| < \epsilon$. Hence for $n > N_1$

$$\sum_{j=1}^{\infty} [\log [1 - W_j(n)] + W_j(n)] = \sum_{j=1}^{\infty} \sum_{k=2}^{\infty} \left[\frac{W_j(n)}{k} \right]^k = \frac{v^2}{2} + \sum_{j=1}^{\infty} \sum_{k=3}^{\infty} \left[\frac{W_j(n)}{k} \right]^k .$$

But for $k \geq 3$ and $n > \max [N_\epsilon, N_1]$

$$\left| \sum_{j=1}^{\infty} [W_j(n)]^k \right| \leq \epsilon^{k-2} \sum_{j=1}^{\infty} |W_j(n)|^2 \ ;$$

so

$$\left| \sum_{j=1}^{\infty} [\log [1-W_j(n)] + W_j(n)] - \frac{v^2}{2} \right| \leq \left[\sum_{j=1}^{\infty} |W_j(n)|^2 \right] \frac{\epsilon}{1-\epsilon} = \frac{\epsilon v^2}{1-\epsilon},$$

and since ϵ may be made arbitrarily small, the lemma is proved.

Thus we have demonstrated equation (2.28).

Now in order to show that (2.29) holds we first note that

$$(2.32)\qquad |K_n(v)| = |\ell(n)|/[\prod_j (1-W_j^2(n))^{\frac{1}{4}} [\ell(n)^2 + v^2]^{\frac{1}{2}}] ,$$

where $W_j^2(n)$ is defined in (2.31). Next we must apply the well known result that if $\Sigma\, a_i < M$ and $a_i > 0$, then

$$\prod_1^\infty (1+a_i) \geq (1+\max_i a_i)^{M/\max_i a_i} .$$

Applying this result and letting $a_i = -W_i^2(n)$ we obtain that for all n,

$$|K_n(v)| \leq \frac{1}{\left[1+ (\frac{v}{\ell(n)})^2\right]^{\frac{1}{2}} \left[1 - \min_j W_j^2(n)\right]^{v^2/(-4 \min_j W_j^2(n))}}$$

$$\leq \frac{1}{\left[1 - \min_j W_j^2(n)\right]^{v^2/(-4 \min_j W_j^2(n)}} .$$

Now letting " ↓" denote monotone decreasing in n, we have that, for fixed v and for all n,

$$\max_j [-W_j^2(n)/v^2] \downarrow 0 .$$

Thus we have that for fixed v and for all $n \geq 1$

$$[1 - \min_j W_j^2(1)]^{v^2/(-4 \min_j W_j^2(1))} \leq [1 - \min_j W_j^2(n)]^{v^2/(-4 \min_j W_j^2(n))} ,$$

and the right hand side monotonely increases to $\exp(v^2/4)$. Furthermore, from the definition of $W_j^2(n)$, there is a unique b such that for all n

$$W_b^2(n) = \min_j W_j^2(n) \ .$$

Thus we may write

$$v^2/-4 \min_j W_j^2(1) = (\sum_k \zeta_k^2/4\ \zeta_b^2)|_{n=1}$$

which implies that

$$|K_n(v)| \leq 1/[1-W_b^2(1)]^{(\sum_K \zeta_K^2/4\zeta_b^2)|_{n=1}}$$

$$= 1/[1+v^2 c_1]^{c_2} ,$$

where c_1 and c_2 are non-negative constants. Hence we have proved (2.29) and thus we have that, for Case I, $R_1/d_n \to 1$.

Case II - The finite dimensional case. Assume that there exists an $N > 0$ such that $\sigma_i^2 = 0, \forall i > N$. Let us also assume, without loss of generality, that $\sigma_i^2 > 0, \forall i \leq N$.

Under the finite dimension assumptions we no longer have $\Sigma\zeta_i^2 \to \infty$, which enabled us to obtain asymptotic normality of the integrand $K_n(v)$. Instead, for the finite case we have

$$\lim_n \sum_{i=1}^{N} \zeta_i^2(n) = \frac{N}{4} \ .$$

Furthermore, it is shown in Section 3 that, for the finite dimensional case,

$$v_1 \sim \frac{1}{2 \log n}$$

or equivilantly, $s_1 \to \frac{1}{2}$. Therefore,

$$\ell(n) \sim \sqrt{N} \log n \to \infty .$$

Also

$$W_j(n) \to \frac{-iv}{\sqrt{N}}$$

so that

$$\prod_{1}^{N} (1 - W_j(n)) \quad e^{W_j(n)} \to (1 + \frac{iv}{\sqrt{N}})^{N} \ e^{-iv\sqrt{N}} \quad .$$

Substituting into (2.23) and using the definition of d_n from (2.17) we obtain

$$\frac{R_1}{d_n} \sim \int_{-\infty}^{\infty} \frac{e^{iv\sqrt{N}/2}}{2\sqrt{\pi}\ (1+\frac{iv}{\sqrt{N}})^{N/2}\ (1+\frac{iv}{\ell(n)})} \, dv$$

(2.33)

$$\to \int_{-\infty}^{\infty} \frac{e^{iv\sqrt{N}/2}}{2\sqrt{\pi}\ (1+\frac{iv}{\sqrt{N}})^{N/2}} \, dv = \frac{\sqrt{\pi}\ N^{(N-1)/2}\ e^{-(1+N/2)}}{2^{(N/2-1)}\ \Gamma(N/2)} ,$$

which is just C_N as defined in (2.20). For the finite case we need not show domination since the limit in (2.33) can be looked upon as the pointwise convergence at the point $x = \sqrt{N}/2$ of the weighted sum of N+2 independent chi square distribution functions in which the coefficients of the first N variables become asymptotically equal to $1/2\sqrt{N}$ and the coefficients of the last two variables go to zero. Hence we have proved (2.17) of the theorem for Case II, the finite dimensional case.

Proof of (2.18): Now in order to evaluate R_2 we let $s_2 \equiv nv_1$. Then from Lemma 2.1 we know that $s_2 \in (0,\infty)\ \forall n$. This enables us to rewrite (2.6) in the form

$$R_2 = \frac{A(1-p_o)}{2\pi i} \int_{s_2-i\infty}^{s_2+i\infty} \frac{\Phi_2(w)\ \exp((wU_n)/n}{w}\, dw \quad . \tag{2.34}$$

Now performing the transformation $n(1+2s) = 2w$ we obtain from (2.9) and (2.10) that

$$\Phi_2\left[\frac{n(1+2s)}{2}\right] = \frac{\Phi_1(s)}{\left[\prod_i (1+n\sigma_i^2)\ e^{-n\sigma_i^2}\right]^{\frac{1}{2}}} \quad . \tag{2.35}$$

Using this we can now rewrite (2.34) in the form of (2.21), namely

$$R_2 = \frac{A(1-p_o)}{2\pi i} \int_{s_1-i\infty}^{s_1+i\infty} \frac{\Phi_1(s)\ e^{sU_n+1/2\ \log K}}{s+\frac{1}{2}}\, ds \quad . \tag{2.36}$$

where K is the constant defined in (2.2). Here we once again perform the substitution $s = s_1(1+iv/\ell(n))$ to obtain

$$R_2 = \frac{Bp_o}{2\pi} \int_{-\infty}^{\infty} \frac{\exp\ (k_n(s_1(1+iv/\ell(n))+1/2))}{2\sqrt{\sum_i \zeta_i^2 + iv}}\ dv \ ,$$

where $k_n(v)$ is defined in (2.15). Moreover, for the infinite dimensional case, we have that both $\ell(n)$ and $\sqrt{\Sigma\zeta_i^2}$ become infinite as n increases. Thus, the same reasoning that was used to obtain the asymptotic type I risk may be used to obtain the type II risk, (2.18).

The finite dimensional asymptotic R_2 is obtained in the same manner that the finite dimensional R_1 was obtained.

2.3 SOME ASYMPTOTIC EVALUATIONS OF R_1.

2.3.1 Introduction. In this section the results obtained in Section 2.2 are applied to some special cases of the covariance matrix, Σ.

In Section 2.3.2 "order" estimates of R_2 and "asymptotic" estimates of the relationship between R_1 and R_2 are obtained for the class of priors for which the diagonal elements of the covariance matrix Σ are of the form $\sigma_i^2 = 1/i^{1/2+\delta}$, $\delta > 0$, $i = 1,2,\ldots$ At the same time a similar case, that in which $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^{1/2+\delta}$, $\delta > 0$, $i = 1,2,\ldots$, will also be considered. It should be noted that the low quality of the estimates ("order" estimates) in this section is not due to a weakness of the results in Section 2.2 but rather to the "rough" integral approximations of sums that was needed in order to obtain estimates for the whole class (i.e., $\delta > 0$). In order to "heuristically" demonstrate that this is, in fact, the case a "best-asymptotic" evaluation of R_1 is obtained in Section 2.4 for the special case in which $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$, $i = 1,2,\ldots$ The values obtained from this particular evaluation of R_2 compare favorably for sample sizes as small as $n = 32$ with the exact evaluation of R_2 which is also obtained for this special case in Section 2.4.

In Section 2.3.3 the case of $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/2a^i$, $a > 1$, $i = 0, 1, 2, \ldots$, which arises in Spectral Theory, is considered and "order-estimates" of R_2 are obtained.

Finally, in Section 2.3.4 results are obtained for the finite dimensional case and compared with the results obtained in [10].

2.3.2 Asymptotic Evaluation of R_2 and R_1/R_2 for the Class of Priors Characterized by the ("Single Roots") $\sigma_i^2 = 1/i^{1/2+\delta}$, $\delta > 0$, $i = 1,2,\ldots$, and by the ("Double Roots") $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^{1/2+\delta}$, $\delta > 0$, $i = 1, 2, \ldots$. In order that we may evaluate R_2 we must first find a value for v_1 which asymptotically satisfies (2.12). We shall break up the solution for v_1 into two parts. First, v_1 is obtained for the case in which $\delta > 1/2$. Then an argument is made whereby the solution is analytically continued to the case in which $0 < \delta \leq 1/2$.

For the case $\delta > 1/2$ and for both the "single root" and "double root" cases of Σ, (2.12) may be asymptotically reduced to the form

$$(2.37)\quad f_n(\delta, v_1(n,\delta)) = \sum_i{}' \frac{n}{i^{1/2+\delta+2v_1(n,\delta)}} - \log \prod_i \left(1 + \frac{n}{i^{1/2+\delta}}\right) + \log K = 0,$$

where $v_1(n,\delta)$ denotes the solution of (2.37) for a particular n and δ. Now applying the results of Lemmas 2.1 and 2.4 we have that $v_1(n,\delta) \downarrow v_1(\delta)$, where $0 \leq v_1(\delta) < 1/4$, $\forall \delta > 1/2$. Furthermore, $\exists$ $f(\delta, v_1(\delta))$ such that

$$f_n(\delta, v_1(n,\delta)) \to f(\delta, v_1(\delta)) .$$

Now since, from Lemma 2.3, we know that $nv_1(n,\delta) \to \infty$ and $v_1(n,\delta)$ is bounded, we may use an integral approximation to show that

$$(2.38)\qquad \sum_i n/(i^{1/2+\delta}+2v_1(n,\delta)n) \sim \frac{[2v_1(\delta)]^{-(2\delta-1)/(2\delta+1)}\pi n^{2/(1+\delta)}}{(1/2+\delta)\ \sin[2\pi/(1+2\delta)]}$$

and that

$$(2.39)\qquad \log K + \log \prod_i (1+n/i^{(1/2+\delta)}) \sim \frac{\pi n^{2/(1+2\delta)}}{\sin[2\pi/(1+2\delta)]}$$

(see [3], p. 118). Substituting these results into (2.37) we obtain that for $\delta > 1/2$

$$f(\delta,v_1(\delta)) = \frac{\pi n^{2/(1+2\delta)}}{\sin[2\pi/(1+2\delta)]}\left(1 - \frac{2v_1(\delta)^{-(2\delta-1)/(2\delta+1)}}{(1/2+\delta)}\right) = 0,$$

so that clearly for $\delta > 1/2$

$$(2.40)\qquad v_1(\delta) = \tfrac{1}{2}(1/2+\delta)^{-(2\delta+1)/(2\delta-1)} .$$

Now in the general case, $0 < \delta$, (2.37) must be written in the form

$$(2.41)\sum_i \frac{2n^2 v_1(n,\delta)}{(i^{1/2+\delta}+2nv_1(n,\delta))i^{1/2+\delta}} - \log\left[K\pi(1+n/i^{1/2+\delta})\ \exp(-n/i^{1/2+\delta})\right] = 0$$

It should be clear by inspection of (2.41) that for each n, $v_1(n,\delta)$ is analytic $\forall\ 0 < \delta$. Furthermore, by Lemma 2.4 for fixed $\delta > 0\ \exists\ v_1(\delta)$ such that $v_1(n,\delta) \downarrow v_1(\delta)$. Finally, since $v_1(n,\delta)$ is uniformly bounded we have that $v_1(n,\delta)$ converges uniformly to $v_1(\delta)$ on every compact subset of $(0,\infty)$. Thus, $v_1(\delta)$ must also be analytic. But now, since we have obtained $v_1(\delta)$ for $\delta > 1/2$ we may apply an analytic continuation argument and claim that (2.40) is in fact the limit of

$v_1(n,\delta)$, the solution of (2.41), for all $\delta > 0$. It now follows from (2.19) that

$$\frac{R_2}{R_1} \to \frac{(1-2v_1)}{2v_1} = \frac{-[1-(1/2+\delta)^{-(2\delta+1)/(2\delta-1)}]}{(1/2+\delta)^{-(2\delta+1)/(2\delta-1)}} .$$

