T095
13/2/85

100

RESTRICTED COLLECTION

ESTIMATION OF A COMMON MEAN AND RECOVERY OF INTER-BLOCK INFORMATION

CHUNI GOPAL BHATTACHARYA

STATISTICAL INSTITUTE LIBRARY
T 095
13 FEB 1985
CALCUTTA

Thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the award of

Doctor of Philosophy

NEW DELHI

1981

RESTRICTED COLLECTION

ACKNOWLEDGMENTS

I am deeply indebted to Professor S.K. Mitra, my thesis supervisor, for the inspiration and masterly guidance I received from him.

I am grateful to Prof. H.K. Nandi and to Prof. K.R. Shah for the benefit I derived from my close association with them. I thank Prof. K.R. Shah for his kind permission to include in this thesis a piece of work I did jointly with him.

I had some very helpful discussions with Prof. B.L.S. Prakasa Rao. I record my gratitude for him and for all who helped me and encouraged me in my work.

I thank the authorities of the Indian Statistical Institute for providing me with the facilities for carrying out this work.

Finally, I thank Sh. V.P. Sharma for typing the manuscript with utmost care.

INDIAN STATISTICAL INSTITUTE LIBRARY
7095
13 FEB 1985
CALCUTTA

RECTANGLE COLLECTION

CONTENTS

NOTATIONS

$L(A)$		column space of the matrix A
$L_*(A)$	=	null space of the matrix A
$\|A\|$		determinant of the square matrix A
A^-		g-inverse of the matrix A
rank A	=	rank of the matrix A
dim(V)		dimension of the vector space V
I_n	=	identity matrix of order n
0	=	matrix having all elements equal to zero
$\mathrm{Diag}(A_1,\ldots,A_n)$	=	partioned matrix having blocks of square matrices $A_1,\ldots,A_n$ along the diagonal and all non-diagonal blocks equal to 0
$\underset{\sim}{1}_n$	=	n-vector having all components equal to 1
$\underset{\sim}{x}^\delta$	=	diagonal matrix whose diagonal elements are the components of the vector $\underset{\sim}{x}$ in the same order as in $\underset{\sim}{x}$.
$\underset{\sim}{x}^{-\delta}$	=	inverse of the matrix x^δ

CHAPTER 1

INTRODUCTION

The problem of combining several estimates of an unknown quantity to obtain an estimate of improved precision arises in many spheres of application of statistics. To begin with let us consider the following simple model:

$$y_i = \mu + \epsilon_i \ , \quad i = 1,\ldots,k$$

where μ is an unknown quantity and ϵ_i's are errors with a common mean zero and a common variance σ^2. If we make no further assumption about the distribution of ϵ_i's, the Gauss-Markoff theorem tells us that among all unbiased linear combinations of y_i's, the least square estimator $\bar{y} = \Sigma y_i/k$ of μ has the minimum variance. If ϵ_i's are jointly normally distributed, then the least square estimator is also the maximum likelihood estimator and has minimum variance in the class of all unbiased estimators [Rao (1952)]. Under the assumption of normality the estimator enjoys yet another property that it is admissible in the class of all estimators with respect to any loss function which is monotonic increasing function of the absolute error [Blyth (1951)]. All these important results admit of immediate extension to the case when ϵ_i's are correlated and have unequal variances provided we know the relative values of the elements of the dispersion matrix of $\underset{\sim}{\epsilon} = (\epsilon_1,\ldots,\epsilon_k)'$ i.e. $V(\epsilon) = \sigma^2 H$, where H is a known matrix. In this case an well known modification of the ordinary least squares procedure provides an estimator with all the properties stated above. In many cases it is not unreasonable to assume that y_i's have independent normal distributions but it is unreasonable to assume that the relative values of the variances of y_i's are known. For an example, suppose that two laboratories have made separate determinations y_i's of the same physical or chemical quantity. It is easy to conceive situations where it is unreasonable to assume that the two laboratories do not differ in precision. In general, the relative precisions are not known but can be estimated from the current or previous data. Thus in the above example each laboratory may provide us with an estimated standard error s_i for the estimate y_i of μ. The problem of obtaining a good combined estimator in such practical situations is not straightforward. The mathematical model generally assumed for the problem is as follows:

(i) $y_i \sim N(\mu, \sigma_i^2)$, $i = 1, \ldots, k$, are independent

(ii) $s_i^2/\sigma_i^2 \sim \chi^2_{m_i}$, $i = 1, \ldots, k$, are independent

The problem of estimating μ of the above model, traditionally known as the weighted mean problem has been of considerable interest to both theoreticians and practitioners of statistics. The theoretical interest arises because of the difficulty of eliminating the unknown variance parameters from inference about μ[See Hinkley (1979) for a recent discussion]. The problem has been treated at length in the literature starting with the papers by Bartllet (1936, 1937) and is of keen interest even today. Our contribution on this problem is presented in chapters 2 and 3. In chapter 2, we consider the problem of estimating the common mean of K normal distributions. The special case of estimating the common mean of two normal distributions is considered in chapter 3. To avoid repetetions. we have omitted in chapter 3 derivation of all results which are derived in chapter 2 unless a simpler approach can be used in the special case considered in chapter 3.

A somewhat similar problem arises in the analysis of a block design under the Eisenhart model III (blocks random, error random) [Eisenhart (1947)] where certain treatment contrasts admit two independent estimates commonly known as the intra-block and the inter-block estimates. The so called problem of recovery of inter-block information seeks to combine these to obtain improved estimators of these contrasts. Yates (1939, 1940) confining himself to special designs was the first to observe this and suggest a method of recovery of inter-block information. His idea has been gradually extended by others and many alternative procedures have been proposed. It should be noted that although similar methods are applicable, the experimental design problem differs from the weighted mean problem in several important aspects and requires seperate treatment. We mention here two main aspects. Firstly, we have to consider an appropriate method of reduction of the data, which is not quite obvious as in the weighted mean problem. Secondly, we are required to estimate not just a single parameter but all estimable parametric functions of several parameters. Our contribution on this problem is presented in chapters 4 and 5. Chapter 4 contains practically all our results on this problem. Chapter 5 contains some additional results concerning the use of modified estimators suggested by Yates (1939) and Stein (1966). Results obtained in this chapter are also useful for the common mean problem treated in chapter 2, if one has a-priori knowledge concerning the variance ratio similar to that we have in the design problem as explained in the introduction of chapter 5.

There has been a good deal of work on the problem of point estimation of the common mean of two normal distributions together with the problem of use of recovery of inter-block information for point estimation of treatment differences in block designs. Since the probability distribution of these estimators are not easily tractable, comparatively very little has been done on the interval estimation of these parameters. Meier (1953 [See also Cochran (1954)], appears to be the first contributor on this problem. Following his work, there has been many useful contributions by others. Our contribution on this problem, which is closely related to that in Brown and Cohen (1974), is presented in chapter 6.

The problem of combining two or more independent unbiased estimators has been so far studied, extensively only in the normal case. The only contributors on this problem in the non-normal case appears to be Hogg (1960) and Cohen (1976). In chapter 7, we improve Cohen's results to add more practical value to his estimator.

In order to keep the material in the text close to the subject matter, we have presented the derivation of some inequalities, used at several places of the text, in the appendix. We believe that theorems A1 and A2 of the appendix would be of general interest for mathematical statistics.

We shall refrain from giving a survey of the literature which is vast and discuss the work by others only when it is strictly necessary for understanding our own work. Each chapter contains an introduction where due references are given to all important contributions along with a brief summary of our own work in that chapter.

CHAPTER 2

ESTIMATION OF THE COMMON MEAN OF SEVERAL NORMAL POPULATIONS

2.1 Introduction

Consider k independent random samples of sizes $n_1,\dots,n_k$ respectively from k normal populations having a common unknown mean μ and unknown variances, $\sigma_1^2,\dots,\sigma_k^2$. The problem is to estimate μ on the basis of the combined sample. Let x_{ij} denote the j-th observation in the ith sample; $\underset{\sim}{x}_i = (\underset{\sim}{x}_{i1},\dots,\underset{\sim}{x}_{in(i)})'$ where for the sake of simplicity we have written n(i) for n_i where it appears as a subscript; $\underset{\sim}{y} = (\underset{\sim}{x}_1',\dots,\underset{\sim}{x}_k')$. Then, our model is

$$\underset{\sim}{y} = A\mu + \underset{\sim}{\varepsilon} \tag{2.1.1}$$

where $A = \underset{\sim}{1}_n$; $n = \sum_{i=1}^{k} n_i$; $\underset{\sim}{\varepsilon} \sim N(0, \Sigma)$;

$$\Sigma = \mathrm{diag}(\sigma_1^2 I_{n(1)},\dots,\sigma_k^2 I_{n(k)}). \quad \text{Let } \underset{\sim}{\sigma} = (\sigma_1^2,\dots,\sigma_k^2).$$

It is well known that a minimal sufficient statistic for $(\mu,\underset{\sim}{\sigma})$ is given by $(X_1,\dots,X_k;\ S_1,\dots,S_k)$, where $X_i = \sum_{j=1}^{n(i)} x_{ij}/n_i$; $S_i = \sum_{j=1}^{n(i)} (x_{ij} - X_i)^2/n_i$ [To simplify notation we have written X_i, S_i in place of the usual notation $\bar{x}_i$, s_i^2 for the sample mean and sample variance]. It is also well known that $X_1,\dots,X_k;\ S_1,\dots,S_k$ are independently distributed and that $X_i \sim N(\mu, \bar{\sigma}_i^2)$, $S_i/\bar{\sigma}_i^2 \sim \chi^2_{n(i)}$ where $\bar{\sigma}_i^2 = \sigma_i^2/n_i$; $m_i = n_i - 1$. Clearly, for any $i \neq j$, $E(X_i - X_j) = 0$ but $\mathrm{Prob}(X_i \neq X_j) > 0$. Hence the minimal sufficient statistic is incomplete and do not lead us to UMVUE. It may be noted that (in view of lemma 2.2.1 which we shall prove in the next section), any estimator of μ can be expressed in the form : $\hat{\mu} = \sum_{i=1}^{k} \phi_i X_i$, where ϕ_i's are measureable functions of $\underset{\sim}{y}$ such that $\sum_{i=1}^{k} \phi_i = 1$. We shall denote the estimator by $\hat{\mu}(\underset{\sim}{\phi})$, where $\underset{\sim}{\phi} = (\phi_2,\ldots,\phi_k)$. For convenience define, $\rho_i = \sigma_i^2/\sigma_1^2$, $\eta_i = \bar{\sigma}_i^2/\bar{\sigma}_1^2$, $\gamma_i = \bar{\sigma}_i^{-2}/\sum_{i=1}^{k} \bar{\sigma}_i^{-2}$ $\underset{\sim}{\rho} = (\rho_2,\ldots,\rho_n)$, $\underset{\sim}{\eta} = (\eta_2,\ldots,\eta_k)$; $\underset{\sim}{\gamma} = (\gamma_2,\ldots,\gamma_k)$. It is well known that BLUE of μ is given by $\hat{\mu}(\underset{\sim}{\gamma})$, if ρ is known. But no optimal solution is apparent, in the present problem, in *which ρ is unknown.*

Among the various solutions proposed in the literature those which are applicable for any $k \geq 2$ and based on well-defined principles are: (i) the maximum likelihood estimator [Bartlett (1936)] and its modifications; one, by Bartlett (1936, 1937) [see also Neyman and Scott (1948)] and another, by Kalbfleisch and Sprott (1950); (ii) Partially Bayes estimator [Cox (1975)], (iii) MINQUE estimator[J.N.K. Rao and Subrahmaniam (1971), (iv) the so called uniformly better estimators [Brown and Cohen (1974), Norwood and Hinkelmann (1977), Shinozaki (1978), Bhattacharya (1979 , 1980)]. An estimator which has been in long use is $\bar{\mu} = \hat{\mu}(\underset{\sim}{\phi})$, where $\underset{\sim}{\phi}$ is given by $\phi_i = (m_i/S_i)/\sum_{i=1}^{k}(m_i/S_i)$, $i = 2,\ldots,k$. This has been studied intensively by many authors [Cochran (1937, 1954), Meir (1953), Cochran and Caroll (1953), Williams (1967, 1975), Bement and Williams (1969), Norwood and Hinkelman (1977) and Sinha (1979), who erroneously calls it M.L.E.]. Yates and Cochran (1938) and Cochran (1954) proposed modification of this depending on preliminary tests of hypotheses concerning the unknown $\underset{\sim}{\sigma}$.

In section 2 we present some general results concerning unbiasedness and variance which apply to all estimators proposed in the literature/present work. In section 3 and 4 we propose estimators of the form $\hat{\mu}(\underset{\sim}{\phi})$, where $\underset{\sim}{\phi}$ is related to an appropriate estimate $\hat{\theta}$ of σ in the same way as $\underset{\sim}{\gamma}$ is related to $\underset{\sim}{\sigma}$. In section 3, $\hat{\theta}$ is obtained by an application of the marginal likelihood procedure formulated by Fraser (1968) and Kalbfleisch and Sprott (1970).

In Section 4, $\hat{\theta}$ is obtained by an application of theory of MIVQUE in Rao (1971). In Section 5 we offer a class of estimators better than X_1. The result is an extension of a similar result in Brown and Cohen (1974). Section 6 is devoted to some studies leading to the class of estimators better than each X_i in Shinozaki (1978) [see also Bhattacharya(1979)]. We improve an important intermediate result in Shinozaki (1978) and provide an alternative proof of his final result, which we believe to be more elegant.

2.2 Unbiasedness and Variances of a General Class of Estimators.

In this section we shall consider a very general class of estimators. We first prove a lemma concerning the general form of all estimators mentioned in the introduction.

Lemma 2.2.1 Any estimator $\hat{\mu}$ of μ can be expressed in the form: $\hat{\mu} = \hat{\mu}(\phi)$

where $$\hat{\mu}(\underset{\sim}{\phi}) = X_1 + \underset{\sim}{\phi}'\underset{\sim}{\delta} = \sum_{i=1}^{k} \phi_i X_i \quad , \qquad (2.2.1)$$

ϕ_i's are measurable functions such that $\sum_{i=1}^{k} \phi_i = 1$; $\phi = (\phi_2, \ldots, \phi_k)'$; $\delta = (\delta_2, \ldots, \delta_k)'$, $\delta_i = X_i - X_1$.

Proof We can write $\hat{\mu}$ in the form:

$$\hat{\mu} = X_1 + \Psi_i(X_i - X_1), \; i = 2,\ldots,k \qquad (2.2.2)$$

where $\Psi_i = (\hat{\mu}-X_1)/(X_i-X_1)$ is measurable since X_i-X_1 has a continuous distribution. Summing both sides of (2.2.2) over i from 2 to k and dividing the result by (k-1), we have

$$\hat{\mu} = X_1 + \sum_{i=2}^{k} \Psi_i^*(X_i - X_1) \qquad (2.2.3)$$

where $\Psi_i^* = \Psi_i/(k-1)$. The result follows from (2.2.3) by taking

$$\phi_i = \Psi_i^* \quad \text{if} \quad i \geq 2$$

$$= 1 - \sum_{i=2}^{k} \psi_i^* \text{ if } i = 1.$$

Let $\underset{\sim}{\varepsilon}_i$ be a column – vector of m_i ortho-normal contrasts of the elements of $\underset{\sim}{x}_i$; $\underset{\sim}{\delta}$ be as in lemma 2.2.1; $\underset{\sim}{\varepsilon}_* = (\underset{\sim}{\varepsilon}_1' \vdots \ldots \vdots \underset{\sim}{\varepsilon}_k' \vdots \underset{\sim}{\delta}')'$ and let $\underset{\sim}{\varepsilon}$ be as in (2.1.1) of the previous section. We shall say that $\underset{\sim}{\phi}$ is even in $\underset{\sim}{\varepsilon}$ if $\underset{\sim}{\phi}(\underset{\sim}{\varepsilon}) = \underset{\sim}{\phi}(-\underset{\sim}{\varepsilon})$ a.s.; and odd in $\underset{\sim}{\varepsilon}$ if $\phi(\underset{\sim}{\varepsilon}) = -\phi(-\varepsilon)$ a.s. Let Φ_0 denote the class of all functions of $\underset{\sim}{\varepsilon}_*$ which are measurable and even in $\underset{\sim}{\varepsilon}_*$. We shall confine ourselves to the class of estimator $\{\hat{\mu}(\underset{\sim}{\phi}), \underset{\sim}{\phi} \in \Phi_0^{k-1}\}$ where Φ_0^k denotes the artesian product of Φ_0 taken k times, and $\hat{\mu}(\phi)$ is as defined in (2.2.1). Note that any linear zero function of y is a linear and hence an odd function of $\underset{\sim}{\varepsilon}$; in particular, each element of $\underset{\sim}{\varepsilon}_*$ is an odd function of $\underset{\sim}{\varepsilon}$. In addition to this simple observation, we shall use the following useful result pointed out by Kakwani (1967).

<u>Lemma 2.2.2</u> An estimator whose expectation exists is unbiased provided its deviation from the true value is of the form $f(\varepsilon)$, where (i) ε has a distribution which is symmetric about zero, (ii) $f(\varepsilon)$ is an odd function of $\underset{\sim}{\varepsilon}$

<u>Theorem 2.2.1</u> Let $\underset{\sim}{\phi} \in \Phi_0^{k-1}$ and assume the $E[\hat{\mu}(\phi)]$ exists. Then

(i) $\hat{\mu}(\underset{\sim}{\phi})$ is unbiased for μ.

(ii) $V[\hat{\mu}(\phi)] = \bar{\sigma}_1^2 [\gamma_1 + E[\underset{\sim}{\delta}'(\underset{\sim}{\phi} - \underset{\sim}{\gamma})]^2]$ (2.2.4)

<u>Proof</u> (i) follows from lemma 2.2.2 since $\hat{\mu}(\underset{\sim}{\phi}) - \mu = X_1 - \mu + \underset{\sim}{\delta}'\underset{\sim}{\phi}$ is an odd function of $\underset{\sim}{\varepsilon}$

(ii) We can write $\hat{\mu}(\underset{\sim}{\phi})$ in the form

$$\hat{\mu}(\underset{\sim}{\phi}) = \hat{\mu}(\underset{\sim}{\gamma}) + \underset{\sim}{\delta}'(\underset{\sim}{\phi} - \underset{\sim}{\gamma}) \quad (2.2.5)$$

Note that $\mu(\underset{\sim}{Y})$ is U.M.V.U.E. and that $\underset{\sim}{\varepsilon}_*$ is a vector of zero functions with finite variances. Hence by the result of Stein (1950), $\text{Cov}[\underset{\sim}{\varepsilon}_*, \hat{\mu}(\underset{\sim}{Y})] = \underset{\sim}{0}$. Note also that $\underset{\sim}{\varepsilon}_*$ and $\hat{\mu}(\underset{\sim}{Y})$ are jointly normal and hence $\hat{\mu}(\underset{\sim}{Y})$ is independent of $\underset{\sim}{\varepsilon}_*$. Thus, the second term on the r.h.s. of (2.2.5) is independent of the first term since it is a measurable function of $\underset{\sim}{\varepsilon}_*$ which has the desired property. Hence,

$$V[\hat{\mu}(\underset{\sim}{\phi})] = V[\hat{\mu}(\underset{\sim}{Y})] + E[\underset{\sim}{d}'(\underset{\sim}{\phi} - \underset{\sim}{Y})]^2$$

The result follows by observing that, $V[\hat{\mu}(\underset{\sim}{Y})] = \bar{\sigma}_1^2 \, Y_1$.

Let us now turn our attention to the class of translation invariant and scale preserving estimators which can be seen to be equivariant in the sense of Berk (1967) and Wijsman (1967). Let R denote the set of all real numbers and let R^n denote the Cartesian product of R taken n times. Consider the group G of transformations on the set R^n defined by

$$G = \{g_{\alpha\beta} \mid g_{\alpha\beta}\, \underset{\sim}{x} = \alpha \underset{\sim}{1} + \beta \underset{\sim}{x},\ \underset{\sim}{x} \in R^n,\ \alpha \in R,\ \beta \in R,\ \beta \neq 0\}$$

Let $T = T(\underset{\sim}{y})$ be a statistic. Then following Zacks (1970), T is said to be an equivariant estimator of μ iff

$$T(g_{\alpha\beta}\, \underset{\sim}{y}) = \alpha + \beta T(\underset{\sim}{y}),\ \forall\, g_{\alpha\beta} \in G$$

From this definition and the proof of lemma 2.4.1 it is easy to see that any estimator $\hat{\mu}$ in the present problem is equivariant iff it is of the form $\hat{\mu}(\underset{\sim}{\phi})$ where $\underset{\sim}{\phi}$ is a measurable function of $(y_1,\ldots,y_n)$, such that it is completely invariant under G. Hence $\underset{\sim}{\phi}$ is a measurable function of any maximal invariant function of $(y_1,\ldots,y_n)$ under G. A maximal invariant function of $(y_1,\ldots,y_n)$ under G is clearly given by $(y_3-y_1)/(y_2-y_1)\ \ldots\ (y_n-y_1)/(y_2-y_1)$ and it is easily seen to be an odd function of $\underset{\sim}{\varepsilon}$. Hence using theorem 2.2.1 we have

<u>Theorem 2.2.2</u> Any equivariant estimator of μ whose expectation exists is unbiased and has a variance given by formula (2.2.4).

Zacks (1970) considered the class of equivariant estimators based on the minimal sufficient statistics in the special case k = 2. Results concerning unbiasedness and variance of such estimators in Zacks (1970) [who used an enitrely different approach] follow from our theorem 2.2.2. For the same special case Brown and Cohen (1974) and Khatri and Shah (1974) considered the more general class of estimator $\hat{\mu}(\phi)$ where ϕ is a measurable function of (S_1, S_2, W) and $W = (X_2-X_1)^2$. Clearly this class of estimators is a subclass of $\{\hat{\mu}(\phi), \phi \varepsilon \Phi_0\}$. Results in Brown and Cohen (1974) and Khatri and Shah (1974) concerning unbiasedness and variance of estimators belonging to this class can be deduced from our Theorem 2.2.1. Arguments used by these authors are essentially same as in Zacks (1970) except for an interesting innovation which leads to an elegant expression for the variance. It can be verified that all estimators of μ considered in the literature/present work belong to the class $\{\hat{\mu}(\phi) | \phi \varepsilon \Phi_0\}$ [in fact with the exception of some minimax estimators in Cohen and Sackrowitz (1974) for the special case k = all are equivariant] and hence unbiased, in view of our theorem 2.2.1. In particular, the M.L.E. is unbiased. As far as the author is aware unbiasedness of the M.L.E. may not have been noticed earlier. The use of M.S.E. for M.L.E. in Levy (1970) is an indication that he considered that M.L.E. may not be unbiased.

2.3 <u>Estimator of μ Based on an Estimator of γ from a Marginal Likelihood of $\underset{\sim}{\sigma}$</u>

In this section we shall obtain an estimate ϕ of $\underset{\sim}{\gamma}$ from a marginal likelihood of $\underset{\sim}{\sigma}$. We then propose to estimate μ by $\hat{\mu}(\phi)$. Let $\bar{y}$ denote the mean of the combined sample and let ε_* be as defined in the previous section.

