FIXED-SIZE CONFIDENCE REGIONS FOR MULTIPARAMETER ESTIMATION

By

AHMED DARWESH ABID

Bachelor of Science
King Saud University
Riyad, Saudi Arabia
1976

Master of Science
Purdue University
West Lafayette, Indiana
1980

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1985

FIXED-SIZE CONFIDENCE REGIONS FOR

MULTIPARAMETER ESTIMATION

Thesis Approved:

N. Mukhopadhyay
Thesis Adviser

Dale E. Alspach

W. W. Sanders

Ronald W. McNew

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I extend special appreciation to my major adviser, Dr. N. Mukhopadhyay, for his guidance, advice, and understanding, without his guidance this work would have not been done. Appreciation is also expressed to Dr. D. E. Alspach, Dr. R. W. McNew, and Dr. W. D. Warde for their help as advisory committee members. I would also like to extend my appreciation to all my teachers.

Sincere appreciation is also extended to my parents, brothers, and sisters for their encouragement. I would like to express my gratitude to my wife, Mariam, my son, Amr, and my daughter, Bushra for their support and patience during my study in the United States of America.

Sincere thanks are extended to Mrs. Janet Sallee for typing this dissertation.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

In this study we discuss in Chapters II and III two separate problems about constructing fixed-size confidence regions for multiparameter estimation. We have reviewed the relevant literature separately at the beginning of each chapter.

The second chapter deals with the problem of constructing a fixed-size ellipsoidal confidence region for the difference of the mean vectors of two independent multinormal populations. We have assumed that the covariance matrices of the first and second populations are respectively given by $\sigma_1^2 H$ and $\sigma_2^2 H$, where σ_1^2 and σ_2^2 are both unknown. Here, H is assumed to be a known positive definite matrix. The three cases namely, (i) $\sigma_1 = \sigma_2$ and equal sample sizes, (ii) $\sigma_1 \neq \sigma_2$ and equal sample sizes, and (iii) $\sigma_1 \neq \sigma_2$ and unequal sample sizes have been dealt with separately. We propose both two-stage and sequential procedures for each problem and study various exact and asymptotic properties of these procedures through several Theorems.

In Chapter III, we present the problem of constructing a fixed-size ellipsoidal confidence region for regression parameters in a general linear model under Gauss-Markoff set up. Here, we propose two-stage, modified two-stage, sequential, and three-stage procedures to tackle this problem. Again, we study various exact and asymptotic properties of these procedures.

We also report numerical results in the form of tables to study the moderate sample behaviors of the proposed procedures for both these problems.

The Chapter IV contains general comments and the summary of our findings for both sets of problems.

In what follows, [x] will always stand for the largest integer smaller than x. This notation has been primarily used in defining the two-stage, modified two-stage, and three-stage procedures.

CHAPTER II

FIXED-SIZE CONFIDENCE REGIONS FOR THE DIFFERENCE OF THE MEANS OF TWO MULTINORMAL POPULATIONS

2.1. Introduction and Review

Let $\{\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_r, \ldots\}$ and $\{\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_s, \ldots\}$ be two independent sequences of independent and identically distributed (i.i.d.) multivariate random variables where each $\underset{\sim}{X}$ is distributed as $N_p(\underset{\sim}{\mu}_1, \sigma_1^2 H)$ and each $\underset{\sim}{Y}$ is distributed as $N_p(\underset{\sim}{\mu}_2, \sigma_2^2 H)$ with $\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2 \in \mathbb{R}^p$ and $0 < \sigma_1, \sigma_2 < \infty$. We assume that $\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2, \sigma_1^2$ and σ_2^2 are all unknown parameters. Here, H is assumed to be a known p x p positive definite matrix. Having recorded r observations on $\underset{\sim}{X}$'s and s observations on $\underset{\sim}{Y}$'s, we wish to construct a confidence region for the difference of the mean vectors, namely $\underset{\sim}{\mu} = \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$. Given $d \in (0,\infty)$ and $\alpha \in (0,1)$, we propose to consider the region

$$\mathcal{R}_{r,s} = \{\underset{\sim}{\mu} \in \mathbb{R}^p : (\bar{\underset{\sim}{T}}_{r,s} - \underset{\sim}{\mu})' H^{-1} (\bar{\underset{\sim}{T}}_{r,s} - \underset{\sim}{\mu}) \le d^2\} \qquad (2.1.1)$$

where $\bar{\underset{\sim}{X}}_r = r^{-1} \sum_{i=1}^{r} \underset{\sim}{X}_i$, $\bar{\underset{\sim}{Y}}_s = s^{-1} \sum_{j=1}^{s} \underset{\sim}{Y}_j$ and $\bar{\underset{\sim}{T}}_{r,s} = \bar{\underset{\sim}{X}}_r - \bar{\underset{\sim}{Y}}_s$. We now require that

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{r,s}) \ge 1 - \alpha, \qquad (2.1.2)$$

since the confidence coefficient associated with the region $\mathcal{R}_{r,s}$ is given by $P(\underset{\sim}{\mu} \in \mathcal{R}_{r,s})$. Now, we have

$$P(\underset{\sim}{\mu} \varepsilon\, R_{r,s}) = P\{(\bar{\underset{\sim}{T}}_{r,s} - \underset{\sim}{\mu})' H^{-1} (\bar{\underset{\sim}{T}}_{r,s} - \underset{\sim}{\mu}) \leq d^2\}$$

$$= P\left\{\chi^2_{(p)} \leq \left(\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s}\right)^{-1} d^2\right\} \qquad (2.1.3)$$

$$= F\left\{\left(\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s}\right)^{-1} d^2\right\} .$$

where $\chi^2_{(p)}$ stands for a chi-square random variable with p degrees of freedom and $F(\cdot)$ is the distribution function of $\chi^2_{(p)}$. We now obtain the positive number "a" such that

$$F(a) = 1 - \alpha . \qquad (2.1.4)$$

Therefore, from (2.1.2), (2.1.3), (2.1.4) and the monotonicity property of the distribution function $F(\cdot)$, it follows that r and s must satisfy the inequality

$$\left(\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s}\right)^{-1} d^2 \geq a \qquad (2.1.5)$$

which implies

$$\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s} \leq \frac{d^2}{a} . \qquad (2.1.6)$$

Let us define the usual unbiased estimators for σ_1^2, σ_2^2 as

$$U_r^2 = \{p(r-1)\}^{-1} \sum_{i=1}^{r} (\underset{\sim}{X}_i - \bar{\underset{\sim}{X}}_r)' H^{-1} (\underset{\sim}{X}_i - \bar{\underset{\sim}{X}}_r)$$

and

$$V_s^2 = \{p(s-1)\}^{-1} \sum_{j=1}^{s} (\underset{\sim}{Y}_i - \bar{\underset{\sim}{Y}}_s)' H^{-1} (\underset{\sim}{Y}_i - \bar{\underset{\sim}{Y}}_s) ,$$

respectively, for $r \geq 2$, $s \geq 2$. We will also consistently write [x] for the largest integer smaller than x.

In Sections 2.2 - 2.4, we will describe solutions for various separate cases, namely (i) $\sigma_1^2 = \sigma_2^2$ and $r = s$, (ii) $\sigma_1^2 \neq \sigma_2^2$ and $r = s$, and (iii) $\sigma_1^2 \neq \sigma_2^2$ and $r \neq s$. In each case, we propose a two-stage procedure and also a sequential procedure. For each procedure we discuss several important exact and asymptotic (as $d \to 0$) properties.

In Section 2.5, we report numerical results in order to study moderate sample size performances of all the procedures proposed in earlier sections of this chapter.

In the case of $p = 1$, the basic problem we consider here is known as the sequential analogue of the ordinary Behrens-Fisher problem. Various authors, e.g. Robbins et al. (1967), Mukhopadhyay (1976, 1977), and Ghosh and Mukhopadhyay (1980) proposed sequential procedures to estimate $\underset{\sim}{\mu}$ when $p=1$. Also, some two-stage procedures were considered in Chapman (1950), Scheffé (1970), and Ghosh (1975a) in order to obtain fixed-width confidence intervals for $\underset{\sim}{\mu}$ when $p = 1$. The present work is the natural and useful generalization of the results obtained in Al-Mousawi (1984). In this regard, one is also referred to the solutions of Chatterjee (1959, 1960) and Srivastava (1967, 1971) for the one-sample proplem. Our results are expected to lead to much better understanding of the sequential analogue of the multivariate Behrens-Fisher situations.

2.2. The Case of Two Equal Covariance Matrices and Equal Sample Sizes

Suppose that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ where $\sigma^2(>0)$ is unknown and we take $r = s = n$. Utilizing (2.1.3) in this case, we would have

$$P(\underset{\sim}{\mu} \varepsilon \mathcal{R}_{n,n}) = F(\frac{nd^2}{2\sigma^2}) , \qquad (2.2.1)$$

and from (2.1.6) we get

$$n \geq \frac{2a\sigma^2}{d^2} . \qquad (2.2.2)$$

Had σ^2 been known, the required optimal fixed sample size would then be given by

$$C = \frac{2a\sigma^2}{d^2} . \qquad (2.2.3)$$

But, since σ^2 is unknown, we will consider two procedures in order to determine the sample size N as a suitable random variable for estimating C.

2.2.1. A Two-Stage Procedure

We start with $m(\geq 2)$ observations from each population, and then define

$$N = \max\{m,[2b\ S_m^2/d^2] + 1\} , \qquad (2.2.4)$$

where $S_m^2 = \frac{1}{2}(U_m^2 + V_m^2)$, and $b = pb'$ where b' is the upper $100\alpha\%$ point of the F-distribution with degrees of freedom p, $2p(m-1)$.

Thus, from the samples $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_N$ we compute $\bar{\underset{\sim}{T}}_{N,N}$ and propose the corresponding region $\mathcal{R}_{N,N}$ as in (2.1.1). Some of the properties of this procedure are listed in Theorem 2.2.1.

<u>Theorem 2.2.1.</u> <u>For the procedure in (2.2.4)</u>, <u>for all</u> $\underset{\sim}{\mu} \in \mathbb{R}^p$, <u>and</u> $\sigma \in (0,\infty)$ <u>we have</u>:

(a) $P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) \geq 1 - \alpha$ <u>for all</u> $d > 0$,

(b) $E(N/C) \to \frac{b}{a}$ <u>as</u> $d \to 0$,

(c) $\mathrm{Var}(N)\{p(m-1)(2b\sigma^2/d^2)^{-2}\} \to 1$ <u>as</u> $d \to 0$, <u>and</u>

(d) $P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) \to 1 - \alpha$ <u>as</u> $d \to 0$.

<u>Proof</u>: To prove Part (a) first notice that

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) = \sum_{n=m}^{\infty} P\{\underset{\sim}{\mu} \in \mathcal{R}_{N,N}, N = n\}$$

$$= \sum_{n=m}^{\infty} P\{\underset{\sim}{\mu} \in \mathcal{R}_{n,n}, N = n\} .$$

Now, the event $\{N = n\}$ depends only on S_m^2, and the event $\{\underset{\sim}{\mu} \in \mathcal{R}_{n,n}\}$ depends only on $\bar{\underset{\sim}{T}}_{n,n}$ for every fixed $n \geq m$. But, we know that S_m^2 and $\bar{\underset{\sim}{T}}_{n,n}$ are independent for every fixed $n \geq m$, and thus we can write

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) = \sum_{n=m}^{\infty} P\{\underset{\sim}{\mu} \in \mathcal{R}_{n,n}\}P\{N = n\}$$

$$= \sum_{n=m}^{\infty} F\left(\frac{nd^2}{2\sigma^2}\right) P\{N = n\}$$

$$= E\{F\left(\frac{Nd^2}{2\sigma^2}\right)\} . \tag{2.2.5}$$

However, we have $N \geq \frac{2bS_m^2}{d^2}$ which implies $\frac{Nd^2}{2\sigma^2} \geq \frac{bS_m^2}{\sigma^2}$. Thus,

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) \geq E\{F(\frac{bS_m^2}{\sigma^2})\} .$$

Now, let $Z \sim \chi^2_{(p)}$ and let it also be independent of S_m^2. Then, we can write

$$E\{F(\frac{bS_m^2}{\sigma^2})\} = E\{\underset{\sim}{P}\{Z \leq \frac{bS_m^2}{\sigma^2} \mid S_m^2\}\}$$

$$= P\{Z/(S_m^2/\sigma^2) \leq pb'\} = 1 - \alpha ,$$

by the choice of b', since $\frac{Z}{p} \div \frac{S_m^2}{\sigma^2} \sim$ F-distribution with p,2p(m-1) degrees of freedom.

To prove Part (b), we consider the basic inequality,

$$\frac{2b\ S_m^2}{d^2} \leq N \leq \frac{2b\ S_m^2}{d^2} + m , \tag{2.2.6}$$

and then we divide all throughout by C. Now, taking expectations on all sides leads to the required result.

To prove Part (c), we again use the basic inequality (2.2.6); and we obtain

$$\left(\frac{2b\ S_m^2}{d^2}\right)^2 \leq N^2 \leq \left(\frac{2b\ S_m^2}{d^2}\right)^2 + 2m\left(\frac{2b\ S_m^2}{d^2}\right) + m^2 .$$

However, $2p(m-1)S_m^2/\sigma^2 \sim \chi^2_{(2p(m-1))}$ and so we have

$$E(S_m^2) = \sigma^2 \quad \text{and} \quad E(S_m^4) = \sigma^4\{(pm-p)^{-1} + 1\} .$$

Therefore, we can write

$$\left(\frac{2b\sigma^2}{d^2}\right)^2 \{(pm-p)^{-1}+1\} \le E(N^2) \le \left(\frac{2b\sigma^2}{d^2}\right)^2 \{(pm-p)^{-1} + 1\} + 2m\left(\frac{2b\sigma^2}{d^2}\right) + m^2, \tag{2.2.7}$$

and

$$- m^2 - 2m\left(\frac{2b\sigma^2}{d^2}\right) - \left(\frac{2b\sigma^2}{d^2}\right)^2 \le - \{E(N)\}^2 \le - \left(\frac{2b\sigma^2}{d^2}\right)^2 . \tag{2.2.8}$$

Combining (2.2.7) and (2.2.8), we get

$$1-2mp(m-1) \left(\frac{d^2}{2b\sigma^2}\right) - m^2p(m-1) \left(\frac{d^2}{2b\sigma^2}\right)^2 \le \text{Var}(N)\{p(m-1)(2b\sigma^2/d^2)^{-2}\} \le 1$$

$$+ 2mp(m-1)\left(\frac{d^2}{2b\sigma^2}\right) + m^2p(m-1)\left(\frac{d^2}{2b\sigma^2}\right)^2 ,$$

and now taking the limit as $d \to 0$ on all sides Part (c) follows.

To prove Part (d), we take the limit as $d \to 0$ in (2.2.5) and apply the dominated convergence theorem to write

$$\lim_{d\to 0} P\{\underset{\sim}{\mu} \in \mathcal{R}_{N,N}\} = \lim_{d\to 0} E\{F\left(\frac{Nd^2}{2\sigma^2}\right)\}$$

$$= E\{F\left(\lim_{d\to 0} \frac{Nd^2}{2\sigma^2}\right)\} .$$

From the inequality (2.2.6), it follows that

$$\lim_{d\to 0} \frac{Nd^2}{2\sigma^2} = \frac{b\ S_m^2}{\sigma^2} \quad \text{w.p.1} ,$$

and thus we have

$$\lim_{d\to 0} P(\underset{\sim}{\mu} \varepsilon \mathcal{R}_{N,N}) = E\{F(\frac{b\, S_m^2}{\sigma^2})\} .$$

This was earlier shown to be equal to $(1 - \alpha)$. This completes the proof of Theorem 2.2.1.

Remark 2.1: Part (a) tells us that the procedure (2.2.4) is "exactly consistent" in the Mukhopadhyay (1982) sense, while Part (d) shows that this procedure is also "asymptotically consistent". In Part (b), we have the limiting ratio $\frac{b}{a}$ which is always larger than one, that is to say that the procedure (2.2.4) oversamples in estimating C even asymptotically.

2.2.2. A Sequential Procedure

We start with $m(\geq 2)$ observations from each population, and then define the following stopping rule:

$$N = \inf\{n \geq m: \; n \geq \frac{2a\, S_n^2}{d^2}\} . \tag{2.2.9}$$

When we stop, we have the samples $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_N$. We compute $\bar{\underset{\sim}{T}}_{N,N}$ and propose the region $\mathcal{R}_{N,N}$ as in (2.1.1). Some of the properties of this procedure are stated in Theorems 2.2.2 and 2.2.3.