The special cases $\delta = 1/2$ and $\delta = 3/2$, which will subsequently be considered, yield

$$\left.\frac{R_2}{R_1}\right|_{\delta=1/2} \to (e-1) \quad \text{and} \quad \left.\frac{R_2}{R_1}\right|_{\delta=3/2} \to 3 .$$

In order to evaluate R_2, defined in (2.18), we must first evaluate $k_n(v_1)$ and $k_n''(v_1)$, defined in (2.15) and (2.16). Substituting the asymptotic expressions obtained in (2.38), (2.39) and (2.40) into (2.15) we obtain

$$k_n(v_1) \sim C_\delta n^{2/(1+2\delta)} \tag{2.42}$$

where

$$C_\delta = \frac{\pi}{\sin[2\pi/(1+2\delta)]}\left(v_1 - \frac{(2v_1)^{2/(1+2\delta)}}{2}\right) .$$

Next, using an integral approximation for the following sum, we obtain for $b_n \to \infty$,

$$\sum_{i=1}^{\infty} 1/(i^{1/2+\delta}+b_n)^2 \sim \frac{2(2\delta-1)\pi b_n^{-[4\delta/(2\delta+1)]}}{(2\delta+1)^2 \sin[2\pi/(2\delta+1)]} .$$

Substituting this approximation into the expression for $k_n''(v_1)$ we obtain

$$k_n''(v_1) \sim D_\delta n^{2/(2\delta+1)} \tag{2.43}$$

where

$$D_\delta \equiv \frac{-4\pi(2\delta-1)(2v_1)^{-[4\delta/(2\delta+1)]}}{(2\delta+1)^2\sin[2\pi/(2\delta+1)]} .$$

Now substituting (2.42) and (2.43) into (2.18), the expression for R_2, we obtain

$$(2.44) \qquad R_2 \sim \frac{Bp_o \exp[C_\delta n^{2/(1+2\delta)}]}{v_1\sqrt{-2\pi D_\delta}} n^{-1/(2\delta+1)} .$$

If one substitutes the appropriate value of δ into (2.44), the following special cases, which will also be considered in the next section, are obtained:

Case (a): $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$, $i = 1,2,\ldots(\delta = 3/2)$

$$R_2 \sim \frac{Bp_o e^{-\pi\sqrt{n}/4} n^{-1/4}}{4\pi}$$

Case (b): $\sigma_i^2 = 1/i^2$, $i = 1, 2,\ldots(\delta = 3/2)$

$$R_2 \sim \frac{Bp_o e^{-\pi\sqrt{n}/8} n^{-5/4}}{2\sqrt{2}\,\pi}$$

Case (c): $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i$, $i = 1, 2,\ldots(\delta = 1/2)$

$$R_2 \sim Bp_o \sqrt{\frac{e}{2\pi}}\; e^{-n/e} n^{-3/2}$$

2.3.3 Asymptotic Evaluation of R_2 and R_1/R_2 for the Class of Prior Distributions Characterized by $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/2a^i$, $a > 1$, $i = 0, 1, 2,\ldots$

In order that we may evaluate R_2 we must first find a value for v_1 which asymptotically satisfies (2.12) which in this case reduces to

$$(2.45) \qquad \log\left[\sqrt{K}\prod_{i=0}^{\infty}(1+n/(2a^i))\right] = \sum_{i=0}^{\infty} n/(2a^i+2v_1n) .$$

Now using the Mean Value Theorem it can be shown that for any $m > 0$, and $0 < b < 1$,

$$(2.46) \qquad \frac{\log(1+mb)}{m\log(1/b)} \leq \sum_j \frac{b^j}{1+mb^j} \leq \frac{\log(1+m)}{m\log(1/b)} .$$

Setting $m = nv_1$ and $b = 1/a$ in (2.46) we obtain that

$$(2.47) \qquad \sum_{i=0}^{\infty} n/(2a^i+2v_1n) \sim \log n/2v_1 \log a .$$

Furthermore, if we integrate with respect to n the three terms in (2.46) between 0 and n, and set $b^i = 1/2a^i$ we obtain the result that

$$\log\left[\sqrt{K}\prod_{i=0}^{\infty}(1+n/(2a^i))\right] \sim \log^2 n/(2\log a) .$$

Substituting the preceeding equation and (2.47) into (2.45) we obtain the result that

$$v_1 \sim 1/\log n .$$

Applying (2.19) we have that

$$R_2 \sim \frac{\log n}{2} R_1 .$$

Now in order to find an order asymptotic result for R_2 we need to evaluate $k_n(v_1)$, which is defined in (2.15). For the problem under consideration $k_n(v_1)$ becomes

$$k_n(v_1) = -\log \prod_i [1+2nv_1/(2a^i)] + 2v_1 \log (K\pi[1+n/(2a^i)])$$

$$\sim -\frac{\log^2 (2nv_1)}{2\log a} + \frac{2v_1 \log^2 n}{2\log a}$$

$$\sim \frac{-\log^2 (2n/\log n)}{2\log a} + \frac{\log n}{\log a} .$$

Furthermore

$$k_n''(v_1) = -2 \sum_i [n/(2a^i+2v_1 n)]^2$$

$$\sim -2\left[\frac{\log (nv_1)}{v_1^2 \log a} + \frac{1}{v_1^2 \log a}\right]$$

$$\sim \frac{-2(\log \log n) \log^3 n}{\log a} .$$

Substituting these values into (2.18) yields

$$R_2 \sim Bp_o \frac{\exp\left[\frac{-\log^2 (2n/\log n)}{2\log a} + \frac{\log n}{\log a}\right]}{[4\pi \log n(\log \log n)/\log a]^{1/2}} . \tag{2.48}$$

2.3.4 The Asymptotic Evaluation of R_1 and R_1/R_2 for the Case in which $\sigma_i^2 = 0,\ i > N$.

In the finite dimensional case, the asymptotic solution of

$$\log K \prod_1^N (1+n\sigma_i^2) = \sum_1^N \frac{n\sigma_i^2}{1+2n\sigma_i^2 v_1}$$

is

$$v_1 \sim \frac{N}{2(\log K + N \log n + \sum_1^N \log \sigma_i^2)}$$

$$\sim \frac{1}{2\log n} \Rightarrow \frac{R_2}{R_1} \sim \log n \ ,$$

the result obtained in [10].

$$R_1 \sim \frac{C}{n^{1/2}(\log n)^{1/2}} \ .$$

2.4 EXACT EVALUATION OF SOME BAYES RISKS.

2.4.1 Introduction. In Section 2.3 the techniques developed in Section 2.2 were used to evaluate the Bayes risk for certain classes of the prior covariance matrix, Σ. Some of these cases were chosen because their Bayes risk could be evaluated exactly, thus permitting an assessment of the error introduced through the utilization of Laplace's asymptotic technique. Let us now perform these exact evaluations.

In Section 2.4.2 we shall consider the exact evaluation of R_1, R_2 and R_1/R_2 for the case in which $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$, $i = 1, 2, \ldots$ The numeric results obtained indicate that a careful application of the asymptotic technique yields useful estimates of the type II risk for sample sizes as small as $n = 32$.

In Section 2.4.3 a technique for evaluating the type II risk for the case $\sigma_i^2 = 1/i$ is considered.

In Section 2.4.4 corroborative results are obtained for the case $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i$ and in Section 2.4.5 some exact estimates are obtained for the case in which $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/2a^i$, $a > 1$, $i = 0, 1, 2, \ldots$

2.4.2 Exact Evaluation of R_1 and R_2 in the Case $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$, $i = 1, 2, \ldots$ Recall that

$$R_1 = Bp_o P[Y_1 \geq \log K \prod_i (1+n\sigma_i^2)]$$

where

$$Y_1 \equiv \sum_i W_i^2 \lambda_i$$

$$\lambda_i \equiv \frac{n\sigma_i^2}{1+n\sigma_i^2}$$

and

$$W_i \sim N[0,1] .$$

Therefore, in this case

$$\Phi_{Y_1}(s) = E\left(e^{-sY_1}\right) = \prod_i \left[\frac{1}{1+2s\lambda_i}\right] = \frac{\sinh\pi\sqrt{n}}{\pi\sqrt{n}} \left[\frac{\pi\sqrt{n(1+2s)}}{\sinh\pi\sqrt{n(1+2s)}}\right] .$$

Furthermore, we know that the derivative of the Jacobi theta function with respect to v is

$$\theta_3'(1/2,v) \equiv -2 \sum_{j=1}^{\infty} (j^2\pi^2)e^{-j^2\pi^2 v}(-1)^j$$

and satisfies the relationship (see [3], p. 77)

$$\int_0^{\infty} e^{-vt}\theta_3'(1/2,t)dt = \frac{\sqrt{v}}{\sinh\sqrt{v}} .$$

Making the appropriate substitutions we get

$$\Phi_{Y_1}(s) = \int_0^{\infty} e^{-sy} \frac{e^{-y/2}\theta_3'(1/2,y/2n\pi^2)}{2n\pi^2} \frac{\sinh\pi\sqrt{n}}{\pi\sqrt{n}} dy .$$

Hence the p.d.f. of Y_1 is

$$f_n(y) = \frac{e^{-y/2}\theta_3'(1/2, y/2n\pi^2)}{2n\pi^2} \frac{\sinh\pi\sqrt{n}}{\pi\sqrt{n}} ,$$

and hence, we may show that

$$(2.49) \qquad R_1 = \frac{2Bp_o}{\sqrt{K}} \sum_{j=1}^{\infty} (-1)^j \frac{[(\sqrt{K}\, \sinh\pi\sqrt{n})/\pi\sqrt{n}]^{-j^2/n}}{1+n/j^2} .$$

Using a similar substitution procedure we can show that (2.6) becomes

$$(2.50) \qquad R_2 = A(1-p_o)P[\Sigma\, W_i^2\sigma_i^2 \leq (\log\, K\pi(1+n\sigma_i^2))/n]$$

$$= A(1-p_o)\left[1 + 2\sum_{j=1}^{\infty} (-1)^j \left[\frac{\sqrt{K}\, \sinh\pi\sqrt{n}}{\pi\sqrt{n}}\right]^{-j^2/n}\right] .$$

In order to obtain some idea of the sample size needed for accurate asymptotic estimates and in order to observe the accuracy of the "order-asymptotic" expression for R_2 obtained in (2.44), some numeric calculations were made. In addition to numerically evaluating the expression for R_1 and R_2, obtained in (2.44), (2.49) and (2.50), the computer was also used to obtain numeric solutions for $v_1(n)$ in (2.12). This value of $v_1(n)$ was then substituted into (2.15), (2.16) and (2.18) in order to obtain what we define as the "best asymptotic solution" for R_2. These results are tabulated in Table 2.1. The closeness between the numeric values for (2.50) and (2.18) should give the reader some idea of how accurate the Laplace asymptotic solution can be.

n	Exact R_1 (2.49)*	Exact R_2 (2.50)	Best Asymptotic R_2 (2.18)	Order Asymptotic R_2 (2.44)
16	3.1×10^{-2}	6.779×10^{-2}	8.003×10^{-2}	1.6×10^{-3}
32	8.8×10^{-3}	2.048×10^{-2}	2.297×10^{-2}	3.8×10^{-4}
128	1.2×10^{-4}	2.955×10^{-4}	3.124×10^{-4}	3.3×10^{-6}
512	1.8×10^{-12}	4.982×10^{-8}	5.123×10^{-8}	3.1×10^{-10}
1024	1.3×10^{-11}	3.493×10^{-11}	3.556×10^{-11}	1.7×10^{-13}

Table 2.1, Values of R_1, R_2 for $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i^2$, $i = 1,2,\ldots$

*The number in parentheses refers to an equation.

2.4.3 Exact Evaluation of R_2 for the Case in which $\sigma_i^2 = 1/i^2$, $i = 1, 2, \ldots$ This problem can be shown to be identical to that of finding the lower tail probability of the Cramer Von Mises statistic. Using the results obtained in [1], p. 202, it can be shown that

$$(2.51) \qquad R_2 = A(1-p_o)P\left[\sum_i \frac{W_i^2}{\pi^2 i^2} \leq z\right]$$

$$= \frac{A(1-p_o)}{\pi\sqrt{z}} \sum_{j=0}^{\infty} (-1)^j \binom{-1/2}{j} (4j+1)^{1/2} e^{-\frac{(4j+1)^2}{16z}} K_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16z}\right],$$

where

$$z = \frac{\log K\pi(1+n\sigma_i^2)}{\pi^2 n}$$

and

$$K_{\frac{1}{4}}(t) = \frac{\sqrt{\pi}\,t^{1/4}}{2^{1/4}\Gamma(3/4)} \int_0^\infty e^{-t\cosh\theta}(\sinh\theta)^{1/2}d\theta \sim \frac{\sqrt{\pi}e^{-t}}{\sqrt{2}} \quad ;$$

and clearly this series converges rapidly as $n \to \infty$. However, since this problem is similar to that in the preceding section, no numeric calculations were made. If, however, one approximates (2.51) with its first term one obtains

$$R_2 \sim Ce^{-\pi\sqrt{n}/8} \cdot n^{3/16} \,,$$

which is order-asymptotic to the expression obtained in Case (b) of Section 2.3.2.

2.4.4 Exact Solution for R_1, R_2 for $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i$, $i = 1, 2, \ldots$ In this section we shall obtain the following:

$$(2.52) \qquad R_2 = A(1-p_o)\,\exp[-(\Gamma(n+1)/\sqrt{K})^{1/n}] \underset{\infty}{\sim} A(1-p_o)e^{-n/e}$$

$$(2.53) \qquad R_1 = A(1-p_o)\,\exp[-(\Gamma(n+1)/\sqrt{K})^{1/n}] \sum_{j=1}^{\infty} \frac{[\Gamma(n+1)/\sqrt{K}]^{j/n} n!}{(n+j)!} .$$

Furthermore, since

$$(2.54) \qquad \lim_{n\to\infty} \sum_{j=1}^{\infty} \frac{[\Gamma(n+1)/\sqrt{K}]^{j/n} n!}{(n+j)!} = \frac{1}{e-1}$$

we obtain that $R_1/R_2 \to 1/(e-1)$.