It is easy to see that the transformation from $\underset{\sim}{y} \to (\bar{y}, \underset{\sim}{\epsilon}_*')'$ is one to one; the likelihood of $(\mu, \underset{\sim}{\sigma})$ is product of two independent factors; one given by the density of $\bar{y}$ (which depends on both μ and $\underset{\sim}{\sigma}$) and the other given by the density of $\underset{\sim}{\epsilon}_*$(which depends solely on σ). The likelihood of σ as given by the density of $\underset{\sim}{\epsilon}_*$ is

$$L_1 \propto |\Sigma_*|^{-\frac{1}{2}} \prod_{i=1}^{k} \sigma_i^{-m_i} \exp\{-(\sum_{i=1}^{k} S_i/\sigma_i^2 + \underset{\sim}{\delta}' \Sigma_*^{-1} \underset{\sim}{\delta})/2\} \qquad (2.3.1)$$

where $\Sigma_* = \sigma_1^{-2} (\underset{\sim}{1}_{k-1} 1_{k-1}' + \underset{\sim}{\eta}^{\delta})$; $\underset{\sim}{1}_k$ = the column vector consisting of all elements equal to 1; $\underset{\sim}{\eta}^{\delta}$ = the diagonal matrix whose diagonal elements are the components of the column vector η in the same order as in $\underset{\sim}{\eta}$; $|\Sigma|$ stands for the determinant of Σ. It is easy to see that L_1 is maximum with respect to σ_1^2 when

$$\hat{\sigma}_1^2 = \frac{1}{m+k-1}\left(S_1 + \sum_{i=2}^{k} S_i/\eta_i + \underset{\sim}{\delta}' H^{-1} \underset{\sim}{\delta}\right) \text{ where } H = \underset{\sim}{1}_{k-1} 1_{k-1}' + \eta^{\delta}$$

Hence, the maximum value of L_1 with respect to σ_1^2 is

$$L_2 \propto |H|^{-\frac{1}{2}} \prod_{i=2}^{k} \eta_i^{-m_i/2} (S_1 + \sum_{i=2}^{k} S_i/\eta_i + \underset{\sim}{\delta}' H^{-1} \underset{\sim}{\delta})^{-(m+k-1)/2}$$
$$(2.3.2)$$

where $m = \sum_{i=1}^{k} m_i$. Note that $\underset{\sim}{\eta}^{\delta} = \gamma_1 \underset{\sim}{\gamma}^{-\delta}$; $\gamma_1 = 1 - \sum_{i=2}^{k} \gamma_i$. Let $\alpha = 1 + \underset{\sim}{1}_{k-1}' \underset{\sim}{\eta}^{-\delta} \underset{\sim}{1}_{k-1}$. It is easy to see that $\alpha = \gamma_1$ and hence

$$|H| = \alpha |\underset{\sim}{\eta}^{\delta}| = \gamma_1^{k-1} \prod_{i=1}^{k} \gamma_i^{-1}$$

$$H^{-1} = \underset{\sim}{\eta}^{-\delta} - \alpha^{-1} \underset{\sim}{\eta}^{-\delta} \underset{\sim}{1}_{k-1} \underset{\sim}{1}_{k-1}' \underset{\sim}{\eta}^{-\delta} = \gamma_1^{-1} (\gamma^{\delta} - \underset{\sim}{\gamma} \underset{\sim}{\gamma}')$$

Also

STATISTICAL INSTITUTE LIBRARY
No. T 095
13 FEB 1985
CALCUTTA

$$\prod_{i=2}^{k} \eta_i^{-m_i/2} = \gamma_1^{-m/2} \prod_{i=1}^{k} \gamma_i^{m_i/2}$$

Hence we have

$$L_2 \propto \prod_{i=1}^{k} \gamma_i^{(m_i+1)/2} \left[\sum_{i=1}^{k} \gamma_i S_i + \underset{\sim}{\delta}' B_* \underset{\sim}{\delta}\right]^{-(m+k-1)/2} \qquad (2.3.3)$$

where

$$B_* = \underset{\sim}{\gamma}^{\delta} - \underset{\sim}{\gamma}\,\underset{\sim}{\gamma}'.$$

We have

$$\hat{\mu}(\underset{\sim}{\gamma}) = \sum_{i=1}^{k} \gamma_i X_i \qquad (2.3.4)$$

Hence

$$\underset{\sim}{\delta}'\underset{\sim}{\gamma} \quad \sum_{i=2}^{k} \delta_i \gamma_i = \hat{\mu}(\underset{\sim}{\gamma}) - X_1 \qquad (2.3.5)$$

and

$$\underset{\sim}{\delta}'\gamma^{\delta}\underset{\sim}{\delta} = \sum_{i=2}^{k} \gamma_i \delta_i^2 = \sum_{i=1}^{k} \gamma_i [X_i - \hat{\mu}(\underset{\sim}{\gamma}) + \hat{\mu}(\underset{\sim}{\gamma}) - X_1]^2$$

$$= \sum_{i=1}^{k} \gamma_i [X_i - \hat{\mu}(\underset{\sim}{\gamma})]^2 + [X_1 - \hat{\mu}(\underset{\sim}{\gamma})]^2$$

Since the product term vanishes in view of (2.3.4). Using (2.3.5) and (2.3.6)

$$\underset{\sim}{\delta}' B_* \underset{\sim}{\delta} = \sum_{i=1}^{k} \gamma_i [X_i - \hat{\mu}(\underset{\sim}{\gamma})]^2$$

Using this we have from (2.3.3)

$$\log L_2 = \text{const.} + \frac{1}{2}\left[\sum_{i=1}^{k} (m_i+1) \log \gamma_i - (m+k-1) \log \sum_{i=1}^{k} \gamma_i T_i(\underset{\sim}{\gamma})\right]$$

where $T_i(\underset{\sim}{\gamma}) = S_i + [X_i - \hat{\mu}(\underset{\sim}{\gamma})]^2$. Note that $\gamma_1 = 1 - \sum_{i=2}^{k} \gamma_i$ and regard L_2 as a function of $(\gamma_2, \ldots, \gamma_k)$. Then differentiating $\log L_2$ w.r.t. γ_i

and equating to zero for $i = 2,\ldots,k$, we arrive at the following equation for the required estimate $\underset{\sim}{\phi} = (\phi_2,\ldots,\phi_k)$ of $\underset{\sim}{\gamma}$:

$$-(m_1+1)/\phi_1 + (m_i+1)/\phi_i - (m+k-1)(T_i^* - T_1^*)/T^* = 0 \qquad (2.3.7)$$

where

$$\phi_1 = 1 - \sum_{i=2}^{k} \phi_i ;\ T_i^* = T_i(\underset{\sim}{\phi}) ;\ T^* = \sum_{i=1}^{k} \phi_i\ T_i^*.$$

Note that (2.3.7) holds for $i = 1$ also. Hence multiplying both sides of (2.3.7) by ϕ_i and summing over i from 1 to k,

$$-(m_1+1)/\phi_1 + 1 + (m+k-1)T_1^*/T^* = 0 \qquad (2.3.8)$$

(2.3.7) and (2.3.8) finally gives,

$$(m_i+1)/\phi_i = 1 + (m+k-1)T_i^*/T^* \ ;\ i = 1,\ldots,k \qquad (2.3.9)$$

Since the r.h.s. of (2.3.9) is positive it is clear that each $\phi_i > 0$ and then we also have that each $\phi_i < 1$ in view of the condition $\Sigma\phi_i = 1$, which is satisfied by any solution of (2.3.9). It can be seen that the expression (2.3.1) from which we derived our estimate $\underset{\sim}{\phi}$ of $\underset{\sim}{\gamma}$ is a marginal likelihood of $\underset{\sim}{\sigma}$ in the sense of both Fraser (1968) and Kalbfleisch and Sprott (1970). It is interesting to observe in this connection that the expression (2.3.2) obtained by maximizing the expression (2.3.1) w.r.t. of σ_1^2 is a marginal likelihood of $\underset{\sim}{\rho}$ in the same sense [see Shaarawi et. al. (1975) for details]. Accordingly the expression (2.3.3) which is equivalent to (2.3.2) if we consider the reparametrization $\underset{\sim}{\rho} \to \underset{\sim}{\gamma}$, is a marginal likelihood of γ. Hence, the derived estimate of $\underset{\sim}{\phi}$ of $\underset{\sim}{\gamma}$ may also be regarded as one based on a marginal likelihood of $\underset{\sim}{\gamma}$.

2.4 Estimator of μ based on LBQUE (with invariance) of $\underset{\sim}{\sigma}$

In this section we shall obtain the locally best quadratic unbiased estimator $\hat{\underset{\sim}{\sigma}}$ of $\underset{\sim}{\sigma}$ subject to the condition of invariance under translation of μ[see Rao (1971b)]. We then propose to estimate μ by $\hat{\mu}(\underset{\sim}{\phi})$, where $\underset{\sim}{\phi}$ is related to $\hat{\underset{\sim}{\sigma}}$ in the same way as $\underset{\sim}{\gamma}$ is related to $\underset{\sim}{\sigma}$. Our model is as given by (2.1.1) Note that in this model $\underset{\sim}{\varepsilon} = \sum_{i=1}^{k} U_i \, \xi_i$ where $\underset{\sim}{\xi}_i \sim N(0, \sigma_i^2 \, I_{n(i)})$ are mutually independent; $U_i = (0,\ldots,0,I_{n(i)},0,\ldots,0)'$ is an $n \times n_i$ matrix.

Let

$$\underset{\sim}{\sigma}_0 = (\alpha_1,\ldots,\alpha_k);\; H_i = U_i U_i';\; T_i = \alpha_i H_i;\; T_* = \sum_{i=1}^{k} T_i;$$

$$R = T_*^{-1} - T_*^{-1} A (A' T^{-1} A)^{-} A' T_*^{-1};\quad Q_i = y' R \, T_i R y;\; \underset{\sim}{Q} = (Q_i)_{k\times 1}$$

$$S_{ij} = \operatorname{tr} R \, T_i \, R T_j;\; S = (S_{ij})_{k\times k} \, .$$

Note that $\underset{\sim}{\varepsilon}$ has a multivariate normal distribution. Hence, according to the theory in Rao (1971a,b) the LBQUE (with invariance) at $\underset{\sim}{\sigma} = \underset{\sim}{\sigma}_0$ is $\hat{\sigma}$ given by

$$S \, \underset{\sim}{\sigma}_0^{-\delta} \;\; \hat{\underset{\sim}{\sigma}} \; = \; \underset{\sim}{Q} \tag{2.4.1}$$

It is easy to see that

$$H_i = \operatorname{diag}\,(0,\ldots,0,\; I_{n(i)},0,\ldots,0) \tag{2.4.2}$$

$$T_*^{-1} = \sum_{i=1}^{k} \alpha_i^{-1} H_i;\; A' T_*^{-1} A = \alpha_0, \quad \text{where} \quad \alpha_0 = \sum_{i=1}^{k} n_i \, \alpha_i^{-1} ;$$

$$R = T_*^{-1} B, \text{ where } \quad B = [I - \alpha_0^{-1} \, (\alpha_1^{-1} \, \underset{\sim}{1}_n \, \underset{\sim}{1}'_{n(1)}, \ldots, \alpha_k^{-1} \, \underset{\sim}{1}_n \underset{\sim}{1}'_{n(k)})] \tag{2.4.3}$$

It is also easy to see that, $T_i T_*^{-1} = H_i$

$$T_*^{-1} T_i T_*^{-1} = \alpha_i^{-1} V_i ; B\underset{\sim}{y} = \underset{\sim}{y} - \alpha_0^{-1} \sum_{i=1}^{k} n_i \alpha_i^{-1} X_i$$

$$BH_i = H_i - \alpha_0^{-1}(0,\ldots,0,\alpha_i^{-1} \underset{\sim}{1}_n \underset{\sim}{1}'_{n(i)}, 0,\ldots,0)$$

Let $\underset{\sim}{\gamma}_0$ be related to $\underset{\sim}{\sigma}_0$ in the same way as $\underset{\sim}{\gamma}$ is related to $\underset{\sim}{\sigma}$; $\tilde{\underset{\sim}{\gamma}}_0 = (\gamma_{01}, \gamma_0')'$, where γ_{0i}'s are related to $\underset{\sim}{\gamma}_0$ in the same way as γ_i's are related to $\underset{\sim}{\gamma}$; $\hat{\underset{\sim}{\varepsilon}}_i = \underset{\sim}{x}_i - \underset{\sim}{1}_{n(i)} \hat{\mu}(\gamma_0)$; $\bar{Q}_i = S_i + [X_i - \hat{\mu}(\gamma_0)]^2$, Note that

$$\gamma_{0i} = n_i \alpha_i^{-1}/\alpha_0 \quad , \hat{\underset{\sim}{\varepsilon}}_i' \hat{\underset{\sim}{\varepsilon}}_i - n_i \bar{Q}_i \tag{2.4.5}$$

Straightforward calculations using (2.4.2) - (2.4.5), give

$$Q_i = \alpha_i^{-1} (B\underset{\sim}{y})' H_i B\underset{\sim}{y} = \alpha_i^{-1} \hat{\underset{\sim}{\varepsilon}}_i \hat{\underset{\sim}{\varepsilon}}_i = \alpha_0 \gamma_{0i} \bar{Q}_i ;$$

$$S_{ij} = \operatorname{tr} BH_i BH_j = n_i - 2\gamma_{0i} + \gamma_{0i}^2 \quad \text{if } i = j$$

$$= \gamma_{0i} \gamma_{0j} \quad \text{if } i \neq j$$

Hence,

$$Q = \alpha_0 \tilde{\underset{\sim}{\gamma}}_0^{\delta} \bar{\underset{\sim}{Q}} \tag{2.4.6}$$

$S = D + \tilde{\underset{\sim}{\gamma}}_0 \tilde{\underset{\sim}{\gamma}}_0'$ where $D = \underset{\sim}{n}^{\delta} - 2\tilde{\underset{\sim}{\gamma}}_0^{\delta}$; $\underset{\sim}{n} = (n_1,\ldots,n_k)$. Then

$$S^{-1} = D^{-1} - \alpha_*^{-1} D^{-1} \tilde{\underset{\sim}{\gamma}}_0 \tilde{\underset{\sim}{\gamma}}_0' D^{-1} \quad \text{where } \alpha_* = 1 + \tilde{\underset{\sim}{\gamma}}_0' D^{-1} \tilde{\underset{\sim}{\gamma}}_0 \tag{2.4.7}$$

Using (2.4.6) and (2.4.7), (2.4.1) gives

$$\hat{\underset{\sim}{\theta}} = \alpha_0 \underset{\sim}{\sigma}_0^{\delta} S^{-1} \tilde{\underset{\sim}{\gamma}}^{\delta} \bar{\underset{\sim}{Q}} = \underset{\sim}{n}^{\delta} \tilde{\underset{\sim}{\gamma}}_0^{-\delta} S^{-1} \tilde{\underset{\sim}{\gamma}}_0^{\delta} \bar{\underset{\sim}{Q}} = \underset{\sim}{n}^{\delta} D^{-1} (\bar{\underset{\sim}{Q}} - Q_* \underset{\sim}{1}_k) \tag{2.4.8}$$

where $Q_* = \alpha_*^{-1} \underset{\sim}{1}_k' \tilde{\underset{\sim}{\gamma}}_0^{\delta} D^{-1} \tilde{\underset{\sim}{\gamma}}_0^{\delta} \bar{\underset{\sim}{Q}}$ Then $\hat{\underset{\sim}{\theta}}^{\delta} = \underset{\sim}{n}^{\delta} D^{-1} (\bar{\underset{\sim}{Q}}^{\delta} - Q_* I)$. Hence it is easy to see that

$$\tilde{\phi}^{\delta} = D(\tilde{Q}^{\delta} - Q_* I)^{-1}/\text{tr } D(\tilde{Q}^{\delta} - Q_* I)^{-1} \quad (2.4.9)$$

where $\tilde{\phi} = (\phi_1, \phi')'$, from which ϕ may be obtained simply by dropping the first component.

J.N.K. Rao and Subrahamaniam (1971) obtained the MINQUE estimator proposed in Rao (1970) where $\sigma_o \; \alpha \; 1_k$ according to the more general formulation of the MINQUE theory in Rao (1971) which gives the estimator obtained here. It can be seen that for $\tilde{\gamma}_o = n^{\delta}/n$, which is equivalent to $\sigma_o \alpha \; 1_k$, our estimator of σ given by (2.4.8) agrees with that obtained by these authors, as is to be expected.

2.5 Estimators better than the first sample mean

We shall be concerned only with unbiased estimators and judge the merit of an estimator by its variance. Brown and Cohen (1974) proposed a class of estimators which are better than X_1 for all σ. We offer a more general class and prove

Theorem 2.5.1 Assume that $m_i \geq 5$ for every $i = 2,\ldots,k$. Let a_i, c_i, $i = 2,\ldots,k$ be arbitrary sequences of positive numbers such that

$$a_i < \text{Min}[1, 2\, c_i(m_i-4)/(m_i+2)] \quad (2.5.1)$$

Let

$$\begin{aligned} b_i &= a_2 \quad \text{if } i = 2 \\ &= a_i(1 - \sum_{j=2}^{i-1} b_j) \quad \text{if } i > 2 \end{aligned} \quad (2.5.2)$$

and,

$$\begin{aligned} \hat{\mu}_r &= X_1 \quad \text{if } r = 1 \\ &= X_1 + \sum_{i=2}^{r} \phi_i(X_i - X_1) \quad \text{if } r > 1 \end{aligned}$$

where

$$\phi_i = b_i S_1/[S_1 + c_i S_i]$$

Then

(i) $\hat{\mu}_r$ is unbiased for μ

(ii) $\hat{\mu}_k$ is better than $\hat{\mu}_r$, for every $r < k$. In particular $\hat{\mu}_k$ is better than X_1.

<u>Proof</u> (i) It is clear that $\hat{\mu}_r$ is equivariant and hence is unbiased in view of Theorem 2.2.2.

(ii) To prove (ii) if suffices to consider only $k \geq 3$, since for $k = 2$, it follows from theorem 3.4.2 which we shall prove in next Chapter. Assume therefore that $k \geq 3$. We have

$$V(\hat{\mu}_k) = E[(1 - \sum_{i=2}^{k} \phi_i)^2 \bar{\sigma}_1^2 + \sum_{i=2}^{k} \phi_i^2 \bar{\sigma}_i^{2}] .$$

Hence,

$$V(\hat{\mu}_k) - V(\hat{\mu}_{k-1}) = \bar{\sigma}_1^2 \; E[(1+\eta_k)\phi_k^2 - 2(1 - \sum_{i=2}^{k-1} \phi_i) \; \phi_k]. \quad (2.5.3)$$

It is easy to see (by induction) that

$$1 - \sum_{i=2}^{r} b_i = \prod_{i=2}^{r} (1-a_i), \; r = 2,\ldots,k \quad (2.5.4)$$

Note that from (2.5.1)

$$0 < a_i < 1, \; i = 2,\ldots,k \quad (2.5.5)$$

(2.5.4) and (2.5.5) imply

$$0 < \sum_{i=2}^{r} b_i < 1, \quad r = 2,\ldots,k \quad (2.5.6)$$

Then (2.5.2), (2.5.5) and (2.5.6) imply

$$0 < b_i < 1, \quad r = 2,\ldots,k \tag{2.5.7}$$

Note also that ϕ_i is non-negative in view of (2.5.7). Hence, from (2.5.3) we have

$$V(\hat{\mu}_k) - V(\hat{\mu}_{k-1}) \leq \bar{\sigma}_1^2 E[(1+\eta_k)\phi_k^2 - 2(1-\sum_{i=2}^{k-1} b_i)\,\phi_k\,] \tag{2.5.8}$$

In view of (2.5.2)

$$\phi_k = a_k(1-\sum_{i=2}^{k-1} b_i)\Psi_k$$

where

$$\Psi_k \quad S_1/[S_1 + c_i\,S_i]$$

Hence, (2.5.8) can be written as

$$V(\hat{\mu}_k)-V(\hat{\mu}_{k-1}) \leq (1-\sum_{i=2}^{k-1} b_i)^2\,\bar{\sigma}_1^2\,E[(1+\eta_k)\,a_k^2\,\Psi_k^2 - 2\,a_k\,\Psi_k] \tag{2.5.9}$$

In view of the formula (3.3.1) (which we shall prove in the next chapter),

$$E[(1+\eta_k)\,a_k^2\,\Psi_k^2 - 2\,a_k\,\Psi_k] \qquad V(\ell) - V(X_1) \tag{2.5.10}$$

where

$$\ell = X_1 + a_k\,\Psi_k\,(X_k - X_1)$$

The estimator ℓ is better than X_1 in view of theorem 3.4.2 and the condition (2.5.1) satisfied by a_k. Hence

$$V(\ell) - V(X_1) \leq 0 \tag{2.5.11}$$

and the desired result follows from (2.5.9), (2.5.10) and (2.5.11).

2.6 Estimators Better than Each Sample Mean

We shall now consider another class of estimators which is perhaps more useful and important in the sense that members of this class would be better than each X_i under suitable conditions.

Let $L = \{1,2,\ldots, k\}$ and let N be any non empty subset of L. Consider

$$\hat{\mu}_N = \sum_{i \in N} \phi_i(N) X_i \tag{2.6.1}$$

where

$$\phi_i(N) = c_i\, S_i^{-1} \Big/ \sum_{j \in N} c_j\, S_j^{-1} \tag{2.6.2}$$

and $c_1, c_2, \ldots, c_k$ are positive constants to be suitably chosen. We prove

Theorem 2.6.1 Let $L = \{1,2,\ldots,k\}$, $L' = \{1,2,\ldots,k-1\}$. Let N be any non-empty subset of L and let $\hat{\mu}_N$ be as defined by (2.6.1). Assume that $m_k \geq 5$. Then

(i) $\hat{\mu}_N$ *is unbiased for* μ

(ii) $\hat{\mu}_L$ is better than $\hat{\mu}_{L'}$ iff

$$c_k/c_i \leq 2(m_k-4)/(m_i+2) \quad \text{for every} \quad i \in L' \tag{2.6.3}$$

Proof (i) The proof is similar to that for part (i) of Theorem 2.5.1 .

(ii) Note that $V(\hat{\mu}_N) = ET(N)$, where

$$T(N) = \sum_{i \in N} \sigma_i^{-2}\, \phi_i^2(N) \tag{2.6.4}$$

Hence, $V(\hat{\mu}_L) = ET$, $V(\hat{\mu}_{L'}) = ET'$

where

$$T = \sum_{=}^{k} \sigma_i^{-2}\, \phi_i^2\,, \quad T' = \sum_{=}^{k-1} \sigma_i^{-2}\, \phi_i^{2'}$$

and ϕ_i, ϕ_i' stand for $\phi_i(L)$ and $\phi_i(L')$ respectively. Hence we have to prove that

$$E(T-T') \leq 0 \quad \text{for all } \underset{\sim}{\sigma} \tag{2.6.5}$$

Let $w_i = c_i S_i^{-1}$, $w = \sum_{i=1}^{k} w_i$, $w' = \sum_{i=1}^{k-1} w_i$. Then, from (2.6.2)

$$w_i = \phi_i w = \phi_i' w' = \phi_i'(w-w_k), \; i = 1,2,\ldots,k-1 \tag{2.6.6}$$

Dividing (2.6.6) by w we have

$$\phi_i = \phi_i'(1-\phi_k), \quad i = 1,2,\ldots,(k-1) \tag{2.6.7}$$

Squaring both sides of (2.6.7), then multiplying both sides by $\bar{\sigma}_i^2$ and finally adding the results, we have

$$T - \phi_k^2 \; \bar{\sigma}_k^2 = T' \; (1-2 \; \phi_k + \phi_k^2) \tag{2.6.8}$$

From (2.6.8) it is easy to see that,

$$T - T' = \phi_k^2 \; (\bar{\sigma}_k^2 + T') - 2 \; \phi_k \; T' \tag{2.6.9}$$

Let $\beta = 1/ \sum_{i=1}^{k-1} \bar{\sigma}_i^{-2}$ and note that

$$\beta \leq T' \quad \text{for all } \underset{\sim}{\sigma} \tag{2.6.10}$$

Note also that $\phi_k \leq 1$ and hence the right hand side of (2.6.9) is non-increasing in T. Hence

$$T - T' \leq \phi_k^2 \; (\bar{\sigma}_k^{-2} + \beta) - 2 \; \phi_k \beta \tag{2.6.11}$$

Hence (2.6.5) holds if

$$E\phi_k^2(\bar{\sigma}_k^{-2} + \beta) \leq E2\beta\phi_k \quad \text{for all } \underset{\sim}{\sigma} \tag{2.6.12}$$

Let $f = (1 + \bar{\sigma}_k^{-2}/\beta)\phi_k$. Then (2.6.12) is equivalent to

$$2\, Ef/Ef^2 \geq 1 \quad \text{for all } \underset{\sim}{\sigma} \tag{2.6.13}$$

It is easy to see that f can be written in the form $f = 1/[p \sum_{i=1}^{k-1} (q_i d_i^{-1} z_i/z_k) + 1 - p]$ where $z_i = \bar{\sigma}_i^{-2}/S_i$, $p = \bar{\sigma}_k^2/(\bar{\sigma}_k^2 + \beta)$; $q_i = \beta/\bar{\sigma}_i^{-2}$, $d_i = c_k/c_i$

Note that $z_i^{-1} \sim \chi^2_{m_i}$ and hence in view of theorem A.2 given in the appendix

$$\inf_p Ef/Ef^2 = \min (1, cE\Psi/E\Psi^2)$$

where $c = Ez_k/Ez_k^2 = m_k - 4$; $\Psi = 1/\sum_{i=1}^{k-1} (q_i\, d_i^{-1}\, z_i)$. Hence (2.6.13) holds if

$$2c\; E\Psi/E\Psi^2 \geq 1, \text{ for every } \underset{\sim}{q} \tag{2.6.14}$$

where $\underset{\sim}{q} = (q_1, \ldots, q_{k-1})$. Note that $q_i \geq 0$ for every i and $\sum_{i=1}^{k-1} q_i = 1$.

Hence using theorem A.2 once again, we have

$$\inf_{\underset{\sim}{q}} E\Psi/E\Psi^2 = \min_{1 \leq i \leq k-1} (a_i d_i^{-1})$$

where

$$a_i = Ez_i^{-1}/Ez_i^{-2} = 1/(m_i + 2).$$

Hence (2.6.14) holds if

$$2ca_i d_i^{-1} \geq 1 \quad \text{for every} \quad i \in L' \tag{2.6.15}$$

(2.6.15) is equivalent to (2.6.?) and thus we have proved the sufficiency of (2.6.3). To prove the necessity observe that equality holds in (2.6.10) almost sure for all $\underset{\sim}{\sigma} \in \Lambda_i$ where

$$\Lambda_i = \{\underset{\sim}{\sigma} | \beta = \bar{\sigma}_i^{-2} \quad \text{for some} \quad i \in L'\}$$

Hence equality holds a.s. in (2.6.11) for all $\underset{\sim}{\sigma} \in \Lambda_1$ and this implies that (2.6.5) holds only if

$$2\,Ef/Ef^2 \geq 1 \quad \text{for all} \quad \underset{\sim}{\sigma} \in \Lambda_1 \tag{2.6.16}$$

It is clear from our previous analysis that inf Ef/Ef^2 is either 1 or the value of Ef/Ef^2 at some point of Λ_1 once it is noted that $\underset{\sim}{\sigma} \in \Lambda_1 \iff q_i$ equals 1 for some i and 0 for all others. Hence (2.6.16) $\iff$ (2.6.13) and the proof is complete.

Theorem 2.6 is an improvement of a similar result in Shinozaki (1978) [see also Bhattacharya (1979)] obtained under the condition:

$$(m_k+2)/[2(m_i-4)] \leq c_k/c_i \leq 2(m_k-4)/(m_i+2) \tag{2.6.17}$$

which is more stringent than (2.6.3). Our condition (2.6.3) being both necessary and sufficient leaves no scope of further improvement.

<u>Remark 2.6.1</u> For k = 2 theorem 2.6.1 reduces to the result in theorem 3.4.1 of the next chapter.

<u>Remark 2.6.2</u> In view of remark 2.6.1 we see that for a given i, (2.6.3) is necessary and sufficient for $\hat{\mu}_{\{i,k\}}$ to be uniformly better than $\hat{\mu}_{\{i\}}$. Hence we have the following alternative statement of theorem 2.6.1

$\hat{\mu}_L$ is uniformly better than $\hat{\mu}_{L'}$ iff $\hat{\mu}_{\{i,k\}}$ is uniformly better than $\hat{\mu}_{\{i\}}$ for every $i \in L'$.

The following corollary is a simple consequence of theorem 2.6.1.

<u>Corollary 2.6.1</u> Suppose the elements of L can be arranged in the form $\{i_1, i_2, \ldots, i_k\}$ such that all elements of M appear before those of $L-M$; The condition

$$c_j/c_i \leq 2(m_j-4)/m_i+2) \tag{2.6.18}$$

hold for every (i,j) such that $j \in L-N$ and $i = i_s$, $j=i_t$ for some $s < t$. Then $\hat{\mu}_L$ is better than $\hat{\mu}_N$. In particular $\hat{\mu}_L$ is better than X_r if (2.6.18) holds for every (i,j) such that $i = i_s$ $j = i_t$ for some $s < t$

We shall now obtain a necessary condition for $\hat{\mu}_L$ to be uniformly better than $\hat{\mu}_N$. We first prove

Lemma 2.6.1 (i) $\hat{\mu}_L$ is better than $\hat{\mu}_N \Rightarrow$ (ii) $\hat{\mu}_M$ is better than $\hat{\mu}_N$ for every M such that $N \subsetneq M \subseteq L$.