Theorem 2.2.2: For the procedure in (2.2.9), for all $\underset{\sim}{\mu} \varepsilon \mathbb{R}^p$, and $\sigma \varepsilon (0,\infty)$ we have:

(a) $E(N) \leq C + m + 2$ for all $d > 0$,

(b) $N/C \to 1$ w.p.1 as $d \to 0$.

(c) $E(N/C) \to 1$ <u>as</u> $d \to 0$,

(d) $P(\underset{\sim}{\mu} \in R_{N,N}) \to 1 - \alpha$ <u>as</u> $d \to 0$, <u>and</u>

(e) $\frac{\sqrt{p}(N-C)}{\sqrt{C}} \xrightarrow{L} N(0,1)$ <u>as</u> $d \to 0$.

<u>Proof</u>: To prove Part (a), notice from (2.2.9) that we have

$$N \le \frac{2a\ S_{N-1}^2}{d^2} + m \ ,$$

which implies

$$(N-2)(N-m) \le \frac{2a}{d^2}\ (N-2)\ S_{N-1}^2 \ .$$

Now, first assume that $E(N) < \infty$. Then, by Wald's first equation, we have

$$\begin{aligned} \frac{2a}{d^2}\ E(N)\ \sigma^2 &\ge E\{(N-2)(N-m)\} \\ &\ge E\{N^2 - (m+2)N\} \\ &\ge \{E(N)\}^2 - (m+2)E(N) \ , \end{aligned}$$

which leads to

$$E(N) - (m+2) \le C \ ,$$

that is

$$E(N) \le C + m+2 \ ,$$

assuming, of course, $E(N) < \infty$. The case of "$E(N) = \infty$" may be tackled as follows: Define $N_k = \min(N,k)$, and we have

$$E(N_k) \le C+m+2 \text{ for all } k = 1,2,\ldots .$$

Since $N_k \uparrow N$ w.p.1 as $k \to \infty$, by the Montone Convergence Theorem we conclude that $E(N) \le C + m + 2$.

To prove Part (b), we consider the basic inequality

$$\frac{2a\, S_N^2}{d^2} \le N \le \frac{2a\, S_{N-1}^2}{d^2} + m , \tag{2.2.10}$$

and then dividing by C and taking limits as $d \to 0$ throughout this inequality we obtain $\frac{N}{C} \to 1$ w.p.1 as $d \to 0$.

To prove Part (c), we note from Part (a) that

$$\limsup_{d\to 0} E(N/C) \le 1 ,$$

and from Part (b) and Fatou's Lemma together we get

$$\liminf_{d\to 0} E(N/C) \ge 1 .$$

This implies $E(N/C) \to 1$ as $d \to 0$.

To prove Part (d), we first note the event $\{N = n\}$ and the event $\{\mu \in \mathcal{R}_{n,n}\}$ are independent, and thus we obtain

$$P(\mu \in \mathcal{R}_{N,N}) = E\{F(\frac{Nd^2}{2\sigma^2})\} . \tag{2.2.11}$$

Then, from Part (b) we can easily obtain

$$\frac{Nd^2}{2\sigma^2} \to a \quad \text{w.p.1}, \quad \text{as } d \to 0 .$$

Now, combining this and the dominated convergence theorem we have

$$\lim_{d\to 0} P(\underset{\sim}{\mu} \in R_{N,N}) = E\{F(\lim_{d\to 0} \frac{Nd^2}{2\sigma^2})\} = F(a) = 1 - \alpha,$$

by the choice of a.

To prove Part (e), we first use Part (b) and Anscombe's (1952) results (Theorems A1.1 and A1.2 in the Appendix) to conclude

$$\frac{\sqrt{pN}\ (S_N^2 - \sigma^2)}{\sigma^2} \xrightarrow{L} N(0,1) \quad \text{as} \quad d \to 0,$$

and

$$\frac{\sqrt{pN}\ (S_{N-1}^2 - \sigma^2)}{\sigma^2} \xrightarrow{L} N(0,1) \text{ as } d \to 0.$$

Then, using the theorem of Ghosh and Mukhopadhyay (1975) (Theorem A2 in the Appendix), we have

$$\frac{\sqrt{p}\ (N-C)}{\sqrt{C}} \xrightarrow{L} N(0,1) \text{ as } d \to 0.$$

This completes the proof of Theorem 2.2.2.

Before we state and prove the next stronger version of our result, let us discuss some basic notations borrowed from non-linear renewal theoretic results of Woodroofe (1977) (Section A4 in the Appendix). The sequential procedure (2.2.9) can be equivalently stated as follows:

$$N = \inf\{n \geq m: \frac{2p(n-1)S_n^2}{\sigma^2} \leq \frac{2p(n-1)n}{C}\}$$

$$= \inf\{n \geq m: \sum_{i=1}^{n-1} Z_i \leq \frac{2p(n-1)n}{C}\} , \qquad (2.2.12)$$

where $Z_1, Z_2, \ldots$ are i.i.d. $\chi^2_{(2p)}$ random variables. The condition (2.5) in Woodroofe (1977) (Condition C3 in Section A4) is easily shown to be satisfied. Also, one can readily see that (2.2.12) had the same form as Woodroofe's (1977) equation (1.1) (C1 in A4) with his $\alpha = 2$, $\beta = 1$, $c = \frac{2p}{C}$, $\mu = 2p$, $\tau^2 = 4p$, $\lambda = C$, $a = p$, $L(n) = 1 + n^{-1}$, and starting sample size $(m-1)$. The constant ν given in (2.4) of Woodroofe (1977) (C3 in A4) would have to be evaluated as

$$\nu = p+1 - \sum_{n=1}^{\infty} n^{-1} E\{(W_n-4np)^+\} ,$$

where $W_n = \sum_{i=1}^{n} Z_i$, and $(x)^+ = \max(0,x)$. Since Z_i's are i.i.d. $\chi^2_{(2p)}$ it follows that $W_n \sim \chi^2_{(2np)}$. Thus, we can write

$$E\{(W_n-4np)^+\} = \frac{2^{-np}}{\Gamma(np)} \int_{4np}^{\infty} (\omega-4np)\, \omega^{np-1}\, e^{-\omega/2}\, d\omega .$$

Let $G(.,.)$ be the incomplete gamma function defined by

$$G(a^*,b^*) = \int_{b^*}^{\infty} t^{a^*-1}\, e^{-t}\, dt, \qquad (2.2.13)$$

for $a^*, b^* > 0$.

Then it follows that

$$E\{(W_n-4np)^+\} = \{\Gamma(np)\}^{-1} \{2(2np)^{np}\, e^{-2np} - 2np\, G(np, 2np)\} .$$

Let us write

$$\eta = (2p)^{-1} (\nu-2) - 1 . \qquad (2.2.14)$$

Theorem 2.2.3: For the procedure in (2.2.9), for all $\underset{\sim}{\mu} \varepsilon \mathbb{R}^p$, and $\sigma \varepsilon (0,\infty)$ we have as $d \to 0$:

(a) $E(N) = C + \eta + o(1)$ if $m > 1 + p^{-1}$, and

(b) $P(\underset{\sim}{\mu} \varepsilon R_{N,N}) = 1-\alpha + \frac{1}{2}(d/\sigma)^2 \{\eta + \frac{1}{4p}(p-a-2)\} f(a) + o(d^2)$,

if (i) $m \geq 4$ for $p=1$, (ii) $m \geq 2$ for $p=2,3,\ldots$.

where the number η is defined in (2.2.14), and $f(\cdot)$ is the p.d.f. of $\chi^2_{(p)}$.

Proof: Part (a) follows directly from theorem 2.4 of Woodroofe (1977) (Theorem A4.1 in the Appendix) with the number η coming from (2.2.14).

To prove Part (b), we recall from (2.2.11) that we have

$$P(\underset{\sim}{\mu} \varepsilon R_{N,N}) = E\{F(\frac{Nd^2}{2\sigma^2})\} .$$

Using Taylor's expansion for the function $F(\cdot)$ at the point a, we have

$$F(\frac{Nd^2}{2\sigma^2}) = F(a) + (\frac{Nd^2}{2\sigma^2} - a) F'(a) + \frac{1}{2}(\frac{Nd^2}{2\sigma^2} - a)^2 F''(W),$$

where W is a suitable random variable between a and $\frac{Nd^2}{2\sigma^2}$. This implies

$$F(\frac{Nd^2}{2\sigma^2}) = 1-\alpha + \frac{1}{2}(d/\sigma)^2 \{(N-C)f(a) + \frac{a}{2} \frac{(N-C)^2}{C} f'(W)\},$$

where $f(\cdot)$ is the p.d.f of a $\chi^2_{(p)}$ random variable. Hence, we get

$$P(\underset{\sim}{\mu} \in R_{N,N}) = 1-\alpha + \frac{1}{2}(d/\sigma^2)\ \{f(a)E(N-C) + \frac{a}{2}\, E\{N^* f'(W)\}\}, \qquad (2.2.15)$$

where $N^* = \frac{(N-C)^2}{C}$. It is clear that $W \to a$ in probability as $d \to 0$. Now, let $h(x;p) = e^{-x/2}\, x^{(p/2)-1}$. Then, $h(x;p)$ attains its maximum at $x = p-2$ for every fixed $p > 2$. Also, for $x > 0$, we can write

$$f'(x) = -k_1\, h(x;p) + k_2 h(x;p-2),$$

where $k_1 = \{2^{(p/2)+1}\, \Gamma(p/2)\}^{-1}$ and $k_2 = \{(p/2)-1\}\{2^{p/2}\, \Gamma(p/2)\}^{-1}$. We now consider several separate cases for p, namely $p > 4$, $p = 1,2,3$ and 4.

Case 1: Let $p > 4$. Then,

$$|N^* f'(W)| \le N^*\{|-k_1 h(p-2;p)| + |k_2 h(p-4;p-2)|\} .$$

Notice that the two terms inside the brackets are bounded. Also, Woodroofe's (1977) Theorem 2.3 (Theorem A4.2 in the Appendix) implies that N^* is uniformly integrable if $m>1+p^{-1}$. Thus, $N^* f'(W)$ is also uniformly integrable if $m>1+p^{-1}$. Now, from Part (e) of Theorem 2.2.2, it follows that $pN^* \xrightarrow{L} \chi^2_{(1)}$ as $d \to 0$. Since $W \to a$ in probability as $d \to 0$, $pN^* f'(W) \to f'(a)\ \chi^2_{(1)}$ as $d \to 0$. Hence, we obtain $E\{N^* f'(W)\} = p^{-1} f'(a) + o(1)$ as $d \to 0$. Thus, (2.2.15) and the identity $af'(a) = \frac{1}{2}(p-a-2)f(a)$ immediately lead to Part (b).

Case 2: Let $p = 4$. Then,

$$|N^*f'(W)| = N^*| - \{2^3\Gamma(2)\}^{-1} h(W;4) + \{2^2\Gamma(2)\}^{-1}h(W;2)|$$

$$\leq N^*\{\frac{1}{8} h(2;4) + \frac{1}{4}\} ,$$

where the quantity inside the brackets is bounded positive constant. Therefore, $N^*f'(W)$ is again uniformly integrable if $m > 1 + p^{-1}$, and we obtain the same result as in Case 1.

Case 3: Let $p = 3$. Then,

$$|N^*f'(W)| = N^*| - \{2^{5/2} \Gamma(3/2)\}^{-1} e^{-W/2} W^{1/2} + \frac{1}{2}\{2^{3/2} \Gamma(3/2)\}^{-1} \times e^{-W/2} W^{-1/2} |$$

$$= N^*\{2^{5/2} \Gamma(3/2)\}^{-1}| - e^{-W/2} W^{1/2} + e^{-W/2} W^{-1/2}|$$

$$= |V|, \text{ say.}$$

Let A be the event $N > \frac{1}{2} C$. Write $V = VI(A) + VI(A^c)$, where $I(\cdot)$ is the indicator function. Then,

$$\lim_{d\to 0} E(V) = \lim_{d\to 0} E\{VI(A)\} + \lim_{d\to 0} E\{VI(A^c)\}, \text{ if the limits exist.}$$

Now,

$$|VI(A)| = N^*\{2^{5/2} \Gamma(3/2)\}^{-1} |-e^{-W/2} W^{1/2} + e^{-W/2} W^{-1/2}|I(A)$$

$$\leq N^*\{2^{5/2} \Gamma(3/2)\}^{-1} \{e^{-W/2} W^{1/2} I(A) + e^{-W/2} \times W^{-1/2} I(A)\}.$$

Since W is between a and $\frac{Nd^2}{2\sigma^2}$ and A is the event $N > \frac{1}{2} C$, we have $\frac{Nd^2}{2\sigma^2} > \frac{1}{2}$ a on the set A. Thus $W > \frac{1}{2}$ a on the set A, and we obtain

$$|VI(A)| \le N^*\{2^{5/2}\,\Gamma(3/2)\}^{-1}\,\{e^{-1/2} + (\tfrac{1}{2}a)^{-1/2}\}\ .$$

Hence, $|VI(A)|$ is uniformly integrable if $m > 1 + p^{-1}$. Also, $I(A) \to 1$ in probability as $d \to 0$. Thus, we have

$$E\{VI(A)\} = p^{-1}f'(a) + o(1) \text{ as } d \to 0.$$

On the other hand, we know that $N \le \frac{1}{2}C$ on the set A^c and thus,

$$E\{|VI(A^c)|\} = \int_{A^c} |N^* f'(W)|\,dP$$

$$\le \{2^{5/2}\,\Gamma(3/2)\}^{-1}$$

$$\times \left|\int_{A^c} - N^* e^{-W/2}\,(W^{1/2} - W^{-1/2})\,dP\right|$$

$$\le \{2^{5/2}\,\Gamma(3/2)\}^{-1}\,\{\int_{A^c} N^*\,dP + \int_{A^c} N^* W^{-1/2}\,dP\}\ .$$

Again, we have W between a and $\frac{Nd^2}{2\sigma^2}$, and $\frac{Nd^2}{2\sigma^2} \le \frac{1}{2}a$ on the set A^c. Thus, we have $W > \frac{Nd^2}{2\sigma^2}$ which implies $W^{-1/2} < (\frac{Nd^2}{2\sigma^2})^{-1/2}$. Therefore,

$$E\{|VI(A^c)|\} \le \{2^{5/2}\,\Gamma(3/2)\}^{-1}\,\{\int_{A^c} C(1-\tfrac{N}{C})^2\,dP + \int_{A^c} C(1-\tfrac{N}{C})^2$$

$$\times (\frac{Nd^2}{2\sigma^2})^{-1/2}\,dP\}$$

$$\le \{2^{5/2}\,\Gamma(3/2)\}^{-1}\,\{C\int_{A^c} dP + a^{-1/2}$$

$$\times C\int_{A^c} (C/N)^{1/2}\,dP\}$$

$$\le \{2^{5/2}\,\Gamma(3/2)\}^{-1}$$

$$\times \{CP(N \le \tfrac{1}{2}C) + a^{-1/2}\,C^{3/2}\,P(N \le \tfrac{1}{2}C)\}\ .$$

From Lemma 2.3 of Woodroofe (1977) (Lemma A4 in the Appendix), we have for $0 < \gamma < 1$,

$$P(N \le \frac{1}{2} C) = o(C^{-3(m-1)}) + o(C^{-\gamma r/2}),$$

as $d \to 0$ where $E(Z_1^r) < \infty$ with suitable $r \ge 2$. Thus, one can readily see that for $m > 1 + \frac{1}{2}$, $\lim_{d\to 0} E\{VI(A^c)\} = 0$. This leads to Part (b) for $p = 3$, since now we can write

$$E(V) = p^{-1} f'(a) + o(1) \text{ as } d \to 0.$$

Case 4: Let $p = 2$. Then,

$$|N^* f'(W)| = N^* |-\frac{1}{4} e^{-W/2}| \le N^* .$$

Since, N^* is uniformly integrable for $m > 1 + p^{-1}$ it follows that $|N^* f'(W)|$ is uniformly integrable. This leads to Part (b) for $p = 2$, as in Case 1.