<u>Proof of (2.52)</u>: The cannonical product of $\Gamma(n+1)$ may be expressed as

$$\prod_i (1+\tfrac{n}{i})\, e^{-n/i} = \left[\frac{e^{-\gamma n}}{\Gamma(n+1)}\right]$$

where γ is Euler's constant. It follows that we may write (2.6) as

$$R_2 = A(1-p_o)P\left[\sum_i (W_i^2-1)\sigma_i^2 \le \frac{2}{n} \log\left[\frac{\sqrt{K}e^{-\gamma n}}{\Gamma(n+1)}\right]\right]$$

and letting

$$Y = \sum_i (W_i^2-1)\sigma_i^2$$

we obtain

$$\Phi_2(s) = E(e^{-Ys}) = \frac{1}{\prod_i (1+2s/i)e^{-2s/i}} = e^{2\gamma s}\Gamma(2s+1) .$$

But (see [5], p. 66)

$$\Gamma(2s+1) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-s(2\gamma+y)}\, \exp[-e^{-(\gamma+y/2)}]\, e^{-(\gamma+y/2)} dy ,$$

so it follows that the p.d.f. of Y is

$$f(y) = \tfrac{1}{2}e^{-(\gamma+y/2)}\, \exp[-e^{-(\gamma+y/2)}] \qquad -\infty < y < \infty .$$

And letting $Z = \gamma + \frac{Y}{2}$, we get the c.d.f. of Z

$$F(z) = e^{-e^{-z}} \qquad -\infty < z < \infty .$$

It follows that

$$R_2 = A[1 - p_o] \exp-\left[\frac{\Gamma(n+1)}{\sqrt{K}}\right]^{1/n} ,$$

the desired result.

<u>Proof</u> <u>of</u> <u>(2.53)</u>: Since, similarly, (2.4) may be written as

$$R_1 = Bp_o P\left[y \geq \frac{2}{n} \log \left[\frac{\sqrt{K}\, e^{-\gamma n}}{\Gamma(n+1)}\right]\right]$$

where

$$Y = \sum_i \left(\frac{Xi^2}{1+n\sigma_i^2} - 1\right) \sigma_i^2 ,$$

we obtain

$$\Phi(s) = E\left[e^{-sY}\right] = \frac{\pi(1 + \frac{n}{i})\, e^{-n/i}}{\pi(1 + \frac{n+2s}{i})} \exp\left(- \frac{n+2s}{i}\right)$$

$$= \frac{\Gamma(n+2s+1)}{\Gamma(n+1)e^{-2s\gamma}}$$

$$\frac{\Gamma(n+2s+1)}{\Gamma(n+1)e^{-2s\gamma}} = \int_{-\infty}^{\infty} \frac{e^{-2(z-\gamma)s}\, e^{-(n+1)z}\, e^{-e^{-z}}\, dz}{\Gamma(n+1)} ;$$

so $Z = \frac{Y}{2} + \gamma$ has p.d.f.

$$f(z) = \frac{e^{-(n+1)z}\, e^{-e^{-z}}}{\Gamma(n+1)} \qquad -\infty < z < \infty .$$

Now letting $h \equiv \left[\frac{\Gamma(n+1)}{\sqrt{K}}\right]^{1/n}$

we may write

$$R_1 = Bp_o P[Z \geq -\log h]$$

$$= Bp_o\left[1 - \sum_{j=0}^{n} \frac{h^j e^{-h}}{j!}\right]$$

$$= Bp_o \sum_{j=n+1}^{\infty} \frac{h^j e^{-h}}{j!} ,$$

which on substituting the defining expressions for h and K reduces to (2.53), the desired result.

Proof of (2.54): Since as $n \to \infty$ $h \sim n/e$ and

$$\frac{n!}{(n+j)!} \sim \frac{1}{(n+j)^j} ,$$

we obtain

$$\sum_{j=1}^{\infty} \frac{h^j n!}{(n+j)!} \sim \sum_{j=1}^{\infty} \left[\frac{1}{e}\right]^j / ((1+j/n)^j) \rightarrow \sum_{j=1}^{\infty} (\frac{1}{e})^j = \frac{1}{e-1} .$$

Some numeric calculations were made to asses the accuracy of the order-asymptotic estimate of R_2 obtained in Case (c) of Section 2.3.2. The results are listed in Table 2.2. Once again the A, B, p_o terms have been ignored.

n	R_1, (2.53)	R_2, (2.52)	R_2, (Case (c))	R_2/R_1, (2.52)/(2.53)
1	2.6×10^{-1}	3.68×10^{-1}	4.5×10^{-1}	.718
16	7.0×10^{-9}	1.11×10^{-3}	2.9×10^{-5}	.632
128	6.1×10^{-22}	1.02×10^{-21}	1.6×10^{-24}	.595

Table 2.2. Values of R_1, R_2 for $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/i$, $i = 1,2,\ldots.$

2.4.4. Exact Evaluation of R_2 the case $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/2\, a^i$, $a > 1$, $i = 0, 1, 2\ldots$.

We may reduce (2.6) to the form

$$R_2 = A(1-p_o)\, P\left[\sum_i W_i^2 \sigma_i^2 \leq \frac{\log K \prod_i [1+n\sigma_i^2]}{n}\right] ;$$

and letting $Y = \sum_i W_i^2 \sigma_i^2$, we get

$$\Phi_Y(s) = E\left[e^{-Ys}\right] = \frac{1}{\prod_{\nu=0}^{\infty}\left(1 + \frac{s}{a^{\nu}}\right)} .$$

Using residue calculus it can be shown that this is the generating function of the c.d.f.

$$F(x) = -C\left[e^{-x} + \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu} e^{-a^{\nu}x}}{\prod_{i=1}^{\nu}(a^i - 1)}\right] \qquad x \geq 0$$

$$= 0 \qquad x < 0$$

where

$$C = \prod_{\nu=1}^{\infty}(1 - a^{-\nu})^{-1}$$

(see [7], p. 350), and in our case

$$x = \frac{\log \sqrt{K}\prod_i(1+n\sigma_i^2)}{n} = \frac{2}{n}\log \sqrt{K}\prod_{i=0}^{\infty}\left(1 + \frac{n}{2a^i}\right) .$$

Thus in this case we wish to evaluate

$$(2.55) \quad R_2 = A(1-p_o)\left[1 - C\left[\left[\prod_{i=0}^{\infty}\left(1 + \frac{n}{2a^i}\right)\right]^{-2/n} + \sum_{\nu=1}^{\infty}\frac{(-1)^{\nu}\left[\prod_{i=0}^{\infty}\left(1 + \frac{n}{2a^i}\right)\right]^{\frac{-2a^{\nu}}{n}}}{\prod_{i=1}^{\nu}(a^i - 1)}\right]\right] .$$

Using a similar procedure it can also be shown that

$$(2.56) \quad R_1 = Bp_oP[\Sigma W_i^2\lambda_i \geq \log K \prod_i(1+n\sigma_i^2)]$$

$$= \frac{-Bp_o 2C}{\sqrt{K}} \quad \frac{\left[\sqrt{K}\prod_{i=0}^{\infty}\left(1 + \frac{n}{2a^i}\right)\right]^{-2/n}}{n+2} + \sum_{\nu=1}^{\infty}\frac{(-1)^{\nu}\left[\sqrt{K}\prod_{i=0}^{\infty}\left(1 + \frac{n}{2a^i}\right)\right]^{\frac{-2a^{\nu}}{n}}}{\prod_{j=1}^{\nu}(a^j - 1)(2 + \frac{n}{a^{\nu}})} .$$

For $A = B = 1$, $p_o = 1/2$ and $a = 2$, some numeric calculations were made. The results are listed in Table 2.3. Equation (2.48) refers to the order-asymptotic result for R_2 obtained in Section 2.3.3.

n	R_1,(2.56)	R_2,(2.55)	R_2/R_1,(2.55)/(2.56)	R_2,(2.48)
8	1.1×10^{-1}	2.0×10^{-1}	1.85	3.8×10^{-1}
128	4.7×10^{-4}	1.3×10^{-3}	2.78	2.2×10^{-3}
2048	1.0×10^{-9}	4.13×10^{-9}	4.05	3.0×10^{-9}

Table 2.3 Values of R_1, R_2 for $\sigma_{2i}^2 = \sigma_{2i-1}^2 = 1/2a^i$, $a = 2$, $i = 0, 1, 2, \ldots$

CHAPTER III

QUADRATIC LOSS

3.1 INTRODUCTION

Let us now extend our results to the case of quadratic loss. Let $X^1, \ldots, X^n$ be infinite dimensional independent normal variables with mean vector θ and covariance matrix I.

Let the null hypothesis be $\theta = 0$ and the alternative be $\theta \neq 0$. Let

$$p_o = \text{probability that } \theta = 0, \text{ and}$$

$$(1-p_o)P(\theta) = \text{prior distribution over } \{\theta, \theta = 0\},$$

where

$$P(\theta) \sim N[0,\Sigma] \,,$$

$\Sigma = \{\sigma_{ij}\}$ and $i,j = 1,\ldots,\infty$.

The losses associated with the type I and type II errors are respectively 1 and $\theta'A\theta$. _It is also assumed that_ A _and_ Σ _are positive definite_. In section (3.2) we assume

$$\text{(3.1)} \qquad Tr[A] < \infty \text{ and } Tr[\Sigma^2] < \infty \,.$$

In section (3.4) the $Tr[A] < \infty$ condition is relaxed.

In the previous chapter we developed asymptotic expressions for the type I and type II Bayes risks, R_1 and R_2 for the case of constant loss. In this chapter these results are used to obtain similar results for the case of quadratic loss.

In Section 3.2 a two stage asymptotic estimate of R_1 is obtained.for the case in which $Tr[A] < \infty$ and $Tr[\Sigma^2] < \infty$.

In Section 3.3 the asymptotic relation between R_1 and R_2 is developed. For the case $Tr[\Sigma] = \infty$, the relationship is asymptotically identical to that obtained in the constant loss case (see Theorem 2.1). If $Tr[\Sigma] < \infty$, relationships similar to those obtained for constant loss are still obtained.

In Section 3.4 a good upper bound for R_1 is obtained. This upper bound is introduced because it is easy to calculate, and requires $Tr[A\Sigma] < \infty$, $Tr[\Sigma^2] < \infty$, is derived using a clever technique and, finally, will permit a subsequent estimation of the accuracy of the two stage estimate of R_1.

In Section 3.5 the asymptotic evaluation of the upper bound for R_1 is obtained for the case in which Σ is diagonal and $\sigma_i^2 = 1/i^2$, $i = 1, 2, \ldots$. There are two cases of loss matrix considered, one in which $Tr[A] < \infty$ and one in which $Tr[A] = \infty$. In each case the asymptotic results differ but the order-asymptotic results are those which were obtained in Section 2.3 for the constant loss case.

Let us now restate the problem for the case of quadratic loss. Let $X \equiv (X_1/\sqrt{n}, X_2/\sqrt{n}, \ldots)$ be the mean vector of $X^1, \ldots, X^n$. Then $X_i \sim N[\sqrt{n}\theta_i, 1]$, $i = 1, 2, \ldots$ and $X \sim N[\theta, I/n]$. We shall use $f(x|\theta)$ to denote the density function of X. Furthermore, since the i^{th} mean $X_i/\sqrt{n}$ is sufficient for θ_i, we may determine the Bayes procedure by finding that set of x such that $L_1(x) = L_2(x)$ where

(3.2) $$L_1(x) = \frac{f(x|\theta=0)p_o}{p_o f(x|\theta=0) + (1-p_o)\int f(x|\theta)P(\theta)d\theta}$$, and

(3.3) $$L_2(x) = \frac{(1-p_o)\int \theta'A\theta\ f(x|\theta)P(\theta)d\theta}{p_o f(x|\theta=0) + (1-p_o)\int f(x|\theta)P(\theta)d\theta} .$$

Since the denominator is the same for each term, the problem may be reduced to finding

(3.4) $$\left\{x: \frac{f(x|\theta=0)}{\int \theta'A\theta\ f(x|\theta)P(\theta)d\theta} = \frac{(1-p_o)}{p_o} \equiv \frac{1}{\sqrt{K}} \right\} .$$

This in turn is equivalent to the problem of finding

(3.5) $$S_x = \left\{x: e^{-n/2|x|^2} = \frac{1}{\sqrt{K}} \int_\theta \frac{\theta'A\theta\exp[-\frac{1}{2}[\theta'\Sigma^{-1}\theta+n(x-\theta)'(x-\theta)]]d\theta}{|2\pi\Sigma|^{\frac{1}{2}}} \right\} .$$

Next, before proceeding we shall simplify the quadratic loss problem in the following manner. Since Σ is positive definite, one may obtain an orthogonal G which has the property that Σ = GDG, where D is diagonal and the i^{th} diagonal term of D is defined as σ_i^2 [if Σ is diagonal to begin with, then $\sigma_{ii} = \sigma_i^2$]. If, for the case in which Σ is not diagonal, one replaces Σ with D and A with G'AG, it can be shown that the acceptance-rejection boundary obtained in (3.5) will be identical to that of the boundary for the original problem. Furthermore, since $Tr[A] < \infty$ we have that

$Tr[G'AG\ ] < \infty$. <u>Since this transformation will make subsequent derivations simpler, we shall, without loss of generality, assume that Σ is diagonal</u>.