Proof We have $T(L) = T(M)$ a.s. for $\underset{\sim}{\sigma} \in \Lambda_2$ where $\Lambda_2 = \{\underset{\sim}{\sigma} | \sigma_i^{-2} = \infty$ for all $i \in (L-M)\}$; and $T(N)$ is defined by (2.6.4). Hence, $V(\hat{\mu}_L) = V(\hat{\mu}_M)$ for every $\underset{\sim}{\sigma} \in \Lambda_2$. The desired result follows from this since $V(\hat{\mu}_M)$ does not depend on σ_i^2 if $i \notin M$

The following theorem is a simple consequence of Lemma 2.6.1 and theorem 2.6.1.

Theorem 2.6.2 $\hat{\mu}_L$ is better than $\hat{\mu}_N$ only if the condition (2.6.18) holds for every pair (i,j) where $i \in N$ and $j \in L-N$. In particular, $\hat{\mu}_L$ is better than X_i only if (2.6.18) holds for every pair (i,j) where $j \neq i$.

Proof Suppose $\hat{\mu}_L$ is better than $\hat{\mu}_N$. Then by lemma 2.6.1 $\hat{\mu}_{N \cup \{j\}}$ is better than $\hat{\mu}_N$ for every $j \in L-N$ and the stated condition must hold by Theorem 2.6.1 .

Combining corollary 2.6.1 and Theorem 2.6.2 we can arrive at the following important result due to Shinozaki (1978) in a direct and more elegant manner.

Theorem 2.6.3 $\hat{\mu}_L$ is better than each X_r iff (2.6.18) holds for every $i \neq j$. In fact if the stated condition holds then $\hat{\mu}_M$ is better than $\hat{\mu}_N$ for every (M,N) such that $N \subsetneq M \subseteq L$.

Remark 2.6.3 The necessity part of theorem 2.6.3 could also be proved using arguments similar to that in Graybill and Deal (1959), as suggested in Shinozaki (1978). But our approach through Lemma 2.6.1 is simpler.

CHAPTER 3

ESTIMATION OF THE COMMON MEAN OF TWO NORMAL POPULATIONS

3.1 Introduction

In the previous chapter we considered the general case of estimating the common mean of several normal distributions. In this chapter we consider the special case of estimating the common mean of two normal distributions and obtain some additional results. We shall follow the same notation as in the previous chapter, unless otherwise stated. The vectors $\underset{\sim}{\rho}, \underset{\sim}{\eta}, \underset{\sim}{\gamma}, \underset{\sim}{\phi}$ which in the present case have only one component will be written as scalars: ρ, η, γ, ϕ. Subscripts to the later symbols would signify specific choices of the vectors which they represent.

Among the various estimators addressed to this special case, those based on well defined principles are: (i) Maximum likelihood estimator [Yates (1939a)] which is also the Bayes estimator with respect to the improper prior proportional to $\sigma_1^{-2}\ \sigma_2^{-2}$ [Box and Tiao (1973)]; (ii) Bayes and fiducial equivariant estimators [Zacks (1970)]; (iii) Minimax estimators [Cohen and Sackrowitz (1974)]; (iv) the so called uniformly better estimators [Graybill and Deal (1959), Brown and Cohen (1974), Cohen and Sackrowitz (1974), Khatri and Shah (1974), Bhattacharya (1980)]. The work of Graybill and Deal (1959) is addressed to the special case of $\bar{\mu}$ (defined in section 2.1). Zacks (1966) proposed modifications of this depending on a preliminary test of hypothesis concerning the unknown value of ρ. Similar but more flexible estimators have been studied by Gurland and Mehta (1969).

In section 2 we present a direct and simpler derivation of the estimator obtained in section 2.4, for the special case considered here.

In section 3, we present some useful results for comparing estimators belonging to the class $\{\hat{\mu}(\phi), \phi \in \Phi\}$. In section 4 we offer a class of estimators better than X_1. We unify the two similar classes in Brown and Cohen (1974) and Khatri and Shah (1974) and improve the Brown-Cohen results.

Section 5 is devoted to some studies concerning estimators better than both X_1 and X_2. We discuss and clarify some misunderstanding in the literature concerning the pioneering work by Graybill and Deal (1959) on this topic. We also remove the restriction: $n_1 = n_2$ in Cohen and Sackrowitz (1974) and extend their results.

3.2 Estimator of μ based on LBQUE of $(\bar{\sigma}_1\ \bar{\sigma}_2^2)$ [with invariance]

The estimation procedure which we consider in this section is same as in section 2.4 where the more general case of k-samples, was treated using the MINQUE theory. However, for the special case considered here we offer a direct and simpler derivation of the resulting estimator. An orthogonal basis for the residuals of the model (2.4.1) in the present case is given by:

(i) m_1 orthogonal contrasts within the first sample;

(ii) m_2 orthogonal contrasts within the second sample;

(iii) The difference between the two sample means.

We shall consider unbiased estimators of $(\bar{\sigma}_1^2, \bar{\sigma}_2^2)$ which are quadratic forms in these residuals [see Rao (1970) for justification]. Note that for any product term of the quadratic form its expectation as well as covariance with a square term is zero. Hence such terms do not contribute anything to the expectation but has a positive contribution to the variance of the quadratic form. Therefore we need to confine ourselves only to quadratic forms of the diagonal type, which can be written in the form:

$$Q^* = a S_1 + b S_2 + cW \quad \text{where } W = (X_2 - X_1)^2$$

and we give equal weight to the squared residuals belonging to the same group in consideration of the fact that these have the same expectations and same variances. We have

$$EQ^* = a\, m_1 \bar{\sigma}_1^2 + b\, m_2 \bar{\sigma}_2^2 + c(\bar{\sigma}_1^2 + \bar{\sigma}_2^2) \qquad (3.2.1)$$

$$V(Q^*) = \bar{\sigma}_1^4 [(a^2 m_1 + c^2) + 2c^2\eta + (b^2 m_2 + c^2)\eta^2] \qquad (3.2.2)$$

Let $\underset{\sim}{\sigma}_0 = (\alpha_1, \alpha_2)$ and η_0, γ_0 be related to $\underset{\sim}{\sigma}_0$ in the same way as η, γ are related to $\underset{\sim}{\sigma}$. The required estimator $(\hat{\sigma}_1^2\ \hat{\sigma}_2^2)$ of $(\bar{\sigma}_1^2, \bar{\sigma}_2^2)$ can be obtained by minimizing (3.2.2) for the specified value of $\underset{\sim}{\sigma} = \underset{\sim}{\sigma}_0$, subject to the condition of unbiasedness which in view of (3.2.1) is given by

$$a\, m_1 + c = 1; \quad b\, m_2 + c = 0$$

in case of $\bar{\sigma}_1^2$ and $a\, m_1 + c = 0$; $b\, m_2 + c = 1$ is case of $\bar{\sigma}_2^2$. After straight-forward calculations which are omitted, we get,

$$\hat{\bar{\sigma}}_1^2 = \lambda[\{\eta_0^2 + m_2 (1+\eta_0)^2\}S_1 - S_2 + m_2 W]$$

$$\hat{\bar{\sigma}}_2^2 = \lambda[-\eta_0^2 S_1 + \{1+m_1(1+\eta_0)^2\} S_2 + m_1 \eta_0^2 W]$$

where $\lambda = 1/[m_2 + m_1\eta_0^2 + m_1 m_2(1+\eta_0)^2]$. Hence the required ϕ (which is related to $\hat{\bar{\underset{\sim}{\sigma}}}$ in the same way as γ related to $\underset{\sim}{\sigma}$) is

$$\phi = \frac{[m_2 + (1-\gamma_0)^2]S_1 - \gamma_0^2 S_2 + m_2\gamma_0^2 W}{m_2 S_1 + m_1 S_2 + [m_2\gamma_0^2 + m_1(1-\gamma_0)^2]W}$$

It can be seen that the result obtained here is in greement with that in section 2.4.

part from the case $\gamma_0 = (m_1+1)/(m_1+m_2+2)$ when we get the usual MINQUE olution [see section 2.4) the cases of special interest are $\gamma_0 = 0$, $_0 = 1$, $\gamma_0 = \frac{1}{2}$, appropriate when η is believed to be very large, very small and in the vicinity of 1 respectively.

3.3 Some results for comparing two estimators

Let Φ be the class of all measurable functions of (S_1, S_2, W), where $W = (X_2 - X_1)^2$. We shall consider estimators belonging to the class $\{\hat{\mu}(\phi) \mid \phi \in \Phi\}$ and obtain some elementary but useful results for comparing two such estimators. Our criterion would be mean square error. We shall use the following result obtained independently by Brown and Cohen (1974), and, Khatri and Shah (1974).

<u>Theorem 3.3.1</u> Let $\phi \in \Phi$ and assume that $E\,\hat{\mu}(\phi)$ exists. Let W_* be such that W_*/EW is a chisquare variable with 3 degrees of freedom distributed independently of (S_1, S_2). Let ϕ_* be the expression obtained from ϕ by replacing W by W_*. Then

(i) $\hat{\mu}(\phi)$ is unbiased for μ

(ii) $V[\hat{\mu}(\phi)] = \bar{\sigma}_1^2 \, [1 + E(\phi_*^2/\gamma - 2\phi_*)]$ (3.3.1)

Since the estimators under consideration are unbiased we shall say that the estimator $\hat{\mu}(\phi_1)$ is better than $\hat{\mu}(\phi_2)$ for all $\rho \in \Omega$, where Ω is a given subset of the positive half of the real line, if

$$V[\hat{\mu}(\phi_1)] \leq V[\hat{\mu}(\phi_2)] \text{ for all } \rho \in \Omega$$

with strict inequality for at least one value of $\rho \in \Omega$. Then for comparing the estimators we have the following useful results which follow from (3.3.1) easily.

<u>Theorem 3.3.2</u> Let $\phi_1, \phi_2 \in \Phi$ and let ϕ_{1*}, ϕ_{2*} be related to ϕ_1, ϕ_2 respectively in the same way as ϕ_* is related to ϕ. Then $\hat{\mu}(\phi_1)$ is better than $\hat{\mu}(\phi_2)$ for all $\rho \in \Omega$ iff, for estimating γ

$$\text{M.S.E. } (\phi_{1*}) \leq \text{M.S.E. } (\phi_{2*})$$

or all $\rho \in \Omega$.

<u>Theorem 3.3.3</u> Let $\phi \in \Phi$. Then $\hat{\mu}(\phi)$ is better than X_1 for all $\rho \in \Omega$ iff

$$2\nu_\Omega(\phi) \geq 1$$

where

$$\nu_\Omega(\phi) = \inf_{\rho \in \Omega} E\bar{\phi}/E\bar{\phi}^2$$

$$\bar{\phi} = \phi_*/\gamma$$

From theorem 3.3.3 we have

<u>Corollary 3.3.1</u> If $\phi = a\Psi$, where a is a positive constant to be suitably chosen and $\Psi \in \Phi$; then $\hat{\mu}(\phi)$ is better than X_1 for all $\rho \in \Omega$ iff

$$a \leq 2\nu_\Omega(\Psi)$$

It should be noted that in the absence of any a - priori knowledge otherwise Ω will be taken to be $(0,\infty)$. For the sake of simplicity we shall denote $\nu_{(0,\infty)}(\Psi)$ by $\nu(\Psi)$. It is easy to see that,

$$\nu(\Psi) = \inf_{\gamma \in (0,1)} E\bar{\Psi}/E\bar{\Psi}^2$$

3.4 Estimators Better than the First Sample Mean

We shall now consider a class of unbiased estimators which have smaller variance than the first sample mean for all $\rho > 0$. Such estimators are important when the first sample has a special significance and one would not like to use the second sample unless its use leads to improvemunt over X_1 for all $\rho > 0$. Another point is that since it is known that the combined estimator is better than X_1, $V(X_1)$ in such cases serves as a lower bound for the variance of the combined estimator, actual value of which is, in general difficult to compute.

The first demonstration of an estimator with the desired property is due to Graybill and Deal (1959). In fact under appropriate conditions given by them (which are both necessary and sufficient) their estimator is better than both sample means, and thus has a more stringent property which we shall discuss more elaborately in the next section. An apparent defect in this estimator is that it does not utilize the information on variances contained in the difference between the sample means. Recently Brown and Cohen (1974) and Khatri and Shah (1974) have come up with estimators which utilize the difference between the sample means and possess the desired property under consideration in this section. Both Brown and Cohen (1974) and Khatri and Shah (1974), in fact consider a whole family of estimators depending on a single parameter and while the former required an upper bound on their parameter the latter required a lower bound. The two families of estimators mentioned above can be treated as particular cases of a two parameter family which we propose to study in order to unify the results obtained in these two papers. Our unified approach would guarantee the desired property by a single condition on the two parameters, which is equivalent to that obtained by Khatri and Shah (1974) for a subclass considered by them and is an improvement of the results in Brown and Cohen (1974) for the subclass considered by these authors. The upper bound set on their parameters by Brown and Cohen (1974) was somewhat crude and involved a complicated expression which called for a table given in their paper. The improved upper bound of the Brown - Cohen parameter, which we shall obtain, is a simple expression for which no table is required and is moreover the best possible, as we shall show. Consider the estimator $\hat{\mu}_1 = \hat{\mu}(\phi_1)$ where

$$\phi_1 = a\, S_1/[S_1 + d(S_2+W)]$$

and a, d are positive constants to be suitably chosen. Here $\phi_1 = a\Psi$

where $\Psi = S_1/[S_1 + d(S_2+W)]$. Let

$$V_1 = S_1/\sigma_1^2,\ V_2 = S_2/\sigma_2^2,\ V_3 = W_*/(\sigma_1^2+\sigma_2^2),\ V_4 = V_2+V_3,\ u = V_3/V_4$$

$$S_1 = \sigma_1^2 V_1, \; S_2 + W_* = \sigma_1^2 (\eta + u) V_4 - (\sigma_1^2/\gamma)[p(u) - q(u)\gamma)] V_4$$

where $p(u) = 1$, $q(u) = 1-u$. Hence $\bar{\Psi} = \Psi_*/\gamma$ can be written as

$$\bar{\Psi} \;=\; V_1[\gamma\, V_1 + dh(u,\gamma)V_4]^{-1} \qquad (3.4.1)$$

where $h(u,\gamma) = p(u) - q(u)\gamma$. Note that $V_1 \sim \chi^2_{m_1}$, $V_4 \sim \chi^2_{m_2+3}$, $u \sim \beta(\frac{3}{2}, \frac{m_2}{2})$ and that V_1, V_4, u are mutually independent.

Since V_1 is almost sure positive (3.4.1) is equivalent to

$$\bar{\Psi} \;=\; 1/[\gamma + dh(u,\gamma)V]$$

where $V = V_4/V_1$.

It can be seen that with the match-up $u \sim x$, $V \sim y$, $\bar{\Psi}$ matches-up with f of theorem A.4 and satisfies all conditions of part A of that theorem provided $EV^{-2} < \infty$, a condition which is equivalent to $m_2 \geq 2$. The support of u is $S = (0,1)$; $S_* = \{s | s \in S;\, q(s) > 0\} = \{0 < s < 1\}$. We have,

$a_0 = EV^{-1}/EV^{-2} = (m_2-1)/(m_1+2)$; $\delta_1 = \inf_{s \in S} p(s) = 1$

$$\delta_2 \;=\; \inf_{s \in S_*} p(s)/q(s) = 1;$$

$$\delta_3 \;=\; \inf_{s \in S_*} h(s;1) = 0$$

$$\delta_5 \;=\; \max(\delta_2,\, d\, a_0\, \delta_3) = 1;$$

$$\pi_1 \;=\; \min\,(da_0\delta_1, \delta_5) = \min(1, da_0).$$

Hence by theorem A.4

$$v(\Psi) \geq \min\,(1, da_0)$$

Note that $[E\bar{\Psi}/E\bar{\Psi}^2]_{\gamma=0} = da_0$

Hence $\inf_{\rho > \rho_0} E\bar{\Psi}/E\bar{\Psi}^2 = \inf_{\gamma < \gamma_0} E\bar{\Psi}/E\bar{\Psi}^2 \leq da_0$ for any $\rho_0 \geq 0$ where γ_0 is related to ρ_0 in the same way as γ is related to ρ. Hence using corollary 3.3.1. we have

<u>Theorem 3.4.1</u> Assume that $m_2 \geq 2$ and let $a_o = (m_2-1)/(m_1+2)$, then

(i) $\hat{\mu}_1$ is better than X_1 for all $\rho > 0$ if

$$a \leq 2 \min(1, da_o) \qquad (3.4.2)$$

(ii) $\hat{\mu}_1$ is better than X_1 for all $\rho > \rho_o$ (for some given $\rho_o \geq 0$) *only if*

$$a \leq 2\, da_o \qquad (3.4.3)$$

Note that if either $a \leq 2$ or $da_o \leq 1$ then (3.4.2) is equivalent to (3.4.3). Hence theorem 3.4.1 gives

<u>Corollary 3.4.1</u> Assume that $m_2 \geq 2$ and that either $a \leq 2$ or $da_o \leq 1$ then a sufficient condition for $\hat{\mu}_1$ to be better than X_1 for all $\rho > 0$ is given by (3.4.3). *Conversely the same condition is also necessary for* $\hat{\mu}_1$ to be better than X_1 for all $\rho > \rho_o$ for some given $\rho_o \geq 0$.

Consider, now the estimator $\hat{\mu}_2 = \hat{\mu}(\phi_2)$, where $\phi_2 = a\, S_1/(S_1 + d\, S_2)$. This can be treated in the same way as $\hat{\mu}_1$. Thus we have

<u>Theorem 3.4.2</u> Theorem 3.4.1 (Corollary 3.4.1) holds word by word for $\hat{\mu}_2$ provided the assumption $m_2 \geq 2$ in this theorem (Corollary) is replaced by $m_2 \geq 5$ and the expression for a_o is replaced by $a_o = (m_2-4)/(m_1+2)$.

<u>Remark 3.4.1</u> The analogue of part (i) of theorem 3.4.1 contained in theorem 3.4.2 above can be improved as follows: $\hat{\mu}_2$ is better than X_1 for all $\rho > 0$ iff $a \leq 2 \min(1, da_o)$. Note that in this case $[E\bar{\Psi}/E\bar{\Psi}^2]_{\gamma=1} = 1$ in addition to $[E\bar{\Psi}/E\bar{\Psi}^2]_{\gamma=0} = da_o$. *Since* $\nu(\Psi) \geq \min(1, da_o)$ *as in the case* of $\hat{\mu}_1$, this implies $\nu(\Psi) = \min(1, da_o)$. Hence the improved result. An alternative and perhaps more elegant approach in this case is given by theorem A.2. Observe that in this case $\bar{\Psi} = 1/[\gamma + (1-\gamma)\, dV_*]$, where $V_* = V_2/V_1$. Note that $\bar{\Psi}$ is of the same form as f of Theorem A.2. (with $k = 2$, and $x_1 = 1, x_2 = dV_*$) and satisfies all conditions of that theorem provided $EV_*^{-2} < \infty$, a condition which is equivalent to $m_2 \geq 5$. Hence, by that theorem $\nu(\Psi) = \min[1, da_o]$, if $m_2 \geq 5$

In the discussion which follows we shall write $\hat{\mu}_1$ and $\hat{\mu}_2$ in the more elaborate form $\hat{\mu}_1(a,d)$ and $\hat{\mu}_2(a,d)$ respectively, to reflect their dependence on the constants a, d. It can be seen that these two estimators include as particular cases the estimators T_a, $T_a(1)$ of Brown and Cohen (1974) and μ^*, μ^{**} of Khatri and Shah (1974). In fact $T_a = \hat{\mu}_1[a,(m_1-1)/(m_2+2)]$; $T_a(1) = \hat{\mu}_2[a,(m_1-1)/(m_2-1)]$; $\hat{\mu}^* = \hat{\mu}_1(1,d)$; $\mu^{**} = \hat{\mu}_2(1,d)$.
We recall that the values of a_o for $\hat{\mu}_1$ and $\hat{\mu}_2$ are $(m_2-1)/(m_1+2)$ and $(m_2-4)/(m_1+2)$ respectively. Hence the values of da_o for T_a and $T_a(1)$ are $(m_1-1)(m_2-1)/[(m_1+2)(m_2+2)]$ and $(m_1-1)(m_2-4)/[(m_1+2)(m_2-1)]$, respectively. It can be seen that in both cases $da_o < 1$. Hence the following two corollaries follow from our corollary 3.4.1 and theorem 3.4.2 respectively.

<u>Corollary 3.4.3</u> T_a is better than X_1 for all $\rho > 0$ iff $a \leq 2(m_1-1)(m_2-1)/[(m_1+2)(m_2+2)]$.

<u>Corollary 3.4.4</u> $T_a(1)$ is better than X_1 for all $\rho > 0$ iff $a \leq 2(m_1-1)(m_2-4)/[(m_1+2)(m_2-1)]$.
These two corollaries are readily seen to be improvements of theorems 2.1 and 2.2, respectively of Brown and Cohen (1974).

For the Khatri-Shah estimators we have a = 1 which satisfies the condition $a \leq 2$. Hence our corollary 3.4.1 and theorem 3.4.2 give

<u>Corollary 3.4.5</u> $\hat{\mu}^*$ is better than X_1 for all $\rho > 0$ iff $d \geq (\frac{1}{2})(m_1+2)/(m_2-1)$

<u>Corollary 3.4.6</u> μ^{**} is better than X_1, for all $\rho > 0$ iff $d \geq (\frac{1}{2})(m_1+2)/(m_2-4)$

These results are same as in Khatri and Shah (1974), who used a completely different method of proof.

Cohen and Sackrowitz (1974) obtained another estimator which is better than X_1 for all $\rho > 0$. They assumed $m_1 = m_2$. We shall remove this restriction and derive the results using the general approach given by our corrollaries.

To begin with we refer to Olkin and Pratt (1958) where it is shown (except for the differences in context and notation) that the unique unbiased estimator of $\rho_* = (\bar{\sigma}_1^2 - \bar{\sigma}_2^2)/[(\bar{\sigma}_1^2 + (p-1)\ \bar{\sigma}_2^2]$ based on (S_1,S_2) is given by

$$H(S_1,S_2,m_1,m_2)=1-\frac{p}{p-1}\ {}_2F_1(1,1-m_2/2;\ m_1/2;z),\ \text{for}\ \ 0 \le z \le 1$$

$$= \frac{p}{p-1}\ {}_2F_1(1,1-m_1/2;m_2/2;\ 1/z),\ \text{for}\ z \ge 1$$

where $z = S_1/[(p-1)S_2]$ and the function ${}_2F_1$ is the well known hypergeometric function. Note that for $p = 2$, we have $(1+\rho_*)/2 = 1/(1+\eta) = \gamma$ and hence the unique unbiased estimator of γ based on $(S_1\ S_2)$ is given by,

$$G = G(S_1,S_2\ ;m_1,m_2) = 1 - {}_2F_1\ (1,1-m_2/2;m_1/2;S_1/S_2)\quad \text{if}\quad S_1 \le S_2$$

$$= {}_2F_1(1,1-m_1/2;m_2/2;\ S_2/S_1)\quad \text{if}\quad S_1 \ge S_2 \qquad (3.4.4)$$

As is to be expected, the above expression for G agrees with that [denoted by G(z)] in Cohen and Sackrowitz (1974) when $m_1 = m_2$, the case considered by these authors [Note that they write z for our $\Sigma(x_{2i}-X_2)^2/\Sigma(x_{1i}-X_1)^2$ which is same as S_2/S_1 in the particular case $m_1 = m_2$].
To see this use the well known formula [Lebedev (1972) p. 243, formula (9.2.15]

$${}_2F_1(\alpha-1\ ,\ \beta+1;\gamma;z) - {}_2F_1(\alpha,\beta;\gamma;z) = \frac{(\alpha-\beta-1)}{\gamma}\ {}_2F_1(\alpha,\beta+1;\ \gamma+1;z)$$

Consider the estimator, $\hat{\mu}_3 = \hat{\mu}(\phi_3)$ where $\phi_3 = a\ G$ and G defined in (3.4.4), Here $\phi_3 = a\Psi$ where $\Psi = G$, Note that Ψ is independent of W and hence $\Psi_* = \Psi$; furthermore Ψ is unbiased for γ. Hence,

$$E\bar{\Psi} = E(\Psi/\gamma) = 1$$

Hence $E\bar{\Psi}/E\bar{\Psi}^2 = 1/E\bar{\Psi}^2 = 1/E\bar{G}^2$ where $\bar{G} = G/\gamma$ and hence

$$\nu(\Psi) = \inf_{\gamma}\ (1/E\bar{G}^2) = 1/\sup_{\gamma} E\bar{G}^2$$

Hence, in view of theorem 3.3.2 we have

Theorem 3.4.3 The estimator $\hat{\Omega}$ is better than X_1 for all $\rho > 0$ iff $a \leq A(m_1,m_2)$ where

$$A(m_1, m_2) = 2/\sup_{\gamma} G^2 \qquad (3.4.5)$$

It is not easy to evaluate $A(m_1;m_2)$ but a non trivial lower bound of it can be obtained in the following way.

Let $m = \min(m_1,m_2)$. Split S_1 into m_1 components u_i, $i = 1,2,\ldots,m_1$ and S_2 into m_2 components v_i, $i = 1,2,\ldots,m_2$ so that we have $u_i/\bar{\sigma}_1^2, i=1,2,3,\ldots,m_1$ $v_i/\bar{\sigma}_2^2$, $i = 1,2,\ldots,m_2$, identically and independently distributed chi-square variables with 1 d.f. each.

Now consider the problem of estimating $\bar{\sigma}_1^2, \bar{\sigma}_2^2$ on the basis of u_i's and v_i's and observe that (S_1,S_2) is then a sufficient statistic for $\bar{\sigma}_1^2$, $\bar{\sigma}_2^2$.

It is easy to see that for every integer r ranging from 3 to m-1

$$G_r = [(2r-2)/(m-r)] \sum_{i=r+1}^{m} u_i^2 / \sum_{i=1}^{r} (u_i+v_i)^2$$

is unbiased for γ, and we have

$$E\, G_r^2 = \gamma^2 c(r) \qquad (3.4.6)$$

where

$$c(r) = (r-1)(m-r+2)/[(m-r)(r-2)] \qquad (3.4.7)$$

Since G is the *unique unbiased estimator of* γ *based on* (S_1,S_2), we have $E(G_r|S_1,S_2) = G$ *for* $r = 3,4,\ldots,m-1$. *Hence by Rao – Blackwell theorem,*

$$EG^2 \leq EG_r^2 \quad \text{for} \quad r = 3,\ldots,m-1 \qquad (3.4.8)$$

Dividing both sides of (3.4.8) by γ^2 and using (3.4.6) we have,

$$EG^2 \leq \min_{3 \leq r \leq m-1} c(r) \qquad (3.4.9)$$

Now (3.4.7) gives

$$c'(r) = [-1/\{(r-2)(r-1)\} + 2/\{(m-r)(m-r+2)\}]c(r)$$

where the prime stands for derivation with respect to r. From this, it can be seen that c(r) is a convex function of r for $r \in [3,m-1]$ and hence, $\text{Min}\{c(3),\ldots,c(m-1)\} = c_*(m)$, where,

$$c_*(m) = \text{Min}\{c[([\alpha]), c([\alpha]+1)\} \qquad (3.4.10)$$

; $[\alpha]$ = the integral part of α; and α is the positive root of the equation

$$\alpha^2+2(m-2)\alpha-m(m+2) + 4 = 0.$$

In conjunction with (3.4.5) and (3.4.9) this implies $A(m_1,m_2) \geq 2/c_*(m)$. In view of theorem 2.4.3 we have thus proved.