Case 5: Let $p = 1$. Then,

$$|N^* f'(W)| = N^* \{2^{3/2} \Gamma(1/2)\}^{-1} | e^{-W/2} W^{-1/2} + e^{-W/2} W^{-3/2}| .$$

Again, let A denote the event $N > \frac{1}{2} C$. Then

$$|VI(A)| \le N^* \{2^{3/2} \Gamma(\tfrac{1}{2})\}^{-1} \{e^{-W/2} W^{-1/2} I(A) + e^{-W/2} W^{-3/2} I(A)\}$$

$$\le N^* \{2^{3/2} \Gamma(\tfrac{1}{2})\}^{-1} \{(a/2)^{-1/2} + (a/2)^{-3/2}\},$$

where the quantities inside the brackets in the right hand side are positive constants. Hence, $|VI(A)|$ is uniformly integrable if $m > 2$, which in turn implies that $E\{VI(A)\} = p^{-1} f(a) + o(1)$ as $d \to 0$. Again with

$K^* = \{2^{3/2}\,\Gamma(\frac{1}{2})\}^{-1}$, we can write

$$E\{|VI(A^c)|\} = \int_{A^c} |V|\,dP$$

$$\leq K^* \{\int_{A^c} N^* e^{-W/2}\, W^{-1/2}\, dP + \int_{A^c} N^*\, e^{-W/2}\, W^{-3/2}\, dP\}$$

$$\leq K^* \{\int_{A^c} N^*\, W^{-1/2}\, dP + \int_{A^c} N^*\, W^{-3/2}\, dP\}.$$

Also, $W^{-1/2} \leq (\frac{Nd^2}{2\sigma^2})^{-1/2}$ on the set A^c, and so we obtain

$$E\{|VI(A^c)|\} \leq K^* \int_{A^c} C(1 - \frac{N}{C})^2 (\frac{Nd^2}{2\sigma^2})^{-1/2}\, dP + K^* \int_{A^c} C(1 - \frac{N}{C})^2$$

$$\times\ (\frac{Nd^2}{2\sigma^2})^{-3/2}\, dP$$

$$\leq K^* a^{-1/2}\, C \int_{A^c} (C/N)^{1/2}\, dP + k^* a^{-3/2}\, C \int_{A^c} (C/N)^{3/2}\, dP$$

$$\leq k^* a^{-1/2}\, C^{3/2}\, P(N \leq \frac{1}{2} C) + k^* a^{-3/2}\, C^{5/2}\, P(N < \frac{1}{2} C)\ .$$

In order to make both $C^{3/2}\, P(N \leq \frac{1}{2} C)$ and $C^{5/2}\, P(N \leq \frac{1}{2} C)$ converge to zero as $d \to 0$, the same basic techniques used at the end of Case 3 would now lead us to the sufficient condition that $\frac{5}{2} - (m-1) < 0$, that is we need $m > \frac{7}{2}$. Earlier, we found the condition $m > 2$. Thus, for $m \geq 4$, we have $\lim_{d\to 0} E\{VI(A^c)\} = 0$. Hence, for $p = 1$ we have

$$E\{N^* f'(W)\} = p^{-1} f'(a) + o(1) \text{ as } d \to 0 \text{ if } m \geq 4.$$

This completes the proof of Theorem 2.2.3.

Remark 2.2: Part (a) of this theorem shows that the sequential procedure (2.2.9) is indeed "asymptotically second order efficient" in the Ghosh-Mukhopadhyay (1981) sense, since we have $\lim_{d \to 0} E(N-C) = \eta$.

2.3 The Case of Two Covariance Matrices Being Unequal But Sample Sizes are Equal

Let $\sigma^2 = \sigma_1^2 + \sigma_2^2$ and $r = s = n$. Utilizing (2.1.3) in this case, we would have

$$P(\underset{\sim}{\mu} \in R_{n,n}) = F\left(\frac{nd^2}{\sigma^2}\right) , \tag{2.3.1}$$

and from (2.1.6) we would obtain $n \geq \frac{a\sigma^2}{d^2}$. If σ_1 and σ_2 were known, the required optimal fixed sample size would have been $C = \frac{a\sigma^2}{d^2}$. But σ_1 and σ_2 are unknown, and so we will now consider two procedures for determining the sample size N as a suitable random variable, and this N will estimate the unknown C.

We define

$$Z_n^2 = \{p(n-1)\}^{-1} \sum_{i=1}^{n} (\underset{\sim}{X}_i - \underset{\sim}{Y}_i - \bar{\underset{\sim}{T}}_{n,n})' H^{-1} (\underset{\sim}{X}_i - \underset{\sim}{Y}_i - \bar{\underset{\sim}{T}}_{n,n}), \tag{2.3.2}$$

for $n \geq 2$.

2.3.1 A Two-Stage Procedure

We start with $m(\geq 2)$ observations from each population, and define the following stopping rule

$$N = \max\{m, [bZ_m^2/d^2] + 1\}, \tag{2.3.3}$$

where $b = pb'$ and b' is the upper $100\alpha\%$ point of the F-distribution with degrees of freedom $p, p(m-1)$. Thus, from the samples $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_N$ we compute, $\bar{\underset{\sim}{T}}_{N,N}$ and propose the region $\mathcal{R}_{N,N}$ as in (2.1.1). Some of the properties of this procedure are listed in Theorem 2.3.1.

Theorem 2.3.1: For the procedure (2.3.3), for all $\underset{\sim}{\mu} \in \mathbb{R}^p$, and $\sigma_1, \sigma_2 \in (0,\infty)$ we have:

(a) $P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) \geq 1-\alpha$ for all $d > 0$,

(b) $E(N/C) \to \frac{b}{a}$ as $d \to 0$,

(c) $\mathrm{Var}(N) \{\frac{p(m-1)}{2} (\frac{d^2}{b\sigma^2})^2\} \to 1$ as $d \to 0$, and

(d) $P(\underset{\sim}{\mu} \in \mathcal{R}_{N,N}) \to 1-\alpha$ as $d \to 0$.

We omit its proof for brevity as it follows along the lines of proof given for Theorem 2.2.1.

2.3.2 A Sequential Procedure

We start with $m(\geq 2)$ observations from each population, and then define the following stopping rule:

$$N = \inf\{n \geq m: n \geq \frac{aZ_n^2}{d^2}\} . \qquad (2.3.4)$$

When we stop, we have the samples $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_N$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_N$. We compute $\bar{\underset{\sim}{T}}_{N,N}$ and propose the region $\mathcal{R}_{N,N}$ as defined in (2.1.1). Some of the properties of this procedure are listed in Theorems 2.3.2 and 2.3.3.

Theorem 2.3.2: For the procedure (2.3.4), for all $\underset{\sim}{\mu} \varepsilon \mathbb{R}^p$, *and* $\sigma_1, \sigma_2 \varepsilon (0,\infty)$, *we have:*

(a) $E(N) \leq C + m + 2$ *for all* $d > 0$,

(b) $N/C \to 1$ *w.p.1 as* $d \to 0$,

(c) $E(N/C) \to 1$ *as* $d \to 0$,

(d) $P(\underset{\sim}{\mu} \varepsilon R_{N,N}) \to 1-\alpha$ *as* $d \to 0$, *and*

(e) $\dfrac{\sqrt{2p}\,(N-C)}{\sqrt{C}} \xrightarrow{L} N(0,1)$ *as* $d \to 0$.

We omit its proof for brevity. We can easily construct a proof along the lines of proof of Theorem 2.2.2.

Theorem 2.3.3: For the procedure (2.3.4), for all $\underset{\sim}{\mu} \varepsilon \mathbb{R}^p$, *and* $\sigma_1, \sigma_2 \varepsilon (0,\infty)$, *we have as* $d \to 0$:

(a) $E(N) = C + \eta + o(1)$, *if* $m > 1 + 2p^{-1}$;

(b) $P(\underset{\sim}{\mu} \varepsilon R_{N,N}) = 1-\alpha + (d/\sigma)^2\{\eta + \frac{1}{2p}(p-a-2)\}f(a)$

$+ o(d^2)$, *if* (i) $m \geq 7$ *for* $p = 1$, (ii) $m \geq 3$ *for* $p = 2$ or 3 *and* (iii) $m \geq 2$ *for* $p \geq 4$.

Here $f(\cdot)$ *is as in Theorem 2.2.3, and* $\eta = \frac{1}{p}(\nu-2)-1$ *where* ν *is given by*

$$\nu = \frac{p}{2} + 1 - \sum_{n=1}^{\infty} \{n\Gamma(nP/2)\}^{-1}\{2(np)^{np/2}\, e^{-np} - npG(\tfrac{np}{2}, np)\},$$

$G(.,.)$ *being defined in* (2.2.13).

We omit its proof for brevity, as it can be given along the lines of proof of Theorem 2.2.3.

2.4 The Case of Two Covariance Matrices Being Unequal and Unequal Sample Sizes

In this case again, the confidence coefficient associated with the region $R_{r,s}$ is the same as in (2.1.3). Our objective is to minimize $(r+s)$ such that $\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s} \le \frac{d^2}{a}$. Using Lagrange's multiplier λ, we have the equation

$$(r+s) + \lambda\left(\frac{\sigma_1^2}{r} + \frac{\sigma_2^2}{s} - \frac{d^2}{a}\right) . \tag{2.4.1}$$

We find that (2.4.1) is minimized for

$$r = r^* = \frac{a\,\sigma_1}{d^2}(\sigma_1 + \sigma_2) , \tag{2.4.2}$$

$$s = s^* = \frac{a\,\sigma_2}{d^2}(\sigma_1 + \sigma_2) . \tag{2.4.3}$$

Had σ_1, σ_2 been known, the optimal fixed sample sizes would have been r^* and s^* from the $\underset{\sim}{X}$'s and $\underset{\sim}{Y}$'s respectively, and the total optimal fixed sample size would then turn out to be

$$n^* = r^* + s^* = \frac{a}{d^2}(\sigma_1 + \sigma_2)^2 . \tag{2.4.4}$$

We note that

$$\frac{r^*}{s^*} = \frac{\sigma_1}{\sigma_2} . \tag{2.4.5}$$

But, since σ_1, σ_2 are actually unknown, we will consider two procedures for determining (R,S) as random variables in order to estimate (r^*, s^*).

2.4.1 A Two-Stage Procedure

We start with $m(\geq 2)$ observations from each population, and then define the following stopping rule: Let $N = R + S$ with

$$R = \max\{m, [hU_m^2/d^2]+1\}, \tag{2.4.6}$$

$$S = \max\{m, [hV_m^2/d^2]+1\}, \tag{2.4.7}$$

where h is a suitable constant such that

$$E\left\{F\left(\frac{h}{\frac{\sigma_1^2}{U_m^2} + \frac{\sigma_2^2}{V_m^2}}\right)\right\} = 1-\alpha . \tag{2.4.8}$$

The reader may note that "h" depends only on m, p and α.

Thus, when we stop, we would have the samples $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_R$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_S$. We compute $\bar{\underset{\sim}{T}}_{R,S}$ and then propose the corresponding region $\mathcal{R}_{R,S}$ as in (2.1.1). Some of the properties of this procedure are listed in Theorem 2.4.1.

<u>Theorem 2.4.1:</u> <u>For the procedure (2.4.6) and (2.4.7), for all</u> $\underset{\sim}{\mu} \in \mathbb{R}^p$, <u>and</u> $\sigma_1, \sigma_2 \in (0,\infty)$ <u>we have:</u>

(a) $P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) \geq 1-\alpha$ <u>for all</u> $d > 0$;

(b) $E(R/r^*) \to \dfrac{h\sigma_1}{a(\sigma_1 + \sigma_2)}$,

$$E(S/s^*) \to \frac{h\sigma_2}{a(\sigma_1 + \sigma_2)} \text{ , \underline{and}}$$

$$E(N/n^*) \to \frac{h(\sigma_1^2 + \sigma_2^2)}{a(\sigma_1 + \sigma_2)^2} \quad \underline{\text{as}} \quad d \to 0;$$

(c) $\mathrm{Var}(R)\{\frac{p(m-1)}{2} (\frac{d^2}{h\sigma_1^2})^2\} \to 1$, <u>and</u>

$$\mathrm{Var}(S)\{\frac{p(m-1)}{2} (\frac{d^2}{h\sigma_2^2})^2\} \to 1 \ \underline{\text{as}}\ d \to 0; \ \underline{\text{and}}$$

(d) $P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) \to 1-\alpha$ <u>as</u> $d \to 0$.

<u>Proof</u>: To prove Part (a), we first note the event $\{R=r, S=s\}$ and the event $\{\underset{\sim}{\mu} \in \mathcal{R}_{r,s}\}$ are independent, and thus we obtain

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) = E\left\{F\left\{\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} d^2\right\}\right\} .$$

However, we have $R \geq \frac{hU_m^2}{d^2}$ and $S \geq \frac{hV_m^2}{d^2}$ which imply that

$$\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S} \leq \frac{d^2}{h} \frac{\sigma_1^2}{U_m^2} + \frac{\sigma_2^2}{V_m^2} ,$$

that is

$$\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} d^2 \geq h\left(\frac{\sigma_1^2}{U_m^2} + \frac{\sigma_2^2}{V_m^2}\right)^{-1} .$$

Thus, we obtain

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) \geq E\left\{F\left\{h\left(\frac{\sigma_1^2}{U_m^2} + \frac{\sigma_2^2}{V_m^2}\right)^{-1}\right\}\right\} = 1-\alpha,$$

by the choice of h.

We omit proofs of Parts (b), (c) and (d) for brevity as they follow along the lines of proofs given for Parts (b), (c) and (d) of Theorem 2.2.1.

2.4.2 A Sequential Procedure

We start with $m(\geq 2)$ observation from each population. Then, if at any stage we have taken $r(\geq m)$ observations on $\underset{\sim}{X}$'s and $s(\geq m)$ observations on $\underset{\sim}{Y}$'s, we take the next observation, if needed,

(a) on $\underset{\sim}{X}$'s if $\frac{r}{s} \leq \frac{U_r}{V_s}$,

(b) on $\underset{\sim}{Y}$'s if $\frac{r}{s} > \frac{U_r}{V_s}$.

The motivation seems to be clear when one looks at (2.4.5). We now propose four more or less equivalent stopping rules, easily motivated from (2.1.6) and (2.4.2) - (2.4.4).

R_1^*: The stopping time $N = N(d)$ is the smallest positive integers $n(\geq 2m)$ such that if $R=r$ observations on $\underset{\sim}{X}$'s and $S=s$ observations on $\underset{\sim}{Y}$'s have been taken with $n = r+s$ such that,

$$n \geq \frac{a}{d^2}(U_r + V_s)^2 . \tag{2.4.9}$$

R_2^*: The same, with (2.4.9) replaced by

$$\frac{U_r^2}{r} + \frac{V_s^2}{s} \leq \frac{d^2}{a} \tag{2.4.10}$$

R_3^*: The same, with (2.4.9) replaced by

$$r \geq \frac{aU_r}{d^2}(U_r + V_s) \text{ and } s \geq \frac{a\ V_s}{d^2}(U_r + V_s) \qquad (2.4.11)$$

R_4^*: The same, with (2.4.9) replaced by

$$r^2 \geq \frac{an}{d^2} U_r^2 \text{ and } s^2 \geq \frac{an}{d^2} V_s^2 \qquad (2.4.12)$$

These rules are of the same form as those of the rules defined in Mukhopadhyay (1976). Using any particular one of these rules, we finally obtain R observations on $\underset{\sim}{X}$'s and S observations on $\underset{\sim}{Y}$'s, namely $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_R$ and $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_S$. We compute $\bar{\underset{\sim}{T}}_{R,S}$ and propose the corresponding region $\mathcal{R}_{R,S}$ as in (2.1.1). Some of the properties of these rules are listed in Theorems 2.4.2 and 2.4.3.

<u>Theorem 2.4.2:</u> <u>For the procedure defined by</u> R_4^*, <u>for all</u> $\underset{\sim}{\mu} \in \mathbb{R}^p$, $\sigma_1, \sigma_2 \in (0,\infty)$ <u>and</u> $d > 0$, <u>we have</u>:

(a) $E(R) \leq r^* + m + \sigma_1 D$,

(b) $E(S) \leq s^* + m + \sigma_2 D$, <u>and</u>

(c) $E(N) \leq n^* + 4m$,

where $D = 2(am/d^2)^{1/2}$.