One should also note that (3.5) must be treated as a symbolic limit (since $|2\pi\Sigma| = 0$). In fact, (3.5) is evaluated for a finite number of dimensions and then a limit taken. The notation needed to indicate this, however, is unnecessarily confusing, and hence not used. Next, let

$$\Lambda \equiv n\Sigma(I+n\Sigma)^{-1} . \tag{3.6}$$

Since we have assumed Σ diagonal and positive definite, we also have that Λ is diagonal and positive definite. We denote the i^{th} diagonal element of Λ by λ_i. If we define the norm $||M||$ of a matrix M as $\sup_i |\lambda_i|$, where the λ_i are the eigenvalues of M, then $||\Lambda|| < 1$ for any n.

Finally, before proceeding with the derivation, one should also note that the conditions assumed in (3.1) make it possible that $Tr[\Sigma] = \infty$. This possibility also arose in the constant loss case and necessitated the normalization of expressions containing Σ with an exponential term; i.e., instead of writing $|I+n\Sigma|$, which equals ∞ if $|\Sigma| = \infty$, one must write

$$\lim_{(\cdot)}[|I+n\Sigma| e^{-nTr[\Sigma]}] = |(I+n\Sigma)\ e^{-n\Sigma}| ;$$

$(\cdot)$ denotes the limit is taken in terms of the dimension of the parameter space.

Let us now continue with the derivation. If we make use of the identity

$$\theta'\Sigma^{-1}\theta + n(X-\theta)'(X-\theta) = (\theta-\Lambda X)'n\Lambda^{-1}(\theta-\Lambda X) + nX'[I-\Lambda]X ,$$

and perform the transformation $Y = \sqrt{n}X$, we may evaluate the integral in (3.5) to obtain that the Bayes critical region is

$$(3.7)\quad S = \left\{y: C_n \le e^{\frac{1}{2}[y'\Lambda y - n\mathrm{Tr}[\Sigma]]}[(\Lambda y)'A(\Lambda y)+\mathrm{Tr}[A\Lambda]]\right\},$$

where

$$(3.8)\quad C_n = n\sqrt{K}\prod_i [(1+n\sigma_i^2)e^{-n\sigma_i^2}]^{\frac{1}{2}};$$

and, since Λ is diagonal,

$$(3.9)\quad \mathrm{Tr}[A\Lambda] = \sum_{i=1}^{\infty} a_{ii}\lambda_i .$$

Thus, we may express the type I risk as

$$(3.10)\quad R_1 = p_o P[Y \in S \mid \theta = 0] .$$

Next, let us define

$$(3.11)\quad U = Y'\Lambda Y - n\mathrm{Tr}[\Sigma], \qquad U \ge -n\mathrm{Tr}[\Sigma]$$

$$V = (\Lambda Y)'A(\Lambda Y), \qquad V \ge 0$$

and

$$(3.12)\quad S_1^* \equiv \left[\{(u,v): C_n \le e^{u/2}[v+\mathrm{Tr}[A\Lambda]]\} \cap \{u: u > -n\mathrm{Tr}[\Sigma]\} \cap \{v: v > 0\}\right] .$$

It should be noted that even though U and V are dependent, the infinite dimensional nature of Y and the assumption that A and Σ are positive-definite are sufficient conditions to guarantee that the joint distribution of (U,V) actually has a nonzero p.d.f. for all $(U,V) \in S_1^*$. We may now express the type I risk as

$$R_1 = p_o P[(U,V) \in S_1^* | \theta = 0].$$

Similarly, if we denote the complement of the critical region by $\bar{S}$, we may integrate out θ and express the type II risk as

$$(3.13)\quad R_2 = (1-p_o) \int_{\bar{S}} \frac{[(\Lambda y)'A(\Lambda y)+Tr[A\Lambda]]\exp[(-\tfrac{1}{2})y'[I-\Lambda]y-nTr(\Sigma)/2]}{[|2\pi(I+n\Sigma)| \quad \exp(-n\ Tr(\Sigma))]^{\frac{1}{2}}} dy$$

The bivariate Laplace transforms that will be needed to obtain the asymptotic risks are defined as follows. In order to obtain R_1 we shall define

$$(3.14)\qquad \Phi_1(s,t) \equiv E(e^{-Us-Vt}|\theta=0)$$

$$= \frac{e^{snTr(\Sigma)}}{|I + 2s\Lambda+2t\Lambda A\Lambda|^{\frac{1}{2}}} \quad ,$$

where $Re(s) < 0$, $Re(t) < 0$

[The reader is reminded that for the case $Tr(\Sigma) = \infty$ one must take the appropriate limit in the parameter space in order to make (3.14) a valid expression.]

Next, in order to obtain the expression for R_2 we shall define the bivariate Laplace transform

$$(3.15)\quad \Phi_2(g,t) \equiv \int \frac{e^{-gU(y)-t\ V(y)}\ e^{-\frac{1}{2}[y'[I-\Lambda]y+n\ Tr(\Sigma)]}}{(|2\pi(I+n\Sigma)|\exp[-n\ TR(\Sigma)])^{\frac{1}{2}}} dy \quad .$$

Here we require

$$Re(g) > 0 \text{ and } Re(t) < 0.$$

Note that by inspection of (3.14) and (3.15) we have the subsequently useful result that

$$\Phi_2(g,t) = \Phi_1(g-\tfrac{1}{2},t)/C_n \tag{3.16}$$

3.2 TWO STAGE ASYMPTOTIC ESTIMATE OF R_1.

We first apply a two dimensional version of Laplace's asymptotic formula which was used in Chapter II Section 2.2. For the two dimensional case we obtain

$$R_1 = \int_{S^*} \int_{t^*-i\infty}^{t^*+i\infty} \int_{s_1-i\infty}^{s_1+i\infty} \Phi_1(s,t)e^{su+tv}\, ds\, dt\, du\, dv. \tag{3.17}$$

In order that the above integral exist we must impose the conditions s_1, $t_1 < 0$ and

$$0 < ||-2\, s_1\, \Lambda\, -2t^*\, \Lambda A \Lambda|| < 1. \tag{3.18}$$

The later condition is sufficient to guarantee that the path of integration stops away from the singularities of $\Phi_1(s,t)$. The two conditions in (3.18) are sufficient to guarantee a subsequently required condition that $0 < ||-t^*\Lambda A\Lambda|| < 1$.

Next we utilize the Fubini Theorem to show that it is possible to interchange the order of integration in (3.18) This procedure was implicitly performed in the constant loss case and led to the constant loss version of (3.17), namely (2.21). If we could continue to follow the constant loss procedure, which lead to an asymptotic expression for R_1, we would next integrate over S_1^* and then proceed to obtain a bivariate version of the asymptotic normal expression for R_1 (see (2.28) for the 1 dimensional version of this procedure). This, however, will not work for the quadratic loss case due to the fact that the integrand in the t direction does not become asymptotically normal for the case in which $Tr[A] < \infty$. Because of this difficulty a simple bivariate asymptotic expression for R_1 could be obtained. Instead, Laplaces asymptotic technique, which was used for the constant loss case, is only applied to the integral over s. The integral over u is evaluated directly and an asymptotic expression for the integral over v and t is obtained. The main result of this section, an asymptotic expression for R_1 is stated in Theorem 3.1. While this estimate is only carried out for the case $Tr[A] < \infty$ it is conjuctured that similar proceedures could also be used to obtain estimates of R_1 for the case in which one only required $Tr[A\Sigma] < \infty$ and $Tr[\Sigma^2] < \infty$. Let us now carry out this program in detail. First let us justify the interchange of the order of integration i.e., let us show that (3.17) is absolutely integrable. To do this we first note that

$$\int_{S_1^*} \int_{s_1-i\infty}^{s_1+i\infty} \int_{t^*-i\infty}^{t^*+i\infty} |\Phi(s,t)\ e^{su+tv}| dsdtdudv$$

$$= \left[\int_{S_1^*} e^{s_1 u + t^* v} dudv \right] \left[\int_{s_1-i\infty}^{s_1+i\infty} \int_{t^*-i\infty}^{t^*+i\infty} |\Phi(s,t)| dsdt \right] .$$

Now since $s_1, t_1 < 0$ we have that

$$\int_{S_1^*} e^{s_1 u + t^* v} du dv < \infty \quad .$$

Thus in order to interchange the order of integration in (3.17) we need only show that

$$(3.19) \qquad \int_{s_1 - i\infty}^{s_1 + i\infty} \int_{t^* - i\infty}^{t^* + i\infty} |\Phi(s,t)| ds dt < \infty \ .$$

We shall show that this holds by finding an integrable function which dominates the integrand of (3.19).

First, for reasons which will become subsequently clear we choose s_1 such that

$$(3.20) \qquad Tr[B_o^{-1}\Lambda - n\Sigma] = 2 \log [C_n / Tr(\Lambda A \Lambda) + Tr(A\Lambda))]$$

$$(3.21) \qquad B_t = I + 2s_1\Lambda + 2(1+2s_1)t\Lambda A\Lambda$$

<u>Lemma 3.1</u> <u>The value of</u> s_1 <u>which satisfies (3.20) has the properties that</u>, <u>given</u> $\epsilon > 0$, <u>for sufficiently large</u> <u>n</u>,

(a) $-\frac{1}{2} < s_1 < -\frac{1}{4} + \epsilon$ and is monotone decreasing in n

(b) for $n \to \infty$, $(1+2s_1)n \uparrow \infty$,

<u>Proof</u>: $Tr(A\Lambda) \to Tr(A)$, $Tr[\Lambda A\Lambda] \to Tr[A] < \infty$ and since all of the matricies which occur in (3.20) are diagonal we may apply the results

obtained for the bounds of s_1 in the constant loss case [see Lemma 2.1, 2.2, 2.3, and 2.4]. The conclusion of these lemmas also proves Lemma 3.1.

Corrolary 3.1

Condition (a) of Lemma 3.1 in conjunction with (3.18) imply that given $\epsilon > 0$, for sufficiently large n,

$$\tfrac{1}{2} + 2\epsilon > \| -2t^{*}\Lambda A \Lambda \| > 0 \ .$$

Next let us perform the transformation

$$2t = (1+2s_1)(t_1+i\tau)$$

$$s = s_1 + i\sigma\, \gamma(n)$$

where

$$\gamma(n) = 2/[\mathrm{Tr}[B_{t_1}^{-1}\Lambda]^2]^{\frac{1}{2}}$$

and

$$\text{(3.22)} \qquad B_{t_1} = I + 2s_1\Lambda + 2(1+2s_1)\, t_1 \Lambda A \Lambda \ .$$

Then we may rewrite (3.19) as

$$\text{(3.23)} \qquad A_n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \frac{d\sigma\, d\tau}{|I + 2i[\sigma P + \tau Q]|^{\frac{1}{2}}} \right| ,$$

where

$$A_n \equiv \frac{2\gamma(n)}{(1+2s_1)B_{t_1} e^{-n\Sigma}|^{\frac{1}{2}}} < \infty$$

$$P \equiv \gamma(n)^{-1} B_{t_1}^{-\frac{1}{2}} \Lambda \ B_{t_1}^{-\frac{1}{2}}$$

$$Q \equiv (1+2s_1) B_{t_1}^{-\frac{1}{2}} \Lambda A \Lambda \ B_{t_1}^{-\frac{1}{2}}$$

For future reference we note that P,Q are symmetric,

$$Tr[P]^2 = 1, \ ||P|| \to 0$$

$$\text{and } Tr[B_{t_1}^{-1}\Lambda]^2 \underset{n}{\to} \infty \ .$$

Furthermore, using an unpublished result obtained by H. Rubin [which is proved in the appendix] we obtain that

$$Tr[Q] \to Tr[(I+2t_1 \ A\ )^{-1} \mathbf{A}]$$

and since we may choose a constant such that $-1 < \ 2t_o \ ||A|| < 2 \ t_1 ||A|| < 0$ we have that there exists a C' such that for all n, $Tr[Q] < C'$.

Next, in order to simplify the argument, let us transform the problem into polar coordinates by letting

$$\sigma = \rho \cos \theta$$

$$\tau = \rho \sin \theta \ .$$

Then the integrand in (3.23) may be expressed as

$$\left|\frac{\rho}{|I+i\rho R(\theta)|^{\frac{1}{2}}}\right| \tag{3.24}$$

where

$$R(\theta) \equiv (\cos\theta)\,P + (\sin\theta)Q \quad .$$

Clearly, for $|\rho| < k < \infty$, k uniformly bounds (3.24). Thus we need only show that (3.24) is dominated by an integrable function for large ρ .

<u>Lemma 3.2</u> <u>There exists an N, $\alpha > 0$ such that for $n > N$</u>

$$\left|\frac{\rho}{|I+i\rho R(\theta)|^{\frac{1}{2}}}\right| = 0\left[\frac{1}{\rho^{1+\alpha}}\right] \tag{3.25}$$

<u>Proof</u>:

First, noting that P,Q are symmetric we define $|r_k(\theta)|$ to be the k^{th} largest characteristic root of $R(\theta)$ in magnitude. Since for all n, $Tr[Q] < C'$, there exists a $0 < C < \infty$ such that $Tr[Q]^2 < c^2$ and $||Q|| < c$. Next we break the proof up into two mutually exhaustive cases and prove the lemma for each case.

<u>Case</u> (a) $\{\theta: \underline{|\sin\theta| \geq 1/(4(1+c))}\}$.