<u>Theorem 3.4.4</u> Assume that $m = \text{Min}(m_1,m_2) \geq 4$ and let $c_*(m)$ be as defined in (3.4.10). Then $\hat{\mu}_3$ is better than X_1 for all $\rho > 0$ if $a \leq 2/c_*(m)$.

The above theorem reduces to theorem 2.1 of Cohen and Sackrowitz (1974) if we replace $2/c_*(m)$ by $A_*(m_1,m_2)$ defined below :

$$\begin{aligned} A_*(m_1,m_2) &= 2/c[(m+1)/2] = 2(m-3)/(m+3) \text{ if } m \text{ is odd} \\ &= 2/c[(m+2)/2] = 2(m-2)^2/[m(m+2)] \text{ if } m \text{ is even} \end{aligned} \qquad (3.4.11)$$

and consider the special case $m_1 = m_2$. Cohen and Sackrowitz (1974) claims that for integral values of r ranging from 3 to m-1, c(r) is minimum at $r = (m+1)/2$ if m is odd and $r = (m+2)/2$ if m is even [see section 2 of their paper; note that they write n for our m+1 and, instead of c(r) consider an expression which is an increasing function of c(r)]. We find that their claim is true iff $m \leq 15$. As such our theorem 3.4.4 is not only a generalization but also an improvement of their theorem 2.1.

3.5 Estimators Better than Both Sample Means

A good combined estimator should have the property that it is better than both X_1 and X_2 for all possible values of ρ. If we have a priori knowledge that $\eta \geq 1$, estimators which were shown to be better than X_1 are automatically better than both X_1 and X_2. By interchanging the role of the two samples in the combined estimator, a similar result can be obtained for the case $\eta \leq 1$. When nothing is known about ρ it is natural to look for a combined estimator which is better than both X_1 and X_2 for all $\rho > 0$.

A general procedure for constructing such estimators is given by the following lemma.

<u>Lemma 3.5.1</u> Let $\phi = \phi(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ and $\phi^* = \phi^*(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ satisfy the condition

$$\phi(\underset{\sim}{x}_1, \underset{\sim}{x}_2) + \phi^*(\underset{\sim}{x}_2, \underset{\sim}{x}_1) = 1 \qquad (3.5.1)$$

Let F be a symmetric subset of the set of positive real numbers in the sense that $\rho \in F \Rightarrow 1/\rho \in F$. Let E denote the set of all ordered pairs of natural numbers such that $\hat{\mu}(\phi)$ is better than X_1 for all $\rho \in F$ if (only if) $(n_1, n_2) \in E$. Let E* be defined similarly in relation to ϕ^*. Then $\hat{\mu}(\phi)$ is better than both X_1 and X_2 for all $\rho \in F$ if (only if) $(n_1, n_2) \in E_0$ where $E_0 = E \cap PE^*$ and PE* stands for the set obtained from E* by permuting the co-crdinates of each pair in E* in the reverse order. If further $\phi = \phi^*$ we have

$$E_0 = E \cap PE$$

<u>Proof</u> Let $\tilde{\mu}(\underset{\sim}{x}_1, \underset{\sim}{x}_2) = \hat{\mu}(\phi)$ and $\tilde{\mu}^*(\underset{\sim}{x}_1, x_2) = \hat{\mu}(\phi^*)$. It is easy to verify that the condition (3.5.1) is equivalent to

$$\tilde{\mu}(\underset{\sim}{x}_1, x_2) = \tilde{\mu}^*(\underset{\sim}{x}_2, \underset{\sim}{x}_1)$$

Since F is symmetric $\tilde{\mu}^*(\underset{\sim}{x}_2, \underset{\sim}{x}_1)$ and hence $\tilde{\mu}(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ is better than X_2 for all $\rho \in F$ if (only if) $(n_1, n_2) \in PE^*$. Hence the result is obvious.

Lemma 3.5.1 may be applied to the estimators $\hat{\mu}_2$ and $\hat{\mu}_3$ of the previous section to construct estimators with the desired property. For this we require $a = 1$ in both cases. Then $\hat{\mu}_2$ reduces to μ^{**} of Khatri and Shah (1974), who have already obtained the result obtainable in this case. We state their result with the object of following it up with a clarification of some misunderstanding which exists in the literature concerning a similar result by Graybill and Deal (1959).

<u>Theorem 3.5.1</u> (Khatri and Shah) The estimator $\hat{\mu}_2$ with $a = 1$ is better than both X_1 and X_2 for all $\rho > 0$ iff

$$(m_1+2)/[(2(m_2-4)] \leq d \leq 2(m_1-4)/(m_2+2) \qquad (3.5.2)$$

If we take $d = m_1/m_2$ we see that (3.5.2) holds iff

$$\min[(m_1-2)(m_2-8), (m_1-8)(m_2-2)] \geq 16 \qquad (3.5.3)$$

Thus, we see that $\hat{\mu}_2$ with $a = 1$ and $d = m_1/m_2$ is better than both X_1 and X_2 iff (3.5.3) holds. This is the result obtained by Graybill and Deal (1959).

The weaker statement in theorem 1 of their paper to the effect that $\hat{\mu}_2$ with $a = 1$ and $d = m_1/m_2$ is better than both X_1 and X_2 if $\min(m_1,m_2) > 9$, has led to some trivial claims of improvement. For example, refer to Norwood and Heinkelmann (1977) who adopted the estimator of Graybill and Deal for estimating the common mean of k normal distributions claimed that their result for $k = 2$ is different from that of Graybill and Deal and Correction of the latter result by Hultquist as quoted by Clement and Williams (1969). In actual fact, the result of Norwood and Hinkelmann (if we take $k = 2$), is no different from what Graybill and Deal actually proved.

We now consider the class of estimators with the desired property obtainable from $\hat{\mu}_3$. The case $n_1 = n_2$ has been considered by Cohen and Sackrowitz (1974). Here we consider the general case where n_1 and n_2 need not be equal. For $\hat{\mu}_3$ with $a = 1$, we have $\phi_3 = G$. Let $A(m_1,m_2)$ and $A_*(m_1,m_2)$ be defined as in (3.4.5) and (3.4.11) respectively. Take $\phi^* = \phi$. Then from (3.4.4) it is clear that (ϕ,ϕ^*) satisfies the condition of lemma 3.5.1. By theorem 3.4.3, the set E of lemma 3.5.1 consists of all pairs

(n_1,n_2) such that $A(m_1,m_2) \geq 1$, Hence the set E_0 of that lemma consists of all pairs (n_1,n_2) such that

$$\min[A(m_1,m_2),\ A(m_2,m_1)] \geq 1 \qquad (3.5.4)$$

Hence, we have

<u>Theorem 3.5.2</u> The estimator $\hat{\mu}_3$ with $a = 1$ is better than both X_1 and X_2 for all $\rho > 0$ iff (3.5.4) holds.

As we have observed earlier, it is not easy to evaluate $A(m_1,m_2)$ but $A(m_1,m_2) \geq A_*(m_1,m_2)$, provided $\min(m_1,m_2) \geq 4$. Since $A_*(m_1,m_2)$ is a symmetric function of (m_1,m_2), we have the following generalization of the result contained in remark 2.2. of Cohen and Sackrowitz (1974).

<u>Theorem 3.5.3</u> The estimator $\hat{\mu}$ with $a = 1$ is better than both X_1 and X_2 if $A_*(m_1,m_2) \geq 1$ which is satisfied if $\min(m_1,m_2) \geq 9$.

CHAPTER 4

ESTIMATION OF TREATMENT EFFECTS IN BLOCK DESIGNS

with recovery of inter-block information under the assumption of normality

4.1. Introduction

An experimental design is an allocation of elements of a set of treatment one on each of a set of experimental units. If v denotes the number of treatments and n denotes the number of experimental units, then the design is specified by an nxv matrix X (called the design matrix) whose (i,j)th element x_{ij} is 1 if the ith experimental unit receives the jth treatment and zero otherwise. Often the experimental units are divided into a number of groups called blocks. If b denotes the number of blocks, the relationship of the experimental units to the blocks is specified by an nxb matrix Z whose (i,j)th element is 1 if the ith experimental unit belongs to the jth block and 0 otherwise. A design with a block structure for the experimental units is known as a block design. The matrix $N = X'Z$ is called the incidence matrix of the design. The (i,j)th element of N gives the number of experimental units of the jth block, which receive the ith treatment. The ith element of the vector $\underset{\sim}{k} = \underset{\sim}{1}_v' N$ where $\underset{\sim}{1}_v$ = a column vector with v elements each equal to 1, gives the number of experimental units in the ith block (called the size of the ith block). The ith element of the vector $\underset{\sim}{r} = N\underset{\sim}{1}_b$ gives the number of experimental units receiving the ith treatment (called the replications of the ith treatment). A block design is called binary if each n_{ij} is either 0 or 1; equireplicate if r_i = const. for all i; proper if k_i = const. for all i. If a block design is equiré-plicate, then the common value of r_i's will be denoted by r; if proper, the common value of k_i's will be denoted by k. The matrix NN' which plays a very important role in the analysis of equireplicate proper block designs is called the association matrix. Some writers [e.g. Tocher (1952)] use the term 'concurrence matrix' for what we have termed as 'association matrix'.

The general additive model for the observations from an experiment using a block design can be written as

$$Y = \underset{\sim}{1}_n\mu + X\underset{\sim}{\tau} + Z\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}_* \qquad (4.1.1)$$

where

μ = general effect

$\underset{\sim}{\tau}$ = the vector of treatment effects

$\underset{\sim}{\beta}$ = the vector of block effects

$\underset{\sim}{\varepsilon}_*$ = the vector of individual effects of the experimental units.

As advocated by Fisher (1935) the treatments are generally allocated to the experimental units at random subject to the restrictions imposed by the design. One can analyze the experiment solely on the basis of the randomization theory. But we shall not consider this approach. For the purpose of estimation only, it is customary to use the Gauss-Markoff theory with ε_* assumed to be a random vector such that $E\varepsilon_* = 0$, $V(\varepsilon_*) = \sigma_1^2 I_n$ or $(\underset{\sim}{\beta}, \underset{\sim}{\varepsilon})$ assumed to be a pair of random vectors such that $\mathrm{cov}(\underset{\sim}{\beta}, \underset{\sim}{\varepsilon}_*) = 0$ $E\underset{\sim}{\varepsilon}_* = 0$, $V(\varepsilon_*) = \sigma_1^2 I_n, E\underset{\sim}{\beta} = 0$, $V(\underset{\sim}{\beta}) = \sigma_2^2 I_b$]. The additional assumption of normality of ε_*[or $(\underset{\sim}{\beta}, \underset{\sim}{\varepsilon}_*)$] is generally introduced if one is also interested in testing of hypothesis. Although we are concerned here only with the estimation problem, the assumption of normality would play a crucial role in the derivation of some of our results. We assume that $(\mu, \underset{\sim}{\tau})$ is fixed and $(\underset{\sim}{\beta}, \underset{\sim}{\varepsilon}_*)$ is a pair of random vectors such that

$$(\underset{\sim}{\beta}', \underset{\sim}{\varepsilon}_*')' \sim N[\underset{\sim}{0}, \mathrm{diag}(\sigma_2^2 I_b, \sigma_1^2 I_n)] \qquad (4.1.2)$$

model (4.1.1) can then be written as

$$Y = A\underset{\sim}{\theta} + \underset{\sim}{\epsilon} \tag{4.1.3}$$

$A = (1_n|X)$; $\underset{\sim}{\theta} = (\mu, \tau')'$; $\underset{\sim}{\epsilon} \sim N(0, \sigma_1^2 H)$; $H = I+\rho ZZ'$; $\rho = \sigma_2^2/\sigma_1^2$, $\underset{\sim}{\tau}$, σ_1^2, ρ are unknown.

Following the well accepted definition [Bose (1944)] we shall say that linear function of treatment effects is estimable if there exosts a linear estimator which is unbiased for it. Our problem is to estimate an arbitrarily given estimable linear function of $\underset{\sim}{\tau}$. If ρ is known the least square theory leads us to UMVUE. But if ρ is not known no optimal solution is apperent.

A set of minimal sufficient statistics has been obtained by some authors [Graybill and Weeks (1959); Roy and Shah (1962)] in special cases, where it is known to be incomplete, when ρ is unknown. The general idea behind all solutions proposed in the literature [with some exceptions e.g. see last paragraph of Stein (1966) and remark on this in Shah (1975)] is to use an estimate $\hat{\rho}$ in place of ρ in the optimal solution for the known ρ case. Among the various methods of estimation of ρ proposed in the literature the most notable ones are:Maximum likelihood procedure [Roy and Shah (1962), Hartley and Rao (1967)] which leads to M.L. estimators concerning $\underset{\sim}{\tau}$; Marginal likelihood procedure formulated by Fraser (1968) and Kalbfleisch and Sprott (1970) [see e.g. Nelder (1968), Patterson and Thompson (1971), Shaarawi et al (1975)]; Method based on LBQUE (with invariance) of $(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$ formulated by Rao (1971) [see e.g. Roy and Shah (1962), Shah and Puri (1976)]; Cunningham-Henderson-Thompson method [Cunningham and Henderson (1968), Thompson (1969)]; Analysis of variance (ANOVA) method [Yates (1939b, 1940), Nair (1944),Rao (1947), Cunningham and Henderson (1968)]; Tocher's method (Tocher (1952)]; Ad-hoc procedures leading

As the so called uniformly better estimators [Yates (1939b); Graybill and Deal (1959); Seshadri (1960a,b), Shah (1964), Stein (1966), Brown and Cohen (1974), Khatri and Shah (1974), Bhattacharya (1980)]. A method of estimating treatment contrasts, which has been in long use is that proposed by Yates (1939b) [using $\hat{\rho}$ by ANOVA method mentioned above]. His method as extended by Rao (1947) has been discussed/studied in several papers [Sprott (1956, 1957), Fraser (1957), Graybill and Weeks (1959), Graybill and Seshadri (1960), Roy and Shah (1962), Shah (1964), Khatri and Shah (1974, 1975), Shaarawi et al. (1975), Bhattacharya (1978)]. Box and Tiao (1973) proposed Bayes estimators concerning $\underset{\sim}{\tau}$, with respect to the improper prior $\propto \sigma_1^{2}(1+k\rho)^{-1}$, in the case of a proper block design. Shaarawi et al (1975) proposed a marginal procedure for obtaining estimators concerning $\underset{\sim}{\tau}$, which they point out to be same as those of Box and Tiao (1973), provided one is willing to relax the condition: $\rho > 0$ in Box and Tiao (1973). It can be seen that the procedure proposed by Shaarawi et al. (1975) leads to estimators which are same as those given by the maximum likelihood procedure.

In section 2 we present some basic ideas needed for our work. Section 3 is concerned with estimation of ρ. There we present two methods which we believe to be new and useful. One of these is an application of the theory of MIVQUE in Rao (1971) and contains similar results in Roy and Shah (1962) and Shah and Puri (1976); the other one is an extension of a method due to Tocher (1952). In subsequent sections, we confine ourselves only to proper block designs. Section 4 gives a canonical reduction leading to a set of minimal sufficient statistics. We treat both cases (i) when ρ is unknown (ii) when ρ is known, although we are interested in case (i) only. Our results fully extend the earlier results [see Graybill and Weeks (1959), Roy and Shah (1962)] applicable only in special cases. In section 5, We present a general approach to the problem of recovery of inter-block information based on the minimal sufficient statistics obtained in section 4. We also present some useful results for comparing different procedures. The ideas proposed and the results obtained apply to all procedures in the literature/present work. To summarize the work of the next two sections, let us say (with a natural motivation) that a procedure is good if for every treatment contrast estimable from intra-block analysis, it provides an estimator which is better than the intra-block estimator [Note that all estimators considered are unbiased and we judge the merit of an estimator by its variance]. *In section 6 we offer several classes of estimators* which are good in the above sense. This section contains unification/extension/improvement relating to works of several authors in this area. Section 7 is devoted to a study of the well known Yates-Rao procedure. We give an expression of the procedure in terms of the minimal sufficient statistics and establish unbiasedness of the procedure (both truncated and untruncated). These results fully extend earlier results [see Graybill and Weeks (1959), Roy and Shah (1962), Khatri and Shah (1974, 1975)] obtained earlier in special cases. We also provide some criteria which can be applied to a wide class of designs to examine if the procedure is good or not. These results unify and extend similar results in Shah (1964) and Bhattacharya (1978).

4.2 Preliminaries

In this section we present some basic ideas related to the analysis *of a block design needed for our work.*

(1) Intra and interblock contrast : Let $L(A)$ and $L_*(A)$ stand for the column space and null space respectively of the matrix A. A contrast of the observations $\underset{\sim}{Y}$ is a linear function $\underset{\sim}{\ell}'\underset{\sim}{Y}$ such that $\ell \in L_*(\underset{\sim}{1}_n') = 0$. It is called an inter-block contrast if $\ell \in L(Z)$ and an intra-block contrast if $\ell \in L_*(Z')$. It is clear that (i) the set of all contrasts of $\underset{\sim}{Y}$ span a vector space of dimension (n-1); (ii) the set of intra and inter-block contrasts seperately span orthogonal subspaces of it of dimensions n-b and b-1 respectively [since rank $Z = b$ and $\underset{\sim}{1}_n \in L(Z)$]; (iii) the vector space in (i) is the direct sum of the two subspaces in (ii). Other useful properties of these contrasts which hold for the model (4.1.3) and can be easily verified are: (iv) a given intra-block contrast is uncorrolated with a given linear function of $\underset{\sim}{Y}$ iff these are orthogonal; (v) any intra-block contrast is uncorrelated with any inter-block contrast. There is, generally, no relationship between orthogonality and uncorrelatedness among the inter-block contrasts. However, we have: (vi) if the block design is proper then a given inter-block contrast is uncorrelated with a given linear function of Y iff these two are orthogonal.

(2) Intra-block analysis : The model for the so called intra-block analysis is given by (4.1.1) where $(\mu,\underset{\sim}{\tau},\underset{\sim}{\beta})$ are fixed and $\varepsilon_* \sim N(\underset{\sim}{0},\sigma_1^2 I_n)$. It is well known [Chakravarty (1962)] that for this model the reduced normal equations for $(\underset{\sim}{\tau},\underset{\sim}{\beta})$ is given by

$$C\,\hat{\underset{\sim}{\tau}} = \underset{\sim}{Q}\ ;\ D\,\hat{\underset{\sim}{\beta}} = \underset{\sim}{P}$$

where

$$C = \underset{\sim}{r}^{\delta} - N\underset{\sim}{k}^{-\delta}N'\ ;\quad D = \underset{\sim}{k}^{\delta} - N'\underset{\sim}{r}^{-\delta}N$$

$$\underset{\sim}{Q} = \underset{\sim}{T} - N\underset{\sim}{k}^{-\delta}\underset{\sim}{B}\ ;\ \underset{\sim}{P} = \underset{\sim}{B} - N\underset{\sim}{r}^{-\delta}\underset{\sim}{T}$$

$\underset{\sim}{T} = X'\underset{\sim}{Y}$ = the vector of treatment totals

$\underset{\sim}{B} = Z'\underset{\sim}{Y}$ = the vector of block totals.

It is also well known that a linear function $\underset{\sim}{p}'\underset{\sim}{\tau}$ of treatment effects is estimable iff $\underset{\sim}{p} \in L(C)$ for which a necessary condition is that $\underset{\sim}{1}_v' \underset{\sim}{p} = 0$. Similarly, a linear function $\underset{\sim}{q}' \underset{\sim}{\beta}$ of block effects is estimable iff $\underset{\sim}{q} \in L(D)$ for which a necessary condition is that $\underset{\sim}{1}_b' \underset{\sim}{q} = 0$. A linear function of treatment effects satisfying the condition $\underset{\sim}{1}_v' \underset{\sim}{p} = 0$ is called a treatment contrast. A block contrast is defined similarly. A design is said to be connected if rank $C = v-1$, a condition which is necessary and sufficient for every treatment contrast to be estimable. As shown in Chakravarty (1963) the above definition of connectedness due to him is equivalent to the original definition of Bose (1947). The condition rank $C = v-1$ is equivalent to rank $D = b-1$ and is also necessary and sufficient for every block contrast to be estimable. We have

$$\underset{\sim}{Q} \sim N(C\underset{\sim}{\tau}\ ,\ C\sigma_1^2)\ .$$

This shows that the distribution of $\underset{\sim}{Q}$ does not depend on $\underset{\sim}{\beta}$. Hence or directly we can see that the distribution of $\underset{\sim}{Q}$ under model (4.1.3) is same as above and does not depend on ρ.

(3) <u>Inter-block analysis</u> : The idea of inter-block analysis is due to Yates (1939b, 1940) who realized the possibility of obtaining estimates concerning treatment effects from block totals, under the assumption that $\underset{\sim}{\beta}$ is also a random variable as assumed in (4.1.2). Some writers e.g. Ogawa (1974) uses the term inter-block analysis synonimously with what we shall call combined intra and inter-block analysis and treat later in this section. In the literature, the most general treatment of inter-block analysis appears to be that in Tocher (1952), who restricted himself to proper block designs with a non-singular association matrix. The treatment which follows is applicable to any block design. From our model (4.1.3), it follows that the model for block totals is given by

$$\underset{\sim}{B} \sim N\ (Z'A,\ Z'HZ\ \sigma_1^2)$$

This will be referred to as the model for inter-block analysis. It can be seen that $Z'A = (\underset{\sim}{k}|N')$, $Z'HZ = k^{\delta}\ (I+\rho k^{\delta})$. Hence, the normal equation for $\underset{\sim}{\theta}$ is given by

$$(k|N')'\underset{\sim}{k}^{-\delta}(I+\rho\underset{\sim}{k}^{\delta})^{-1}(\underset{\sim}{k}|N')\underset{\sim}{\hat{\theta}} = (k|N')'\ \underset{\sim}{k}^{-\delta}(I+\rho\underset{\sim}{k}^{\delta})^{-1}\ \underset{\sim}{B}$$

From this, the equation for estimating $\underset{\sim}{\tau}$ is given by

$$\tilde{C}\ \hat{\tau} = \underset{\sim}{\tilde{Q}} \qquad (4.2.1)$$

where

$$\tilde{C} = N\underset{\sim}{k}^{-\delta}(I+\rho\underset{\sim}{k}^{\delta})^{-1}N' \quad \frac{N(I+\rho\underset{\sim}{k}^{\delta})^{-1}\ 1_b 1_b'(I+\rho\underset{\sim}{k}^{\delta})^{-1}N'}{1_b'(I+\rho\underset{\sim}{k}^{\delta})^{-1}\ \underset{\sim}{k}}$$

$$\underset{\sim}{\tilde{Q}} = N\underset{\sim}{k}(I+\rho\underset{\sim}{k}^{\delta})^{-1}\underset{\sim}{B} \quad \frac{N(I+\rho\underset{\sim}{k}^{\delta})^{-1}\ 1_b 1_b'(I+\rho\underset{\sim}{k}^{\delta})^{-1}\underset{\sim}{B}}{1_b'(I+\rho\underset{\sim}{k}^{\delta})^{-1}\ \underset{\sim}{k}}$$

It can be seen that a linear function $\underset{\sim}{p}'\underset{\sim}{\tau}$ of treatment effects is estimable iff $p \in L(\tilde{C})$ for which a necessary condition is that $1_v'p = 0$. We have

$$\underset{\sim}{\tilde{Q}} \sim N(\tilde{C}\ \underset{\sim}{\tau},\ \tilde{C}\sigma_1^2).$$

It can be seen that $\underset{\sim}{Q}$ is a vector of intra-block contrasts, whereas $\underset{\sim}{\tilde{Q}}$ is a vector of inter-block contrasts. Hence, it follows as a consequence of the property (v) of such contrasts in (I) that intra-block and inter-block analyses provide independent sets of estimates concerning treatment effects.

The case of a proper block design is simpler and of special interest. In this case the equation (4.2.1) is equivalent to:

$$\bar{C}\ \underset{\sim}{\hat{\tau}} = \underset{\sim}{\bar{Q}}$$

where

$$\bar{C} = NN'/k - \underset{\sim}{r}\ \underset{\sim}{r}'/n; \quad \underset{\sim}{\bar{Q}} = NB/k - G\underset{\sim}{r}/n$$

$$G = \underset{\sim}{1}_b'\ \underset{\sim}{B}$$

we have,

$$\bar{\underset{\sim}{Q}} \sim N(\bar{C}\,\tau\,,\,\bar{C}\,\sigma_*^2)$$

where $\sigma_*^2 = \rho_*\sigma_1^2$ and $\rho_* = 1+k\rho$. Neither $\bar{C}$ nor $\bar{\underset{\sim}{Q}}$ depends on ρ. Hence, in this special case, inter-block estimates are obtainable without the knowledge of ρ. In general, however, solutions of equation (4.2.1) depend on ρ and the inter-block analysis poses difficulties of the same nature as we encounter with a complete analysis of model (4.1.3).