<u>Proof</u>: From the definition of the procedure in R_4^* we obtain

$$(R-1)^2 \leq \frac{a}{d^2}(N-1)\ U_{R-1}^2,$$

on the set $R > m$. But, since $(R-m)^2 \leq (R-1)^2$ and $\frac{a}{d^2}(N-1) \leq \frac{a}{d^2} N$

we obtain the following inequality

$$(R-m)^2 \leq \frac{a}{d^2} N \, U_{R-1}^2 .$$

which implies that

$$\frac{(R-2)(R-m)^2}{N} \leq \frac{a}{d^2} (R-2) \, U_{R-1}^2 .$$

From Wald's 1st equation (Govindarajulu (1981), page 43) we have $E\{(R-2)U_{R-1}^2\} \leq \sigma_1^2 E(R)$ and using convexity argument and Jensen's inequality we also have

$$\frac{E(R-2)\{E(R-m)\}^2}{E(N)} \leq E\{\frac{(R-2)(R-m)^2}{N}\} .$$

Therefore, we have

$$\{E(R-m)\}^2 \leq \frac{a \, \sigma_1^2}{d^2} E(N) . \tag{2.4.13}$$

In the same way we can obtain

$$\{E(S-m)\}^2 \leq \frac{a \, \sigma_2^2}{d^2} E(N) . \tag{2.4.14}$$

Notice, that $N-2m = (R-m) + (S-m)$. Therefore, from (2.4.13) and (2.4.14) we obtain

$$\{E(N-2m)\}^2 \leq \frac{a}{d^2} E(N)(\sigma_1 + \sigma_2)^2 = n^* E(N) .$$

But, $\{E(N)\}^2 - 4m \, E(N) \leq \{E(N-2m)\}^2$. Hence, we have

$$E(N) \le n^{*} + 4m \ .$$

Of course, we assumed thus far that $E(N) < \infty$. In case $E(N) = \infty$, we can use a truncation technique similar to the one we used in the proof of Part (a) in Theorem 2.2.2. This proves Part (c). To prove Part (a), we have from (2.4.13):

$$E(R-m) \le \left(\frac{a}{d^2}\right)^{1/2} \sigma_1 \, E^{1/2}(N) \ .$$

From Part (c) we have $E^{1/2}(N) \le (n^{*} + 4m)^{1/2}$, and we also know that $(n^{*} + 4m)^{1/2} \le n^{*1/2} + 2m^{1/2}$. Therefore, we have

$$E(R-m) \le \left(\frac{a}{d^2}\right)^{1/2} \sigma_1 (n^{*1/2} + 2m^{1/2}) \ . \qquad (2.4.15)$$

From (2.4.2) and (2.4.4) we obtain

$$\left(\frac{a}{d^2}\right) \sigma_1 \, n^{*1/2} = r^{*} \ .$$

From this and (2.4.15) we have

$$E(R) \le r^{*} + m + \sigma_1 D,$$

where $D = 2\left(\frac{am}{d^2}\right)^{1/2}$.

This proves Part (a). In the same way we can prove Part (b). This completes the proof of Theorem 2.4.2.

<u>Theorem 2.4.3:</u> <u>For the procedures defined by</u> $R_1^{*} - R_4^{*}$, <u>for all</u> $\underset{\sim}{\mu} \in \mathbb{R}^p$, <u>and</u> $\sigma_1, \sigma_2 \in (0,\infty)$ <u>we have as</u> $d \to 0$:

(a) $N/n^* \to 1$ w.p.1,

(b) $E(N/n^*) \to 1$, and

(c) $P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) \to 1-\alpha$.

Proof: Let N_i denote the total sample size required by the rule R_i^*. Notice that

$$N_1 \le N_2 \le N_3 \le N_4 \quad \text{w.p.1} . \tag{2.4.16}$$

Now, from R_1^* we obtain the inequality

$$N_1 \ge \frac{a}{d^2}(U_R + V_S)^2 .$$

since $U_R \to \sigma_1$ and $V_S \to \sigma_2$ w.p.1 as $d \to 0$, we obtain

$$\liminf_{d\to 0} N_1/n^* \ge 1 \quad \text{w.p.1} . \tag{2.4.17}$$

From R_4^*, again we obtain

$$(R-m)^2 \le \frac{a}{d^2} N\, U_{R-1}^2 \quad \text{and} \quad (S-m)^2 \le \frac{a}{d^2} N\, U_{S-1}^2 .$$

Which implies

$$N_4 \;=\; R+S \le \frac{a}{d^2}\left(\frac{S}{R}+1\right) U_{R-1}^2 + \frac{a}{d^2}\left(\frac{R}{S}+1\right) U_{S-1}^2 + 4m . \tag{2.4.18}$$

From the Lemma of Robbin et al. (1967) (Lemma A3 in the Appendix) we obtain as $d \to 0$:

$$\frac{R}{S} \to \frac{\sigma_1}{\sigma_2} \quad \text{w.p.1} . \tag{2.4.19}$$

But, we know that $U^2_{R-1} \to \sigma^2_1$ and $V^2_{S-1} \to \sigma^2_2$ w.p.1 as $d \to 0$. Therefore, from these facts together with (2.4.18), and (2.4.19) we obtain

$$\limsup_{d\to 0} N_4/n^* \leq 1 \ . \tag{2.4.20}$$

Now, from (2.4.16), (2.4.17), and (2.4.20) Part (a) follows.

To prove Part (b), we first have from Part (a) and Fatou's lemma

$$\liminf_{d\to 0} E(N_i^*/n^*) \geq 1 \ ,$$

$i = 1, 2, 3, 4$.

Then, from Part (c) of Theorem 2.4.2 we obtain

$$\limsup_{d\to 0} E(N_4/n^*) \leq 1 \ .$$

This implies Part (b).

To prove Part (c), notice that the events $\{R=r, S=s\}$, and $\{\underset{\sim}{\mu} \in \mathcal{R}_{r,s}\}$ are independent for all fixed $r \geq m$ and $s \geq m$. From this we have

$$P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) = E\left\{F\left\{\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} d^2\right\}\right\}$$

Now, we have the basic inequality

$$\frac{a}{d^2} N \ U_R^2 \leq R^2 \leq \frac{a}{d^2} N \ U_{R-1}^2 + 2mR,$$

which implies

$$\frac{a}{d^2}(\frac{S}{R} + 1)U_R^2 \leq R \leq \frac{a}{d^2} (\frac{S}{R} + 1) \ U_{R-1}^2 + 2m.$$

Dividing this inequality by r^*, and then taking the limit throughout as $d \to 0$, and also using (2.4.19) we obtain

$$\frac{R}{r^*} \to 1 \quad \text{w.p.1} \quad \text{as } d \to 0 .$$

Now,

$$\frac{d^2}{a}\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} = \frac{d^2}{a}\left(\frac{R\,S}{S\sigma_1^2 + R\sigma_2^2}\right) = \frac{R}{r^*}\left\{\frac{\sigma_1(\sigma_1 + \sigma_2)}{\sigma_1^2 + \frac{R}{S}\sigma_2^2}\right\} .$$

Therefore, we conclude

$$\frac{d^2}{a}\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} \to 1 \quad \text{w.p.1} \quad \text{as } d \to 0 ,$$

which implies

$$d^2\left(\frac{\sigma_1^2}{R} + \frac{\sigma_2^2}{S}\right)^{-1} \to a \quad \text{w.p.1} \quad \text{as } d \to 0 .$$

Hence, using the dominated convergence theorem, we obtain

$$\lim_{d \to 0} P(\underset{\sim}{\mu} \in \mathcal{R}_{R,S}) = F(a) = 1 - \alpha ,$$

by the choice of "a".

This completes the proof of Theorem 2.4.3.

2.5 Moderate Sample Size Performance

In this section, we present numerical results obtained through simulations using PROC MATRIX from the SAS package. The subsections 2.5.1, 2.5.2 and 2.5.3 respectively present results of our simulation studies for the case of equal covariance matrices and equal sample size as discussed in Section 2.2, the case of unequal covariance matrices and equal sample sizes as discussed in Section 2.3, and the case of unequal covariance matrices and unequal sample sizes as discussed in Section 2.4.

In Equation (2.4.8), we introduced a constant "h", which depends on the starting sample size m, the dimension parameter p, and the confidence coefficient (1-α). From (2.4.8) we may recall that

$$E\left\{F\left\{h\left(\frac{\sigma_1^2}{U_m^2}+\frac{\sigma_2^2}{V_m^2}\right)^{-1}\right\}\right\} = 1-\alpha .$$

Let us define

$$q(x) = x^{\gamma} (1-x)^{-\gamma+1} e^{-x/(1-x)}$$

for $0<x<1$ where $\gamma = p(m-1)/2$.

We know that $p(m-1)U_m^2/\sigma_1^2 \sim \chi^2_{(p(m-1))}$, $p(m-1)V_m^2/\sigma_2^2 \sim \chi^2_{(p(m-1))}$, and then making some simple transformations we can easily show that

$$E\left\{F\left\{h\left(\frac{\sigma_1^2}{U_m^2}+\frac{\sigma_2^2}{V_m^2}\right)^{-1}\right\}\right\} = \{\Gamma(\gamma)\}^{-2} \int_0^1 \int_0^1 F\{\frac{h}{\gamma} (\frac{1-x}{x}+\frac{1-y}{y})^{-1}\} q(x)q(y)dx\, dy. \tag{2.5.1}$$

We use FORTRAN Language on an IBM 3081D computer system with WATFIV compiler, and utilize the subroutine called DMLIN from IMSL (1982) in order to numerically evaluate the integral in (2.5.1). Using this subroutine we calculate the values of h for p = 2,3,4,5, m = 2(1)10, and α = 0.10, 0.05 and 0.01. But, due to some peculiarities of the integrand, this subroutine fails to evaluate the integral (2.5.1) for p = 5, and m $\geq$ 15. So, we tried to find a simpler integral which is equivalent to (2.4.8).

We can show that

$$E\left\{F\left\{h\left(\frac{\sigma_2^2}{U_m^2}+\frac{\sigma_2^2}{V_m^2}\right)^{-1}\right\}\right\} = P\left(\frac{1}{F_1}+\frac{1}{F_2}\le\frac{h}{p}\right),$$

where (F_1, F_2) has a bivariate F-distribution of Kimball (1951) with degrees of freedom $(p,p(m-1))$ and $(p,p(m-1))$. From the joint density of (F_1, F_2) and a simple transformation thereafter we can write

$$P\left(\frac{1}{F_1}+\frac{1}{F_2}\le\frac{h}{p}\right) = c(p,m,h)\int_0^1\int_0^1 \frac{(y-y^2)^{\frac{pm}{2}-1}\, x^{\frac{p}{2}-1}}{\left\{1+\frac{hx}{p(m-1)}(y-y^2)\right\}^{p(m-\frac{1}{2})}}\,dx\,dy, \quad (2.5.2)$$

where

$$c(p,m,h) = \frac{h^{p/2}\,\Gamma\{p(m-\frac{1}{2})\}\{p(m-1)\}^{-p/2}}{\Gamma(p/2)\,\{\Gamma\{\frac{p(m-1)}{2}\}\}^2}.$$

We used the old subroutine DMLIN and the integral in (2.5.2) to calculate the values of h for p=2,3,4,5, m = 15(5)40(10)80, 100, and α = 0.10, 0.05, 0.01.

The values of h for p = 2,3,4,5, α = 0.10, 0.05, 0.01, and m = 2(1) 10(5)40(10)80,100 and when $m \to \infty$ are reported in Table V. The values of h when p=1 can be obtained from Ghosh's (1975b) table.

Let us now explain the way we carry out the simulations. In any particular table we used a particular "rule" to determine the sample sizes N or (R,S) depending on what case we are considering. In all the cases we take H=I, the identity matrix. If p=2, we take $\underset{\sim}{\mu}_1' = (1\ 2)$ and $\underset{\sim}{\mu}_2' = (0\ 0)$, and if p=3 we take $\underset{\sim}{\mu}_1' = (1\ 2\ 3)$ and $\underset{\sim}{\mu}_2' = (0\ 0\ 0)$. Then, if p=2 we have

$\underset{\sim}{\mu}' = (\mu_1 \mu_2) = (1\ 2)$, and if p=3 we have $\underset{\sim}{\mu}' = (1\ 2\ 3)$. In the case of $\sigma_1=\sigma_2$ we take $\sigma_1=\sigma_2=1$, and in the case of $\sigma_1 \neq \sigma_2$, we take $\sigma_1=1$ and $\sigma_2=2$. A particular "rule" is replicated k times, the ith replicate giving rise to observed values of N and $\bar{\underset{\sim}{T}}_{N,N}$ (or (R,S) and $\bar{\underset{\sim}{T}}_{R,S}$) as, say, n(j) and $\bar{\underset{\sim}{T}}_{n(j),n(j)}$ (or (r(j), s(j)) and $\bar{\underset{\sim}{T}}_{r(j),s(j)}$) respectively depending on what case we are considering. Then, we estimate E(N) and μ_i (or (E(R), E(S)) and μ_i) by $\bar{n} = k^{-1} \sum_{j=1}^{k} n(j)$ and $\hat{\mu}_i = k^{-1} \sum_{j=1}^{k} \bar{T}_{in(j),n(j)}$ (or $\bar{r} = k^{-1} \sum_{j=1}^{k} r(j)$, $\bar{s} = k^{-1} \sum_{j=1}^{k} s(j)$ and $\hat{\mu}_i = k^{-1} \sum_{j=1}^{k} \bar{T}_{ir(j),s(j)}$) respectively depending on what case we are considering. We also compute the corresponding standard errors

$$SD(\bar{n}) = \{(k^2-k)^{-1} \sum_{j=1}^{k} (n(j)-\bar{n})^2\}^{\frac{1}{2}} \quad \underline{\text{and}}$$

$$SD(\hat{\mu}_i) = \{(k^2-k)^{-1} \sum_{j=1}^{k} (\bar{T}_{in(j),n(j)} - \hat{\mu}_i)^2\}^{\frac{1}{2}}, \quad i = 1,2,\ldots,p,$$

or

$$SD(\bar{r}+\bar{s}) = \{(k^2-k)^{-1} \sum_{j=1}^{k} (r(j) + s(j) - \bar{r}-\bar{s})^2\}^{\frac{1}{2}} \quad \underline{\text{and}}$$

$$SD(\hat{\mu}_i) = \{(k^2-k)^{-1} \sum_{j=1}^{k} (\bar{T}_{ir(j),s(j)} - \hat{\mu}_i)^2\}^{\frac{1}{2}}, \quad i = 1,2,\ldots,p$$

depending on what case we are in. We consider j=1,...,k. While using a particular rule, we also estimate the coverage probability of the region $\mathcal{R}_{N,N}$ (or $\mathcal{R}_{R,S}$), say, by c.p. where

$$\text{c.p.} = \text{relative frequency of} \sum_{i=1}^{p} (\bar{T}_{in(j),n(j)} - i)^2 \leq d^2,$$

or

$$\text{c.p.} = \text{relative frequency of } \sum_{i=1}^{p} (\bar{T}_{ir(j),s(j)} - i)^2 \le d^2$$

from all the replicates for j = 1,2,...,k depending on the case. Here we are considering 95% confidence regions only, that is, we keep $\alpha = 0.05$ fixed and d is computed for given C or n^* which depends on what situation we are considering. For all the cases and rules, we took k = 500.

2.5.1 Moderate Sample Size Performances for the Problem of Section 2.2

For the two-stage and the sequential procedures defined by (2.2.4) and (2.2.9), we present results when p=2 and 3, m=5 and 10, and C = 10, 40, 70, 100. The Table I summarizes our findings for the two-stage procedure. The Table II summarizes our findings for the sequential procedure.

Remark 2.3: From Table I, we notice that $\bar{n}$ is always larger than C, however, almost always the estimated coverage (c.p.) exceeds the target which is 0.95. The result gets better in the sense of less oversampling as m increases and this is generally expected. From Table II we notice that both $\bar{n}$ and c.p. are close to C and 0.95 respectively for the sequential procedure.