Since $Tr[Q] < \infty$ for all n, and converges to the finite trace of a positive definite matrix, we have the result [assuming that the rank of A is greater than 5] that there are at least 5 q_k which must also converge to positive limits. Thus we may choose a $q_o > o$ such that for all n,

$$\|P\| \leq \frac{q_o}{8(1+c)} < \frac{q_k}{8(1+c)} \quad , \quad k = 1 \ldots 5 \ .$$

Now by consideration of the four cases $\cos\theta \gtrless 0$, $\sin\theta \gtrless 0$ it can be shown that

$$(3.26) \qquad |r_k(\theta)| \geq |\sin\theta| \; q_k - \|P\| \ .$$

Thus for case (a) we have the result that

$$|r_k| > \frac{1}{4(1+c)} \left[q_k - \frac{q_o}{2} \right] > \frac{q_o}{8(1+c)} \equiv \eta > 0$$

But then

$$\left| \left| I + i\,\rho R(\theta) \right|^{\frac{1}{2}} \right| \geq (1+\eta^2\rho^2)^{\frac{5}{4}} \ .$$

Thus, for case (a) we have demonstrated (3.25) ($\alpha = \frac{1}{2}$).

<u>case</u> (b) $\{\theta : \underline{|\sin\theta| < 1/(4(1+c))}\}$.

Using considerations similar to those needed to obtain (3.26) it can be shown that

$$(3.27) \qquad \|R(\theta)\| \leq |\sin\theta| \; \|Q\| + \|P\| \ .$$

Next, using the property that $\|P\| \to 0$, we choose N such that for $n > N$

$$\|P\| < 1/(8(1+c)) \ .$$

Thus for $n > N$ we obtain the result that

$$\|R(\theta)\| \le \frac{c}{4(1+c)} + \frac{1}{8(1+c)} \le \frac{1}{4} \quad .$$

Next we note that if $|\sin\theta| < 1/(4(1+c)) < \frac{1}{4}$, then

$$|\cos\theta|^2 > 1 - [\tfrac{1}{4}]^2 \; .$$

From this it follows that

$$\begin{aligned} \mathrm{Tr}[R(\theta)]^2 &= \sin^2\theta + \mathrm{Tr}[Q]^2 + 2\sin\theta\cos\theta\,\mathrm{Tr}[PQ] \\ &\quad + \cos^2\theta\,\mathrm{Tr}[P]^2 \\ &\ge [\,|\cos\theta|\sqrt{\mathrm{Tr}[P]^2} - |\sin\theta|\sqrt{\mathrm{Tr}[Q]^2}\,]^2 \\ &\ge [1-1/16 - c/(4(1+c))]^2 \\ &\ge 9/16 \quad . \end{aligned}$$

But now we may utilize the matrix form of the result that was applied in the preceeding chapter subsequent to (2.32) namely,

$$\begin{aligned} &\|I + i\rho R(\theta)|^{\frac{1}{2}}| \\ &= \prod_{k=1}^{\infty} (1+\rho^2 r_k^2(\theta))^{\frac{1}{4}} \\ &\ge (1+\rho^2\|R(\theta)\|^2)^{\mathrm{Tr}[R(\theta)]^2/(4\|R(\theta)\|^2)} \\ &\ge (1+\rho^2/16)^{9/4} \end{aligned}$$

Thus for case (b) we have demonstrated (3.25) ($\alpha = 5/2$).

Thus we have proved Lemma 3.2. Thus the order of integration in (3.17) may be interchanged.

We may now state the main result of this section.

Theorem 3.1 *If* $Tr[A] < \infty$ *and* $Tr[\Sigma^2] < \infty$, *then as* $n \to \infty$

$$R_1 \sim \Phi(s_1,0)\left[\frac{C_n[1+2s_1]}{s_1 \ell(n)}\right]^{2s_1} \cdot d \ .$$

C_n *is defined in* (3.8), s_1 *satisfies* (3.20) *and*

(3.28) $$\ell(n) = 2\,[Tr[B_o^{-1}\Lambda]^2]^{\frac{1}{2}} \to \infty \ .$$

B_o *is defined in* (3.21).

Furthermore, if

(a) $(1+2s_1) \to 0$, *then* $d = \sqrt{\pi/2}\ (\tfrac{1}{2})\ Tr[A]$.

If

(b) $(1+2s_1) \not\to 0$, *then* $d = \sqrt{\pi/2}\ [-s_1]^{2s_1-1}\int_0^\infty [2v+(1+2s_1)Tr[A]]^{-2s_1-1}dF(v)$,

where $dF(v)$ *is the pdf of the random variable* $\tfrac{1}{2}X'AX$, *and* $X \sim N[0,I]$.

Proof:

In order to determine the asymptotic expression for R_1 we first integrate (3.17) over u to obtain

$$R_1 = \frac{1}{(2\pi i)^2}\int_0^\infty \int_{t^*-i\infty}^{t^*+i\infty} \int_{s_1-i\infty}^{s_1+i\infty} \frac{e^{tv}\Phi(s,t)}{s}\left[\frac{C_n}{v+Tr(A\Lambda)}\right]^{2s} ds\,dt\,dv \ .$$

Next, performing the substitution

$$t = (1+2s_1)(t_1+i\tau)$$
$$v = \nu/(1+2s_1)$$
$$s = s_1+i\,\sigma/\ell(n) \ ,$$

where $\ell(n)$ is defined in (3.28), we obtain

$$R_1 = A_n \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_n(\nu)\; J_n(\tau,\nu)\; S_n(\tau,\nu,\sigma)\; d\tau d\sigma d\nu \tag{3.29}$$

where

$$A_n = \frac{[C_n(1+2s_1)]^{2s_1}\Phi(s_1,0)}{s_1\ell(n)}$$

$$H_n(\nu) = [\quad + (1+2s_1)\,\mathrm{Tr}[A\Lambda]]^{-2s_1}$$

$$J_n(\tau,\nu) = \frac{1}{2\pi}\;\frac{e^{\nu(t_1+i\tau)}}{|I+2(1+2s_1)(t_1+i\tau)B_o^{-1}\Lambda A\Lambda|^{\frac{1}{2}}} \tag{3.30}$$

$$S_n(\tau,\nu,\sigma) = \frac{1}{2\pi}\;\frac{e^{\ell(\sigma,\tau,\nu)}}{1+i\sigma/(s_1\ell(n)}$$

$$\ell(\sigma,\tau,\nu) = \log\left[\frac{\Phi(s_1+i\sigma/\ell(n),\;(1+2s_1)(t_1+i\tau))}{\Phi(s_1,\;(1+2s_1)(t_1+i\tau))}\right]$$

$$+\; 2(s-s_1)\;\log\left[\frac{C_n(1+2s_1)}{\nu+(1+2s_1)\,\mathrm{Tr}(A\Lambda)}\right].$$

Thus we have reduced the proof of Theorem 3.1 to showing that

$$\int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_n(\nu)\; J_n(\tau,\nu)\; S_n(\tau,\nu,\sigma)\; d\tau d\sigma d\nu \to d$$

To do this we shall use the dominated convergence theorem.

First we show domination. In order to do this we write

$$(3.31)\qquad |H_n(\nu)\ J_n(\tau,\nu)\ S_n(\tau,\nu,\sigma)|$$

$$= \frac{e^{\nu t_1}}{(2\pi)^2} \frac{[\nu+(1+2s_1)\mathrm{Tr}[A\Lambda]]^{-2s_1}}{[1+[\sigma/(s_1\ell(n))]^2]^{\frac{1}{2}}\Phi(s_1,t_1)} \cdot \frac{1}{||I+i(\sigma P+\tau Q)|^{\frac{1}{2}}|}$$

$$= A_n \frac{e^{\nu t_1}[\nu+(1+2s_1)\mathrm{Tr}[A\Lambda]]^{-2s_1}}{||I+i(\sigma P+\tau Q)|^{\frac{1}{2}}|}$$

Here A_n is defined in (3.23).

Now, clearly, the domination argument that was used to show that the order of integration could be interchanged can now also be used to show that the (σ,τ) variables can be dominated by an integrable function. Next, in order to dominate the v variables, we note that since $\mathrm{Tr}[A] < \infty$ we may choose t_1 so that it is bounded away from zero. Let $t_o \equiv \sup_n(t_1) < 0$. Next, recalling that s_1 is monotone decreasing in n we define $s_o = s_1\big|_{n=n_o}$. Thus since $\mathrm{Tr}[A\Lambda] \uparrow \mathrm{Tr}[A]$ we have that

$$(3.32)\qquad e^{\nu t_1}\ [\nu+(1+2s_1)\mathrm{Tr}[A\Lambda]]^{-2s_1}$$

$$\leq\ e^{\nu t_o}\ [\nu+(1+2s_o)\mathrm{Tr}[A]]\ .$$

But now combining (3.31) and (3.32) we have dominated the integrand in (3.29) and need only show that it converges to a limit.

Lemma 3.3 For fixed τ, v and σ,

(a) $$S_n(\tau,\nu,\sigma) \to \frac{e^{-\sigma^2/4}}{2\pi} \equiv S(\sigma)$$

(b) $$J_n(\tau,\nu) \to \frac{1}{2\pi} \frac{e^{\nu(t_1+i\tau)}}{|I+2(t_1+i\tau)A|^{\frac{1}{2}}} \equiv J(\tau,\nu)$$

(c) case (1) $(1+2s_1) \to 0$

$$H_n(\nu) \to \nu = H(\nu)$$

(c) case (2) $(1+2s_1) \not\to 0$

$$H_n(\nu) \to [\nu+(1+2s_1)Tr(A)]^{-2s_1}$$

Proof: (c) is obvious. In order to prove (a) we shall show that

$$(3.33)\ \ell(\sigma,\tau,\nu) = [\ell(\sigma,\tau,\nu) - \ell(\sigma,it_1,\nu)] + \ell(\sigma,it_1,\nu)$$
$$\to \qquad 0 \qquad\qquad -\sigma^2/4 \quad ,$$

$\ell(\sigma,\tau,\nu)$ defined in (3.30), we must first show that $\ell(n) \to \infty$, where $\ell(n)$ is defined in (3.28). Now since Λ is diagonal and $n(1+2s_1) \to \infty$, we may write

$$Tr[B_o^{-1}\Lambda]^2 = \sum_{i=1}^{\infty} [\lambda_i/(1+2s_1\lambda_i)]^2$$

$$= \sum_{i=1}^{\infty} [n\sigma_i^2/(1+(1+2s_1)n\sigma_i^2)]^2$$

$$\to \infty .$$

Thus $\ell(n) \to \infty$.

Next let us show that

(3.34) $$\ell(\sigma, it_1, \nu) \to -\sigma^2/4 .$$

To do this we first write

$$\ell(\sigma, it_1, \nu) = \log\left[\frac{\Phi(s_1+i\sigma/\ell(n),0)}{\Phi(s_1,0)}\right] + 2(i\sigma/\ell(n))\log\left[\frac{C_n(1+2s_1)}{\nu+(1+2s_1)\mathrm{Tr}(A\Lambda)}\right].$$

Next, noting that

$$\Phi(s_1+i\sigma/\ell(n),0)/\Phi(s_1,0) = 1/\left|I+(2i\sigma/\ell(n))B_o^{-1}\Lambda\right|^{\frac{1}{2}},$$

we obtain

$$\ell(\sigma, it_1, \nu) = -\tfrac{1}{2}\log\left[\prod_k [1-W_k(n)]\exp[W_k(n)]\right]$$

$$- \frac{i\sigma}{\ell(n)}\left[\mathrm{Tr}[B_o^{-1}\Lambda - n\Sigma] - 2\log\left[\frac{C_n(1+2s_1)}{\nu+(1+2s_1)\mathrm{Tr}[A\Lambda]}\right]\right],$$

where

$$W_k(n) = \frac{-2i\sigma}{\ell(n)} \cdot \frac{\lambda_k}{1+2s_1\lambda_k}$$

and

$$\sum_k W_k^2(n) = -\sigma^2.$$

But if we substitute in the s_1 condition defined in (3.20) and follow the procedures of Lemma 2.5, we obtain, as $n \to \infty$, that

$$\ell(\sigma,it,\nu) = -\tfrac{1}{2}\log\left[\prod_k [1-W_k(n)]\exp[W_k(n)]\right]$$

$$-\frac{2i\sigma}{\ell(n)}\log\left[\frac{\nu+(1+2s_1)\mathrm{Tr}[A\Lambda]}{\mathrm{Tr}[\Lambda A\Lambda]+(1+2s_1)\mathrm{Tr}[A\Lambda]}\right] \to \sigma^2/4 \ .$$

Thus we have proved (3.34).

Now we shall demonstrate that for fixed σ,τ, and ν

(3.35) $$\ell(\sigma,\tau,\nu) - \ell(\sigma,it_1,\nu) \to 0 \ .$$

In order to show this we first note that

$$\ell(\sigma,\tau,\nu) - \ell(\sigma,it_1,\nu)$$

$$= \ell(\sigma,\tau,\nu) - \ell(\sigma,it_1,\nu) - \ell(0,\tau,\nu) + \ell(0,it_1,\nu)$$

$$= \int_{s_1}^{s_1+i\sigma/\ell(n)} \int_{t_1}^{t_1+i\tau} \ell_{st}(z,w)\,dz\,dw,$$

$$= \frac{-t\sigma}{\ell(n)} \int_0^1 \int_0^1 \ell_{st}\left(s_1+\frac{\lambda i\sigma}{\ell(n)},\ t_1+\eta it\right) d\eta\, d\lambda \ ,$$

where

$$\ell_{st}(z,w) = (1+2s_1)\mathrm{Tr}[B_w^{-1}\Lambda B_w^{-1}\Lambda A\Lambda]$$

and

$$B_w = [I+2z\Lambda+2(1+2s_1)w\Lambda A\Lambda].$$

Now for fixed σ, t and $0 < \lambda < 1$, $0 < \eta < 1$, if we can show that

$$\ell_{st}(s_1 + i\sigma\lambda/\ell(n), t_1 + \eta it)/\ell(n) \to 0 , \tag{3.36}$$

we will have shown (3.35).