(4) Combined intra and inter-block analysis : We have seen that the intra-block and the inter-block analyses provide us with independent sets of estimates concerning treatment effects. The idea of recovery of inter-block information originally due to Yates (1939b, 1940) is to combine these two sets in order to gain increased precision. Yates restricted himself to special designs. Rao (1947) extended his idea to all proper block designs. The approach in Rao is somewhat different from that of Yates'. This point has been discussed by Sprott (1956, 1957) and Fraser (1957). We observe that a natural way of extending the idea in Yates (1939b, 1940) would be to combine the two linear models given by the equations for the intra-block and inter-block estimates. We have

$$\underset{\sim}{Q}_o \sim N(C_o\underset{\sim}{\tau}\,,\,H_*\sigma_1^2)$$

where

$$\underset{\sim}{Q}_o = (\underset{\sim}{Q}'|\tilde{\underset{\sim}{Q}}')'\;;\quad C_o = (C'|\tilde{C}')';\quad H_* = \text{diag}\,(C,\,\tilde{C})$$

From this model we can find the BLUE of any estimable linear function of $\underset{\sim}{\tau}$ using the unified theory of least squares in Rao (1973). Note that $L(C_o) \subseteq L(H_*)$ and hence, as shown in Rao and Mitra (1971), the normal equation for $\underset{\sim}{\tau}$ is obtainable by minimizing $(\underset{\sim}{Q}_o - C_o\underset{\sim}{\tau})'\,H_*^-(\underset{\sim}{Q}_o - C_o\underset{\sim}{\tau})$, where H_*^- is any g-inverse of H_*. We take

$$H_*^- = \text{diag}(C^-, \tilde{C}^-)$$

and hence obtain the normal equation as

$$C_* \hat{\underset{\sim}{\tau}} = \underset{\sim}{Q}_* \tag{4.2.2}$$

where $C_* = C + \tilde{C}$; $\underset{\sim}{Q}_* = \underset{\sim}{Q} + \underset{\sim}{\tilde{Q}}$. A necessary and sufficient condition for estimability of $\underset{\sim}{p}'\underset{\sim}{\tau}$ is that $\underset{\sim}{p} \in L(C_*)$. It can be seen that $L(C_*) = L_*(\underset{\sim}{1}_v')$. Hence every treatment contrast is estimable. It is important to note that one can arrive at equation (4.2.2) also by a direct application of the Gauss-Markoff theory to the model (4.1.3). Note that for a proper block design the equation (4.2.2) can be written as

$$(C + \rho_*^{-1} \bar{C})\hat{\underset{\sim}{\tau}} = \underset{\sim}{Q} + \rho_*^{-1} \bar{\underset{\sim}{Q}} \tag{4.2.3}$$

which was obtained by Rao (1947).

In the practical problem of recovery of inter-block information ρ is unknown. Alamost all methods proposed in the literature (with exceptions to be mentioned in due course), uses a suitable estimate $\hat{\rho}$ for ρ in equation (4.2.2) which then serves as the basis for estimates concerning $\underset{\sim}{\tau}$.

4.3 Estimation of ρ

Various methods of estimation of ρ proposed in the literature have already been mentioned in Section 1. We shall add to these two more methods which we believe to be new and useful.

(1) Method based on LBQUE (with invariance) of (σ_1^2, σ_2^2): By straightforward application of the theory in Rao (1971a,b) to our model (4.1.3) we find that the quadratic unbiased estimator $(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$ of (σ_1^2, σ_2^2) which is locally best at $(\sigma_1^2, \sigma_2^2) = (\sigma_{o1}^2, \sigma_{o2}^2)$ subject to the condition of invariance under translation of $\underset{\sim}{\theta}$, is given by

$$\hat{\sigma}_1^2 S_{11} + \rho_o^{-1} \hat{\sigma}_2^2 S_{12} = Q_1^*$$

$$\hat{\sigma}^2 S \quad + \rho^{-1} \hat{\sigma}^2 S \quad - Q^*$$

where $\rho_o = \sigma_{o2}^2/\sigma_{o1}^2$;

$S_{11} = \text{tr } R^2$; $S_{12} = S_{21} = \rho_o \text{ tr}(R^2ZZ')$; $S_{22} = \rho_o^2 \text{ tr } (RZZ')^2$;

$Q_1^* = Y'R^2Y$; $Q_2^* = \rho_o Y'RZZ'RY$;

$R = H_o^{-1} - H_o^{-1} X (X' H_o^{-1} X)^{-1} X'H_o^{-1}$; $H_o = I + \rho_o ZZ'$.

Hence we estimate ρ by $\hat{\rho} = \hat{\sigma}_2^2/\hat{\sigma}_1^2$ which is given by

$$\frac{Y'R^2Y}{Y'RZZ'RY} = \frac{\text{tr } R^2 + \hat{\rho}\text{tr}(R^2ZZ')}{\text{tr}(R^2ZZ') + \hat{\rho}\text{tr}(RZZ')^2}$$

Similar result in Roy and Shah (1962) can be obtained from this as a special case. The case $\rho_o = 0$ is of special interest. In this case $R = I - X(X'X)^{-1}X'$ and hence $\text{tr } R^2 = \text{tr}R = n-v$, $\text{tr } R^2ZZ' = \text{tr } Z'RZ = \text{tr}(\underset{\sim}{k}^{\delta} - N'\underset{\sim}{r}^{-\delta}N) = \text{tr}D$; $\text{tr}(RZZ')^2 = \text{tr}(Z'RZ)^2 = \text{tr}D^2$; $Y'R^2Y = Y'RY = Y'Y - \underset{\sim}{T}'\underset{\sim}{r}^{-\delta}\underset{\sim}{T}$; $Y'RZZ'RY = (\underset{\sim}{B} - N'r^{-\delta}\underset{\sim}{T})'(B - N'\underset{\sim}{r}^{-\delta}\underset{\sim}{T}) = \underset{\sim}{P}'\underset{\sim}{P}$. Hence, $\hat{\rho}$ is given by

$$\frac{Y'Y - \underset{\sim}{T}'r^{-\delta}\underset{\sim}{T}}{\underset{\sim}{P}'\underset{\sim}{P}} = \frac{n-v + \hat{\rho}\text{ tr } D}{\text{tr } D + \hat{\rho}\text{ tr } D^2}$$

The result in Shah and Puri (1976) can be obtained from this as a special case.

(2) <u>Generalization of Tocher's Method</u> : Tocher (1952) suggested a method of estimating ρ in the case of connected proper block designs. He was motivated by the following reasoning: If we consider the model for the intra-block analysis, the residual sum of squares of the intra-block analysis is minimum variance quadratic unbiased estimator of $(n-b-v+1)\sigma_1^2$ provided only that errors have a distribution with normal skewness and kurtosis. This is obvious from the fact that under the assumed condition, the variance of any quadratic unbiased estimator of σ_1^2 would be same as in the normal case for which the residual sum of square is known to be U.M.V.U.E.

[Rao (1973) p. 319; see Hsu (1938) and Rao (1952 , 1971b) for other conditions]. Similarly, the best quadratic estimator of a quadratic function of estimable parametric functions of parameters in the linear set up is the corresponding quadratic function of the BLUES corrected for bias. Correction for bias is to be done by subtracting an appropriate multiple of the residual sum of squares. Thus his estimator for ρ is

$$\hat{\rho} = \frac{(n-b-v+1)\ [\hat{\beta}'(I_b - 1_b\ 1_b'/b)\hat{\beta} - \overline{\text{Bias}}\{\hat{\beta}'(I_b - 1_b\ 1_b')\hat{\beta}\}]}{(b-1)\ S_1}$$

where S_I = intra-block error SS;

$\overline{\text{Bias}}$ = The appropriate unbiased estimate of the bias. We drop the assumption that the design is connected and generalize his method as follows:

Let L be the matrix consisting of the columns which constitute a complete set of orthonormal eigenvectors of D corresponding to its non-zero eigenvalues, which we denote in the vector form by ζ. Let $\beta_* = L'\beta$; $\hat{\beta}_* = \zeta^{-\delta} L'P$. Then, under the model for intra-block analysis $\hat{\beta}_*$ is BLUE for β_*. Furthermore we have

$$\hat{\beta}_* \sim N(\beta_*, \zeta^{-\delta}\ \sigma_1^2).$$

Hence

$$E\hat{\beta}_*'\hat{\beta}_* = \beta_*'\ \beta_* + \sigma_1^2\ \text{tr}\ \zeta^{-\delta}$$

Then an estimator of σ_2^2, similar to that of Tocher (1952) is given by $\hat{\beta}_*'\ \hat{\beta}_*$/rank D corrected for bias by subtracting the appropriate multiple of S_1. Hence the estimator for ρ is given by

$$\hat{\rho} = [\ \frac{(n-v-\text{rank } D)\ \hat{\beta}_*'\ \hat{\beta}_*}{S_1} - \text{tr}\ \zeta^{-\delta}\ ]/\ \text{rank } D$$

The estimator clearly reduces to that of Tocher (1952) if the design is connected.

4.4 Canonical Reduction and Minimal Sufficient Statistics

Graybill and Deal (1959) gave a canonical reduction leading to a set of minimal sufficient statistics in the case of B I B designs. Roy and Shah (1962) extended the idea to all connected binary equireplicate proper block designs. We shall extend the idea still further and obtain results which are applicable to any proper block designs. Let

$$\text{rank } C = p$$

$$\text{rank } \bar{C} = q$$

$$\dim[L(C) \cap L(\bar{C})] = s$$

Note that

$$p+q-s = \dim[L(C) + L(\bar{C})] = v-1$$

Since C, $\bar{C}$ are n.n.d. there exist a non-singular matrix M such that

$$M' C M = \text{Diag}(\underset{\sim}{\alpha}^{\delta}, \underset{\sim}{\alpha}_*^{\delta}, 0, 0)$$

$$M' \bar{C} M = \text{Diag}(I_s, 0, I_{q-s}, 0)$$

$$\underset{\sim}{\alpha} = (\alpha_1, \ldots, \alpha_s), \quad \underset{\sim}{\alpha}_* = (\alpha_{s+1}, \ldots, \alpha_p)$$

and α_i is positive for each $i = 1, \ldots p$. Let E be the matrix obtained from $(M^{-1})'$ by deleting the last column. Let E be partitioned in the form $E = (E_0 \vdots E_1 \vdots E_2)$ such that E_0 consists of the first s columns of E; E_1 consists of the next p-s columns and E_2 consists of the last v-1-p columns. It is clear that (i) columns of E constitute a basis of $L(C) + L(\bar{C})$, (ii) columns of E_0 span $L(C) \cap L(\bar{C})$, (iii) columns of $(E_0 \vdots E_1)$ span $L(C)$, (iv) columns of $(E_0 \vdots E_2)$ span $L(\bar{C})$. Let

$$\underset{\sim}{\xi}_0 = E_0' \underset{\sim}{\tau}, \quad \underset{\sim}{\xi}_1 = E_1' \underset{\sim}{\tau}, \quad \underset{\sim}{\xi}_2 = E_2' \underset{\sim}{\tau}$$

and let

$$\underset{\sim}{\xi} \quad (\underset{\sim}{\xi}_0' \vdots \underset{\sim}{\xi}_1' \vdots \underset{\sim}{\xi}_2')'$$

The elements of $\underset{\sim}{\xi}$ will be called canonical contrasts. We observe that ξ_1

is estimable only from intra-block analysis, $\underset{\sim}{\xi}_2$ is estimable only from inter-block analysis, but $\underset{\sim}{\xi}_0$ is estimable from both intra-block and inter-block analysis. Let $F = (F_0, F_1, F_2)$ be a matrix such that F_0, F_1, F_2 are related to M in the same way as U_0, U_1, U_2 are related to $(M^{-1})'$. Let

$$\underset{\sim}{x} = \alpha^{-\delta} F_0' Q , \qquad \underset{\sim}{y} = F_0' \tilde{\underset{\sim}{Q}}$$

$$\underset{\sim}{x}_* = \alpha_*^{-\delta} F_1' Q \qquad \underset{\sim}{y}_* = F_2' \tilde{\underset{\sim}{Q}}$$

It is clear that $\underset{\sim}{x}$, $\underset{\sim}{y}$ are intra-block and inter-block estimates of $\underset{\sim}{\xi}_0$ which is estimable from both analyses; $\underset{\sim}{x}_*$ is the intra-block estimate of $\underset{\sim}{\xi}_1$, which is estimable only from intra-block analysis; and $\underset{\sim}{y}_*$ is the inter-block estimate of $\underset{\sim}{\xi}_2$ which is estimable only from inter-block analysis. It is also easy to see that

$$\begin{aligned} V(\underset{\sim}{x}) &= \alpha^{-\delta} \sigma_1^2 & V(\underset{\sim}{y}) &= \sigma_*^2 \\ V(\underset{\sim}{x}_*) &= \alpha_*^{-\delta} \sigma_1^2 & V(\underset{\sim}{y}_0) &= \sigma_*^2 \\ Cov(\underset{\sim}{x}, \underset{\sim}{x}_*) &= 0 & Cov(\underset{\sim}{y}, \underset{\sim}{y}_*) &= 0 \end{aligned} \qquad (4.4.1)$$

In view of the properties of the intra-block and inter-block contrast discussed in the previous section $(\underset{\sim}{x}, \underset{\sim}{x}_*)$ being an uncorrelated set of intra-block contrasts, must be a set of orthogonal intra-block contrasts. Similarly $(\underset{\sim}{y}, \underset{\sim}{y}_*)$ must be a set of orthogonal inter-block contrasts. Hence, we can have (i) a vector $\underset{\sim}{\varepsilon}_1$ of n-b-p normalized intra-block contrasts which are orthogonal to each other and to $(\underset{\sim}{x}, \underset{\sim}{x}_*)$, (ii) a vector $\underset{\sim}{\varepsilon}_2$ of b-1-q normalized inter-block contrasts which are orthogonal to each other and to $(\underset{\sim}{y}, \underset{\sim}{y}_*)$. By an appeal to the properties of intra-block and inter-block contrasts once more, it follows that G, $\underset{\sim}{x}$, $\underset{\sim}{y}$, $\underset{\sim}{x}_*$, $\underset{\sim}{y}_*$, $\underset{\sim}{\varepsilon}_1$, $\underset{\sim}{\varepsilon}_2$ are uncorrelated. We shall now show that $\underset{\sim}{\varepsilon}_1$, $\underset{\sim}{\varepsilon}_2$ belong to error. It is easy to see that $(G, \underset{\sim}{x}, x_*)$ span the space of all linear functions of Y which are BLUE in the model of the intra-block analysis. Since $\underset{\sim}{\varepsilon}_1$ is uncorrelated with any such function it follows that $\underset{\sim}{\varepsilon}_1$ belongs to error in the intra-block analysis. Similarly, $\underset{\sim}{\varepsilon}_2$ belongs to error in the inter-block analysis. It is now easy to see that $(\underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2)$ belongs to error in the combined analysis. The transofrmation from $\underset{\sim}{Y}$ to$(G, \underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{x}_*, \underset{\sim}{y}_* \underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2)$ is linear and one to one. We have $G, \underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{x}_*, \underset{\sim}{y}_*, \underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2$ independently distributed. Furthermore,

$$G \sim N(n\mu + \underset{\sim}{r}'\tau,\ n\,\sigma_*^2\ )$$

$$\underset{\sim}{x} \sim N[\underset{\sim}{\xi}_0,\ V(\underset{\sim}{x})],\ \underset{\sim}{y} \sim N[\underset{\sim}{\xi}_0, V(\underset{\sim}{y})]$$

$$x_* \sim N[\underset{\sim}{\xi}_1,\ V(\underset{\sim}{x}_*)], \underset{\sim}{y}_* \sim N[\underset{\sim}{\xi}_2,\ V(\underset{\sim}{y}_*)]$$

$$\underset{\sim}{\varepsilon}_1 \sim N(\underset{\sim}{0},\ I_{e_1}\sigma_1^2),\ \underset{\sim}{\varepsilon}_2 \sim N(\underset{\sim}{0},\ I_{e_2}\sigma_*^2)$$

where

$$e_1 = n-b-p; \quad e_2 = b-1-q$$

and $V(\underset{\sim}{x}), V(\underset{\sim}{y}), V(\underset{\sim}{x}_*), V(\underset{\sim}{y}_*)$ are defined by (4.4.1). Let

$$S_1 = \underset{\sim}{\varepsilon}_1'\,\underset{\sim}{\varepsilon}_1,\ S_2 = \underset{\sim}{\varepsilon}_2'\,\underset{\sim}{\varepsilon}_2$$

We have $S_1/\sigma_1^2 \sim \chi^2(e_1)$, $S_2/\sigma_*^2 \sim \chi^2(e_2)$. Then, with the help of the operation described on p. 328 of Lehmann and Scheffe (1950) it follows that $(G, \underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{x}_*, \underset{\sim}{y}_*, S_1, S_2)$ is a set of minimal sufficient statistics for $(\mu, \underset{\sim}{\tau}, \sigma_1^2, \sigma_*^2)$. Since $E(\underset{\sim}{x}-\underset{\sim}{y}) = 0$, whereas $Prob(\underset{\sim}{x} \neq \underset{\sim}{y}) > 0$, it is clear that the minimal sufficient statistics is incomplete and do not lead us to U.M.V.U.E. If ρ is known, we can have further reduction leading to a set of statistics which is both minimal and complete. Let

$$\underset{\sim}{z}(\underset{\sim}{\phi}) \;=\; x + \underset{\sim}{\phi}^{\delta}\underset{\sim}{\varepsilon}_3 \tag{4.4.2}$$

where $\underset{\sim}{\phi}$ is a random vector and $\underset{\sim}{\varepsilon}_3 = \underset{\sim}{y} - \underset{\sim}{x}$. Let $\underset{\sim}{z}_* = \underset{\sim}{z}(\underset{\sim}{\gamma})$ where $\underset{\sim}{\gamma}$ is given by

$$\underset{\sim}{\gamma}^{\delta} \;=\; (I_s + \rho_*\underset{\sim}{a}^{\delta})^{-1} \tag{4.4.3}$$

Clearly $\underset{\sim}{\varepsilon}_3$ is a vector of orthogonal contrasts of the elements of Y belonging to error and is uncorrelated with $\underset{\sim}{z}_*$ which is the best unbiased linear combination of

$\underset{\sim}{z}$ and $\underset{\sim}{y}$. Also $(\underset{\sim}{z}_*, \underset{\sim}{\varepsilon}_3)$ is uncorrelated with $(G, \underset{\sim}{x}_*, \underset{\sim}{y}_*, \underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2)$ since it is a linear function of $(\underset{\sim}{x}, \underset{\sim}{y})$ which has the desired property. The transformation from Y to $(G, \underset{\sim}{z}_*, \underset{\sim}{x}_*, \underset{\sim}{y}_*, \underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2, \underset{\sim}{\varepsilon}_3)$ is linear and one to one. The transformed variables are mutually independent. We have

$$\underset{\sim}{z}_* \sim N[\xi_0, (\underset{\sim}{\alpha}^{\delta} + \rho_*^{-1} I_s)^{-1} \sigma_1^2]$$

$$\underset{\sim}{\varepsilon}_3 \sim N[0, (\underset{\sim}{\alpha}^{-\delta} + \rho_* I_s)\sigma_1^2]$$

The distributions of the remaining variables remain same as before except that we should now write σ_*^2 as $\rho_*\sigma_1^2$. An application of the procedure due to Lehmann and Scheffe (1950) mentioned earlier now shows that a set of minimal sufficient statistics for $(\mu, \underset{\sim}{\tau}, \sigma_1^2)$ is given by $(\underset{\sim}{z}_*, \underset{\sim}{x}_*, \underset{\sim}{y}_*, S_0)$ where $S_0 = S_1 + \rho_*^{-1} S_2 + \underset{\sim}{\varepsilon}_3'(\underset{\sim}{\alpha}^{-\delta} + \rho_* I_s \quad)^{-1} \underset{\sim}{\varepsilon}_3$. Completeness of the minimal sufficient statistics follows from a well known result, concerning exponential families [Lehmann (1959) Theorem I, page 132]. This result is a generalization of a similar result in Roy and Shah (1962).

The particular case of equireplicate proper block design is simpler and of special interest. In this case the matrices NN', C, $\bar{C}$ have the same set of eigenvectors. It is easily seen that $\underset{\sim}{1}_v$ is a common eigenvector and the corresponding eigenvalue is rk for NN' and 0 for both C and C. Moreover, if for a common eigenvector belonging to $L_*(1_v')$ the eigenvalue for NN' is χ, then it follows easily that the corresponding eigenvalue for C is $r-\chi/k$ and that for $\bar{C}$ is χ/k. Let ℓ and m be the multiplicities of the eigenvalues rk and 0 respectively for NN'. Obviously $m = v-t$, where $t = \text{rank } NN'$. Then

$$s = v - \ell - m = t-\ell$$

$$p = v - \ell$$

$$q = v-1-m = t - 1$$

The matrices E_o, E_1, E_2 and the vectors α, α_* can be conveniently obtained from the eigenvectors and eigenvalues of NN'. Let

U_0 = The matrix consisting of columns which constitute a complete set of orthonormal eigenvectors of NN' corresponding to eigenvalues other than 0 and rk.

U_1 = The matrix consisting of columns which constitute a complete set of orthonormal eigenvectors of NN' corresponding to the eigenvalue 0.

U_2 = The matrix consisting of columns which together with $v^{-\frac{1}{2}} \underset{\sim}{1}_v$ constitute a complete set of orthonormal eigenvectors of NN' corresponding to the eigenvalue rk.

Clearly U_o has s columns. Let $\underset{\sim}{\chi} = (\chi_1, \ldots, \chi_s)$, where χ_i is the eigenvalue of NN' corresponding to the eigenvector given by the ith columns of U Let $U = (U_o \vdots U_1 \vdots U_2 \vdots v^{-\frac{1}{2}} \underset{\sim}{1}_v)$. Then it is easy to see that

$$U' \, C \, U \;=\; \mathrm{Diag}(rI_s - \underset{\sim}{\chi}^{\delta}/k,\; rI_m, 0, 0)$$

$$U' \, \bar{C} \, U \;=\; \mathrm{Diag}(\chi^{\delta}/k, 0, rI_{\ell-1}, 0)$$

Hence the matrices $E_o, E_1, E_2, F_o, F_1, F_2$ and the vectors $\underset{\sim}{\alpha}, \alpha_*$ are given by

$$E_o \;=\; U_o(\underset{\sim}{\chi}^{\delta}/k)^{1/2};\; E_1 = U_1;\; E_2 = r^{1/2}\, U_2$$

$$F_o \;=\; U_o(\chi^{\delta}/k)^{-1/2};\; F_1 = U_1;\; F_2 = r^{-1/2}\, U_2 \;,$$

$$\underset{\sim}{\alpha}^{\delta} \;=\; rk\chi^{-\delta} - I_s;\; \underset{\sim}{\alpha}_*^{\delta} = r\, I_m.$$

If the design is connected then we have, $\ell = 1$. Hence it follows that for a connected equireplicate proper block design, s=q=t-1; p=v-1 and U_2 is void.

4.5 A General Approach to Recovery of inter-block Information for Proper Block Designs

From now on we shall confine ourselves only to proper block designs. If ρ_* is known, from the sufficiency and completeness of $(\underset{\sim}{z}_*, \underset{\sim}{x}_*, \underset{\sim}{y}_*, S_o)$ it follows that $\underset{\sim}{z}_*, \underset{\sim}{x}_*, \underset{\sim}{y}_*$ are UMVUE of $\underset{\sim}{\xi}_o$, $\underset{\sim}{\xi}_1$ and $\underset{\sim}{\xi}_2$ respectively. We observe that every treatment contrast can be expressed uniquely as a linear function of ξ_0, ξ_1, ξ_2 and hence in this case the corresponding linear function of z_*, x_*, y_* gives us the UMVUE of that treatment contrast. We also observe x_* and y_* do not depend on ρ_* and are therefore UMVUE of ξ_1, and ξ_2 respectively, even when ρ_* is unknown. Thus even when ρ_* is unknown UMVUE exists for any treatment contrast which can be expressed as a linear function of ξ_1 and ξ_2 and is given by the corresponding linear function of x_* and y_*. Even if we drop the assumption of normality it can be seen that (z_*, x_*, y_*) are BLUE when ρ_* is known and that (x_*, y_*) remain BLUE even when ρ_* is unknown. To see this we only have to observe that (z_*, x_*, y_*) is uncorrelated with $\epsilon_{**} = (\epsilon_1', \epsilon_2', \epsilon_3')$, a vector of n-v of orthogonal contrasts belonging to error which must span the space of all linear functions belonging to error since the dimension of the error space is n-v. Suppose now that ρ_* is not known and an estimate of ρ_* is used in place of ρ_* in equation (4.2.3) to obtain the combined estimator of a given treatment contrast $p'\tau$ (as is the case with almost all methods of recovery of inter-block information proposed in the literature). Then it follows from our previous analysis that this combined estimator must be

$$\widehat{p'\tau} = \ell_0' z(\hat{\gamma}) + \ell_1' x_* + \ell_2' y_* \qquad (4.5.1)$$

where $\hat{\gamma} = (I_s + \hat{\rho}_* \alpha^{\delta})^{-1}$; $z(\phi)$ is as defined by (4.4.2) and ℓ_0, ℓ_1, ℓ_2 are uniquely determined from the representation

$$p'\tau = \ell_0'\xi_0 + \ell_1'\xi_1 + \ell_2'\xi_2 \qquad (4.5.2)$$

Thus all these methods in effect seek to combine the two independent unbiased estimators x and y of ξ_0 but differ from each other in the manner in which this is done. The form of the combined estimator, as given by $z(\hat{\gamma})$ is natural but do not apply to all methods proposed in the literature. As an example, we refer to the method suggested by Stein (1966), in the last paragraph of his paper where differerent estimates of ρ_* are required for purpose of estimating different subsets of ξ_o [see Shah (1975)]. We observe that all estimators of ρ_* proposed in the literature belong to the class Φ of all measurable functions of S_1, S_2, W_1, W_2 ... W_s and that a general form of the estimating equation (for treatment contrasts) which applies to all methods proposed in the literatures is given by (4.5.1) provided $\hat{\gamma}$ appearing there is replaced by ϕ where $\phi \in \Phi^s$ and Φ^s = Cartesian product of Φ taken s times.

We shall now obtain some basic results which apply to all methods proposed in the literature.

Lemma 4.5.1 Assume that $\phi \in \Phi^s$. Then (i) ϕ is uncorrelated with the function $\ell' Y$ provided all elements of ℓ are even functions of ε_* (ii) ϕ is independent of x_*, y_*

Proof We have

$$\ell' Y - E \ell Y = \ell' \varepsilon$$

Also note that elements of ε_{**} are linear functions of ε and that S_1, S_2, W_1, W_2, ..., W_s are even functions of ε_{**}. Hence $\ell'\varepsilon$ and $\phi_1 \ell' \varepsilon$ are odd function of ε. Since the distribution of ε is symmetric about 0, it follows that

$$E\ell' Y = 0; \quad E\phi_i \ell' Y = 0$$

Hence

$$\mathrm{Cov}(\phi_i, \ell' Y) = E(\phi_i \ell' \varepsilon) = 0$$

(ii) This is a simple consequence of the fact that ϕ_i is a measurable function of ε_{**} which is independent of x_*, y_*.

Remark 4.5.1 The result in part (i) of lemma 4.5.1 implies that ϕ is uncorrelated with $z(\phi)$ [See Shearawi et. al. (1975) for a similar result].