2.5.2 Moderate Sample Size Performances for the Problem of Section 2.3

Here we use the "rule" as being the two-stage procedure and the sequential procedure defined by (2.3.3) and (2.3.4) respectively. For both these procedures, we give results for p=2 and 3, m=5 and 10, and C = 10, 40, 70, 100. The Table III summarizes our findings for the two-stage

TABLE I

EQUAL COVARIANCE MATRICES AND EQUAL SAMPLE SIZES:
TWO-STAGE PROCEDURE (2.2.4)

p	m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	10	1.095	13.02	0.20	1.013	2.020	.	0.954
		40	0.547	48.74	0.73	1.014	2.009	.	0.948
		70	0.414	86.27	1.33	0.996	1.996	.	0.958
		100	0.346	125.05	2.01	1.004	1.995	.	0.932
2	10	10	1.095	12.99	0.09	0.931	1.832	.	0.956
		40	0.547	44.46	0.44	0.998	2.008	.	0.950
		70	0.414	76.20	0.79	0.998	1.995	.	0.944
		100	0.346	109.48	1.10	1.003	2.009	.	0.946
3	5	10	1.250	12.56	0.16	1.033	1.993	2.992	0.954
		40	0.625	47.47	0.61	0.985	2.010	2.983	0.952
		70	0.472	79.82	1.02	1.008	2.002	3.005	0.954
		100	0.395	117.16	1.45	0.996	2.006	3.010	0.956
3	10	10	1.250	11.51	0.07	1.005	2.038	2.997	0.970
		40	0.625	43.73	0.37	0.991	2.016	3.002	0.938
		70	0.472	74.66	0.62	1.006	2.002	2.996	0.944
		100	0.395	105.77	0.89	0.998	2.000	2.995	0.932

TABLE II

EQUAL COVARIANCE MATRICES AND EQUAL SAMPLE SIZES: SEQUENTIAL PROCEDURE (2.2.9)

p	m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	10	1.095	10.22	0.11	1.008	2.000	.	0.930
		40	0.547	40.20	0.20	1.000	2.012	.	0.956
		70	0.414	69.82	0.26	1.000	2.007	.	0.932
		100	0.346	99.63	0.31	1.000	2.003	.	0.952
2	10	10	1.095	10.99	0.06	0.997	1.989	.	0.950
		40	0.547	39.93	0.20	1.002	1.994	.	0.940
		70	0.414	70.29	0.26	1.000	2.008	.	0.944
		100	0.346	100.09	0.32	0.999	1.996	.	0.946
3	5	10	1.250	10.17	0.09	1.012	1.985	2.998	0.930
		40	0.625	40.56	0.16	1.017	1.987	3.008	0.932
		70	0.473	70.59	0.22	0.999	1.997	2.992	0.954
		100	0.395	100.42	0.24	1.006	1.994	3.004	0.962
3	10	10	1.250	10.90	0.05	0.968	1.965	3.026	0.958
		40	0.625	40.07	0.17	1.006	2.009	3.002	0.960
		70	0.473	70.27	0.21	0.996	1.991	2.994	0.948
		100	0.395	100.57	0.25	1.003	2.001	2.993	0.962

TABLE III

UNEQUAL COVARIANCE MATRICES AND EQUAL SAMPLE SIZES:
TWO-STAGE PROCEDURE (2.3.3)

p	m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	10	1.731	14.69	0.29	0.999	2.026	.	0.964
		40	0.865	58.61	1.23	0.974	1.992	.	0.950
		70	0.654	102.94	2.22	0.995	1.996	.	0.936
		100	0.547	146.53	3.09	1.003	1.988	.	0.954
2	10	10	1.731	12.82	0.15	0.971	2.042	.	0.958
		40	0.865	47.50	0.68	0.979	2.005	.	0.942
		70	0.654	82.00	1.26	0.995	2.001	.	0.954
		100	0.547	120.10	1.64	1.003	1.992	.	0.958
3	5	10	1.977	13.83	0.24	0.982	1.971	2.981	0.954
		40	0.988	53.95	0.92	1.004	1.974	2.982	0.946
		70	0.747	94.97	1.72	0.984	1.992	2.992	0.962
		100	0.625	134.47	2.28	1.007	2.006	2.996	0.954
3	10	10	1.977	12.23	0.12	0.969	2.037	3.019	0.966
		40	0.988	46.31	0.52	1.030	1.982	2.993	0.946
		70	0.747	79.95	0.93	0.999	2.005	3.003	0.956
		100	0.625	111.91	1.35	0.986	1.993	3.006	0.942

TABLE IV

UNEQUAL COVARIANCE MATRICES AND EQUAL SAMPLE SIZES:
SEQUENTIAL PROCEDURE (2.3.4)

p	m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	10	1.731	9.52	0.14	0.998	1.983	.	0.922
		40	0.865	40.24	0.30	0.992	1.990	.	0.936
		70	0.654	69.35	0.39	0.999	1.983	.	0.956
		100	0.547	99.20	0.45	1.003	2.009	.	0.960
2	10	10	1.731	11.15	0.07	0.984	1.984	.	0.960
		40	0.865	38.97	0.29	1.011	2.002	.	0.938
		70	0.654	69.35	0.38	0.994	1.991	.	0.956
		100	0.547	99.61	0.44	1.018	2.011	.	0.950
3	5	10	1.977	9.83	0.12	1.021	1.976	2.957	0.932
		40	0.988	40.08	0.24	1.008	2.005	3.002	0.934
		70	0.747	69.70	0.30	1.011	1.989	3.006	0.956
		100	0.625	100.64	0.35	1.003	2.004	2.993	0.944
3	10	10	1.977	11.02	0.06	0.955	2.023	3.044	0.956
		40	0.988	40.07	0.25	1.006	2.002	3.005	0.944
		70	0.747	69.75	0.29	0.993	1.992	2.981	0.964
		100	0.625	99.82	0.36	0.979	2.006	2.992	0.936

procedure. The Table IV summarizes our finding for the sequential procedure.

Remark 2.4: From Table III, we notice, as we expect that $\bar{n}$ is always larger than C. The amount of oversampling reduces when we go from p=2 to p=3. The results also get better in this sense as m increases. The estimated coverage probability (c.p.) almost always exceeds the target which is 0.95 for the two-stage procedure. From Table IV, we notice that both $\bar{n}$ and c.p. are very close to C and 0.95 respectively for the sequential procedure, and naturally, the sequential procedure also performs better when m increases.

2.5.3 Moderate Sample Size Performances for the Problem of Section 2.4

We use the "rule" as being the two-stage procedure defined by (2.4.6) and (2.4.7). We give results for p=2 and 3, m=5 and 10, and $n^* = 20, 80, 140, 200$ where $n^* = r^* + s^*$. For each R_i^* (considered in subsection 2.4.2) defining the sequential procedure, we give moderate-sample results for p=2 and 3, m=5 and 10, and $n^* = r^* + s^* = 20, 80, 140, 200$. The Table VI reports the results for the two-stage procedure, and Tables VII, VIII, IX and X report the results for the sequential procedures defined by R_1^*, R_2^*, R_3^* and R_4^*, respectively.

Remark 2.5: From Table VI, we notice that $\bar{s}$ always overestimates s^* by a large margin but $\bar{r}$ is a fairly good estimator for r^*. This is due to the fact that σ_2^2 is four times larger than σ_1^2. The estimated coverage probability (c.p.) always seems to exceed the target which is 0.95. We notice also that the amount of oversampling is reduced when we go from p=2 to p=3. As m increases the two-stage procedure (2.4.6) - (2.4.7) performs better.

TABLE V

THE h-VALUES NEEDED FOR THE TWO-STAGE PROCEDURE DEFINED BY (2.4.6) AND (2.4.7)

m	p					
	2			3		
	α			α		
	0.10	0.05	0.01	0.10	0.05	0.01
2	42.8106	85.5927	412.5142	33.9436	54.8146	155.5587
3	17.8003	26.9637	61.8595	19.5283	26.8801	49.5526
4	13.9247	19.7438	37.7372	16.6165	21.9875	36.5761
5	12.4330	17.1556	30.4958	15.3990	20.0346	31.9482
6	11.6520	15.8475	27.1373	14.7345	18.9935	29.6134
7	11.1736	15.0629	25.2226	14.3170	18.3484	28.2135
8	10.8511	14.5413	23.9911	14.0307	17.9102	27.2824
9	10.6192	14.1700	23.1347	13.8222	17.5933	26.6192
10	10.4445	13.8924	22.5053	13.6637	17.3536	26.1230
15	9.9716	13.1500	20.8670	13.2276	16.6992	24.7896
20	9.7605	12.8229	20.1659	13.0287	16 4047	24.2024
25	9.6410	12.6389	19.7770	12.9165	16.2374	23.8716
30	9.5641	12.5211	19.5298	12.8435	16.1297	23.6618
35	9.5105	12.4391	19.3588	12.7923	16.0542	23.5140
40	9.4711	12.3788	19.2335	12.7545	15.9986	23.4052
50	9.4167	12.2960	19.0621	12.7024	15.9219	23.2557
60	9.3812	12.2418	18.9504	12.6680	15.8716	23.1579
70	9.3560	12.2036	18.8718	12.6439	15.8361	23.0889
80	9.3374	12.1753	18.8136	12.6259	15.8096	23.0376
100	9.3114	12.1359	18.7329	12.6008	15.7729	22.9665
∞	9.2103	11.9829	18.4207	12.5028	15.6295	22.6898

TABLE V (Continued)

m	p = 4, α = 0.10	p = 4, α = 0.05	p = 4, α = 0.01	p = 5, α = 0.10	p = 5, α = 0.05	p = 5, α = 0.01
2	32.8601	48.0374	106.5254	33.7425	46.5578	89.8196
3	21.9035	28.6976	47.6389	24.4461	31.0450	48.3801
4	19.3934	24.7027	38.2284	22.1523	27.5172	40.6595
5	18.2991	23.0206	34.5786	21.1265	25.9827	37.5113
6	17.6889	22.0990	32.6582	20.5467	25.1276	35.8121
7	17.3004	21.5183	31.4770	20.1744	24.5832	34.7505
8	17.0314	21.1193	30.6779	19.9153	24.2064	34.0249
9	16.8343	20.8283	30.1017	19.7245	23.9303	33.4976
10	16.6837	20.6068	29.6784	19.5783	23.7192	33.0971
15	16.2659	19.9962	28.4825	19.1705	23.1334	31.9975
20	16.0744	19.7181	27.9509	18.9825	22.8647	31.4989
25	15.9645	19.5591	27.6490	18.8742	22.7105	31.2142
30	15.8932	19.4561	27.4543	18.8039	22.6105	31.0302
35	15.8432	19.3841	27.3184	18.7546	22.5403	30.9014
40	15.8062	19.3308	27.2181	18.7180	22.4884	30.8062
50	15.7552	19.2573	27.0801	18.6675	22.4167	30.6750
60	15.7216	19.2090	26.9895	18.6342	22.3696	30.5888
70	15.6978	19.1748	26.9255	18.6106	22.3362	30.5279
80	15.6801	19.1494	26.8779	18.5931	22.3113	30.4825
100	15.6555	19.1140	26.8117	18.5687	22.2767	30.4194
∞	15.5589	18.9755	26.5534	18.4727	22.1410	30.1726

TABLE VI

UNEQUAL COVARIANCE MATRICES AND UNEQUAL SAMPLE SIZES: TWO-STAGE PROCEDURE (2.4.6) AND (2.4.7)

p	m	r^*	s^*	d	$\bar{r}$	$\bar{s}$	$SD(\bar{r} + \bar{s})$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	6.7	13.3	1.642	7.33	25.61	0.56	0.950	1.989	.	0.980
		26.7	53.3	0.821	26.55	102.99	2.29	1.006	2.021	.	0.956
		46.7	93.3	0.621	44.17	183.02	4.07	0.993	2.009	.	0.956
		66.7	133.3	0.519	63.35	255.58	5.70	1.007	1.997	.	0.964
2	10	6.7	13.3	1.642	10.01	21.10	0.28	0.974	2.016	.	0.986
		26.7	53.3	0.821	21.21	83.28	1.25	0.988	2.000	.	0.954
		46.7	93.3	0.621	36.67	147.82	2.35	1.005	2.017	.	0.962
		66.7	133.3	0.519	52.43	199.53	3.07	0.992	2.011	.	0.942
3	5	6.7	13.3	1.875	6.47	23.15	0.41	1.016	1.980	3.032	0.972
		26.7	53.3	0.938	23.43	93.74	1.74	0.992	1.988	3.023	0.940
		46.7	93.3	0.709	39.66	159.30	2.98	1.002	2.018	2.978	0.960
		66.7	133.3	0.593	57.42	227.81	4.40	0.992	2.012	3.001	0.930
3	10	6.7	13.3	1.875	10.00	20.39	0.23	1.012	1.992	3.013	0.992
		26.7	53.3	0.938	20.28	79.11	0.91	0.981	1.999	2.996	0.962
		46.7	93.3	0.709	34.49	139.73	1.66	0.995	2.010	2.999	0.964
		66.7	133.3	0.593	49.79	201.16	2.31	1.007	1.993	3.008	0.952

TABLE VII

UNEQUAL COVARIANCE MATRICES AND UNEQUAL SAMPLE SIZES: SEQUENTIAL PROCEDURE R_1^* (2.4.9)

p	m	r^*	s^*	d	$\bar{r}$	$\bar{s}$	SD($\bar{r}$ + $\bar{s}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	6.7	13.3	1.642	6.52	12.00	0.22	1.031	1.942	.	0.930
		26.7	53.3	0.821	26.33	52.47	0.41	1.008	1.994	.	0.922
		46.7	93.3	0.621	46.28	93.01	0.53	0.998	2.019	.	0.942
		66.7	133.3	0.519	66.49	132.95	0.63	1.016	1.993	.	0.952
2	10	6.7	13.3	1.642	10.02	11.68	0.12	0.934	1.945	.	0.948
		26.7	53.3	0.821	26.19	52.23	0.42	0.996	2.004	.	0.938
		46.7	93.3	0.621	49.19	91.96	0.52	1.011	1.997	.	0.942
		66.7	133.3	0.519	66.54	132.61	0.64	1.017	2.018	.	0.950
3	5	6.7	13.3	1.875	6.70	12.54	0.18	1.013	1.977	2.971	0.908
		26.7	53.3	0.938	26.75	52.91	0.32	0.979	1.984	2.989	0.962
		46.7	93.3	0.709	46.71	93.26	0.42	0.991	1.995	2.990	0.952
		66.7	133.3	0.593	66.57	132.47	0.50	0.996	2.008	3.009	0.944
3	10	6.7	13.3	1.875	10.01	11.43	0.10	1.019	1.975	3.007	0.922
		26.7	53.3	0.938	26.78	53.18	0.30	0.988	2.010	3.016	0.958
		46.7	93.3	0.709	46.40	92.59	0.44	1.017	2.021	2.988	0.960
		66.7	133.3	0.593	66.55	132.92	0.50	0.987	2.009	3.011	0.918

TABLE VIII

UNEQUAL COVARIANCE MATRICES AND UNEQUAL SAMPLE SIZES: SEQUENTIAL PROCEDURE R_2^* (2.4.10)

p	m	r^*	s^*	d	$\bar{r}$	$\bar{s}$	$SD(\bar{r} + \bar{s})$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	6.7	13.3	1.642	6.52	12.13	0.21	1.031	1.942	.	0.932
		26.7	53.3	0.821	26.33	52.48	0.41	1.008	1.994	.	0.922
		46.7	93.3	0.621	46.28	93.01	0.53	0.998	2.019	.	0.942
		66.7	133.3	0.519	66.49	132.96	0.63	1.016	1.993	.	0.952
2	10	6.7	13.3	1.642	10.02	12.33	0.12	0.945	1.935	.	0.956
		26.7	53.3	0.821	26.19	52.25	0.42	0.996	2.003	.	0.938
		46.7	93.3	0.621	46.19	91.97	0.52	1.011	1.997	.	0.942
		66.7	133.3	0.519	66.54	132.61	0.64	1.017	2.018	.	0.950
3	5	6.7	13.3	1.875	6.70	12.63	0.18	1.009	1.969	2.973	0.914
		26.7	53.3	0.938	26.75	52.92	0.32	0.979	1.984	2.989	0.962
		46.7	93.3	0.709	46.71	93.26	0.42	0.991	1.995	2.990	0.952
		66.7	133.3	0.593	66.57	132.47	0.50	0.996	2.008	3.009	0.944
3	10	6.7	13.3	1.875	10.01	12.15	0.10	1.013	1.974	3.003	0.924
		26.7	53.3	0.938	26.78	53.20	0.30	0.987	2.011	3.015	0.958
		46.7	93.3	0.709	46.40	92.59	0.44	1.017	2.021	2.988	0.960
		66.7	133.3	0.593	66.55	132.92	0.50	0.987	2.009	3.011	0.918

TABLE IX

UNEQUAL COVARIANCE MATRICES AND UNEQUAL SAMPLE SIZES: SEQUENTIAL PROCEDURE R_3^* (2.4.11)

p	m	r^*	s^*	d	$\bar{r}$	$\bar{s}$	$SD(\bar{r} + \bar{s})$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	6.7	13.3	1.642	6.71	12.66	0.21	1.032	1.955	.	0.936
		26.7	53.3	0.821	26.55	52.89	0.41	1.008	1.991	.	0.924
		46.7	93.3	0.621	46.50	93.47	0.53	0.997	2.018	.	0.946
		66.7	133.3	0.519	66.72	133.32	0.63	1.015	1.993	.	0.958
2	10	6.7	13.3	1.642	10.03	13.59	0.13	0.954	1.944	.	0.966
		26.7	53.3	0.821	26.38	52.68	0.41	0.995	2.004	.	0.938
		46.7	93.3	0.621	46.37	92.42	0.52	1.011	1.997	.	0.940
		66.7	133.3	0.519	66.72	133.04	0.64	1.017	2.018	.	0.950
3	5	6.7	13.3	1.875	6.86	13.15	0.17	1.021	1.986	2.969	0.926
		26.7	53.3	0.938	26.93	53.29	0.32	0.981	1.986	2.987	0.962
		46.7	93.3	2.709	46.87	93.62	0.41	0.990	1.994	2.990	0.948
		66.7	133.3	0.593	66.77	132.83	0.50	0.997	2.008	3.009	0.950
3	10	6.7	13.3	1.875	10.01	13.57	0.11	1.018	1.972	2.998	0.942
		26.7	53.3	0.938	26.96	53.57	0.30	0.985	2.014	3.016	0.954
		46.7	93.3	0.709	46.59	92.99	0.44	1.016	2.022	2.989	0.964
		66.7	133.3	0.593	66.72	133.31	0.50	0.988	2.009	3.010	0.914