Since $(1+2s_1)n \to \infty$, we have the result that

$$(1+2s_1)\ell(n) = 2\left[(1+2s_1)^2 \sum_{i=1}^{\infty} \left[\frac{n\sigma_i^2}{1+(1+2s_1)n\sigma_i^2}\right]^2\right]^{\frac{1}{2}}$$

$$\to \infty .$$

Thus, if we can show that $\exists\ 0 < M < \infty$, such that

$$|(1+2s_1)\ell_{st}(s_1 + \lambda i\sigma/\ell(n), t_1 + \eta it)| < M , \ \forall\ n , \tag{3.37}$$

then we are done. But

$$|\ell_{st}(s_1 + i\sigma\lambda/\ell(n), t_1 + \eta it)| \le (1+2s_1)\ell_{st}(s_1, t_1) \tag{3.38}$$

$$\to \mathrm{Tr}[(I+2t_1 A)^{-2} A] .$$

The convergence to a uniformly bounded limit in the above statement is obtained using the same procedure that was used subsequent to (3.23) to prove that $\mathrm{Tr}[Q] < C'$.

So now we have proved (3.35), and hence have proved (a) of Lemma 3.3.

Finally we must prove (b) of Lemma 3.3.

Let us now define Q_n' as $Q_n' \equiv (1+2s_1)B_o^{-\frac{1}{2}}\,\Lambda A \Lambda B_o^{-\frac{1}{2}}$.

Now since the square root of the positive operator B_o is a self adjoint positive operator we have that Q_n' is self adjoint. Therefore, using the procedure of Theorem A-1, it can be shown that

$$\mathrm{Tr}[Q_n'] \to \mathrm{Tr}[A] < \infty \quad ,$$

and that q_i^n, the eigenvalues of Q, are real, converge to q_j the eigenvalues of A. But these are sufficient conditions to guarantee that for tixed τ, t_1,

$$(3.39) \qquad |I+2(1+2s_1)(t_1+i\tau)B_o^{-1}\Lambda A\Lambda| = \prod_{i=1}^{\infty} (1+2(t_1+i\tau)q_j^n)$$

$$\to \prod_{j=1}^{\infty} (1+2(t_1+i\tau)q_j) \; .$$

Thus we have completely proved Lemms 3.3.

Finally, in order to show convergence for the whole integrand in (3.31), we note that for fixed σ, τ, and ν the three variables in the limit $H(\cdot)$, $J(\cdot)$, $S(\cdot)$ are bounded. Thus, since

$$|J_nS_nH_n-JSH| \le |J_nS_n|\;|H_n-H| + |HJ_n|\;|S_n-S| + |SH|\;|J_n-J|$$

$$\to 0,$$

we have the desired result.

3.3 THE ASYMPTOTIC RELATIONSHIP BETWEEN R_1 AND R_2:

Using the results obtained in (3.16) that $\Phi_2(g,t) = \Phi_1(g-\frac{1}{2},t)/C_n$ we may rewrite the expression for R_2 obtained in (3.13) as

$$(3.40) \quad R_2 = \frac{(1-P_o)}{(2\pi i)^2} \int_{\bar{S}} \int_{s_1-i\infty}^{s_1+i\infty} \int_{t_1-i\infty}^{t_1+i\infty} \frac{\Phi_1(s,t)[Z+Tr[A\Lambda]]}{C_n} e^{(s+\frac{1}{2})w+tz} \, dwdzdsdt$$

where

$$(3.41) \quad \bar{S} = \{(w,z): C_n \geq e^{w/2} [z+Tr[A\Lambda]]\} \cap \{w: w > -n\, Tr(\Sigma)\} \cap \{z: z > 0\} .$$

Now once again Fubini's Theorem can be applied to show that the order of integration in (3.40) may be interchanged. We shall now break up this development into two cases. In the first case, $Tr[\Sigma] = \infty$, a rather elegant transformation permits us to obtain the relationship in a simple closed form. In the second case, $Tr[\Sigma] < \infty$, the result is asymptotically the same as that of the first case and can be proved in a straight forward, albeit messy, manner.

For this reason the case $Tr(\Sigma) < \infty$ is omitted and only an outline of the necessary steps presented.

Theorem 3.2 If $Tr(\Sigma) = \infty$, then

$$(3.42) \quad R_2 \sim \frac{-2s_1}{2s_1+1} R_1$$

where s_1 satisfies (3.20).

Proof: For the case $Tr(\Sigma) = \infty$, (3.41) reduces to

$$\bar{S}_\infty = \{(w,z): C_n \geq e^{w/2}\,[z+Tr[A\Lambda]]\} \cap \{z: z > 0\}.$$

Next, interchanging the order of integration and performing the substitutions

$$[z+Tr[A\Lambda]]\; e^{(s+\frac{1}{2})w}/C_n = e^{sv}\;, \qquad (3.43)$$

$$z = v,$$

we obtain the remarkable result that

$$\int_{\bar{S}_\infty} \frac{[z+Tr(A\Lambda)]}{C_n}\, e^{(s+\frac{1}{2})w+tz}\, dwdz$$

$$= \frac{(-s)}{(s+\frac{1}{2})} \int_{S_1^*} e^{su+tv} dudv \;.$$

Here S_1^* is defined in (3.12). Thus we may rewrite (3.40) as

$$R_2 = \frac{1}{(2\pi i)^2} \int_{S_1^*} \int_{t_1-i\infty}^{t_1+i\infty} \int_{s_1-i\infty}^{s_1+i\infty} \left(\frac{-s}{s+\frac{1}{2}}\right) \Phi_1(s,t)\; e^{su+tv} dsdtdudv.$$

This, of course, is almost the expression for R_1 obtained in (3.17) except for the modification introduced by $(-s/(s+\frac{1}{2}))$. Furthermore, since $s/(s+\frac{1}{2})$ is slowly varying with respect to the rest of the integrand, we may apply the procedure used in Theorem 3.1 to obtain (3.42).

In order to sketch the proof of Theorem 3.2 for the case $Tr(\Sigma)<\infty$ we first note that the acceptance region, defined in (3.41), for the case $Tr(\Sigma) < \infty$, is a compact set whereas for the case $Tr(\Sigma) = \infty$

it is not (see Section 4.2). Therefore when one performs the transformation defined in (3.43) one does not convert $\overline{S}$ into S_1^* but rather a subset of it. It turns out, however, that Theorem 3.2 is still asymptotically valid since the region in which the transforms differ (for the cases $\mathrm{Tr}[\Sigma] = \infty$ vs. $\mathrm{Tr}(\Sigma) < \infty$) is probabilistically negligible. Thus the proof for $\mathrm{Tr}(\Sigma) < \infty$ consists of making the transformation defined in (3.43) and then showing that the excess region obtained in the transform is negligible.

3.4 AN UPPER BOUND FOR R_1. Now if the boundary of the (u,v) region defined in (3.12) were linear, then because of the nature of (u,v) we could reduce our problem to one which could be solved by asymptotically inverting a univariate Laplace transform. Fortunately, the best linear approximation to this (u,v) region, which we shall now develop, enables us to do precisely this. Consider the region $A_n(\alpha)$ defined by a hyperplane which supports the (u,v) region at the point $v = \alpha\,\mathrm{Tr}[A\Lambda]$. Elementary calculations show that this may be expressed as

$$(3.44)\quad A_n(\alpha) = \{(u,v): v + (\tfrac{\alpha+1}{2})\ \mathrm{Tr}(A\Lambda)u \geq a_n\} \cap \{(u,v): u \geq 0 \text{ and } v \geq 0\},$$

where, in order to remain on the curved part of the boundary defined in (3.44), we shall require that $0 \leq \alpha \leq C_n/\mathrm{Tr}[A\Lambda]^{-1}$. Now we define

$$(3.45)\qquad a_n \equiv \mathrm{Tr}[A\Lambda](\alpha + (\alpha+1)b_n)$$

$$b_n \equiv \log[C_n/((\alpha+1)\mathrm{Tr}[A\Lambda])].$$

$$(3.46) \qquad R_1 = p_o P[(U,V) \in S_1^* | \theta = 0]$$

$$\leq p_o P[(U,V) \in A_n(\alpha)] \equiv R_1^\alpha , \quad \alpha > 0 .$$

Thus we have for each n the upper bound for R_1, namely

$$R_1^{\alpha_1} = p_o \inf_{\alpha \in \ell} R_1^\alpha , \quad \ell = \left\{\alpha : \frac{C_n}{Tr[A\Lambda]} - 1 > \alpha > 0\right\} .$$

The univariate Laplace transform needed to obtain the upper bound can be derived from (3.14). To see this, we first note that if we let $Z_\alpha \equiv V + [(\alpha+1)/2]Tr[A\Lambda]$, then we may write

$$(3.47) \qquad R_1^\alpha = P[Z_\alpha \geq a_n | \theta = 0] .$$

Thus the univariate characteristic function needed to obtain an asymptotic expression for R_1^α is

$$\Phi_1^\alpha(q) = E(e^{-qZ_\alpha}) = E\left[e^{-q[((\alpha+1)/2)Tr[A\Lambda]U+V]} | \theta = 0\right] .$$

But this can be directly obtained from (3.14) if we set $s = q[((\alpha+1)/2)Tr[A\Lambda]]$ and $t = q$. Doing so, we obtain

$$(3.48) \qquad \Phi_1^\alpha(q) = \frac{e^{nTr[\Sigma](\frac{\alpha+1}{2})qTr[A\Lambda]}}{|I+q((\alpha+1)Tr[A\Lambda]\Lambda+2\Lambda A\Lambda|^{\frac{1}{2}}} .$$

Next, let us define

$$(3.49) \qquad k_n(q,\alpha) \equiv \log \Phi_1^\alpha(q) + qa_n .$$

Then, using the same reasoning that was used to obtain (2.21) in the constant loss case, we rewrite (3.47) to obtain

$$R_1^\alpha = \frac{p_o}{2\pi i} \int_{q_1^\alpha - i\infty}^{q_1^\alpha + i\infty} \frac{\exp(k_n(q,\alpha))}{q} dq \ , \ q_1^\alpha < 0 \ , \tag{3.50}$$

and q_1^α is to the right of the singularities of $\Phi_1^\alpha(q)$.

Next, let us define

$$B \equiv (\alpha+1)\mathrm{Tr}[A\Lambda]\Lambda + 2\Lambda A\Lambda \ , \tag{3.51}$$

and note that

$$\frac{\partial k_n(q,\alpha)}{\partial(iq)} = i\left[-\tfrac{1}{2}\mathrm{Tr}[(I+qB)^{-1}B - (\alpha+1)n\mathrm{Tr}[A\Lambda]\Sigma] + a_n\right]$$

(3.52)

$$\frac{\partial^2 k_n(q,\alpha)}{\partial^2(iq)} = -\left[-\tfrac{1}{2}\mathrm{Tr}[(I+qB)^{-2}B^2]\right] \quad > 0 \text{ for } -\frac{1}{||B||} < q < 0 \ .$$

Furthermore, if for $-1/||B|| < q < 0$ we also have that $\partial^2 k_n(q,\alpha)/\partial^2(iq) \to \infty$, then for any α and q such that $\partial k_n(q^\alpha,\alpha)/\partial(iq) = 0$ we will have the result that

$$R_1^\alpha \sim \frac{p_o \exp(k_n(q^\alpha,\alpha)}{(-s_1)\sqrt{-2\pi k_n''(q^\alpha,\alpha)}} \cdot C_N \ ,$$

where C_N is defined in equation (2.20).

Now if in addition to obtaining R_1^α we also wish to choose that α which will give the best lower bound of R_1, R_1^α, then we must simultaneously maximize $k_n(q,\alpha)$ with respect to both q and α.

Now if it can be shown that there exists a (q_1,α_1) which for all n satisfy (3.49) and also satisfy the relationship

$$\frac{\partial k_n(q_1,\partial_1)}{\partial\alpha} = \frac{\partial k_n(q_1,\alpha_1)}{\partial q} = 0 ,$$

then we will have that

<u>Theorem 3.3</u>

$$R_1^{\alpha_1} \sim \frac{p_o \exp(k_n(q_1,\alpha_1))}{(-s_1)\sqrt{-2\pi k_n''(q_1,\alpha_1)}} \cdot C_N ,$$

<u>where</u> C_N <u>is defined in</u> (2.20).

<u>Proof</u>: Differentiating $k_n(q,\alpha)$ with respect to α we obtain

$$\text{(3.53)} \qquad \frac{\partial k_n(q,\alpha)}{\partial\alpha} = -\tfrac{1}{2}q\mathrm{Tr}[A\Lambda]\mathrm{Tr}[(I+qB)^{-1}\Lambda-n\Sigma] + qb_n ,$$

where b_n is defined in (3.45).

Next, in order to show that the unconstrained (q_1,α_1) exist for all n we shall obtain an upper and lower bound for α_1. The upper and lower bound for α_1 is then used to obtain a lower bound on q_1 which satisfies (3.49). Then, we can simplify the existence proof by rewriting (3.52) to obtain

$$(3.54) \quad \frac{\partial k_n(q,\alpha)}{\partial\, q} = (\alpha+1)\,Tr[A\Lambda](b_n - \tfrac{1}{2}Tr[(I+qB)^{-1}\Lambda - n\Sigma])$$

$$- Tr[(I+qB)^{-1}\Lambda A\Lambda] + \alpha Tr[A\Lambda]$$

Now, let us assume that we may choose (q_1,α_1) such that

$$\frac{\partial k_n(q_1,\alpha_1)}{\partial\alpha} = \frac{\partial k_n(q_1,\alpha_1)}{\partial q} = 0 \ .$$

Then (3.53), (3.54) may be reduced to

$$(3.55) \qquad Tr[(I+q_1B)^{-1}\Lambda A\Lambda] = \alpha Tr[A\Lambda]$$

$$(3.56) \qquad \tfrac{1}{2}Tr[(I+q_1B)^{-1}\Lambda - n\Sigma] = b_n \ .$$

Special cases of solutions of (3.55), (3.56) for (q_1,α_1) are presented in Section 3.5.