Theorem 4.5.1 Let $\underset{\sim}{\phi} \in \Phi^s$ and let $\underset{\sim}{z}(\underset{\sim}{\phi})$ be as defined by (4.4.2). Let f denote the joint density of $(S_1, S_2, W_1, W_2, \ldots, W_s)$ and let W_{i*} $i = 1, \ldots s$ be such that W_{i*}/EW_i is a chi-square variable with 3 degrees of freedom, distributed independently of $(S_1, S_2, W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_s)$. Let the symbol E_{i*} stand for expectation with respect to the density $f_{i*} = fW_i/EW_i$ and let ϕ_{i*} stand for the expression obtained from ϕ_i by replacing W_i by W_{i*}. Finally let

$$h_i(\phi_i) = \phi_i^2/\gamma_i - 2\phi_i$$

and assume that $E[\underset{\sim}{z}(\underset{\sim}{\phi})]$ exists. Then

(i) $\underset{\sim}{z}(\underset{\sim}{\phi})$ is unbiased for $\underset{\sim}{\xi}_0$

(ii) $$V[z_i(\underset{\sim}{\phi})] = V(x_i)\,[1+E_{i*}h_i(\phi_i)] \qquad (4.5.2i)$$

$$= V(x_i)\,[1+Eh_i(\phi_{i*})] \qquad (4.5.2ii)$$

(iii) $\text{cov}[z_i(\underset{\sim}{\phi}), z_j(\underset{\sim}{\phi})] = 0$

(iv) $\underset{\sim}{z}(\underset{\sim}{\phi})$ is independent of $(\underset{\sim}{x}_*, \underset{\sim}{y}_*)$.

Proof (i) Since $\underset{\sim}{x}$ is unbiased for $\underset{\sim}{\xi}_0$ we have to show that

$$E[\phi_i\, \varepsilon_{3i}] = 0$$

This follows since $E\phi_i\, \varepsilon_{3i}$ is an odd function of $\underset{\sim}{\varepsilon}_{**}$ having a distribution symmetric about zero.

(ii) Write $z_i(\underset{\sim}{\phi})$ in the form

$$z_i(\underset{\sim}{\phi}) = z_{*i} + (\phi_i - \gamma_i)\ \varepsilon_{3i} \qquad (4.5.3)$$

Observe that on the r.h.s. of this the first term is U.M.V.U.E. and that the second term is independent of the first term since it is a measurable function of $\underset{\sim}{\varepsilon}_{**}$, which has the desired property as shown for a similar situation in the proof of Theorem 2.2.1(ii). Hence

$$V[z_i(\underset{\sim}{\phi})] = V(z_{*i}) + E[(\phi_i-\gamma_i)^2 W_i] \tag{4.5.4}$$

and

$$V(z_{*i}) \quad V(x_i)(1-\gamma_i) \tag{4.5.5}$$

also

$$EW_i = V(x_i)/\gamma_i$$

$$E[(\phi_i-\gamma_i)^2 W_i] = [V(x_i)/\gamma_i]E[(\phi_i-\gamma_i)^2 W_i/EW_i] = [V(x_i)/\gamma_i]E_{i*}(\phi_i-\gamma_i)^2$$

$$= V(x_i)[\gamma_i + E_{i*}\, h_i(\phi_i)] \tag{4.5.6}$$

Formula (4.5.2i) now follows from (4.5.4), (4.5.5) and (4.5.6). The Formula (4.5.2ii) follows from this in view of the identity

$$w\, p(w;1) = p(w;3)$$

where $p(w;m)$ denotes the density function of a chisquare variable with m degrees of freedom.

(iii) First observe that if $i \neq j$, then z_{*i} is independent of z_{*j} and $(\phi_j-\gamma_j)\,\varepsilon_{3j}$. Hence

$$\text{cov}[z_i(\underset{\sim}{\phi}),\ z_j(\underset{\sim}{\phi})] = E[(\phi_i-\gamma_i)(\phi_j-\gamma_j)\varepsilon_{3i}\ \varepsilon_{3j}]$$

It is easy to see that the term within square bracket on the right hand side of this is an odd function of ε_{3i} for any given value of the remaining arguments which are independent of ε_{3i}. Since ε_{3i} has a distribution symmetric about zero, it follows that the conditional expectation of this term given the values of all arguments other than ε_{3i} is zero. Hence, the expectation of this term is zero and the proof of (iii) is complete.

(iv) Recall that $(\underset{\sim}{x},\underset{\sim}{y})$ is independent of $(\underset{\sim}{x}_*,\underset{\sim}{y}_*)$. Also by part (ii) of lemma (4.5.1) $\underset{\sim}{\phi}$ is independent of $(\underset{\sim}{x}_*,\ \underset{\sim}{y}_*)$. The result follows since $\underset{\sim}{z}(\underset{\sim}{\phi})$ is a measurable function of $(\underset{\sim}{x},\underset{\sim}{y},\underset{\sim}{\phi})$ which is independent of $(\underset{\sim}{x}_*,\underset{\sim}{y}_*)$.

Remark 4.5.2 Results similar to (i)—(iii) of our theorem (4.5.1) were proved by Roy and Shah (1962) but they confined themselves to a more restricted class of designs as well as to a more restricted choice of ϕ. Our approach is similar except that we use more refined arguments in (ii) and (iii) which enables us to replace a condition required by them [see e.g. condition (6.2) in their paper] by the weaker condition that $E\underset{\sim}{z}(\phi)$ exists. While arguments in Roy and Shah (1962) can only show that the two terms on r.h.s. of (4.5.3) are uncorrelated provided one is willing to assume a condition similar to (6.2) of their paper, we show that these are in fact independent under the milder condition mentioned above. Our formula (4.5.2ii) is essentially equivalent to (2.5) of Khatri and Shah (1974) but our proof is algebraically simpler. The following theorem states an important consequence of theorem (4.5.1).

Theorem 4.5.2 Let $\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2 \in \Phi^s$. Let $\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_1$ and $\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_2$ be expressions obtained from (4.5.1) by replacing $\hat{\gamma}$ appearing there by ϕ_1 and ϕ_2 respectively. Then

$$V[z_i(\underset{\sim}{\phi}_1)] \leq V[z_i(\underset{\sim}{\phi}_2)], \quad \text{for every} \quad i = 1,\dots,s$$

$$\Longleftrightarrow V[\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_1] \leq V[\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_2],$$

for every treatment contrast $\underset{\sim}{p}'\underset{\sim}{\tau}$.

Proof Theorem (4.5.1) implies

$$V\,[\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_j] = \sum_{i=1}^{s} \ell_{oi}^2 V[z_i(\underset{\sim}{\phi}_j)] + \underset{\sim}{\ell}_1' V(x_*)\underset{\sim}{\ell}_1 + \underset{\sim}{\ell}_2' V(y_*)\underset{\sim}{\ell}_2, \; j = 1,2$$

where $\underset{\sim}{\ell}_o$, $\underset{\sim}{\ell}_1$ and $\underset{\sim}{\ell}_2$ are as determined in (4.5.2) and ℓ_{oi} denotes the ith component of $\underset{\sim}{\ell}_o$. Hence the desired result is obvious.

In the following we shall refer to the procedure based on $\underset{\sim}{\phi}$ as procedure $\underset{\sim}{\phi}$. The procedure based on $\underset{\sim}{\phi} = 0$ will be referred to as $\underset{\sim}{\phi}_o$. Let $\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_1$ and $\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_2$ be as defined in Theorem 4.5.2. We shall say that $\widehat{(\underset{\sim}{p}'\underset{\sim}{\tau})}_1$ is better than $(\widehat{\underset{\sim}{p}'\underset{\sim}{\tau}})_2$ for all $\rho_* \epsilon \Omega$ if

$$V[(\widehat{\underset{\sim}{p}'\underset{\sim}{\tau}})_1] \leq V[(\widehat{\underset{\sim}{p}'\underset{\sim}{\tau}})_2] \quad \text{for all} \quad \rho_* \epsilon \Omega$$

The procedure $\underset{\sim}{\phi}_1$ will be said to be better than ϕ_2 for all $\rho_* \epsilon \Omega$ if $(\widehat{\underset{\sim}{p}'\underset{\sim}{\tau}})_1$ is better than $(\widehat{\underset{\sim}{p}'\underset{\sim}{\tau}})_2$ for all treatment contrasts $\underset{\sim}{p}'\underset{\sim}{\tau}$. In view of the above definitions Theorem 4.5.2 is equivalent to

<u>Theorem 4.5.3</u> Let $\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2 \epsilon \Phi^s$. Then $\underset{\sim}{\phi}_1$ is better than $\underset{\sim}{\phi}_2$ iff $z_i(\underset{\sim}{\phi}_1)$ is better than $z_i(\underset{\sim}{\phi}_2)$ for every i.

The following two Theorems which can be easily deduced from theorem 4.5.1 would be useful for application of the result just stated.

<u>Theorem 4.5.4</u> Let $\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2 \epsilon \Phi^s$. Then $z_i(\underset{\sim}{\phi}_1)$ is better than $z_i(\underset{\sim}{\phi}_2)$ for all $\rho_* \epsilon \Omega$ iff for estimating γ_i,

$$\text{M.S.E. } (\phi_{1i*}) \leq \text{M.S.E. } (\phi_{2i*}) \text{ for all } \rho_* \epsilon \Omega$$

<u>Theorem 4.5.5</u> Let $\underset{\sim}{\phi} \epsilon \Phi^s$ then $z_i(\underset{\sim}{\phi})$ is better than the corresponding intra-block estimator for all $\rho_* \epsilon \Omega$ iff

$$2\nu_\Omega(\phi_i) \geq 1$$

where

$$\nu_\Omega(\phi_i) = \inf_{\rho_* \epsilon \Omega} E\bar{\phi}_i / E\bar{\phi}_i^2 \quad ;$$

$$\bar{\phi}_i = \phi_{i*}/\gamma_i .$$

We shall conclude this section with the following corollary of Theorem 4.5.5 which would be used repeatedly in the next two sections.

<u>Corollary 4.5.1</u> Let $\phi_i = a_i \Psi_i$ where a_i is a positive constant to be suitably chosen and $\Psi_i \epsilon \Phi$. Then $z_i(\underset{\sim}{\phi})$ is better than the corresponding intra-block estimator for all $\rho_* \epsilon \Omega$ iff $a_i \leq 2\nu_\Omega(\Psi_i)$.

It should be noted that under the assumption (4.1.2) we must have $\rho_* > 1$. However it can be easily seen that for a proper block design H is p.d. iff $\rho > -1/k$. Hence in the resulting model (4.1.3) which is the basis of our analysis, we may allow $\rho_* = 1 + k\rho$ to assume any positive value. In the following sections where we shall use the above results, Ω will be either $(0,\infty)$ or $(1,\infty)$ and for the sake of simplicity we shall denote $\nu_\Omega(\Psi_i)$ in the two cases by $\nu(\Psi_i)$ and $\nu_*(\Psi_i)$ respectively. Let

$$\gamma_{oi} = 1/(1+\alpha_i)$$

Then it is easy to see that

$$\nu(\Psi_i) = \inf_{\gamma_i \in (0,1)} E\bar{\Psi}/E\bar{\Psi}^2$$

$$\nu_*(\Psi_i) = \inf_{\gamma_i \in (0,\gamma_{oi})} E\bar{\Psi}/E\bar{\Psi}^2$$

4.6 Estimation Procedures Better than ϕ_O

In this section we develop some estimation procedures which are better than the procedure ϕ_O. There are two main reasons for considering such procedures. Firstly it is natural to require that the use of additional information obtainable from the inter-block analysis ought to be made in such a way that under no circumstances it leads to estimators worse than what we could obtain without using it. Secondly estimation procedures which utilize interblock information, generally produce estimators for which variences are difficult to compute or estimate and as such the procedures we are looking for has the advantage that simple and unbiasedly estimable upper bounds of these are provided by those for the procedure ϕ_O. The pioneering work on the construction of estimators with the desired property was done by Yates (1939). Estimate of ρ_*, used by him in this connection was based on the inter-block error sum squares and differs from the usual one based on the adjusted block sum of sequares, recommended by him in the same paper. After a lapse of twenty years interest in the problem was revived by Graybill and Deal (1959), who offered a similar estimator in case of BIBD with appropriate restrictions. The work of Graybill and Deal was quickly followed by a series of papers by Seshadri (1963 a,b), Shah (1964) and Stein (1966) but the results obtained were still applicable only to special designs. The estimates proposed by these three authors displayed some similarity and utilized only the treatment component of the adjusted block sum of squares. Notable contributions in recent years are due to Brown and Cohen (1971) on BIBD and by Khatri and Shah (1974) on connected binary equireplicate proper block designs. We shall unify and extends ideas in these two papers and obtain comparable results for any proper block design. Our result would contain those in Khatri and Shah (1974) and constitute an improvement over the Brown Cohen results. We shall also unify and extend the estimators proposed by Seshadri (1963 a,b), Shah (1964) and Stein (1966) and generalize their results in a similar manner.

Following the approach of the previous section we consider five procedures:

$$\underset{\sim}{\Phi}_\kappa = (\phi_{\kappa 1},\ldots,\phi_{\kappa s}), \quad \kappa = 1,\ldots,5$$

where

$$\phi_{\kappa i} = a_i \Psi_{\kappa i} ,$$

$$\Psi_{1i} = S_1/[S_1 + c_i(S_2 + \sum_{j=1}^{s} W_j)] ,$$

$$\Psi_{2i} = S_1/[S_1 + c_i(S_2 + W_i)] ,$$

$$\Psi_{3i} = S_1/(S_1 + c_i S_2) ,$$

$$\Psi_{4i} + S_1/(S_1 + c_i \sum_{j=1}^{s} W_j) ,$$

$$\Psi_{5i} = S_1/(\sum_{j=1}^{s} c_{ij} W_j) ;$$

a_i, c_i, c_{ij} are positive constants.

It should be pointed out that the constant a_i, c_i, c_{ij} which appear above are to be interpreted as generic constants, that is to say not necessarily the same values of these constants would be employed in different classes. The procedures $\phi_1 - \phi_4$ are related to those in Brown and Cohen (1974) and Patri and Shah (1974). Procedure ϕ_5 is related to those in Seshadri (1963a,b), Shah (1964) and Stein (1966). We first consider ϕ_1. Let

$$V_1 = S_1/\sigma_1^2, \quad V_2 = S_2/\sigma_*^2$$

$$V_{3i} = W_{i*}/[\sigma_1^2/\alpha_i + \sigma_*^2]$$

$$V_{3j} = W_j/[\sigma_1^2/\alpha_j + \sigma_*^2], \; j \neq i$$

$$V_3 = \sum_{j=1}^{s} V_{3j}, \; V_4 = V_2 + V_3$$

$$u_j = V_{3j}/V_4, \quad j = 1,2,\ldots,s$$

$$u = \alpha_i \sum_{j=1}^{s} (u_j/\alpha_j) .$$

Then

$$S_1 = \sigma_1^2 V_1$$

$$S_2 + W_{i*} + \sum_{j \neq i} W_j = \sigma_*^2 V_2 + \sum_{j=1}^{s} (\sigma_1^2/\alpha_j + \sigma_*^2) V_{3j}$$

$$= (\sigma_1^2/\alpha_i)\,(u + \alpha_i\,\rho_*)\,V_4$$

$$= [\sigma_1^2/(\alpha_i\,\gamma_i)][p(u) - q(u)\gamma_i]V_4$$

where $p(u) = 1$; $q(u) = 1-u$. Hence it is easy to see that $\bar{\Psi}$ can be written as

$$\Psi = V_1/[\gamma_i V_1 + d_i h(u,\gamma_i) V_4] \qquad (4.6.1)$$

where

$$d_i = c_i/\alpha_i$$

$$h(u,\gamma) = p(u) - q(u)\gamma$$

Note that $V_1 \sim \chi^2(e_1)$; $V_4 \sim \chi^2(e_2+s+2)$; $u_j \sim \beta[\frac{1}{2}, (e_2+s+1)/2]$ if $j \neq i$ $u_j \sim \beta[3/2,(e_2+s-1)/2]$ if $j = i$; $\sum_{j=1}^{s} u_j \sim \beta[(s+2)/2, e_2/2]$ and that V_1, V_4,u are mutually independent. Since V_1 is almost sure positive, (4.6.1) is equivalent to;

$$\overline{\Psi} = 1/[\gamma_i + d_i\, h(u,\gamma_i)V]$$

where $V = V_4/V_1$. It can be seen that with the matchup $u \sim x, V \sim y$, $\bar{\Psi}$ matches up with f of theorem A4 and satisfies all conditions of part A of that theorem provided $EV^{-2} < \infty$, a condition which is satisfied iff $e_2 + s \geq 3$. The support of u is $S = (0, \alpha_i/\alpha_*)$ where

$$\alpha_* = \min_j \alpha_j; \; S_* = \{t | t \in S; \; q(t) > 0\}$$
$$= \{t | 0 < t < 1\}.$$

We have

$$a_0 = EV^{-1}/EV^{-2} = (e_2+s-2)/(e_1+2);$$

$$\delta_1 = \inf_{t \in S} p(s) = 1; \quad \delta_2 = \inf_{t \in S_*} p(t)/q(t) = 1\;;$$

$$\delta_3 = \inf_{t \in S_*} h(t;1) = 0;$$

$$\delta_5 = \text{Max}(\delta_2, \; d_i a_o \delta_3) = 1\;;$$

$$\pi_1 = \text{Min}(d_i a_o \delta_1, \delta_5) = \min(1, d_i a_o).$$

In view of the above calculation theorem A.4 gives $v(\psi) \geq \pi_1$. Also clearly

$$\inf_{\rho_* > \rho_{*o}} E\bar{\Psi}/E\bar{\Psi}^2 = \inf_{\gamma_1 < 1/(1+\alpha_i \rho_{*o})} E\bar{\Psi}/E\bar{\Psi}^2 \leq [E\bar{\Psi}/E\bar{\Psi}^2]_{\gamma_1=0} = d_i a_o \quad \text{for any } \rho_{*o} > 0.$$

Hence corollary 4.5.5 gives

<u>Theorem 4.6.1</u> Assume that $e_2 + s \geq 3$. Let $a_o = (e_2+s-2)/(e_1+2)$. Then

(i) $z_i(\underset{\sim}{\phi}_1)$ is better than x_i for all $\rho > 0$ if

$$a_i \leq 2 \min(1, d_i a_o) \qquad (4.6.1)$$

(ii) $z_i(\underset{\sim}{\phi}_1)$ is better than x_i for all $\rho_* > \rho_{*o}$ (for some $\rho_{*o} \geq 0$) only if

$$a_i \leq 2\, d_i a_o \qquad (4.6.2)$$

Note that if either $a_i \leq 2$ or $d_i a_o \leq 1$ then (4.6.1) is equivalent to (4.6.2). Hence theorem 4.6.1 gives

Corollary 4.6.1 Assume that $e_2 + s \geq 3$ and that either $a_i \leq 2$ or $d_i a_o \leq 1$ then $z_i(\underset{\sim}{\phi}_1)$ is better than x_i for all $\rho_* > \rho_{*o}$ (for some $\rho_{*o} \geq 0$) iff $a_i \leq 2\ d_i a_o$.

Similar arguments can be applied to each of the procedures ϕ_2, ϕ_3, ϕ_4. Thus, we have

Theorem 4.6.2 Statement of Theorem 4.6.1 (and hence corollary 4.6.1) holds word by word for each of the estimator $z_i(\underset{\sim}{\phi}_2)$, $z_i(\underset{\sim}{\phi}_3)$ and $z_i(\underset{\sim}{\phi}_4)$ provided the assumption $e_2 + s \geq 3$ is replaced by $e_2 \geq 2$, $e_2 \geq 5$, $s \geq 3$ respectively and the expression for a_o is replaced by $a_o = (e_2-1)/(e_1+2)$, $a_o = (e_2-4)/(e_1+2)$, $a_o = (s-2)/(e_1+2)$ respectively.

Remark 4.6.1 In case of $\underset{\sim}{\phi}_3$, the analogoue of part (i) of theorem 4.6.1 contained in theorem 4.6.2 can be improved by using arguments, similar to that in Remark 3.4.1. We have thus the following result: Assume that $e_2 \geq 5$ and let $a_o = (e_2-4)/(e_1+2)$. Then $z_i(\underset{\sim}{\phi}_3)$ is better than x_i for all $\rho_* > 0$ iff

$$a_i \leq 2 \min (1, d_i a_o)$$

It can be seen that the estimators $\hat{\xi}_i$ of Khatri and Shah (1974) and the estimators $\hat{\mu}_a$, $\hat{\mu}_a^{(1)}$, $\hat{\mu}_a^*$ of Brown and Cohen (1974) can be written as

$$\hat{\xi}_i = z_i(\underset{\sim}{\phi}_1) \text{ with } a_i = 1,\ c_i = c$$

$$\hat{\mu}_a = z_i(\underset{\sim}{\phi}_2) \text{ with } a_i = a,\ c_i = e_1 \alpha_i/(e_2+3)$$

$$\hat{\mu}_a^{(1)} = z_i(\underset{\sim}{\phi}_3) \text{ with } a_i = a,\ c_i = e_1 \alpha_i/e_2$$

$$\hat{\mu}_a^* = z_i(\underset{\sim}{\phi}_1) \text{ with } a_i = a, c_i = e_1 \alpha_i/(e_2+s+2)$$

(From the details given in Section 4.4 note that for connected binary equireplicate proper block designs, considered by these authors, $e_2+s = b-1$). We observe that $\hat{\xi}_i$ is a particular case of $z_i(\underset{\sim}{\phi}_1)$ with $a_i \leq 2$. Also $\hat{\mu}_a$, $\hat{\mu}_a^{(1)}$

$\hat{\mu}$ are particular cases of $z_i(\phi_2)$, $z_i(\phi_3)$ and $z_i(\phi_1)$ respectively such that $d_i a_0 \leq 1$ in each case, the values of $d_i a_0$ being $e_1(e_2-1)/[(e_1+2)(e_2+3)]$, $e_1(e_2-4)/[e_2(e_1+2)]$ and $e_1(e_2+s-2)/[(e_1+2)(e_2+s+2)]$ respectively. Hence by the preceding theorems we have

<u>Corollary 4.6.2</u> For all $\rho_* > 0$, (i) ξ_i is better than x_i iff $c \geq (\frac{1}{2})\alpha_i \times (e_1+2)/(e_2+s-2)$, provided $e_2+s \geq 3$.

(ii) $\hat{\mu}_a$ is better than x_i iff $a \leq 2\, e_1(e_2-1)/[(e_1+2)(e_2+3)]$, provided $e_2 \geq 2$.

(iii) $\hat{\mu}_a^{(1)}$ is better than x_i iff $a \leq 2\, e_1(e_2-4)/[(e_1+2)e_2]$ provided $e_2 \geq 5$.

(iv) $\hat{\mu}_a^*$ is better than x_i iff $a \leq 2\, e_1(e_2+s-2)/\{(e_1+2)(e_2+s+2)\}$ provided $e_2+s \geq 3$.

The result (i) of the above corollary is an extension of a similar result by Khatri and Shah (1974). The result (ii) - (iv) are extensions of similar results in Brown and Cohen (1974) who confined themselves to BIBD. We observe that for a BIBD, $s = p = q = v-1$. Hence the results above concerning the Brown - Cohen estimators are readily seen to be improvements of these in section 3 of Brown and Cohen, where the knowledge that $\alpha_i\, \rho_* > 1$ is used to improve the upper limit from $a_{max}(e,e_1) = 2EV^{-1}/EMax(V^{-1},V^{-2})$ to $a^*_{max}(e,e_1) = 2EV^{-1}/EMax[2/\{V(1+V)\},V^{-2}]$ with $V \sim F(e,e_1)$ and $e = e_2+3$, e_2, e_2+v+1 in cases (i), (ii) and (iii) respectively. We shall now concern ourselves with some results and discussion concerning the procedure ϕ_5. Let V_1, V_{3j} ($j = 1,\ldots,s$) and V_3 be defined as before. Let

$$u_j^* = V_{3j}/V_3 \quad (j = 1,2,\ldots,s)$$

$$u^* = \sum_{j=1}^{s} c_{ij}\, u_j^*/\alpha_j, \quad v^* = (\sum_{1}^{s} c_{ij}\, u_j^*)/\alpha_i$$

$$c_{ii} W_{i*} + \sum_{j \neq i} c_{ij} W_j = \sum_{j=1}^{s} c_{ij}(\sigma_1^2/\alpha_j + \sigma_*^2) V_{3j} = \sigma_1^2(u^* + \alpha_i \rho_* v^*) V_3$$

nce, it is easy to see that $\bar{\Psi}$ can be written as

$$\bar{\Psi} = (V_1/V_3)/[\gamma_i u^* + (1-\gamma_i)v^*]$$

that V_1, V_3 are independent of each other and of (u^*, v^*). Furthermore $_1$ $\chi^2(e_1)$, $V_3 \sim \chi^2(s+2)$, $u_j^* \sim \beta(\frac{1}{2}, \frac{s+1}{2})$, if $j \neq i$, $u_i^* \sim \beta(3/2, \frac{s-1}{2})$, $\sum_{j=1}^{s} u_j^* = 1$. Hence $E\bar{\Psi}/E\bar{\Psi}^2 = a_o Ef/Ef^2$, where

$$a_o = E(V_1/V_3)/E(V_1/V_3)^2 = (s-2)/(e_1+2), \qquad (4.6.3)$$

$f = 1/[\gamma_i u^* + (1-\gamma_i)v^*]$. Hence

$$\nu(\Psi) = a_o A_o \qquad (4.6.4)$$

where

$$A_o = \inf_{\gamma_i \in (0,1)} Ef/Ef^2$$

We now need to find A_o for which we would naturally like to appeal to theorem 4.2. Clearly, u^* and v^* are non-negative but in general u^* and v^* are dependent and the result there is not applicable. However, when (i) $c_{ij} = 1$ for every j, we have $u^* = v_1$; $v^* = 1/\alpha_i$ where

$$v_1 = \sum_{j=1}^{s} (u_j^*/\alpha_j) \qquad (4.6.5)$$

Also, when (ii) $c_{ij} = \alpha_j$ for every j, we have $u^* = 1$; $v^* = v_2/\alpha_i$ where

$$v_2 = \sum_{j=1}^{s} (\alpha_j u_j) \qquad (4.6.6)$$

In each of these two cases u^* and v^* are obviously independent (one of the two being a constant).