TABLE X

UNEQUAL COVARIANCE MATRICES AND UNEQUAL SAMPLE SIZES: SEQUENTIAL PROCEDURE R_4^* (2.4.12)

p	m	r^*	s^*	d	$\bar{r}$	$\bar{s}$	SD($\bar{r}$ + $\bar{s}$)	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	c.p.
2	5	6.7	13.3	1.642	6.83	13.11	0.21	1.036	1.959	.	0.938
		26.7	53.3	0.821	26.87	53.23	0.41	1.006	1.990	.	0.922
		46.7	93.3	0.621	46.67	93.82	0.53	0.997	2.018	.	0.946
		66.7	133.3	0.519	66.88	133.64	0.63	1.015	1.992	.	0.956
2	10	6.7	13.3	1.642	10.03	15.15	0.13	0.954	1.933	.	0.974
		26.7	53.3	0.821	26.51	53.01	0.41	0.996	2.004	.	0.938
		46.7	93.3	0.621	46.54	92.75	0.52	1.011	1.998	.	0.940
		66.7	133.3	0.519	66.84	133.37	0.64	1.017	2.018	.	0.948
3	5	6.7	13.3	1.875	6.98	13.57	0.17	1.011	1.984	2.968	0.938
		26.7	53.3	0.938	27.07	53.63	0.32	0.980	1.985	2.988	0.956
		46.7	93.3	0.709	46.99	93.96	0.43	0.990	1.994	2.990	0.948
		66.7	133.3	0.593	66.89	133.12	0.50	0.997	2.008	3.009	0.958
3	10	6.7	13.3	1.875	10.01	15.10	0.11	1.035	1.973	2.994	0.964
		26.7	53.3	0.938	27.09	53.92	0.30	0.985	2.011	3.018	0.954
		46.7	93.3	0.709	46.74	93.33	0.44	1.015	2.020	2.988	0.962
		66.7	133.3	0.593	66.90	133.64	0.50	0.989	2.009	3.011	0.910

Remark 2.6: If we let $\bar{n}_i$ to be the total sample size estimated from using the rule R_i^*, we can immediately readout from Tables VII, VIII, IX and X that $\bar{n}_1 < \bar{n}_2 < \bar{n}_3 < \bar{n}_4$, and this is quite expected. The estimated coverage probability (c.p.) is not so close to the target 0.95 for these sequential procedures for some of the entries.

CHAPTER III

FIXED-SIZE CONFIDENCE REGIONS FOR THE REGRESSION PARAMETERS IN THE GENERAL LINEAR MODEL WITH NORMALITY

3.1 Introduction and Review

We start by formulating the problem. Suppose we have the general linear model given by

$$\underset{\sim}{Y}_n = X_n \underset{\sim}{\beta} + \underset{\sim}{E}_n \tag{3.1.1}$$

where $\underset{\sim}{Y}_n$ is an observed nx1 random vector, X_n is a known nxp matrix of rank p, $\underset{\sim}{\beta}$ is a px1 vector of unknown regression parameters, and $\underset{\sim}{E}_n$ is nx1 random vector of errors distributed as $N_n(\underset{\sim}{0}, \sigma^2 I_n)$, with $\sigma \in (0,\infty)$ being unknown. We assume that $p \geq 2$.

Given two numbers $d \in (0,\infty)$ and $\alpha \in (0,1)$, we propose to consider the following ellipsoidal confidence region for $\underset{\sim}{\beta}$. We define

$$\mathcal{R}_n = \{\underset{\sim}{W} \in \mathbb{R}^p : n^{-1}(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{W})'(X_n' X_n)(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{W}) \leq d^2\} , \tag{3.1.2}$$

where $\hat{\underset{\sim}{\beta}}_n = (X_n' X_n)^{-1} X_n' \underset{\sim}{Y}_n$, with $p > n$. Now, the confidence coefficient associated with this region $\mathcal{R}_n$ is given by

$$P(\underset{\sim}{\beta} \in \mathcal{R}_n) = P\{(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta})' (X_n' X_n)(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}) \leq nd^2\}$$

$$= P\{\frac{1}{\sigma^2} (\hat{\beta}_n - \underset{\sim}{\beta})' (X_n' X_n)(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}) \leq \frac{nd^2}{\sigma^2}\}$$

$$= P\{\chi^2_{(p)} \le \frac{nd^2}{\sigma^2}\}$$

$$= F(\frac{nd^2}{\sigma^2}) \ , \qquad (3.1.3)$$

where $F(\cdot)$ is defined as in Chapter II, that is

$$F(t) = P(\chi^2_{(p)} \le t) \quad \text{for } t > 0 \ .$$

<u>Remark 3.1</u>: In (3.1.2), we take the weight matrix as $\frac{1}{n}(X_n' X_n)$. Since $\frac{1}{n}(X_n' X_n)$ is generally assumed to converge to a positive definite matrix A, say, as $n \to \infty$ in order to study large sample properties in the theory of linear regression analysis. However, we do not make this assumption.

Now, we require the confidence coefficient to be at least $(1-\alpha)$, so we need the sample size n to be at least $\frac{a\sigma^2}{d^2} = C$, say, where $F(a) = 1-\alpha$. This "C" is referred to as the "optimal fixed sample size" required to solve the problem if σ^2 is known. However, C is unknown since σ^2 is unknown and thus we must estimate C by using a suitable positive integer valued random variable N, say. Once we determine N, we propose the corresponding confidence region $\mathcal{R}_N$ for $\underset{\sim}{\beta}$. Naturally, the characteristics for "goodness" of having the region $\mathcal{R}_N$ for $\underset{\sim}{\beta}$ will depend on the "closeness" between N and C.

In Section 3.2, we propose a two-stage procedure along the lines of Stein (1945, 1949), Chatterjee (1959, 1960), and Mukhopadhyay (1982), and study various properties.

Our Section 3.3 deals with a modified two-stage procedure to obtain "asymptotic efficiency" in the Chow and Robbins (1965) sense. This procedure is motivated by the results of Mukhopadhyay (1980, 1982).

In Section 3.4 we present a sequential procedure where we take one sample at a time after we start to get to the stopping stage. Here, we derive second order expansions for E(N) and $P(\underset{\sim}{\beta} \varepsilon R_N)$ using the nonlinear renewal theory of Woodroofe (1977, 1982) as it was carried out in Al-Mousawi (1984).

Our Section 3.5 deals with a three-stage procedure proposed along the lines of Hall (1981) and Al-Mousawi (1984). The motivation behind this procedure can be summarized as follows. After starting the experiment with $m(\geq p+1)$ samples, we estimate a fraction rC of the optimal fixed sample size C by, say, M. Then, depending on the size of M, we decide whether to obtain all the remaining samples of size N-M where N is the estimate of C found in the third stage.

Section 3.6 is devoted to numerical studies found by simulated experiments for all the procedures proposed in previous sections of this chapter. These results help us in exemplifying the moderate sample size behavior of all our proposed sampling techniques.

Gleser (1965, 1966) proposed a sequential procedure to construct a spherical confidence region for the regression parameters without the normality assumption. Albert (1966) and Srivastava (1967, 1971) proposed sequential procedures to construct both spherical and ellipsoidal confidence regions for the regression parameters without the normality assumption. Mukhopadhyay (1974) proposed a sequential point estimation procedure for the regression parameters assuming the loss function to be squared error plus linear cost. Recently, Finster (1983) studied sequential point estimation problems for the regression parameters in a multivariate linear model, and thus these are natural extensions of the work of Mukhopadhyay (1974). In this chapter again we write [x] for the largest integer smaller than x. Let S_n^2 be the usual estimate of σ^2, that is the mean squared error, namely

$$S_n^2 = (n-p)^{-1} (\underset{\sim}{Y}_n - X_n\hat{\underset{\sim}{\beta}}_n)'(\underset{\sim}{Y}_n - X_n\hat{\underset{\sim}{\beta}}_n), \qquad (3.1.4)$$

for $n \geq p+1$ and $p \geq 2$.

3.2 A Two-Stage Procedure

We start with $m(\geq p+1)$ samples, and define

$$N = \max\{m, [bS_m^2/d^2] + 1\}, \qquad (3.2.1)$$

where $b = pb'$, b' being the upper $100\alpha\%$ point of the F-distribution with degrees of freedom p, $m-p$.

If N=m, we stop sampling at the starting stage. Otherwise, we sample the difference N-m at the second stage. Thus, when we stop we have $\underset{\sim}{Y}_N$ as our random vector for the response variable. Then, we compute $\hat{\underset{\sim}{\beta}}_N$ and propose the corresponding region R_N as in (3.1.2). Some of the properties of this procedure are stated in Theorem 3.2.1.

Theorem 3.2.1: For the procedure in (3.2.1), for all $\underset{\sim}{\beta} \in \mathbb{R}^p$, and $\sigma \in (0,\infty)$ we have:

(a) $P(\underset{\sim}{\beta} \in R_N) \geq 1-\alpha$ for all $d > 0$,

(b) $E(N/C) \to \frac{b}{a}$ as $d \to 0$,

(c) $\text{Var}(N) \{\frac{(p-m)}{2} (\frac{d^2}{b\sigma^2})^2\} \to 1$ as $d \to 0$, and

(d) $P(\underset{\sim}{\beta} \in R_N) \to 1-\alpha$ as $d \to 0$.

We omit its proof for brevity, as it can be given along the lines of the proof of Theorem 2.2.1.

Remark 3.2: Part (a) tells us that the procedure (3.2.1) is "exactly consistent" in the Mukhopadhyay (1982) sense, while Part (d) shows that this procedure is also "asymptotically consistent" in the Chow and Robbins (1965) sense. In Part (b), we have the limiting ratio $\frac{b}{a}$ which is always larger than one which means that N overestimate C, even asymptotically.

3.3 A Modified Two-Stage Procedure

Motivated by the results of Mukhopadhyay (1980, 1982), we first choose and fix a real number $\rho\varepsilon(0,\infty)$ and let the starting sample size be

$$m = \max\{p+1, [(a/d^2)^{\frac{1}{1+\rho}}] + 1\} .$$

Then, we define the stopping rule as:

$$N = \max\{m, [bS_m^2/d^2] + 1\} . \tag{3.3.1}$$

The number b remains the same as in (3.2.1). Again if N=m, we stop sampling at the starting stage itself. Otherwise, we sample the difference N-m. We compute $\hat{\underset{\sim}{\beta}}_N$ and propose the corresponding confidence region R_N for $\underset{\sim}{\beta}$ as in (3.1.2). The main point to observe here is that $m \to \infty$ as $d \to 0$, however, $m/C \to 0$ as $d \to 0$. Thus, $b/a \to 1$ as $d \to 0$. Some of the properties of this procedure are listed in Theorem 3.3.1.

Theorem 3.3.1: For the procedure (3.3.1), for all $\underset{\sim}{\beta} \in \mathbb{R}^p$, and $\sigma \in (0,\infty)$ we have:

(a) $P(\underset{\sim}{\beta} \in R_N) \geq 1-\alpha$ for all $d > 0$,

(b) $E(N/C) \to 1$ as $d \to 0$,

(c) $\mathrm{Var}(N)\ \{\frac{(p-m)}{2} (\frac{d^2}{b\sigma^2})^2\} \to 1$ as $d \to 0$, and

(d) $P(\underset{\sim}{\beta} \in R_N) \to 1-\alpha$ as $d \to 0$.

Proof: Parts (a), (c) and (d) follow along the same lines as those discussed in the proof of Theorem 2.2.1. To prove Part (b), we consider the new basic inequality

$$\frac{b\ S_m^2}{d^2} \leq N \leq \frac{b\ S_m^2}{d^2} + \left(\frac{a}{d^2}\right)^{\frac{1}{1+\rho}} + p + 1.$$

Dividing by C and taking expectations on both sides, we now obtain

$$\frac{b}{a} \leq E(N/C) \leq \frac{b}{a} + \sigma^{-2}\left(\frac{d^2}{a}\right)^{\frac{\rho}{1+\rho}} + \frac{p+1}{C}.$$

Then, taking limits as $d \to 0$, we can conclude

$$E(N/C) \to 1 \quad \text{as} \quad d \to 0 .$$

This completes the proof of Theorem 3.2.1.

Remark 3.3: The important thing to note here is that by manipulating the starting sample size m, we can make the two-stage procedure to be "asymptotically first-order efficient" in the Ghosh-Mukhopadhyay (1981) sense, that is we can conclude $\lim_{d\to 0} E(N/C) = 1$ for the modified two-stage procedure (3.3.1).

3.4 A Sequential Procedure

Here, we start with the sample size $m(\geq p+1)$, and define the stopping rule

$$N = \inf\{n \geq m: n \geq \frac{a S_n^2}{d^2}\} \qquad (3.4.1)$$

For all $\underset{\sim}{\beta} \in \mathbb{R}^p$ and $\sigma \in (0,\infty)$, N is a positive integer valued random variable which can be easily shown to be finite with probability one.

Thus, when we stop, we compute $\hat{\underset{\sim}{\beta}}_N$ and propose the corresponding confidence region R_N as defined in (3.1.2). Some of the properties of this procedure are listed in Theorem 3.4.1 and 3.4.2.

<u>Theorem 3.4.1</u>: <u>For the procedure (3.4.1)</u>, <u>for all</u> $\underset{\sim}{\beta} \in \mathbb{R}^p$, <u>and</u> $\sigma \in (0,\infty)$ <u>we have</u>:

(a) $E(N) \leq C + m + p + 1$ <u>for all</u> $d > 0$,

(b) $N/C \to 1$ <u>as</u> $d \to 0$,

(c) $E(N/C) \to 1$ <u>as</u> $d \to 0$,

(d) $P(\underset{\sim}{\beta} \in R_N) \to 1-\alpha$ <u>as</u> $d \to 0$, <u>and</u>

(e) $\frac{(N-C)}{\sqrt{2C}} \xrightarrow{L} N(0,1)$ <u>as</u> $d \to 0$.

We omit its proof for brevity, as it follows along the lines of Theorem 2.2.2 with obvious modifications in various steps.

Theorem 3.4.2: *For the procedure* (3.4.1), *for all* $\underset{\sim}{\beta} \varepsilon \mathbb{R}^p$, *and* $\sigma \varepsilon (0,\infty)$ *we have as* $d \to 0$:

(a) $E(N) = C + \nu - 2 - p + o(1)$ *if* $m \geq p + 3$,

(b) $P(\underset{\sim}{\beta} \varepsilon R_N) = 1-\alpha + (d/\sigma)^2 \{\nu - 3 - \frac{(p+a)}{2}\} f(a) + o(d^2)$, *if* (i) $m \geq p+3$ *for* $p=2$ or $p \geq 4$ *and*

(ii) $m \geq 7$ *for* $p=3$.

Here $f(\cdot)$ *is as in Theorem 2.2.3 and*

$$\nu = \frac{3}{2} - \sum_{n=1}^{\infty} \{\Gamma(\tfrac{n}{2})\}^{-1} \{2n^{n/2} e^{-n} - nG(\tfrac{n}{2}, n)\},$$

the function $G(.,.)$ *being defined in (2.2.13).*

We omit its proof also for brevity. The tedious derivations will follow along the same lines of proof of Theorem 2.2.3.

Remark 3.4: The Part (a) of Theorem 3.4.2 shows that the sequential procedure (3.4.1) is indeed "asymptotically second-order efficient" in the Ghosh-Mukhopadhyay (1981) sense, since we have $\lim_{d\to 0} E(N-C) = \nu-2-p$. The modified two-stage procedure (3.3.1) can be shown to have the property $\lim_{d\to 0} E(N-C) = \infty$ instead.

3.5 A Three-Stage Procedure

Motivated by the results of Hall (1981), we now propose the following three-stage procedure in order to estimate C and thereby estimating $\underset{\sim}{\beta}$ in the end.

We start with a sample of size $m(\geq p+1)$ and fix a real number $r \in (0,1)$ and let,

$$M = \max\{m, [ra\ S_m^2/d^2] + 1\}. \tag{3.5.1}$$

We take fresh samples, if needed, to form an $M\times 1$ vector $\underset{\sim}{Y}_M$ at this stage. Then, we compute $\hat{\underset{\sim}{\beta}}_M$ and later obtain S_M^2. We now define

$$N = \max\{M, [a\ S_M^2/d^2] + 1\}, \tag{3.5.2}$$

and take new samples, if needed, to form $\underset{\sim}{Y}_N$. Once we determine N, we compute $\hat{\underset{\sim}{\beta}}_N$ and propose the corresponding region $\mathcal{R}_N$ as in (3.1.2).