Let us now prove that (q_1,α_1) which satisfy the requirements of (3.52) and (3.44) exist.

Lemma 3.5. _There exists an_ N_2, $(q_1\alpha_1)$ _such that for_ $n \geq N_2$, $(q_1\alpha_1)$ _simultaneously satisfy_ (3.55), (3.56), _and_ (3.50). _In addition_ α_1 _has the property that_

$$(3.57) \qquad 0 < \frac{Tr[\Lambda A\Lambda]}{Tr[A\Lambda]} \leq \alpha_1 \leq \frac{C_n e^{-Tr[\Lambda - n\Sigma]/2}}{Tr[A\Lambda]} - 1 \ .$$

The preceeding condition on α_1 clearly satisfies the restriction placed on α in (3.44).

Proof: The easiest way to show that the lemma is valid is to consider the coordinate system whose axes are α and $q||B||$. Now in order that $q||B|| = 0$ for $\alpha > 0$ we must have $q = 0$. From this it follows that (3.56) intercepts the α axis at α_u where

$$\alpha_u \equiv C_n \exp(-\mathrm{Tr}[\Lambda - n\Sigma]/2)/\mathrm{Tr}[A\Lambda] - 1 .$$

Furthermore, for $-1 < \alpha < \alpha_u$ the value of $q||B||$ which satisfies (3.57) is finite, continuous and $\partial(q||B||)/\partial\alpha > 0$. Next, since $\mathrm{Tr}[\Lambda A\Lambda] \leq ||\Lambda||\mathrm{Tr}[A\Lambda] < \mathrm{Tr}[A\Lambda]$, we have that the α axis of (3.55) has the property that

$$0 < \alpha_L \equiv \mathrm{Tr}[\Lambda A\Lambda]/\mathrm{Tr}[A\Lambda] < 1$$

and that for $\alpha > \alpha_L$ and fixed n the value of $q||B||$ which satisfies (3.55) is finite, (3.55) is continuous, and $\partial(q||B||)/\partial\alpha < 0$.

Now we must show that there exists an N_2 such that for all $n > N_2$, α_u has the property that $\alpha_u > 1 > \alpha_L$. First, by inspection of (3.9) we see that $\mathrm{Tr}[A\Lambda] = o(n)$. Thus, if we can show that eventually

$$\alpha_u > \frac{n\sqrt{K}}{\mathrm{Tr}[A\Lambda]} - 1, \tag{3.58}$$

then we will have shown that eventually $\alpha_u > 1$. In order to demonstrate (3.58) we note that since, for $a > 0$,

$$\log(1+a) > a/(1+a) ,$$

we have that for all n

$$\mathrm{Tr}[\Lambda - n\Sigma] > -2\log\left[\frac{C_n}{n\sqrt{k}}\right] .$$

This condition is equivalent to (3.58). Thus we have the result that the value N_2 defined in the lemma is

$$N_2 = \{n : n\sqrt{k}/\mathrm{Tr}[A\Lambda] = 2\} .$$

Finally, in order to complete this proof of Lemma 3.5, we must show that the value of q which satisfies (3.55), (3.56) also satisfies the requirements of (3.50); i.e., we must show that q_1 is to the right of the singularities of $\Phi_1^{\alpha_1}(q)$, that is we must show that there exists a q_1 with the properties that

(3.59) $$-1 < q_1 ||B|| \Big|_{\alpha=\alpha_1} < 0 ,$$

and satisfying the relationship

(3.60) $$\tfrac{1}{2}\mathrm{Tr}[(I+q_1 B)^{-1}B] = a_n \Big|_{\alpha=\alpha_1} .$$

Now, since we don't know α_1 precisely but do have an upper and lower bound for it, we shall show that for α_1 in this range q_1 does indeed satisfy (3.59).

We begin the proof by noting that if α_1 satisfies (3.58), then for $n > N_2$

(3.60) $$b_n \geq \mathrm{Tr}[\Lambda - n\Sigma]/2 .$$

Next, we note that for $q_1 \in (-1/||B||, 0)$,

$$-\tfrac{1}{2}\mathrm{Tr}[(I+qB)^{-1}B-(\alpha+1)n\mathrm{Tr}[A\Lambda]\Sigma]$$

is continuous and monotone in q. Furthermore, since

$$\lim_{q\downarrow \ell} \mathrm{Tr}[(I+qB)^{-1}B-(\alpha+1)n\mathrm{Tr}[A\Lambda]\Sigma] = \infty > a_n ,$$

where $\ell = [-1/||B||]$, and

$$\tfrac{1}{2}\mathrm{Tr}[(I+qB)^{-1}B-(\alpha+1)n\mathrm{Tr}[A\Lambda]\Sigma]\Big|_{q=0} = (\tfrac{\alpha+1}{2})\mathrm{Tr}[A\Lambda]\mathrm{Tr}[\Lambda-n\Sigma]+\mathrm{Tr}[\Lambda A\Lambda]$$

$$\leq \mathrm{Tr}[A\Lambda]\Big[(\tfrac{\alpha+1}{2})\mathrm{Tr}[\Lambda-n\Sigma]+\alpha\Big]$$

$$\leq \mathrm{Tr}[A\Lambda]\Big[(\alpha+1)b_n+\alpha\Big] = a_n ,$$

we must have a unique q_1 which satisfies (3.59).

Thus we have completed the proof of Lemma 3.5, and hence, Theorem 3.3.

Thus in this section we have obtained an asymptotic estimate of R_1. If one compares this quadratic loss estimate with that which was obtained for the constant loss case, it can be seen that the order results are the same.

In the next section an upper bound for R_1 is obtained. While it has not been done in this thesis it would be worthwhile to compare the asymptotic results which are obtained for the more easily derived upper bound with those asymptotic results which were obtained in the preceeding section.

3.5 ASYMPTOTIC EVALUATION OF UPPER BOUND FOR R_1.

3.5.1 Introduction. In this section the results obtained in Section 3.4 are applied to the case in which the diagonal terms resulting from the orthogonal decomposition of the positive definite matrix Σ are of the form $\sigma_i = 1/i^2$, $i = 1, 2, \ldots$ In order to simplify this application we have also assumed that the positive definite quadratic loss matrix A is diagonal. Thus

$$(3.61) \qquad A = \{a_{ij}\}_{i,j=1}^{\infty} \quad \text{and } a_{ij} = \begin{cases} a_i > 0, & i = j \\ 0, & i \neq j \end{cases} .$$

In Sections 3.5.1 and 3.5.2 asymptotic expressions for R_1 are respectively obtained for the cases in which $\Sigma a_i < \infty$ and $a_i = 1$, $i = 1, 2, \ldots$

The reader should note that these asymptotic evaluations are only being carried out to demonstrate a solution technique. For this reason the estimates for R_1 and R_2 will only be of order accuracy, i.e., the approximation will only be accurate up to the coefficient of the exponential rate of decay of the risk. More precise asymptotic solutions may be obtained with a computer.

3.5.2 Asymptotic Solution for R_1 for a Case in which $Tr[A] < \infty$.

In order to obtain an expression for R_1 we must first obtain a simultaneous asymptotic solution (q_2, α) to the set of equations (3.55) and (3.56). Now using the fact that A is diagonal we may reduce equations (3.55) and (3.56) to the forms

$$(3.62)\qquad \frac{1}{2}\sum_{i=1}^{\infty} \frac{\lambda_i}{1+q_1((\alpha+1)\ \mathrm{Tr}[A\Lambda]\ \lambda_i + 2\lambda_i^2\ a_i} = \log[D_n/((\alpha+1)\ \mathrm{Tr}(A\Lambda))]$$

and

$$(3.63)\qquad \sum_{i=1}^{\infty} \frac{\lambda_i^2\ a_i}{1+q_1((\alpha+1)\ \mathrm{Tr}[A\Lambda]\ \lambda_i + 2\lambda_i^2\ a_i} = \alpha\ \mathrm{Tr}[A\Lambda]$$

where $D_n = C_n e^{n\mathrm{Tr}[\Sigma]/2}$ and

$\mathrm{Tr}[A\Lambda] \to \mathrm{Tr}[A]$.

Now since $\alpha > 0$ we may write that for all i

$$(\alpha+1)\ \mathrm{Tr}(A\Lambda)\ \lambda_i \gg \lambda_i^2\ a_i \quad .$$

Therefore we may make the following approximation of the left hand side of (3.63):

$$\sum_{i=1}^{\infty} \frac{\lambda_i^2\ a_i}{1+q_1((\alpha+1)\ \mathrm{Tr}[A\Lambda]\ \lambda_i + 2\lambda_i^2\ a_i}$$

$$\sim \sum_{i=1}^{\infty} \frac{\lambda_i^2\ a_i}{1+q_1(\alpha+1)\ \mathrm{Tr}[A\Lambda]\ \lambda_i}$$

$$\sim \mathrm{Tr}[A]/(1+q_1(\alpha+1)\ \mathrm{Tr}[A]).$$

Thus (3.63) reduces to

$$(3.64) \qquad \frac{1}{1+q_1(\alpha+1)\ \mathrm{Tr}[A]} \sim \alpha \ .$$

Now in order to reduce equation (3.62) to a solvable form we first note that its left hand side may be reduced as follows. First, applying the asymptotic technique of Section 2.3 we obtain

$$D_n \equiv n \prod_{i=1}^{\infty} (1+n/i^2)^{\frac{1}{2}}$$

$$= n^{\frac{3}{4}} \sqrt{\frac{\sinh(\pi\sqrt{n})}{2\pi}} \sim e^{\pi\sqrt{n}/2} \quad ,$$

and therefore

$$\log [D_n/(\alpha+1)\ \mathrm{Tr}[A\Lambda]] \sim \pi\sqrt{n}/2 \ .$$

Finally, using the result that

$$\sum_{j=1}^{\infty} 1/(j^2+x^2) = [\pi \coth(\pi x) - 1/x]/2x$$

$$\sim \pi/2x \quad ,$$

we can approximate the right hand side of (3.62) to obtain

$$\frac{1}{2}\mathrm{Tr}[[I+q_1B]^{-1}\Lambda] \sim \frac{1}{2} \Sigma\lambda_i/(1+q_1(\alpha+1)\ \mathrm{Tr}[A]\ \lambda_i)$$

$$= \frac{1}{2} \sum_{i=1}^{\infty} n/(i^2+n(1+q_1(\alpha+1)\ \mathrm{Tr}[A]))$$

$$\tfrac{1}{2}\mathrm{Tr}[(I+q_1B)^{-1}\Lambda] \sim \pi\sqrt{n}/(4\sqrt{1+q_1(\alpha+1)\,\mathrm{Tr}[A]})$$

Combining the above results, we may now asymptotically reduce equation (3.62) to the form

$$\frac{1}{2\sqrt{1+q_1(\alpha+1)\,\mathrm{Tr}(A)}} \sim 1 \quad . \tag{3.65}$$

Solving (3.65) and (3.64) simultaneously we now obtain that

$$\alpha \sim 4$$

$$q_1 \sim 3/(20\,\mathrm{Tr}[A])$$

We can now use this "approximate-simultaneous-asymptotic-solution" to obtain a rough expression for R_1. First, since $q_1(\alpha+1)\,\mathrm{Tr}(A) \sim -3/4$ we may reduce $\Phi_1^{\alpha}(q_1)$, defined in (3.47), to the form

$$\Phi_1^{\alpha}(q_1) = \prod_{i=1}^{\infty} (1+q_1(\alpha+1)\,\mathrm{Tr}(A\Lambda)\,\lambda_i + 2\lambda_i^2\,a_i)^{-\frac{1}{2}}$$

$$\sim \prod_{i=1}^{\infty} (1+q_1 n(\alpha+1)\,\mathrm{Tr}(\Lambda)/(i^2+n))^{-\frac{1}{2}}$$

$$= \prod_{i=1}^{\infty} \left[\frac{1+n/(4i^2)}{1+n/i^2}\right]^{-\frac{1}{2}} \sim e^{\pi\sqrt{n}/4} \quad .$$

Furthermore we may approximate $q_1\,a_n$ by $q_1\,a_n \sim q_1\mathrm{Tr}[A](\alpha+1)\log D_n$

$$\sim -\tfrac{3}{4}(\pi\sqrt{n}/2) \quad .$$

Finally, since we are only concerned with obtaining order asymptotics we may approximate (3.60) with

$$\inf_{\alpha} R_1^{\alpha} \sim \exp(k_n(q_1,\alpha_1))$$

$$= \exp(k_n(q_1,\alpha_1)) \sim e^{-\pi\sqrt{n/8}} .$$

3.5.3 Asymptotic Solution for R_1 for a case in which $Tr[A] = \infty$.

Let us now consider the case in which A, the quadratic loss matrix is diagonal and $a_i = 1$, $i = 1,2....$ In this case

$$Tr[A\Lambda] = Tr(\Lambda)$$

$$\sim \frac{\pi\sqrt{n}}{2} .$$

Using these results, plus the fact that for all i, $\lambda_i^2 = 0((\alpha+1)Tr[\Lambda]\lambda i$, we may approximate the left hand side of (3.63) with

$$\sum_{i=1}^{\infty} \lambda_i^2/(1+q_1\ (\alpha+1)\ Tr[\Lambda]\ \lambda_i + 2\lambda_i^2)$$

$$\sim \sum_{i=1}^{\infty} \lambda_i^2/(1+q_1(\alpha+1)\ Tr[\Lambda]\ \lambda_i)$$

$$= \frac{1}{q(\alpha+1)} - \frac{1}{q(\alpha+1)\ Tr[\Lambda]} \sum_{i=1}^{\infty} \frac{n}{i^2 + n(1+q_1(\alpha+1)\ Tr[\Lambda]}$$

$$\sim \frac{1}{q_1(\alpha+1)} [1 - 1/\sqrt{1+q_1(\alpha+1)\ Tr[\Lambda]}\ ] .$$

Thus in this case (3.63) asymptotically becomes

$$(3.66) \qquad \frac{1}{q_1(\alpha+1)} \left[1 - 1/\sqrt{1+q(\alpha+1)\ \mathrm{Tr}[\Lambda]}\right] \sim \alpha\ \mathrm{Tr}\ \Lambda$$

By a similar argument we may asymptotically simplify (3.62) to obtain

$$(3.67) \qquad \frac{1}{2\sqrt{1+q_1(\alpha+1)\ \mathrm{Tr}[\Lambda]}} \sim 1 ,$$

which we may write as

$$q_1(\alpha+1)\ \mathrm{Tr}[\Lambda] \sim -\frac{3}{4} .$$

Simultaneously solving (3.67) and (3.66) we obtain

$$\alpha \sim 4/3$$

$$q_1 \sim -9/(28\ \mathrm{Tr}[\Lambda]) \sim -9/(14\ \pi\sqrt{n}) .$$

Following the procedure of the previous example we obtain

$$\Phi_\alpha^1(q_1) \sim e^{\pi\sqrt{n}/4}$$

and

$$q_1 a_n \sim q_1\ \mathrm{Tr}[\Lambda](\alpha+1)\ \log Un \sim -3\ \pi\sqrt{n}/8 .$$

Ignoring the non-order terms, we obtain

$$\inf_\alpha R_1^\alpha \sim e^{-\pi\sqrt{n}/8} .$$

Hence, for the two cases considered, we have shown that the upper bound for the type I risks obtained for the quadratic and the asymptotic evaluation of the constant loss are of the same asymptotic order. Differences between the case for non-finite and finite A begin to appear in the "non-order" asymptotic terms which may easily be obtained from "expansions" used in the theory and by computer solution.