Now by theorem A.2 we have

$$A_o = \min[1/\alpha_i, M_1] \text{ in case (i)} \quad (4.6.7)$$

$$A_o = \min(1, M_2/\alpha_i] \text{ in case (ii) where}$$
$$M_j = Ev_j^{-1}/Ev_j^{-2}, \quad j = 1, 2$$

Another case which is also simple and is of special interest concerns the class of designs for which (iii) $\alpha_j = \alpha_o$ for all j. In this case $v^* = v_3/\alpha_o$ where

$$v_3 = \sum_{j=1}^{s} c_{ij} u_j \quad (4.6.8)$$

Then $f = \alpha_o/v_3$ and hence it is obvious that

$$A_o = (1/\alpha_o) M_3 \text{ in case (iii) where}$$

$$M_3 = Ev_3^{-1}/Ev_3^{-2} \quad (4.6.9)$$

To complete the evaluation of A_o in the three cases we need to find computable expressions for M_1, M_2, M_3. For this we shall use the following result in Ruben (1962).

Lemma 4.6.1 Let $x_1, \ldots, x_s$ be independent chisquare variables with $m_1, \ldots, m_s$ d.f. respectively. The density function of

$$y = \sum_{j=1}^{s} d_j^* x_j, \; d_j > 0 \; \forall \; j \quad (4.6.10)$$

is given by

$$\sum_{j=0}^{\infty} f_j \, g\left(\frac{1}{2p_*} ; \frac{m_*}{2} + j\right)$$

where

$$m_* = \sum_{j=1}^{s} m_j ;$$

satisfies

$$\max_j |1 - p_*/d_j^*| < 1 ;$$

f_j's are given by

$$\sum_{j=0}^{\infty} f_j z^j = \prod_{j=1}^{s} (p_*/d_j^*)^{m_j/2} [1-(1-p_*/d_j^*)z]^{-m_j/2}, \tag{4.6.11}$$

$$|z| \leq \min_j |1-p_*/d_j^*| ;$$

and $g(\cdot,\cdot)$ stands for the gamma density. Further the f_j's may be determined from the relations $f_o = \sum_{j=1}^{s} (p/d_j)^{m_j/2}$; $f_{j+1} = \sum_{\kappa=0}^{j} f_{j-\kappa}\, g_\kappa/[2(j+1)]$ where $g_t = \sum_{\kappa=1}^{s} m_\kappa (1-p/d_\kappa)^{t+1}$

Consider, $\quad y = V_3 v_1 = \sum_{j=1}^{s} V_{3j}/\alpha_j \qquad (4.6.12)$

his is of the form (4.6.10) with

$$\begin{aligned} s &= s \\ m_j &= 1 \quad \text{if} \quad j \neq i \\ &= 3 \quad \text{if} \quad j = i \\ d_j^* &= 1/\alpha_j \end{aligned} \tag{4.6.13}$$

hen using Lemma 4.6.1,

$$Ey^{-1} = F_1^{(1)}/p_1 \; ; \; Ey^{-2} = F_2^{(1)}/p_1^2 \tag{4.6.14}$$

where

$$F_1^{(1)} = \sum_{j=0}^{\infty} f_j^{(1)}/(s-2j);$$

$$F_2^{(1)} = \sum_{j=1}^{\infty} f_j^{(1)}/[(s-2j)(s-2j-2)] \tag{4.6.15}$$

where $f_j^{(1)}$ stands for the expression obtained from f_j (of lemma 4.6.1) with $s;m_1,\ldots,m_s;d_1^*,\ldots,d_s^*$ as given by (4.6.13) and $p_* = p_1$ satisfying $0 < p_1 < \min_j(1/\alpha_j)$. Using (4.6.14) and the fact that V_3, v_1 are independent it is easy to see from (4.6.12) that

$$M_1 = (EV_3^{-2}/EV_3^{-1})(Ey^{-1}/Ey^{-2}) = (p_1 F_1^{(1)}/F_2^{(1)})/(s-2) \qquad (4.6.16)$$

Let p_2, p_3 satisfy

$$0 < p_2 < \min_j(\alpha_j); \ 0 < p_3 < \min_j(c_{ij})$$

Let $f_j^{(2)}$, $f_j^{(3)}$ be obtained from $f_j^{(1)}$[defined immediately after (4.6.15)] by replacing the argument p_* by p_2 and p_3 respectively and argument d_j^* for each j by $d_j^* = \alpha_j$ and $d_j^* = c_{ij}$ respectively. Let $(F_1^{(2)}, F_2^{(2)})$ and $(F_1^{(3)}, F^{(3)})$ be obtained from $(F_1^{(1)}, F_2^{(1)})$ [defined in (4.6.15)] by replacing $f_j^{(1)}$ by $f_j^{(2)}$ and $f_j^{(3)}$ respectively. Then, in a similar manner, we find

$$M_2 = (p_2 F_1^{(2)}/F_2^{(2)})/(s-2); \quad M_3 = (p_3 F_1^{(3)}/F_2^{(3)})/(s-2) \qquad (4.6.17)$$

Now let

$$\begin{aligned} A_{1i} &= 2\, a_o \,\mathrm{Min}(1/\alpha_i, M_1) \ ; \\ A_{2i} &= 2\, a_o \,\mathrm{Min}(1, M_2/\alpha_i) \ ; \\ A_{3i} &= (2a_o/\alpha_o)M_3 \end{aligned} \qquad (4.6.18)$$

where a_o is given by (4.6.3); M_1, M_2 and M_3 are given by (4.6.16) and (4.6.17). Then from (4.6.4), (4.6.7) and (4.6.9), we find the value of $2\nu(\Psi)$ to be A_{1i}, A_{2i} and A_{3i} in the cases (i), (ii) and (iii) respectively. Hence by corollary 4.5.5 we have

<u>Theorem 4.6.3</u> Assume that $s \geq 3$ and let A_{1i}, A_{2i} and A_{3i} be as defined by (4.6.18). Then

(i) $z_i(\phi_5)$ with $c_{ij} = 1$ for all j is better than x_i for all $\rho_* > 0$ iff $a_i \leq A_{1i}$.

(ii) $z_i(\phi_5)$ with $c_{ij} = \alpha_j$ is better than x_i for all $\rho_* > 0$ iff $a_i \leq A_{2i}$.

(iii) if $\alpha_j = \alpha_0$ for all $j, z_i(\phi_5)$ is better than x_i for all $\rho_* > 0$ iff $a_i \leq A_{3i}$.

<u>Remark 4.6.2</u> With obvious modifications of the condition: $s \geq 3$ and the formulae (4.6.16) and (4.6.17) theorem 4.6.3 (iii) holds even when some c_{ij}'s are zero.

It can be seen that the estimator proposed by Seshadri (1963a,b), Shah (1964) (untruncated form) and Stein (1966) for the recovery of interblock information for special designs, where $\alpha_j = \alpha_0$ for every j, respectively. It can be seen that $a_{2i} \leq A_{3i}$ iff $e_1 \geq 2$; $a_{3i} \leq A_{3i}$ iff $(e_1-2)(s-4) \geq 8$ and $a_{4i} = A_{3i}/2$ is always less than A_{3i}. Hence by part (iii) of theorem 4.6.3, the result in those papers follow (Note that Seshadri, Shah and Stein write t-1, p and p, respectively for our s; and f, e_o and n, respectively for our e_1). It should be noted, however, that though the proof here is simple, the model here is less general than in Stein (1966) who does not require the normality of block effects. Also unlike Shah (1964), we consider the untruncated form of the estimators. That in our model, all these results hold for the truncated form also follows from theorem 5.3.3(ii), which we shall prove in the next chapter.

The upper limits of a_i in theorem 4.6.3 would be generally difficult to compute without the aid of a computer. Hence we now wish to provide upper limits of a_i which can be used easily in practice to ensure that $z_i(\underset{\sim}{\phi}_5)$ is better than x_i for all $\rho_* > 0$. For this we shall use the following lemma.

Lemma 4.6.2 Let $\underset{\sim}{x}$ be a random vector and $\lambda(\underset{\sim}{x})$ be a measurable function of $\underset{\sim}{x}$. Then $E[1/\lambda(\underset{\sim}{x})]/E[1/\lambda(\underset{\sim}{x})]^2 \geq \min\limits_{\underset{\sim}{x}} \lambda(\underset{\sim}{x})$.

Proof The proof is elementary and is omitted.

Let

$$A^*_{1i} = 2a_o M^*_1;\ A^*_{2i} = 2a_o M^*_2/\alpha_i;\ A^*_{3i} = 2a_o M^*_3/\alpha_o \qquad (4.6.19)$$

where

$$M^*_1 = \min_j (1/\alpha_j);$$

$$M^*_2 = \min_j \alpha_j;$$

$$M^*_3 = \min_j c_{ij}$$

From (4.6.5), (4.6.6) and (4.6.8), it is easy to see that

$$M_i^* = \underset{(u_1^*,\ldots,u_p^*)}{\text{Min}}\ v_i \quad , \ i = 1,2,3$$

Hence, using lemma (4.6.2) we see from (4.6.16) and (4.6.17) that

$$M_i \geq M_i^*$$

Using this, (4.6.18) gives :

$$A_{1i} \geq 2a_o \text{ Min } [1/\alpha_i, M_i^*] = 2a_o M_1^* = A_{1i}^* \qquad (4.6.21)$$

since $M_1^* \leq 1/\alpha_i$ by (4.6.20).

$$A_{2i} \geq 2a_o \text{ Min}[1, M_2^*/\alpha_i] = 2a_o M_2^*/\alpha_i = A_{2i}^* \qquad (4.6.22)$$

Since $M_2^*/\alpha_i \leq 1$ by (4.6.20)

$$A_{3i} \geq 2a_o M_3^*/\alpha_o = A_{3i}^* \qquad (4.6.23)$$

In view of (4.6.21), (4.6.22) and (4.6.23) theorem 4.6.3 yields.

Corollary 4.6.3 Assume that $s \geq 3$. Let $A_{si}^*, s = 1,2,3$ be as defined by (4.6.19). Then we have

(i) $z_i(\phi_5)$ with $c_{ij} = 1$ for all j is better than x_i for all $\rho_* > 0$ if $a_i \leq A_{1i}^*$.

(ii) $z_i(\phi_5)$ with $c_{ij} = \alpha_j$ for all j is better than x_i for all $\rho_* > 0$ if $a_i \leq A_{2i}^*$.

(iii) if $\alpha_j = \alpha_o$ for all j, then $z_i(\phi_5)$ is better than x_i for all $\rho_* > 0$ if $a_i \leq A_{3i}^*$.

4.7 Yates - Rao Procedure

As pointed out in the previous section the motivation behind the recovery of inter-block information is not just to use the inter-block information but to use it to improve upon the customary intra-block estimators. It is therefore, desirable that we examine all well known procedures and obtain precise conditions under which the resulting estimators have the desired property. Although theorem 4.5.5 apparently offers a neat theoretical solution to this problem, practical application of this theoretical result

In the present study we shall restrict ourselves to a method which is perhaps the oldest and most widely used. The method was originally proposed by Yates (1939b, 1940) and gradually extended by Nair (1944) to all PBIB designs, by Rao (1947) to all proper block designs and finally by Cunninghem and Henderson (1968) to all block designs. Since we are concerned only with proper block designs, we shall refer to this method as Yates-Rao procedure. The method can be described as follows:

(i) Obtain the estimates $\hat{\sigma}_1^2$ and $\hat{\sigma}_*^2$ of σ_1^2 and σ_*^2 respectively by equating the intra-block error SS(S_1) and adjusted block SS (to be denoted by SS_B) to their respective expectations, (ii) estimate ρ_* by $\hat{\rho}_* = \hat{\sigma}_*^2/\hat{\sigma}_1^2$, (iii) substitute $\hat{\rho}_*$ for ρ_* in equation (4.2.3) and from this obtain the estimate of any desired treatment contrast.

As suggested by Yates, it is customary to modify the estimator $\hat{\rho}_*$ by

$$\tilde{\rho}_* = \hat{\rho}_* \text{ if } \hat{\rho} > 1$$
$$= 1 \text{ otherwise}$$

The reason put forward for this is that under the assumption (4.1.2), ρ_* cannot assume values less than 1. The procedure with or without this modification will be referred to as the truncated and untruncated form respectively of the Yates-Rao procedure. Following the notation of the previous section we first establish the general form of the Yates-Rao procedure.

<u>Theorem 4.7.1</u> The untruncated form of Yates-Rao procedure is given by

$$\underset{\sim}{\phi}_6 = (\phi_{61},\ldots,\phi_{6s});$$
$$\phi_{6i} = S_1/[b_iS_1 + c_i\{S_2 + \sum_{j=1}^{s} (1-\gamma_{0j})W_j\}]$$

where

where

$$b_i = 1 - \alpha_i\gamma_{oo}/[e_2 + s - \gamma_{oo}]$$

$$c_i = e_1\, \alpha_i/[e_2 + s - \gamma_{oo}] \qquad (4.7.1)$$

$$\gamma_{oo} = \sum_{j=1}^{s} \gamma_{oj}$$

(ii) The truncated form of Yates-Rao procedures is given by $\phi_6^* = (\phi_{6i}^*, \ldots, \phi_{6s}^*)$, $\phi_{6i}^* = \phi_{6i}$ if $0 \le \phi_{6i} \le \gamma_{oi}$; $= \gamma_{oi}$ otherwise.

Proof The proof is straightforward once it is noted that the adjusted block sum of squares can be expressed in the form:

$$SS_B \geq S_2 + \sum_{j=1}^{s} [W_j/(\alpha_j^{-1} + 1)]$$

which is a straightforward generalization of statement (2.13) in Roy and Shah (1962).

Remark 4.7.1 The expression for the Yates-Rao estimator given by theorem 4.7.1(ii) is a generalization of a similar expression [obtainable from Roy and Shah (1962)] given in Khatri and Shah (1975) for the special case of connected binary equireplicate proper block designs.

Graybill and Weeks (1959) showed that in case of BIBD, the Yates-Rao procedure is based on the minimal sufficient statistic. The work of Roy and Shah (1962) showed that this is true for all connected binary equireplicate proper block designs. In view of our result in Section 4.4 our Theorem 4.7.1 shows that this is true for all proper block designs.

The question of unbiasedness of Yates-Rao estimators has been examined by Graybill and Weeks (1959), Graybill and Seshadri (1960) and Roy and Shah (1962). Of these the most general result is contained in Roy and Shah (1962) who established the unbiasedness of the Yates-Rao estimators for all connected binary equireplicate proper block designs. In view of our Theorem 4.7.1 it follows from our Theorem 4.5.1 that the Yates-Rao estimators (both truncated and untruncated) are unbiased for any proper block design. Unlike several others proposed in the literature the Yates-Rao procedure utilizes all between block comparisons. So far the only known designs for which the truncated form of it, fails to give uniform improvement over the intra-block estimators are (i) the Linked block designs[introduced by Youden (1951)] with $b \leq 6$, as shown by Shah (1964); there are many such designs [e.g. the symmetrical BIBD with $b = v = 4, k=3$; several others which are not BIBD can be found in Roy and Laha (1966)]. (ii) The asymmetrical BIBD with $v = 4$, $b = 6$, $k = 2$ shown by the Bhattacharya (1978). While the properties of the Yates-Rao procedure remains largely unexplored, the desire to construct estimators better than those by the procedure ϕ_0 defined in section 4.5, has led to several modifications of it, of which a fairly comprehensive account has been given in the previous section. Simulation studies by Shaarawi et.al. (1975) as well as numerical comparisions by Khatri and Shah (1975) show that Yates-Rao estimator compares favourably with that of Khatri and Shah (1974). It is therefore, both interesting and important to examine the conditions under which the Yates-Rao procedure is better than ϕ_0. The question has been resolved by Shah (1964) for all Linked block designs which include all symmetrical BIBD's and by Bhattacharya (1978) for all asymmetrical BIBD's listed in Fisher and Yates (1963) with the exception of one (the BIBD with $v = 5$, $b = 10$, $k = 2$). The results obtained by the author (1978) were applicable only to designs belonging to the D_1-class [defined in Shah (1964)], other than Linked block designs; the asymmetrical BIBD's were treated as special cases. In the present work we extend those results to any proper block design for which

$$b_j \geq 0 \text{ for all } j \tag{4.7.2}$$

where b_j's are as defined in (4.7.1). It can be seen that Linked block designs treated in Shah (1964) but excluded in Bhattacharya (1978) belong to the larger class of designs satisfying (4.7.2) which we treat here. In view of (4.7.1), (4.7.2) is equivalent to

$$\gamma_{**} \geq \gamma_{oo} / (e_2 + s), \quad \text{for all } j \tag{4.7.3}$$

where $\gamma_{**} = \underset{i}{\text{Min}}\, \gamma_{oi}$. We shall also assume that

$$e_2 + s \geq 3 \tag{4.7.4}$$

In all theorems which follow, the conditions (4.7.3) and (4.7.4) are assumed without explicitly mentioning those. For the sake of simplicity we first consider ϕ_6. We treat the two cases (1) $b_i > 0$ (2) $b_i = 0$ separately.

Case (1) $\quad \phi_{6i} = a_i \Psi \quad$ where

$$a_i = 1/b_i \ ; \tag{4.7.5}$$

$$\Psi = S_1[S_1 + (c_i/b_i)\{S_2 + \sum_{j=1}^{s} (1-\gamma_{oj})W_j\}] \ .$$

Let V_1, V_2, V_3, V_4, u_j and V be as defined in the previous section. Let

$u = \sum_{j=1}^{s} \gamma_{oj} u_j$. Note that

$$S_2 + (1-\gamma_{oi})W_{i*} + \sum_{j \neq i} (1-\gamma_{oj})W_j = \sigma_*^2 V_2 + \sum_{j=1}^{s} (1-\gamma_{oj})(\alpha_j^{-1}\sigma_1^2 + \sigma_*^2)V_{3j}$$

$$= \sigma_1^2[\rho_* + w(1-\rho_*)]V_4 = [\sigma_1^2/(\alpha_i\gamma_i)][p(w)-q(w)\gamma_i]V_4$$

where $p(w) = 1-w$, $q(w) = 1-w/\gamma_{oi}$. Hence $\bar{\Psi}$ can be written as

$$\bar{\Psi} = V_1/[\gamma_i V_1 + d_i\, h(w;\gamma_i)V_4] \tag{4.7.6}$$

where $h(w;\gamma_i) = p(w) - q(w)\gamma_i$;

$$d_i = c_i/(\alpha_i b_i) = c/b_i; \quad c = e_1/(e_2+s-\gamma_{oo}). \qquad (4.7.7)$$

Since V_1 is almost sure positive (4.7.6) is equivalent to

$$\bar{\Psi} = 1/[\gamma_i + d_i\, h(w,\gamma_i)V]$$

It can be seen that with the match up $w \sim x$, $V \sim y$, $\bar{\Psi}$ matches up with f of Theorem A.4 and satisfies all conditions of Parts B and C of that theorem if we take $\gamma_o = \gamma_{oi}$. The support of w is $S = (0,\gamma_*)$, where $\gamma_* = \max_j \gamma_{oj}$, $S_* = \{t|t \in S; q(t) > 0\} = \{t|0 < t < \gamma_{oi}\}$. We have

$$a_o = EV^{-1}/EV^{-2} = (e_2+s-2)/(e_1+2); \qquad (4.7.8)$$

$$\delta_1 = \inf_{t\in S} p(t) = 1-\gamma_* ; \quad \delta_2 = \inf_{t\in S_*} p(t)/q(t) = 1;$$

$$\delta_3 = \inf_{t\in S_*} h(t);\cdot) = 0; \quad \delta_4 = \inf_{t\in S_*} h(t;\gamma_{oi}) = 1-\gamma_{oi} ;$$

$$\delta_5 = \max(\delta_2, d_i a_o \delta_3) = 1;$$

$$\pi_1 = \mathrm{Min}(d_i a_o \delta_1, \delta_5) = \mathrm{Min}[1, \quad (1-\gamma_*)] \qquad (4.7.9)$$

$$\pi_2 = \mathrm{Min}(d_i a_o \delta_1, \delta_6) = d_i a_o (1-\gamma_*)$$

since $\delta_6 = \max(\delta_2, d_i a_o \delta_4) \geq d_i a_o \delta_4 \geqslant d_i a_o \delta_1$,

$$\bar{g}_*(0) = E(1-w)^{-1}/E[(1-w)\mathrm{Min}(1-w,1-\gamma_{oi})]^{-1} ; \qquad (4.7.10)$$

$$g_*(1) = [(1-\gamma_{oi})/\gamma_{oi}]\, Ew^{-1}/Ew^{-2}; \; \delta_8 = \mathrm{Min}[\bar{g}_*(0), g_*(1)]$$

In view of the above calculation, an application of Theorem A.4 in the present context gives $\nu(\Psi) \geq \max(\pi_1, d_i a_o \delta_8)$; $\nu_*(\Psi) \geq \mathrm{Max}[\pi_2, d_i a_o \bar{g}_*(0)] = d_i a_o \bar{g}_*(0)$ since $w \leq \gamma_* \Rightarrow \mathrm{Min}(1-w, 1-\gamma_{oi}) \geq 1-\gamma_* \Rightarrow \bar{g}_*(0) \geq 1-\gamma_*$. Note that $[E\bar{\Psi}/E\bar{\Psi}^2] = d_i a_o\, g(0)$, where

$$g(0) = E(1-w)^{-1}/E(1-w)^{-2}. \qquad (4.7.11)$$

Hence

$$\inf_{\rho_* > \rho_{*o}} E\bar{\Psi}/E\bar{\Psi}^2 = \inf_{\gamma_i < 1/(1+\alpha_i\rho_{*o})} E\bar{\Psi}/E\bar{\Psi}^2 \leq da_o g(0), \text{ for any } \rho_{*o} \geq 0$$

In view of corollary 4.5.1 we now conclude that

(i) $z_i(\phi_6)$ is better than x_i for all $\rho_* > 0$ if $a_i \leq 2\max(\pi_1, da_o\delta_8)$, which [in view of (4.7.5), (4.7.7) and (4.7.9)] is equivalent to $A_i \geq 1/2$, where

$$A_i = \text{Max}[\text{Min}\{b_i, ca_o(1-\gamma_*)\},\ ca_o\,\delta_8) \tag{4.7.12}$$

(ii) $z_i(\phi_6)$ is better than x_i for $\rho_* \geq 1$ if $a_i \leq 2da_o\bar{g}(0)$, which is equivalent to $A_{i*} \geq 1/2$, where

$$A_{i*} = ca_o\,\bar{g}(0) \tag{4.7.13}$$

(iii) $z_i(\phi_6)$ is better than x_i for all $\rho_* > \rho_{*o}$ for some given $\rho_{*o} \geq 0$ if $a_i \leq 2da_o g(0)$, which is equivalent to

$$A_{io} \geq 1/2 \quad \text{where} \quad A_{io} = ca_o g(0) \tag{4.7.14}$$

<u>Case (2)</u> $\phi_{6i} = a_i\Psi$ where $a_i = 1/c_i$;

$\Psi = S_1/[S_2 + \sum_{j=1}^{s} (1-\gamma_{oj})W_j]$. Then, $\bar{\Psi} = (V/\alpha_i^{-1})/h(w;\gamma_i)$; $E\bar{\Psi}/E\bar{\Psi}^2 = \alpha_i^{-1}a_o Ef/Ef^2$

where $f = 1/h(w;\gamma_i)$. Hence, using Theorem A.3 we obtain

$$\nu(\Psi) \geqslant \alpha_i^{-1}a_o\delta_8\ ;\ \nu_*(\Psi) \geqslant \alpha_i^{-1}a_o\bar{g}(0)$$

Also in the same way as in the previous case

$$\inf_{\rho_* > \rho_{*o}} E\bar{\Psi}/E\bar{\Psi}^2 \leq \alpha_i^{-1} a_o g(0)$$

In view of corollary 4.5.1 we conclude

(i) $z_i(\phi_6)$ is better than x_i for all $\rho_* \geq 0$ if $a_i \leq 2\alpha_i^{-1}a_o\delta_8$, a condition which is equivalent to $A_i \geq 1/2$

(ii) $z_i(\phi_6)$ is better than x_i for all $\rho_* > 1$ if $a_i \leq 2\alpha_i^{-1}a_o\bar{g}(0)$, a condition which is equivalent to $A_{i*} \geq 1/2$.

(iii) $z_i(\phi_6)$ is better than x_i for all $\rho_* > \rho_{*o}$ for some given $\rho_{*o} \geq 0$ if $a_i \leq 2\alpha_i^{-1} a_o g(0)$, a condition which is equivalent to $A_{io} \geq 1/2$. This completes the analysis for case (2), where the final result is seen to be the same as in case (1). We have thus proved theorem 4.7.2.

<u>Theorem 4.7.2</u> Let A_i, A_{i*}, A_{io} be as defined in (4.7.12) - (4.7.14). Then (i) $z_i(\phi_6)$ is better than x_i for all $\rho_* > 0$ if $A_i \geq 1/2$, (ii) $z_i(\phi_6)$ is better than x_i for all $\rho_* > 1$ if $A_{i*} \geq 1/2$ (iii) $z_i(\phi_6)$ is better than x_i for all $\rho_* > \rho_{*o}$ (for some given $\rho_{*o} \geq 0$) only if $A_{io} \geq 1/2$.

From (4.7.16), observe that, $\operatorname{Min}_i \bar{g}_*(0) = 1-\gamma_*$; $\operatorname{Min} g_*(1) \leq 1-\gamma_*$; and hence $\operatorname{Min}_i \delta_8 = \operatorname{Min}_i g_*(1)$. Let

$$\begin{aligned} A_* &= ca_o(1-\gamma_*) \\ A &= \operatorname{Min}[\operatorname{Max}(b_*, A_* M), A_*] \; ; \; A_0 = A_* M_* \end{aligned} \qquad (4.7.15)$$

where

$$b_* = \operatorname{Min}_i b_i ; M = \operatorname{min}_i g_*(1)/(1-\gamma_*) ; M_* = \operatorname{Min}_i g(0)/(1-\gamma_*) \qquad (4.7.16)$$

Then from (4.7.12) - (4.7.14) we see that $\operatorname{Min}_i A_i = A$; $\operatorname{Min}_i A_{i*} = A_*$; $\operatorname{Min}_i A_{io} = A_o$. Hence theorem 4.7.2 leads to theorem 4.7.

<u>Theorem 4.7.3</u> Let A, A_* and A_o be as defined in (4.7.15). Then (i) ϕ_6 is better than ϕ_o for all $\rho_* > 0$ if $A \geq 1/2$, (ii) ϕ_6 is better than ϕ_o for all $\rho_* > 1$ if $A_* \geq 1/2$, (iii) ϕ_6 is better than ϕ_o for all $\rho_* > \rho_{*o}$ (for some given $\rho_{*o} \geq 0$) only if $A_o \geq 1/2$.