Using representation analogous to those in (2.2.12), we can easily rewrite (3.5.1) - (3.5.2) in the following equivalent fashions: we have

$$M = \max\{m, [ra\ \bar{U}_m/d^2] + 1\}, \tag{3.5.3}$$

$$N = \max\{M, [a\ \bar{U}_M/d^2] + 1\}, \tag{3.5.4}$$

where $\bar{U}_k = (k-p)^{-1} \sum_{i=1}^{k-p} U_i$, $k = m, m+1, \ldots$, the U_i's being i.i.d.$\sigma^2\ \chi^2_{(1)}$.

Some properties of this procedure are listed in Theorems 3.5.1 and 3.5.2.

<u>Theorem 3.5.1: For the procedure in (3.5.1) - (3.5.2), for all</u> $\underset{\sim}{\beta} \in \mathbb{R}^p$, <u>and</u> $\sigma \in (0,\infty)$ <u>we have as</u> $d \to 0$:

(a) $E(N) = C + \frac{1}{2} - 2r^{-1} + o(1)$,

(b) $Var(N) = 2r^{-1}C + o(\lambda)$, <u>and</u>

$$\text{(c)} \quad E|N - E(N)|^3 = o(\lambda^2),$$

where $\lambda = a/d^2$.

Proof: We follow very closely the developments in Hall (1981). We indicate only some of the basic steps assuming $\sigma^2 = 1$. Using (4.1) of Hall (1981), we get

$$\begin{aligned} \lambda E(\bar{U}_M) &= \lambda - r^{-1} \operatorname{Var}(U_1) + o(1) \\ &= \lambda - 2r^{-1} + o(1) . \end{aligned} \tag{3.5.5}$$

Also, $E\{\lambda \bar{U}_M - [\lambda \bar{U}_M]\} = \frac{1}{2} + o(1)$, and this can be justified along the lines of Hall (1981). Let $T = [\lambda \bar{U}_M] + 1$. Then Hall's (1981) equation (4.2) will lead to

$$\begin{aligned} E(N) &= E(T) + o(1) \\ &= 1 + E(\lambda \bar{U}_M) - E\{\lambda \bar{U}_M - [\lambda \bar{U}_M]\} + o(1) . \end{aligned}$$

Using (3.5.5) we have

$$E(N) = \lambda + \frac{1}{2} - 2r^{-1} + o(1). \tag{3.5.6}$$

Notice now that $\lambda\sigma^2 = C$, and this leads to Part (a). Again, by using (4.3) and (4.4) from Hall (1981), we obtain

$$\begin{aligned} \operatorname{Var}(N) &= \operatorname{Var}(T) + o(1) \\ &= r^{-1}\lambda \operatorname{Var}(U_1) + o(1) \\ &= 2r^{-1}\lambda + o(\lambda) . \end{aligned} \tag{3.5.7}$$

In (3.5.7), replacing λ by $\lambda\sigma^2$ we obtain Part (b). We omit the proof of Part (c) as it can be tackled along the similar lines as in Hall (1981). This completes the proof of Theorem 3.5.1.

We now modify the three-stage procedure (3.5.1) - (3.5.2) slightly so as to be able to conclude that the resulting coverage probability, namely, $P(\underset{\sim}{\beta} \in R_N)$ turns out to be $(1-\alpha) + o(d^2)$. In order to achieve that goal, we define:

$$m_1 = 3r^{-1} - \frac{1}{2}(p-a)\, r^{-1} - \frac{1}{2} ,$$

$$M = \max\{m, [ra\, S_m^2/d^2] + 1\} , \tag{3.5.8}$$

and

$$N^* = \max\{M, [\frac{a\, S_M^2}{d^2} + m_1] + 1\} . \tag{3.5.9}$$

Once we determine N^*, we compute $\hat{\underset{\sim}{\beta}}_{N^*}$ and propose the corresponding region R_{N^*} for $\underset{\sim}{\beta}$ as in (3.1.2).

Theorem 3.5.2: For the procedure in (3.5.8) - (3.5.9), for all $\underset{\sim}{\beta} \in \mathbb{R}^p$, and $\sigma \in (0,\infty)$ we have as $d \to 0$:

(a) $P(\underset{\sim}{\beta} \in R_{N^*}) = 1-\alpha + o(d^2)$, and

(b) $E(N^*) = C + r^{-1} - \frac{1}{2}(p-a)r^{-1} + o(1)$.

Proof: We first verify Part (a). In fact, we start working with (M,N) from (3.5.1) - (3.5.2), and at the end we show that N must be modified to N^* defined in (3.5.9) to conclude Part (a).

We start with

$$P(\underset{\sim}{\beta} \in R_N) = E\{F(\frac{Nd^2}{\sigma^2})\}$$

$$= E\{F(\ell N)\},$$

where $\ell = d^2/\sigma^2$. Now we can write

$$E\{F(\ell N)\} = F(\ell E(N)) + \frac{1}{2}\ell^2 E(N - E(N))^2 F''(\ell E(N)) + r_1(d),$$

say, where

$$|r_1(d)| \leq K \ell^3 E\{|N - E(N)|^3\} = o(d^2),$$

by Part (c) of Theorem 3.5.1. Here, K is used as a generic positive constant independent of d. We have used the same K whenever needed.
Again, we can write

$$F(\ell E(N)) = F(a) + \{\ell E(N)-a\}F'(a) + r_2(d), \qquad (3.5.10)$$

where we let

$$r_2(d) = \frac{1}{2}\{\ell E(N)-a\}^2 F''(z),$$

for a suitable positive number z.

Let us now take $\lambda = \lambda(d) = a(1+\varepsilon)/d^2$, and with this choice, we have $|r_2(d)| = o(d^2 + |\varepsilon|)$. Also, we have from Part (a) of Theorem 3.5.1

$$\ell E(N)-a = \ell\{\lambda\sigma^2 + \frac{1}{2} - 2r^{-1} + o(1)\} - a$$

$$= a\varepsilon + \frac{d^2}{\sigma^2}(\frac{1}{2} - 2r^{-1}) + o(d^2). \qquad (3.5.11)$$

Thus, combining (3.5.10) and (3.5.11), we obtain

$$F(\ell E(N)) = F(a) + \{a\varepsilon + \frac{d^2}{\sigma^2}(\frac{1}{2} - 2r^{-1})\} F'(a) + o(d^2) + o(d^2 + |\varepsilon|). \tag{3.5.12}$$

Again, we have from Part (b) of Theorem 3.5.1

$$\frac{1}{2}\ell^2 E\{(N-E(N))^2\} = \frac{a\,d^2}{\sigma^2} r^{-1} + o(d^2). \tag{3.5.13}$$

By combining (3.5.12) and (3.5.13), we get

$$\begin{aligned} E\{F(\ell N)\} &= 1-\alpha + F'(a)\,\{a\varepsilon + \frac{d^2}{\sigma^2}(\frac{1}{2} - 2r^{-1})\} + o(d^2) \\ &+ \{\frac{a\,d^2}{\sigma^2} r^{-1} + o(d^2)\} F''(a) + o(d^2 + |\varepsilon|), \end{aligned}$$

which implies

$$\begin{aligned} E\{F(\ell N)\} &= 1-\alpha + \{a\varepsilon + \frac{d^2}{\sigma^2}(\frac{1}{2} - 2r^{-1})\}F'(a) \\ &+ \frac{ad^2}{\sigma^2} r^{-1} F''(a) + o(d^2) + o(d^2 + |\varepsilon|). \end{aligned}$$

Now, note that

$$a\,F''(a) = \frac{1}{2}(p-2-a)\,F'(a),$$

and thus we get

$$E\{F(\ell N)\} = (1-\alpha) + \{a\varepsilon + \frac{d^2}{\sigma^2}(\frac{1}{2} - 2r^{-1}) + \frac{r^{-1}}{2}(p-2-a)\}\ F'(a)$$

$$+\ o(d^2) + o(d^2 + |\varepsilon|). \qquad (3.5.14)$$

Now, in order to make the second term from the left in (3.5.14) vanish, we choose ε such that

$$C\ \varepsilon = 2r^{-1} - \frac{1}{2}(p-2-a)r^{-1} - \frac{1}{2}$$

$$= \{3 - \frac{(p-a)}{2}\}r^{-1} - \frac{1}{2}$$

$$= m_1 .$$

Hence, we can immediately see from (3.5.14) that

$$P(\underset{\sim}{\beta} \in R_{N^*}) = 1-\alpha + o(d^2) \ \underline{\text{as}} \ d \to 0.$$

This proves Part (a).

For Part (b), simply notice from Part (a) of Theorem 3.5.1 that

$$E(N^*) = C + \frac{1}{2} - 2r^{-1} + \{3 - \frac{(p-a)}{2}\}r^{-1} - \frac{1}{2} + o(1)$$

$$= C + r^{-1} - \frac{1}{2}(p-a)r^{-1} + o(1).$$

This completes the proof of Theorem 3.5.2.

3.6 Moderate Sample Size Performance

In this section, we present numerical results obtained through simulations using PROC MATRIX from the SAS package. The subsections 3.6.1, 3.6.2, 3.6.3 and 3.6.4 respectively present the results of our simulation studies when we used the two-stage, the modified two-stage, the sequential and the three-stage procedures.

Let us now briefly explain the way we carried out the simulated experiments in the computer. In any particular table we use a particular stopping "rule" to determine the sample size N, say. We generate a sequence of random samples $\{E_i,\ i=1,2,\ldots\}$ from $N(0,1)$ and let

$$Y_i = \beta_0 + \beta_1 i + E_i,\ i = 1,2,\ldots .$$

In all our procedures we take $\underset{\sim}{\beta}' = (\beta_0, \beta_1) = (1,0.5)$ and $\sigma = 1$. We fix

$$X_s' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & s \end{pmatrix} \text{ where } s = 1,2,\ldots .$$

A particular "rule" is replicated k times, the jth replicate giving rise to observed values of N and $\hat{\underset{\sim}{\beta}}_N$ as, say, $n(j)$ and $\hat{\underset{\sim}{\beta}}_{n(j)}$ respectively. Then, we estimate $E(N)$ and β_ℓ by $\bar{n} = k^{-1} \sum_{j=1}^{k} n(j)$ and $\hat{\beta}_\ell = k^{-1} \sum_{j=1}^{k} \hat{\beta}_{\ell n(j)}$ respectively, $\ell = 0,1$. We also compute the corresponding standard error, namely

$$SD(\bar{n}) = \{(k^2 - k)^{-1} \sum_{j=1}^{k} (n(j) - \bar{n})^2\}^{\frac{1}{2}},$$

$$SD(\hat{\beta}_\ell) = \{(k^2 - k)^{-1} \sum_{j=1}^{k} (\hat{\beta}_{\ell\,(j)} - \hat{\beta}_\ell)^2\}^{\frac{1}{2}},$$

for $\ell = 0,1$. While using a particular "rule", we also estimate the coverage probability of the region R_N by, say, c.p. where c.p. = Observed relative frequency of

$$(\hat{\beta}_{n(j)} - \beta)'(X'_{n(j)} X_{n(j)}) (\hat{\beta}_{n(j)} - \beta) \leq n(j)d^2,$$

among all the replicates for $j = 1,2,\ldots,k$.

For all the simulations we consider 95% confidence regions only, that is, we keep $\alpha = 0.05$ fixed and d is computed using the relationship $d = (a/C)^{\frac{1}{2}}$. All computations are carried out with k = 500.

3.6.1 Moderate Sample Size Performances of the Two-Stage Procedure

We use the "rule" as being the two-stage procedure of Section 3.2. We give results for m=5 and 10 and C = 10, 40, 70, 100. Table XI summarizes our findings.

Remark 3.5: From Table XI, we notice that $\bar{n}$ is always larger than C, however, almost always the estimated coverage (c.p.) exceeds the target which is 0.95. In the sense of less oversampling, the results get better as m increases. We suggest that m be taken as 10 in the absence of any further knowledge.

3.6.2 Moderate Sample Size Performances of the Modified Two-Stage Procedure

Here, we use the "rule" as being the modified two-stage procedure of Section 3.3. We naturally have to choose $\rho(>0)$ suitably. We may notice that as ρ decreases the starting sample size m increases. We first fix C = 10, 40, 70, 100 and then we select ρ = 0.05, 0.1, 0.3, 0.5. In Table XII, we summarize our findings.

TABLE XI

TWO-STAGE PROCEDURE (3.2.1)

m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\beta}_0$	$\hat{\beta}_1$	c.p.
5	10	0.774	32.95	1.13	0.991	0.503	0.962
	40	0.387	134.23	4.56	0.975	0.503	0.956
	70	0.293	210.30	8.24	1.005	0.501	0.944
	100	0.245	307.23	11.01	1.005	0.500	0.956
10	10	0.774	15.94	0.31	0.985	0.501	0.978
	40	0.387	56.77	1.24	0.991	0.501	0.952
	70	0.293	104.03	2.27	1.000	0.500	0.944
	100	0.245	149.96	3.30	1.010	0.500	0.946

TABLE XII

MODIFIED TWO-STAGE PROCEDURE (3.3.1)

ρ	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\beta}_0$	$\hat{\beta}_1$	c.p.
0.05	10	0.774	16.69	0.34	1.010	0.500	0.966
	40	0.387	45.18	0.43	1.008	0.500	0.960
	70	0.293	75.99	0.59	1.002	0.500	0.966
	100	0.245	105.50	0.69	1.002	0.500	0.948
0.10	10	0.774	16.69	0.34	1.010	0.500	0.966
	40	0.387	45.76	0.51	1.004	0.500	0.948
	70	0.293	75.89	0.66	0.996	0.500	0.956
	100	0.245	106.24	0.80	0.011	0.500	0.946
0.30	10	0.774	24.39	0.69	1.059	0.490	0.970
	40	0.387	48.23	0.75	1.032	0.499	0.940
	70	0.293	79.67	1.00	0.984	0.500	0.944
	100	0.245	110.46	1.14	1.002	0.500	0.944
0.50	10	0.774	30.25	1.11	0.980	0.498	0.970
	40	0.387	55.13	1.07	1.015	0.499	0.954
	70	0.293	87.11	1.40	1.012	0.499	0.966
	100	0.245	116.04	1.63	0.994	0.500	0.950

Remark 3.6: From Table XII, we notice that the average sample size $\bar{n}$ is close to C and the estimated coverage probability (c.p.) is also close to 0.95. The results get better in the sense of less oversampling as ρ decreases, and this is generally expected. For this experiment, $\rho = 0.05, 0.10$ or 0.3 seems to be the right choice. We have also run the same experiment with $\sigma^2 = 0.5$ and 0.25 and we found that the most suitable choice for ρ is 0.3. We recommend using the procedure in practice with $\rho = 0.3$ in the absence of any further knowledge.

3.6.3 Moderate Sample Size Performances of the Sequential Procedure

Here, we use the "rule" as being the sequential procedure of Section 3.4. We give results for m = 5 and 10 and C = 10, 40, 70, 100. Table XIII summarizes our findings.

Remark 3.7: From Table XIII, we notice that both $\bar{n}$ and c.p. are very close to C and 0.95 respectively for the sequential procedure. Naturally, this procedure performs better when m increases.