BIBLIOGRAPHY

[1] Anderson, T., Darling, D. (1952) "Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic processes". Ann. Math. Stat. 19, 293-325.

[2] Bahadur, R. R. (1967) "Rates of convergence of estimates and test statistics", Ann. Math. Stat. 38, 303-324.

[3] Beckenbach, E. F. and Bellman, R. (1965) Inequalities, Springer-Verlag, New York, N.Y.

[4] Chernoff, H. (1952) A measure of asymptotic efficiency for tests of a hypothesis based on sums of observations. Ann. Math. Stat., 23, 493-507.

[5] Hirschman, I., Widder, D. V. (1955) The Convolution Transform, Princeton University Press, Princeton, N. J.

[6] Hsu, L. C. (1951) On the asymptotic evaluation of a class of multiple integrals involving a parameter. Am. J. Math. 73, 625-634.

[7] Karlin, S. (1968) Total Positivity, Stanford University Press, Stanford, California.

[8] Lindley, D. V. (1960) "The use of prior probability distributions in statistical inference and decision theory", Proceedings of the Fourth Berkeley Symposium. University of California Press, Berkeley, 1, 453-468.

[9] Pitmann, E. J. G. (1938) "Significance tests which may be applied to samples from any population III. The Analysis of Variance Test. Biometrika, 29, 322.

[10] Rubin, H. and Sethuraman, J. (1965) Bayes risk efficiency, Sankhya, Series A, 27, 347-356

[11] Rubin, H. and Sethuraman, J. (1965) Probabilities of moderate deviation, Sankhya, Series A, 27, 325-346.

[12] Watson, G. S. (1971) Selected topics in statistical theory, Mathematical Association of America, Williams College, Williamstown, Massachusetts.

[13] Widder, D. V. (1941) The Laplace Transform, Princeton University Press, Princeton, New Jersey.

APPENDIX

In this appendix we shall state and prove an unpublished result due to H. Rubin. In doing so we shall use the notation of Section 3.2 except as noted.

Recall that

$$Q = (1+2s_1)\ B_{t_1}^{-1}\ \Lambda A \Lambda$$

$$= (1+2s_1)(\ +2s_1\Lambda+2(1+2s_1)\Lambda A\Lambda)^{-1}\Lambda A\Lambda\ .$$

<u>Theorem A-1</u>

<u>If</u> $-t_1 \|A\|$ <u>uniformly bounded below</u> 1 <u>and above</u> 0

<u>then</u>

$$\mathrm{Tr}[Q] \underset{n}{\rightarrow} \mathrm{Tr}[(I+2t_1 A)^{-1}A]$$

<u>and</u>

$$q_j^n \underset{n}{\rightarrow} q_j \quad , \quad j = 1,2,\ldots\ .$$

<u>Here we define</u> q_j^n <u>to be the</u> j^{th} <u>largest eigenvalue of</u> Q <u>and</u> q_j <u>to be the</u> j^{th} <u>largest eigen value of</u> $[(1+2t_1 A)^{-1}A]$.

Proof:

Let us define the projection

$$E_m = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} \qquad m = 1,2,\ldots,$$

where E_m is infinite dimensional and I_m is the m dimensional identity matrix. We also define $V_{mn} = E_m \Lambda 1$

Definition A.1 If C_n is positive definite and symmetric then we define $C_n \Rightarrow C$ to mean (a) $Tr[C_n] \to Tr[C]$ and (b) the i^{th} largest characteristic root of C_n converges to the i^{th} largest characteristic root of C. Now in order to prove the Theorem we need the following lemmas.

Lemma A.1. Clearly

$$\| V_{mn} \, A \, V_{mn} \| \le \| A \| \quad .$$

Here, since we are dealing with compact operators we may still use $\| A \|$ to denote the largest eigenvalue of A as well as the norm of the operator A. Next we must show

Lemma A.2

$$V_{mn} \, A \, V_{mn} \Rightarrow A \; [\text{ as } m,n \to \infty].$$

Proof:

Let us define the i^{th} largest eigenvalue of $R \equiv V_{mn} \, A \, V_{mn}$ as r_i. Then we may apply the well known result that for R, real and symmetric

$$\sum_{i=1}^{k} r_i = \max_{X^{(1)}..X^{(k)}} \sum_{i=1}^{k} X^{(i)'} R x^{(i)}$$

$$1 \leq k \leq n$$

where the max is taken over the set of all k orthonormal vectors [see [3], p.77]. Now since $Tr[R] < Tr[A]$

$$\max_{X^{(1)}..X^{(k)}} \sum_{i=1}^{k} (X^{(i)'} R x^{(i)} \leq \max_{X^{(1)}..X^{(k)}} \sum_{i=1}^{k} X^{(i)}, A x^{(i)}$$

$$= \sum_{i=1}^{k} a_i$$

we need only show that for eack k, given $\epsilon > 0$, we eventually have that

(A.1) $$\sum_{1}^{k} r_i > \sum_{1}^{k} a_i - \epsilon .$$

But having shown that $\sum_{1}^{k} r_i \uparrow_{n} \sum_{1}^{k} a_i \uparrow_{k} Tr[A] < \infty \quad \forall k$

we will have shown that $R \Rightarrow A$.

Now in order to show that (A.1) holds we let X_i be the eigenvector associated with a_i. ie $AX_i = a_i X_i$.

Next, given $\delta = \{\delta : \epsilon = \delta \|A\| [\delta + 2]\}$ and for m,n sufficiently large we have that

$$\|V_{mn} X_i - X_i\| < \delta , \; i = 1 .. k ,$$

which implies that

$$\|A \, V_{mn} \, x_i = A_{x_i}\| \leq \|A\| \delta .$$

But then it follows that

$$|X_i' V_{mn} A V_{mn} X_i - X_i' A X i|$$

$$= |(X_i' V_{mn} - X_i')(A V_{mn} X_i - A_{x_i})$$

$$+ 2(X_i')(V_{mn} A x_i - Ax_i)|$$

$$\leq \delta \|A\| [\delta+2] = \epsilon .$$ Thus (A.1) holds.

Thus we have demonstrated lemma (A.2).

Next, if we use the fact that the eigenvalues of the real matrix $[I+2t_1 V_{mn} A V_{mn})^{-1} V_{mn} A V_{mn}$ are the same as those of the real symmetric matrix $A^{\frac{1}{2}} V_{mn} [I+2t_1 V_{mn} A V_{mn}]^{-1} V_{mn} A^{-1}$ one can use the proceedure which was used to demonstrate the proceeding Lemma to show

<u>Lemma A.3</u>

$$[I+2t_1 V_{mn} A V_{mn}]^{-1} V_{mn} A V_{mn} \Rightarrow [I+2t_1 A]^{-1} A.$$

We are now in a position to begin the proof of the theorem. First we note that for any symmetric matrix S we have the result that

(A.2) $$E_m [E_m S E_m + \alpha(I-E_m)]^{-1} E_m \leq S^{-1} , \alpha \neq 0.$$

Furthermore since $s_1 < 0$ and $\Lambda-I < 0$ we have that

(A.3) $$I+2s_1 \Lambda+2t_1(1+2s_1)\Lambda A\Lambda = (1+2s_1)(I+2t_1\Lambda A\Lambda) + 2s_1(\Lambda-I)$$

$$> (1+2s_1)(I+2t_1\Lambda A\Lambda)$$

Now if in (A.2) we let $\alpha = 1$, $S \equiv I + 2s_1\Lambda + 2t_1(1+2s_1)\Lambda A\Lambda$ and combine the result with (A.3) we obtain

$$\text{(A.4)} \qquad E_m(I+2s_1 V_{mn} + 2(1+2s_1)t_1 V_{mn} A V_{mn})^{-1} E_m$$

$$\leq (I+2s_1\Lambda + 2t_1(1+2s_1)\Lambda A\Lambda)^{-1}$$

$$\leq (1+2s_1)^{-1} (I+2t_1\Lambda A\Lambda)^{-1}.$$

Finally, since the eigenvalues of Q are the same as those of $D \equiv (1+2s_1)A^{\frac{1}{2}} \Lambda(I+2s_1\Lambda+2(1+2s_1) \Lambda A\Lambda)^{-1} \Lambda A^{\frac{1}{2}}$, let us consider the inequality

$$D_m \equiv (1+2s_1A^{\frac{1}{2}} V_{mn}(I+2s_1 V_{mn} + 2(1+2s_1)t_1 V_{mn} A V_{mn})^{-1} V_{mn} A^{\frac{1}{2}}$$

$$\leq D \leq A^{\frac{1}{2}} \Lambda(I+2t_1\Lambda A\Lambda)^{-1} \Lambda A^{\frac{1}{2}} \Rightarrow A^{\frac{1}{2}} [I+2t_1A]^{-1} A^{\frac{1}{2}}$$

which follows from Lemma (A.3) and (A.4) . Now if we can show that the eigenvalue of the left and right hand side of (A.5) are asymptotically equal then we will have proved the theorem.

In order to show this we first note that applying lemma A.3 we have that given $\epsilon > 0$ $\exists$ m_o, n_o such that for all (m,n) $m > m_o, n > n_o$ we have the condition that for all i $|a_i = b_{imn}| < \epsilon$ where b_{imn} is the i^{th} eigenvalue of $B \equiv A^{\frac{1}{2}} V_{mn}(I+2t_1 V_{mn} A V_{mn})^{\frac{1}{2}} V_{mn} A^{\frac{1}{2}}$

Furthermore, since $V_m - E_m \underset{n}{\rightarrow} 0$ we have that for fixed m

$$|d_i^m - b_i^m| \leq ||D-B|| \rightarrow 0 \ , \ i = 1,2\ldots$$

More precisely, given $\epsilon > 0$ if $n > n_1(m)$ then $|d_i^m - b_i^m| < \epsilon$.

Now since there exists an m^* such that for $m > m^*$ $n_1(m+1) < n_1(m)$ we may state that for $n > \max[n_o, n_1(m^*)]$ we have that $|a_i - b_i^m| < 2\epsilon \Rightarrow$ $|q_i^n - q_i| < 2\epsilon$, $i = 1,2...$

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)* Purdue University	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Bayes Risk for the Test of Location-The Infinite Dimensional Case

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

Technical Report, June, 1972

5. AUTHOR(S) *(Last name, first name, initial)*

Cohen, Philip

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June, 1972	89	13

8a. CONTRACT OR GRANT NO. N00014-67-A-0226-0008 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Mimeograph Series # 288 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research Washington, D. C.

13. ABSTRACT

Let X_1, X_2 be independent normal with mean vector θ and covariance matrix I. Let the null hypothesis be normal with zero mean and covariance matrix Σ. Quadratic loss ($\theta'A\theta$) and constant loss are considered. Rubin and Sethuraman obtained asymptotic results for the above test in the case of a finite dimensional parameter space. New asymptotic results have been obtained for the case in which the parameter space is infinite dimensional. This development is motivated by the need to extend the Bayes Risk Efficiency analysis to time series problems and to problems in which the alternative hypothesis is a function space. For the case of constant loss, exact results have been obtained if the characteristic roots of Σ are pairs of $1/i$ or $1/i^2$. Exact results have been obtained for the characteristic roots $\sigma_i = i^{-\rho}$, $\rho > .5$, or $\sigma_i = a^i$, $0 < a < 1$. In the latter case one still obtains the Rubin-Sethuraman result that the type II risk, R_2, is asymptotic to $R_1 \log n$ where n is the sample size. In the former case it can be shown that the risks are asymptotically proportional. A generalized expression for the proportionality constant is obtained. Similar asymptotic results are obtained for quadratic loss.

DD FORM 1 JAN 64 1473

unclassified
Security Classification

TABLE OF CONTENTS

BIBLIOGRAPHY

APPENDIX