3.6.4 Moderate Sample Size Performances of the Three-Stage Procedure

We use the "rule" as being the three-stage procedure of Section 3.5 which is defined by (3.5.8) and (3.5.9). We estimate $E(M)$ and $E(N^*)$ by $\bar{m}$ and $\bar{n}^*$ respectively where

$$\bar{m} = k^{-1} \sum_{j=1}^{k} m(j) \text{ and } \bar{n}^* = k^{-1} \sum_{j=1}^{k} n^*(j) .$$

We also compute the standard errors

TABLE XIII

SEQUENTIAL PROCEDURE (3.4.1)

m	C	d	$\bar{n}$	SD($\bar{n}$)	$\hat{\beta}_0$	$\hat{\beta}_1$	c.p.
5	10	0.774	8.88	0.18	1.008	0.494	0.884
	40	0.387	35.42	0.57	1.037	0.496	0.884
	70	0.293	67.40	0.71	0.986	0.502	0.940
	100	0.245	97.41	0.87	0.999	0.500	0.938
10	10	0.774	11.56	0.11	0.966	0.507	0.968
	40	0.387	38.45	0.46	0.998	0.502	0.922
	70	0.293	68.53	0.62	1.012	0.499	0.922
	100	0.245	97.36	0.66	1.000	0.500	0.940

TABLE XIV

THREE-STAGE PROCEDURE (3.5.8) - (3.5.9)

r	m	C	d	$\bar{m}$	SD($\bar{m}$)	$\bar{n}^*$	SD($\bar{n}^*$)	$\hat{\beta}_0$	$\hat{\beta}_1$	c.p.
0.3	5	10	0.774	5.50	0.07	25.26	0.25	1.010	0.501	0.998
		40	0.387	12.98	0.45	46.97	0.84	1.002	0.500	0.934
		70	0.293	20.42	0.77	72.53	1.33	0.996	0.500	0.922
		100	0.245	30.49	1.12	102.11	1.71	0.978	0.501	0.906
0.3	10	10	0.774	10.00	0.00	26.79	0.21	0.986	0.501	0.998
		40	0.387	13.45	0.22	52.12	0.64	0.985	0.501	0.960
		70	0.293	21.29	0.42	77.40	0.98	1.006	0.500	0.934
		100	0.245	29.81	0.67	107.25	1.36	0.997	0.500	0.938
0.5	5	10	0.774	6.83	0.15	17.97	0.22	1.055	0.494	0.982
		40	0.387	20.67	0.68	43.86	0.79	0.991	0.502	0.924
		70	0.293	35.48	1.30	74.16	1.28	0.995	0.499	0.908
		100	0.245	55.83	2.00	106.06	1.65	1.015	0.499	0.932
0.5	10	10	0.774	10.08	0.02	19.71	0.20	0.981	0.502	0.994
		40	0.387	21.32	0.40	46.10	0.59	1.014	0.500	0.952
		70	0.293	35.81	0.80	74.54	0.89	0.989	0.500	0.942
		100	0.245	50.21	1.16	104.62	1.08	0.997	0.500	0.940

$$SD(\bar{m}) = \{(k^2 - k)^{-1} \sum_{j=1}^{k} (m(j) - \bar{m})^2\}^{\frac{1}{2}}, \text{ and}$$

$$SD(\bar{n}^*) = \{(k^2 - k)^{-1} \sum_{j=1}^{k} (n^*(j) - \bar{n}^*)^2\}^{\frac{1}{2}} .$$

In this experiment we also took k = 500. We consider r = 0.3, 0.5, 0.7, C = 10, 40, 70, 100 and m = 5, 10.

While carrying out simulation with r = 0.7, we noticed some instability in the estimated coverage probability (c.p.), with no detectable change in the estimates of the average sample sizes. On the other hand, the average sample sizes seemed to increase for r = 0.3. The results for r = 0.5 seemed to be most stable. The results for r = 0.3 and r = 0.5 are reported in Table XIV.

<u>Remark 3.8</u>: In Table XIV when r = 0.5, we notice that $\bar{n}^*$ and c.p. are very close to C and 0.95 respectively. When r = 0.3 we notice that $\bar{n}^*$ overestimates C. In the absence of any further knowledge, we recommend using the three-stage procedure (3.5.8) - (3.5.9) with r = 0.5 and starting sample size m = 5 or 10.

<u>Remark 3.9</u>: In a particular application, if all our procedures can possibly be implemented, we will suggest using the modified two-stage or the three-stage procedure, simply because these will be less time-consuming. However, the sequential procedure will give the best theoretical results if it can be implemented. The main point to note is that the three-stage procedure can be almost as good. The final recommendation should also consider the structure and design of the particular application. We must also stress that we have the coverage probability to be

at least $(1-\alpha)$ with $\lim_{d \to 0} E(N/C) = 1$ for the modified two-stage procedure of Section 3.3. However, the coverage probability becomes only asymptotically $(1-\alpha)$ for the sequential and three-stage procedures.

CHAPTER IV

CONCLUSIONS

In this study we have presented two different problems in the area of constructing fixed-size ellipsoidal confidence regions for multiparameter estimation. Fixed-size ellipsoidal confidence regions for the difference of mean vectors of two independent multinormal populations have been constructed through two-stage and sequential procedures. For this problem the three separate cases, namely, (i) the covariance matrices being equal with equal sample sizes, (ii) the covariance matrices being unequal with equal sample sizes and (iii) the covariance matrices being unequal with unequal sample sizes have been discussed individually. Our two-stage procedure in these contexts guarantee the exact confidence coefficient to be at least the nominal prescribed level. Next, various first-order and second-order asymptotic properties are also considered for the proposed sequential procedures.

Through simulated experiments, we study the moderate sample behaviors of these procedures, and we notice that these procedures perform very well.

The final choice among those proposed procedures should depend on the goals and the types of results one expects to have in a particular context.

Next, we have dealt with the problem of constructing fixed-size ellipsoidal confidence regions for the regression parameters in the general linear model under Gauss-Markoff set up through two-stage, modified two-stage, sequential and three-stage procedures. The proposed two-stage and modified two-stage procedures guarantee the coverage probability to be at least the preassigned nominal value $(1-\alpha)$. For our sequential and three-stage procedures, the coverage probability is shown to be only asymptotically close to $(1-\alpha)$. Numerical simulations for moderate sample sizes have been used to show practical merits of the proposed statistical procedures. Even though our theoretical results are mostly asymptotic in nature, the numerical results indicated that the performances of all these sampling procedures seem to be excellent, even for moderate sample size. Again, the final choice among those procedures should truly depend on the goals and types of results one expects to have in a particular context.

BIBLIOGRAPHY

Albert, A. (1966). "Fixed size confidence ellipsoids for linear regression parameters." Ann. Math. Stat. 37, 1602-1630.

Al-Mousawi, J. S. (1984). "Fixed size confidence regions for the mean vector of multinormal distribution." Ph.D. dissertation in statistics, Oklahoma State University.

Anscombe, F. J. (1952). "Large sample theory of sequential estimation." Proc. Cambridge Philos. Soc. 48, 600-608.

Chapman, D. G. (1950). "Some two-sample tests." Ann. Math. Stat. 21, 601-606.

Chatterjee, S. K. (1959). "On an extension of Stein's two-sample procedure to multinormal problem." Calcutta Stat. Assoc. Bul. 8, 121-148.

Chatterjee, S. K. (1960). "Some further results on the multinormal extension of Stein's two-sample procedure." Calcutta Stat. Assoc. Bul. 9, 20-28.

Chow, Y. S. and Robbins, H. (1965). "On the asymptotic theory of fixed-width sequential confidence intervals for the mean." Ann. Math. Stat. 36, 457-462.

Finster, M. (1983). "A frequentistic approach to sequential estimation in the general linear model." Jour. Amer. Stat. Assoc. 78, 403-407.

Ghosh, B. K. (1975a). "A two-stage procedure for the Behrens-Fisher problem." Jour. Amer. Stat. Assoc. 70, 457-462.

Ghosh, B. K. (1975b). "On the distribution of the difference of two t-variables." Jour. Amer. Stat. Assoc. 70, 463-467.

Ghosh, M. and Mukhopadhyay, N. (1975). "Asymptotic normality of stopping times in sequential analysis." Unpublished manuscript.

Ghosh, M. and Mukhopadhyay, N. (1980). "Sequential point estimation of the difference of two normal means." Ann. Stat. 8, 221-227.

Ghosh, M. and Mukhopadhyay, N. (1981). "Consistency and asymptotic efficiency of two-stage and sequential procedures." Sankhyā, A, 43, 220-227.

Gleser, L. J. (1965). "On the asymptotic theory of fixed size confidence bounds for linear regression parameters." Ann. Math. Stat. 36, 463-467 (see correction note: Ann. Math. Stat. 37, 1053-1055 (1966)).

Govindarajulu, Z. (1981). The Sequential Statistical Analysis of Hypothesis Testing, Point and Interval Estimation, and Decision Theory. American Sciences Press, Columbus, Ohio.

Hall, P. (1981). "Asymptotic theory of triple sampling for sequential estimation of a mean." Ann. Stat. 9, 1229-1238.

IMSL (1982). International Mathematical and Statistical Libraries. Edition 9 (Revised 1982). Houston, Texas.

Kimball, A. W. (1951). "On dependent tests of significance in the analysis of variance. Ann. Math. Stat. 22, 600-602.

Mukhopadhyay, N. (1974). "Sequential estimation of regression parameters in Gauss-Markoff setup." Jour. Indian Stat. Assoc. 12, 39-43.

Mukhopadhyay, N. (1976). "Sequential estimation of a linear function of means of three normal populations." Jour. Amer. Stat. Assoc. 71, 149-153.

Mukhopadhyay, N. (1977). "Remarks on sequential estimation of a linear function of two means: the normal case." Metrika 24, 197-201.

Mukhopadhyay, N. (1980). "A consistent and asymptotically efficient two-stage procedure to construct fixed width confidence interval for the mean." Metrika 27, 781-784.

Mukhopadhyay, N. (1982). "Stein's two-stage procedure and exact consistency." Scandinavisk Aktuarietdskrif, 110-122.

Robbins, H., Simons, G., and Starr, N. (1967). "A sequential analogue of the Behrens-Fisher problem." Ann. Math. Stat. 38, 1384-1391.

Scheffe, H. (1970). "Practical solutions of the Behrens-Fisher problem." Jour. Amer. Stat. Assoc. 65, 1501-1508.

Srivastava, M. S. (1967). "On fixed-width confidence bounds for regression parameters and mean vector." Jour. Roy. Stat. Soc., Ser. B, 29, 132-140.

Srivastava, M. S. (1971). "On fixed-width confidence bounds for regression parameters." Ann. Math. Stat. 42, 1403-1411.

Stein, C. (1945). "A two-sample test for a linear hypothesis whose power is independent of the variance." Ann. Math. Stat., 16, 243-258.

Stein, C. (1949). "Some problems in sequential estimation", (Abstract). Econometrica, 17, 77-78.

Woodroofe, M. (1977). "Second order approximation for sequential point and interval estimation." Ann. Stat. 5, 984-995.

Woodroofe, M. (1982). "Non Linear Renewal Theory in Sequential Analysis." The Society for Industrial and Applied Mathematics, CBMS 39, Philadelphia.

APPENDIX

A1. Anscombe's (1952) Central Limit Theorem

Let $\{Y_n\}$ be an infinite sequence of random variables. We suppose that there exists a real number θ, a sequence of positive numbers $\{W_n\}$, and a distribution function $F(\cdot)$, such that the following conditions are satisfied:

(C1) $P(\frac{Y_n - \theta}{W_n} \leq x) \longrightarrow F(x)$ as $n \to \infty$, for all x such that $F(\cdot)$ is continuous at x.

(C2) $\{Y_n\}$ is uniformly continuous in probability.

(C3) $\{n_\nu\}$ is an increasing sequence of integers, and $\{N_\nu\}$ is a sequence of a proper random variable such that $\frac{N_\nu}{n_\nu} \to 1$ in probability as $\nu \to \infty$.

Theorem A1.1: Under all the stated conditions (C1)-(C3) we have

$$P(\frac{Y_{N_\nu} - \theta}{W_{n_\nu}} \leq x) \to F(x) \text{ as } \nu \to \infty,$$

for all continuity points x of $F(\cdot)$.

Theorem A1.2: Let $X_1, X_2, \ldots$ be i.i.d., and $Y_n = n^{-1} \sum_{i=1}^{n} X_i$.

Suppose (C1) is satisfied and $\frac{W_n}{W_{n+1}} \to 1$ as $n \to \infty$. Then $\{Y_n\}$ is uniformally continuous in probability.

A2. Distribution of a Stopping Time: A Theorem Due to Ghosh and Mukhopadhyay (1975)

Let $N = N_\nu = \inf\{n \geq n_o : n \geq \psi_\nu T_n\}$,

where

(i) $\{\psi_\nu\}$ is a sequence of positive numbers, and $\psi_\nu \to \infty$ as $\nu \to \infty$.

(ii) $\{T_n\}$ is a sequence of statistics such that $P(T_n > 0) = 1$ for $n = 1,2,\ldots$.

Suppose that

$$(a^*) \quad N_\nu^{\frac{1}{2}} \frac{(T_{N_\nu} - a)}{b} \xrightarrow{L} N(0,1) \quad \text{as} \quad \nu \to \infty, \text{ and}$$

$$(b^*) \quad N_\nu^{\frac{1}{2}} \frac{(T_{N_{\nu-1}} - a)}{b} \xrightarrow{L} N(0,1) \quad \text{as} \quad \nu \to \infty,$$

for some a, b > 0.

Theorem A2: Under the stated conditions we have

$$a^{\frac{1}{2}} \frac{(N_\nu - a\psi_\nu)}{b\psi_\nu^{\frac{1}{2}}} \xrightarrow{L} N(0,1) \quad \text{as} \quad \nu \to \infty.$$

A3. Robbins, Simons, and Starr's (1967) Lemma

Given constants $C_{i,j}$ $(i,j = 1,2,\ldots)$ such that $0 < C_{i,j} \to c^* > 0$ as $i,j \to \infty$ and any integer $n_o \geq 0$ define $i(2n_o) = j(2n_o) = n_o$ and for

$n \geq 2n_o$, let

(I) $i(n+1) = i(n)+1, \quad j(n+1) = j(n) \quad \text{if } \frac{i(n)}{j(n)} \leq C_{i(n),j(n)}$

(II) $i(n+1) = i(n), \quad j(n+1) = j(n)+1 \quad \text{if } \frac{i(n)}{j(n)} > C_{i(n),j(n)}.$

Lemma A3: Under the stated conditions we have

$$\frac{i(n)}{j(n)} \to c^* \text{ as } n \to \infty .$$

A4. Woodroffe's (1977) Nonlinear Renewal Theoretic Results

Suppose we have the following:

(C1) $t_c = \inf\{n \geq n_o(\geq 1): S_n < cn^\alpha L(n)\}$, where $\alpha > 0$, $L(n) = 1 + L_o n^{-1} + o(n^{-1})$, $L_o \in (-\infty, \infty)$, and c is a positive parameter (which is often allowed to approach zero). $S_n = \sum_{i=1}^{n} X_i$; X_1, X_2 ... are i.i.d. positive random variables with the distribution function $F(\cdot)$, such that $E(X_1) = \mu$, $Var(X_1) = \tau^2$ and,

(C2) $F(x) \leq Bx^a$ for some a, $B > 0$.

(C3) $\nu = \frac{\beta}{2\mu}\{(\alpha-1)^2 \mu^2 + \tau^2\} - \sum_{n=1}^{\infty} n^{-1} E\{(S_n - n\alpha\mu)^+\}$,

where,

$\beta = (\alpha-1)^{-1}$, $\lambda = \mu^\beta c^{-\beta}$ and $Y^+ = \max(0,Y)$.

Theorem A4.1: Under conditions C1 - C3, and suppose $E(X_1^r) < \infty$ for some $r \geq 2$. Then we have as $c \to 0$.

$$E(t_c) = \lambda + \beta\nu\mu^{-1} - \beta L_o - \frac{1}{2}\alpha\beta^2\tau^2\mu^{-2} + o(1)$$

if $\quad r(2\alpha-1) > 4$ and $n_o a > \beta$.

Theorem A4.2: Suppose $E(X_1^r) < \infty$ for some $r \geq 2$, and C1 - C3 are satisfied. Then, $|\lambda^{-\frac{1}{2}}(t_c-\lambda)|^2$ is uniformly integrable, if $0 < s < \min\{r, \frac{1}{2}(2\alpha-1)r\}$ and $n_o a > \frac{1}{2}\beta s$.

Lemma A4: Suppose $E(X_1^r) < \infty$ for some $r \geq 2$, and C1 - C3 are satisfied. Then, for $0 < \delta, \gamma < 1$ we have

$$P(t_c \leq \delta\lambda) = 0(c^{n_o a}) + 0(\lambda^{-r\gamma/2})$$

as $c \to 0$.

VITA

Ahmed Darwesh Abid

Candidate for the Degree of

Doctor of Philosophy

Thesis: FIXED-SIZE CONFIDENCE REGIONS FOR MULTIPARAMETER ESTIMATION

Major Field: Statistics

Biographical:

Personal Data: Born in Taif, Saudi Arabia, January 1, 1950.

Education: Graduated from Thakif High School, Taif, Saudi Arabia, in May, 1972; received Bachelor of Science degree in Pure and Applied Mathematics from King Saud University, Riyad, Saudi Arabia, in June, 1976; received Master of Science degree in Applied Statistics from Purdue University, West Lafayette, Indiana, in December 1980; completed requirements for the Doctor of Philosophy degree in Statistics at Oklahoma State University in July, 1985.

Professional Experience: Teaching Assistant, King Saud University, Riyad, Saudi Arabia, 1976-1977.

Professional Organization: The Institute of Mathematical Statistics, the Biometric Society, American Statistical Association, Mu Sigma Rho.