In order to apply theorem 4.7.3, we need computable expressions for $g(0)$ and $g_*(1)$. These can be obtained by the same technique as used for M_1, M_2, M_3 in the previous section. Let p_4, p_5 satisfy the conditions

$$0 < p_4 < 2(1-\gamma_*); \quad 0 < p_5 < 2\gamma_{**}$$

Let $f_j^{(4)}$ stand for the expression obtained from f_j (of lemma 4.6.1) when $p_* = p_4$; $\alpha = s + 1$;

$$m_j = 1 \text{ if } j \neq i, j = 1,2,\ldots,s; \quad d_j^* = 1-\gamma_{oj} \text{ if } j = 1,\ldots,s$$
$$= 3 \text{ if } j = i \qquad\qquad = 1 \text{ if } j = s+1$$
$$= e_2 \text{ if } j = s+1.$$

Let $f_j^{(5)}$ be obtained from $f_j^{(1)}$ by replacing the value of the argument p_* by p_5 and that of d_j for each $j = 1,\ldots,s$ by $d_j = \gamma_{oj}$. Let $(F_1^{(4)}, F_2^{4})$ and $(F_1^{(5)}, F_2^{(5)})$ be obtained from $(F_1^{(1)}, F_1^{(2)})$ given by (4.6.15) by replacing $f_j^{(1)}$ by $f_j^{(4)}$ and $f_j^{(5)}$, respectively. Then, from (4.7.11) and (4.7.10),

$$g(0) = [p_4 F_1^{(4)}/F_2^{(4)}]/(e_2+s-2) \tag{4.7.17}$$

$$g_*(1) = (1-\gamma_{oi})/\gamma_{oi}][p_5 F_1^{(5)}/F_2^{(5)}]/(e_2+s-2)$$

From (4.7.1), it is easy to see that

$$b_* = (e_2 + s - \gamma_{oo}/\gamma_{**})/(e_2+s-\gamma_{oo}) \tag{4.7.18}$$

We are unable to obtain similar expressions for M, M_* but we can use (4.7.17) to compute $g(0)$, $g^*(1)$ for each i and hence M, M_* using (4.7.16). In view of the difficulty in computing M_* and hence A_o consider

$$M_{io} = E(1-w); \quad M^* = \operatorname*{Min}_i M_{oi}/(1-\gamma_*).$$

Then from (4.7.11) we see that $g(0) \leq M_{io}$ for all i and hence (4.7.15) gives

$$M_* \leq M^*$$

It is easy to see that $M_{io} = [e_2+s+2-\gamma_{oo} -2\gamma_{oi}]/(e_2+s+2)$. Hence

$$M^* = (1-\gamma_*)^{-1} [e_2+s+2-\gamma_{oo} - 2\gamma_*]/(e_2+s+2) \tag{4.7.19}$$

From (4.7.15) and (4.7.19) we have $A_o \leq A^*$, where

$$A^* = A_* M^* \tag{4.7.20}$$

Hence part (iii) of Theorem 4.7.3 leads to Corollary 4.7.1.

Corollary 4.7.1 $\underset{\sim}{\phi}_6$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* \geq \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) only if $A^* \geq 1/2$, where A^* is as defined in (4.7.20).

In the special case where $e_2 = 0$ we have [in view of (4.7.3)] $b_i = 0$ for all i and $w = \gamma_*$. Hence, $A = A_0 = A_*$. Therefore Theorem 4.7.3 gives :

Corollary 4.7.2 Assume that $e_2 = 0$. Then, $\underset{\sim}{\phi}_6$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) iff $A_* \geq 1/2$

We now turn our attention to $\underset{\sim}{\phi}_6^*$ (defined in Theorem 4.7.1). Combining the results of theorem 4.7.3 and part (ii) of theorem 5.3.3, we see that $\underset{\sim}{\phi}_6^*$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > 0$ if $A_* \geq 1/2$. On the other hand, it is easy to see that $\underset{\sim}{\phi}_6^*$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) only if $A_0 \geq 1/2$ since $P(\phi_{6i}^* \neq \phi_{6i}) \to 0$ as $\rho_* \to \infty$, which implies that $[E\bar{\phi}_{6i}^*/E\,\bar{\phi}_{6i}^{*2}]_{\gamma_i=0} = [E\bar{\phi}_{6i}/E\bar{\phi}_{6i}^2]_{\gamma_i=0}$, for every i. Hence, we have,

Theorem 4.7.4 $\underset{\sim}{\phi}_6^*$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) (i) if $A_* \geq 1/2$ and (ii) only if $A_0 \geq 1/2$.

Corollary 4.7.3 $\underset{\sim}{\phi}_6^*$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$ only if $A^* \geq 1/2$.

Corollary 4.7.4 Assume that $e_2 = 0$. Then, $\underset{\sim}{\phi}_6^*$ is better than $\underset{\sim}{\phi}_0$ for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) iff $A_* \geq 1/2$.

Remark 4.7.1 Corollary 4.7.4 is a generalization of a similar result in Shah (1964) concerning linked block designs. [see cases (2c) and (2d) of application in special cases to be given shortly].

We shall be now consider application of theorem 4.7.3 and theorem 4.7.4 to some special cases. Our aim is to consider the special features of the designs in each case and derive, if possible, more explicit expressions for the basic quantities required for application from those given earlier in the text for the general case. The quantities which are necessary but not considered have to be obtained with the help of the earlier expression for the general case.

(1) <u>Equireplicate designs</u> We have $s=t-\ell$; $p = v-\ell$; $q = t-1$; $\alpha_j = (rk-\chi_j)/\chi_j$. Hence $e_1 = b(k-1)-(v-\ell)$; $e_2 = b-t$; $\gamma_{oj} = \chi_j/(rk)$

$$\gamma_* = \chi_*/(rk); \ \gamma_{**} = \chi_{**}/(rk); \ \gamma_{oo} = (t_*-rk)/(rk)$$

where χ_* and χ_{**} are the largest and the smallest latent roots of NN' and $t_* = \text{tr } NN'$. We have,

$$ca_o = rke_2(b-\ell-2)/\ (e_1+2)\{rk(b-\ell+1)-t_*\}]$$

$$A_* = e_1(b-\ell-2)(rk-\chi_*)/[(e_1+2)\ \{rk(b-\ell+1)-t_*\}]$$

$$b_* = rk[\chi_{**}(b-\ell)+rk-t_*]/[rk(b-\ell+1)-t_*]$$

$$M^* = [rk(b-\ell+3)-t_*-\chi_*]/[(b-\ell+2)(rk-\chi_*)]$$

Our results are applicable if $\chi_{**} \geq (t_*-rk)/(b-t)$ and $b \geq \ell + 3$.

(1a) <u>Binary equireplicate designs</u> We have, $t_* = vr = bk$. Hence $\gamma_{oo}=(b-r)/r$

$$ca_o = re_1(b-\ell-2)/[(e_1+2)\ \{r(b-\ell+1)-b\}]$$

$$A_* = e_1(b-\ell-2)(rk-\chi_*)/[k(e_1+2)\{r(b-\ell+1)-b\}]$$

$$b_* = r[\chi_{**}(b-\ell)-k(b-r)]/[r(b-\ell+1)-b]$$

$$M^* = [rk(b-\ell+3)-bk-2\chi_*]/[(b-\ell+2)(rk-\chi_*)].$$

Our results are applicable if $\chi_{**} \geq k(b-r)/(b-\ell)$ and $b \geq \ell + 3$.

(1b) <u>Connected binary equireplicate designs</u>: We have $\ell = 1$; $e_1 = b(k-1)-(v-1)$. Hence

$$ce_o = re_1(b-3)/[b(r-1)(e_1+2)]$$

$$A_* = e_1(b-3)(rk-\chi_*)/[(bk(e_1+2)(r-1)]$$

$$b_* = r[\chi_{**}(b-1)-k(b-r)]/[b(r-1)]$$

$$M^* = [rk(b+2)-bk-2\chi_*]/[(b_1+2)(rk-\chi_*)].$$

Our results are applicable if $\chi_{**} \geq k(b-r)/(b-1)$ and $b \geq 4$.

(2) <u>Designs for which</u> γ_{oj} = <u>constant for all j</u> : Let γ_{o*} denote the common value of γ_{oj}. Then $\gamma_* = \gamma_{**} = \gamma_{o*}$; $\gamma_{oo} = s\gamma_{o*}$. Hence,

$$ce_o = e_1(e_2+s-2)/[(e_1+2)\{e_2+s(1-\gamma_{o*})\}]$$

$$b_* = e_2/[e_2+s(1-\gamma_{o*})]$$

$$A_* = ce_o(1-\gamma_{o*})$$

$$M^* = [e_2+(s+2)(1-\gamma_{o*})]/[(e_2+s+2)(1-\gamma_{o*})].$$

We can also simplify the expressions for M and M_*. From (4.6.11) note that the generating function of $f_j^{(4)}$ is given by

$$\sum_{j=0}^{\infty} f_j^{(4)} z^j - p_4^{(e_2+s+2)/2} (1-\gamma_{o*})^{-(s+2)/2}[1-(1-p_4)z]^{-e_2/2}[1-(1-\frac{p_4}{1-\gamma_{o*}})z]^{(s+2} \quad (4.7.21)$$

where p_4 can be suitably chosen subject to the condition : $0<p_4<2(1-\gamma_{o*})$. Let us take $p_4 = 1-\gamma_{o*}$. Then (4.7.21) becomes $\sum_{j=0}^{\infty} f_j^{(4)} z^j = (1-\gamma_{o*})^{e_2/2}(1-\gamma_{o*}z)^{-e_2/2}$

Hence

$$f_j^{(4)} = (1-\gamma_{o*})^{e_2/2}[(e_2/2 + j-1)_{(j)}/j!]\gamma_{o*}^j$$

Using this, (4.7.16) gives

$$M_* = \frac{\sum_{j=0}^{\infty}[(e_2/2+j-1)_{(j)}/\{j!(e_2+s+2j)\}]\gamma_{o*}^j}{(e_2+s-2)\sum_{j=0}^{\infty}[(e_2/2+j-1)_{(j)}/\{j!(e_2+s+2j)(e_2+s+2j-2)\}]\gamma_{o*}^j} \quad (4.7.22)$$

Also, it can be seen that, $w/\gamma_{0*} \sim \beta(\frac{s+2}{2}, \frac{e_2}{2})$. Hence, (4.7.16) gives

$$M = (s-2)/(e_2+s-2)$$

It is easy to see that the assumption (4.7.3) is always satisfied. Hence, our results are applicable provided only $e_2+s \geq 3$.

(2a) D_1-class designs : As defined in Shah (1964), these are connected binary equireplicate designs for which χ_j = constant for all j, a condition which is equivalent to γ_{0j} = constant for all j. We have $\gamma_{0*} = (b-r)/[r(t-1)]$; s, e_1, e_2, ca_0 as in case (1b). Hence

$$A_* = e_1(b-3)(rt-b)/[b(e_1+2)(r-1)(t-1)]$$

$$b_* = r(b-t)/[b(r-1)];$$

$$M = (t-3)/(b-3); \; M^* = [b(r-1)(t-1)+2(rt-b)]/[(b+1)(rt-b)];$$

M_* is given by (4.7.22) where e_2, s and γ_{0*} have the value specified here. Our results are applicable provided only $b \geq 4$.

(2b) Balanced incomplete block designs : These are D_1-class designs for which t = v. Hence, $\gamma_{0*} = (b-r)/[r(v-1)]$;

$$A_* = e_1(b-3)(k-1)/[(e_1+2)(r-1)(v-1)]$$

$$b_* = (r-k)/(r-1) ;$$

$$M = (v-3)/(b-3); \; M^* = [(r-1)(v-1)+2(k-1)]/[(b+1)(k-1)]$$

M_* is given by (4.7.22), where $e_2 = b-v$; $s = v-1$.

(2c) Designs for which $e_2 = 0$: It is easy to see that we must have γ_{0j} constant for all j. Furthermore, $w = \gamma_{0*}$. It then follows that $b_i = 0$ for every i and that $g(0) = g_*(1) = 1-\gamma_{0*}$. Hence we have

Table 4.7.1 : Comparison of ϕ_6 and ϕ_6^* with ϕ_0

[for asymmetrical BIBD's listed in Fisher and Yates (1963)]

Design No	r	v	b	k	A_*	A	A*	A_0	* Conclusion
1	3	4	6	2	3000		.3429	-	III
2	4	5	10	2	.4375	-	.5568	.5474	IV
3	5	6	15	2	.6000	.6000		-	I
4	5	6	10	3	.6176	.5000	-	-	I
5	6	5	10	3	.6222	.6000	-	-	I
6	6	7	21	2	.5294	.5294	-	-	I
7	6	13	26	3	.7302	.6000	-	-	I
8	7	8	28	2	.5435	.5435	-	-	I
9	7	8	14	4	.7432	.5000	-	-	I
10	7	15	35	3	.7356	.6667	-	-	I
11	8	9	36	2	.5500	.5500	-	-	I
12	8	9	18	4	.7701	.5714	-	-	I
13	8	25	50	4	.8262	.5714	-	-	I
14	9	10	45	2	.5526	.5526	-	-	I
15	9	10	30	3	.7217	.7217	-	-	I
16	9	10	18	5	.8077	.5000	-	-	I
17	9	19	57	3	.7347	.6667	-	-	I
18	9	28	63	4	.8232	.6250	-	-	I
19	10	6	15	4	.7619	.6667	-	-	I
20	10	9	18	5	[illegible]	.5554	-	-	I
21	10	11	55	2	.5909	.5909	-	-	I
22	10	21	70	3	.7322	.7322	-	-	I
23	10	41	82	5	.8717	.5556	-	-	II
24	6	[illegible]	12	[illegible]	.6667	[illegible]	-	-	I
25	5	6	20	4	.8158	.6223	-	-	II
26	6	10	15	4	.7579	.4421	-	-	I
27	6	25	30	5	.8816	.7183	-	-	I
28	7	15	21	5	.8333	.5556	-	-	I
29	8	9	12	6	.7738	.5159	-	-	I
30	8	21	28	6	.8782	.6323	-	-	I
31	8	49	56	7	.9399	.8158	-	-	II
32	9	10	15	6	.8088	.4718	-	-	I
33	9	16	24	6	.8608	.5329	-	-	I
34	9	28	36	7	.9071	.6872	-	-	I
35	9	46	69	6	.9106	.5933	-	-	I
36	9	64	72	8	.9540	.8434	-	-	I
37	10	21	30	7	.8889	.5926	-	-	I
38	10	36	45	8	.9256	.7273	-	-	I
39	10	51	85	6	.9063	.5305	-	-	I
40	10	81	90	9	.9637	.8640	-	-	I

* Conclusions are coded as follows:

I Both ϕ_6 and ϕ_6^* are better than ϕ_0 for all $\rho_* > 0$

II ϕ_6 is better than ϕ_0 for all $\rho_* > 1$ and ϕ_6^* is better for all $\rho_* > 0$

III Neither ϕ_6 nor ϕ_6^* is better than ϕ_0 for all $\rho_* > \rho_{*0}$ (if $\rho_{*0} \geq 0$ is given)

IV No conclusion could be made

$$A = A_* = A^* = ce_0(1-\gamma_{0*}) = e_1(s-2)/([s\ (e_1+2)]$$

Using theorem 4.7.3 and theorem 4.7.4 together with corollary 4.7.2 and corollary 4.7.4, we conclude : Both ϕ_6 and ϕ_6^* are better than ϕ_0 for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) iff

$$e_1(s-2)/[s(e_1+2)] \geq 1/2 \text{ or, equivalently}$$

$$(e_1-2)(s-4) \geq 8 .$$

This result is a generalization of a similar result in Shah (1964) concerning Linked designs.

(2d) <u>Linked block designs</u> These are D_1-class designs for which t = b. Hence $e_2 = 0$ in addition to (4.7.3). We have s = b-1. Hence from the result obtained in the previous case, we have:

Both ϕ_6 and ϕ_6^* are better than ϕ_0 for all $\rho_* > \rho_{*0}$ (for some given $\rho_{*0} \geq 0$) iff $(e_1-2)(b-5) \geq 8$. The part of the above result which relates to ϕ_6^* was proved by Shah (1964) in a completely different way. Our result which relates to ϕ_6 as well is stronger.

The actual application of our results to any given design, is a routine exercise. What we have to do can be stated as follows: Examine if the assumptions (4.7.3) and (4.7.4) are satisfied. If not we are unable to conclude anything. If yes, compute A_*. If $A_* < \frac{1}{2}$, compute A^*. If $A^* < \frac{1}{2}$, we conclude that neither ϕ_6 nor ϕ_6^* is better than ϕ_0 for all $\rho_* > \rho_{*0}$ (if $\rho_{*0} \geq 0$ is given). If $A^* \geq \frac{1}{2}$ compute A_0. If $A_0 < \frac{1}{2}$, we conclude the same as in the case: $A^* < \frac{1}{2}$. If $A_* < \frac{1}{2}$ and $A_0 \geq \frac{1}{2}$, we are unable to conclude anything. If $A_* \geq \frac{1}{2}$, compute A. If $A \geq \frac{1}{2}$, we conclude that both ϕ_6 and ϕ_6^* are better than ϕ_0 for all $\rho_* > 0$. If $A_* \geq \frac{1}{2}$ but $A < \frac{1}{2}$, we conclude that ϕ_6 is better than ϕ_0 for all $\rho_* > 1$ and that ϕ_6^* is better for all $\rho_* > 0$. [In this case we are unable to decide if ϕ_6 is better than

ϕ_0 for $\rho_* < 1$]. For illustration, we present a table of A_*, A, A^*, A_0 (computed according to the programme just described) for all asymmetrical BIBD's listed in Fisher and Yates (1963). All entries are easily obtained with the exception of the value of A_0 for design no. 2. This is obtained by the formula : $A_0 = A_* M_*$, where M_* is given by (4.7.7) [Earlier, the author (1978), had to use numerical integration since he had not discovered the above expression at the time]. The conclusions are given in the table and will not be repeated here.

CHAPTER 5

USE OF MODIFIED ESTIMATORS IN RECOVERY OF INTER-BLOCK INFORMATION

5.1 Introduction

In the analysis of a proper block design with recovery of inter-block information, an estimator (say $\hat{\mu}$) of a given treatment contrast (estimable from both intra-block and inter-block analysis) is generally obtained as a weighted average of the intra-block and inter-block estimators of that contrast using suitably chosen random weights. The weight (say ϕ) given to the inter-block estimator can be expressed in the form $\phi = 1/(1+\hat{\eta})$ and one may regard $\hat{\mu}$ as an analogue of the best linear unbiased combination of the intra-block and the inter-block estimators in which the unknown ratio (say η) of the variance of the inter-block estimator to that of the intra-block estimator is replaced by $\hat{\eta}$. η is generally a known multiple (depending on the design) of the ratio (say ρ_*) of the inter-block error variance (per plot) to the intra-block error variance. Under the infinite models generally used in the literature, ρ_* exceeds unity and then, η exceeds a known quantity (say η_0). But the value of $\hat{\eta}$ may turn out to be less than η_0 and in such a case, it is usually recommended that the value of $\hat{\eta}$ be replaced by η_0. Although, this truncation procedure first proposed by Yates (1939) is widely used in practice very little theoretical discussion seems to be available in the literature concerning this and other alternatives to this.

Stein (1966) considered a particular estimator of the same form as $\hat{\mu}$ discussed above with a non-negative ϕ, and proposed a truncation procedure based on ϕ according to which if ϕ turns out to be greater than $\phi_0 = 1/(1+\eta_0)$ the value of ϕ is replaced by ϕ_0. Note that if ϕ is non-negative than $\hat{\eta} < \eta_0 \iff \phi > \phi_0$; therefore the truncation procedure proposed by Yates and Stein are equivalent. Stein (1966) conjectured that his truncation procedure would lead to a better estimator (say $\hat{\mu}_*$) than the original estimator $\hat{\mu}$. Shah (1971) formally proved a result supporting this conjecture in the sense that $\hat{\mu}_*$ is better than μ for all $\eta \geq \eta_0$, under certain assumptions. Shah (1971) did not make the necessary distinction between the modification of $\hat{\mu}$ suggested by Yates and Stein but it can be seen that his condition 2.6, which he felt to be unnecessarily restrictive, was, in fact, necessary for ϕ to be non-negative and hence for the two suggestions to be equivalent, in his case. When the assumption that ϕ is non-negative is not satisfied [e.g. ϕ based on untruncated estimator of ρ by the customary Yates-Rao procedure for some of the treatment contrasts in case of the PBIB design R1 in Bose, Clatworthy and Shrikhande (1954)], one can extend Stein's suggestion in many ways but none of these agree with that suggested by Yates.

Section 2 contains the preliminary notations and results. Section 3 is then devoted to the special case in which ϕ is non-negative (so that Yates' and Stein's suggestions are equivalent). Theorem 5.3.1 improves Shah's results and supports Stein's conjecture in the same sense as in Shah (1971) under the milder assumption that (i) $\phi \in \Phi$, where Φ is as defined in the next section (ii) $\phi \geq 0$ a.s. (iii) $E\hat{\mu}$ exists. [Note that Shah imposes the unnecessary restriction on ϕ by assuming that his $\hat{\rho}^*$ ($\hat{\rho}_*$ in our notation) is of the form (2.3) of his paper and satisfies the condition $V(w_B) < \infty$ in the notation of his paper]. Two more results in support of the truncation procedure are established in theorem 5.3.3 under the same assumption. The first of these tells us that for small η (to be precise $\eta \leq 1 + 2\eta_0$) $\hat{\mu}_*$ is better than the intra-block estimator [illegible] other result asserts that (i) $\hat{\mu}$ is better than the intra-block estimator [illegible] all $\eta \geq 1+2\eta_0$ implies (ii): $\hat{\mu}_*$ is better than the intra-block estimator for all $\eta \geq 0$. Note that if (i) holds then using theorem 5.3.1 one can only assert (iii): $\hat{\mu}_*$ is better than the intra-block estimator for all $\eta \geq 1+2\eta_0$, which is weaker than (ii).

In Section 5.4, we relax the assumption that ϕ is non-negative and generalize the results of section 5.3. Two fundamental results are contained in theorems 5.4.1 and 5.4.2. Theorem 5.4.1 gives us a class of modified estimators for which the result of theorem 5.3.1 holds. It also tells us that any combined estimator which with a positive probability gives to the inter-block estimator a weight which is either negative or in excess of $1/(1+\eta_0)$, is inadmissible with respect to the restricted parameter set $[\eta \geq \eta_0]$. Theorem 5.4.2 gives a class of modified estimators for which the results of theorem 5.3.1 and 5.3.3 both hold. Finally, we consider five modifications of $\hat{\mu}$ emerging from Stein's suggestion in addition to that of Yates. It is shown that three of the five modifications of $\hat{\mu}$ emerging from Stein's suggestions are either inadmissible or almost sure identical with one of the remaining two. We exclude these three from further consideration. Theorem 5.4.3 then makes a theoretical comparison among the remaining three. It is seen that Yates' modification of $\hat{\mu}$ is better than the other two for small η (to be precise for $\eta < 1+2\eta_0$ in one case and $\eta \leq \eta_0$ in the other case) and is worse than each of the other two for large η (to be precise for $\eta > 1+2\eta_0$). Comparison of the other two shows that one of them is better than the other for small η(to be precise $\eta<1+2\eta_0$). It appears that none of the results in theorems 5.3.1 and 5.3.3 with the exception of part (i) of theorem 5.3.3 need hold for the Yates' modification of $\hat{\mu}$ in this general situation. But, theorem 5.4 shows that the results of theorems 5.3.1 and 5.3.3 both hold for the other two emerging from Stein's suggestion. Theorem 5.4.5 which shows that part (i) of theorem 5.3.3 holds for the Yates' modification of $\hat{\mu}$ is an improvement of the result of Shah (1964a): in the same sense in which Theorem 5.3.1 is an improvement of the result of Shah (1971), which was discussed in the previous paragraph.

5.2 Preliminary Notations and Results

Let x, y, S_1, S_2, w_i, $i = 1,2,\ldots,q$ be independent random variables such that $x \sim N(\mu\ \alpha_0\sigma^2)$, $y \sim N(\mu, \beta_0\sigma_*^2)$, $S/\sigma^2 \sim \chi_m^2$, $T/\sigma_*^2 \sim \chi_n^2$, $w_i/(\alpha_i\sigma^2 + \beta_i\sigma_*^2) \sim \chi_1^2$, $i = 1,2,\ldots,q$ where α_i's and β_i's are known constants and $\mu, \sigma^2, \sigma_*^2$ are unknown parameters. Interprete x, y, S_1, S_2, w_i's as follows: x and y as the intra-block and inter-block estimators of a given canonical contrast which is estimable from both intra-block and inter-block analysis, S_1 and S_2 as the intra-block and inter-block error sum of squares, and w_i's as the squared differences between the inter-block and intra-block estimators of other canonical contrasts which are estimable from both intra-block and inter-block analysis. Finally for convenience, define $w_0 = (y-x)^2$. Let Φ be the class all measurable functions of $S_1, S_2, w_0, w_1, \ldots, w_q$.

Consider the estimator $\hat{\mu}$ of μ given by

$$\hat{\mu} = x + \phi(y-x) \tag{5.2.1}$$

where $\phi \in \Phi$. The following theorem is essentially a restatement of the result in (4.5.2).

<u>Theorem 5.2.1</u> Let $\phi \in \Phi$ and let $\hat{\mu}$ be as defined in 5.2.1. Assume that $E\hat{\mu}$ exists. Then,

$$V(\hat{\mu}) = \alpha_0\sigma^2 [1+E_* h(\phi)] \tag{5.2.2}$$

where $h(t) = t^2(1+\eta) - 2t$, $\eta = \beta_0\sigma_*^2/(\alpha_0\sigma^2)$, E_* stands for the expectation with respect to the density $w_0 f/Ew_0$ and f stands for the joint density of $S, T, w_0, w_1, \ldots, w_q$.

Note that the negative part of $h(\phi)$ is bounded in absolute value by 1 and under the assumption that ϕ is measurable, $v(\hat{\mu})$ always exists (finitely or infinitely). The expression (2.2) will be used repeatedly in obtaining the results of section 5.3 and 5.4. In addition, the following lemma will help to make the results of section 5.3 and 5.4 more transparent.

<u>Lemma 5.2.1</u> Let $h(t) = t^2(1+\eta) - 2t$ where $\eta \geq 0$ and let $c = 1/(1+\eta_0)$, where $\eta_0 \geq 0$. Then

a) for every $t < 0$ and $u \in (t, |t|]$, $h(t) > h(u)$ for all $\eta \geq 0$.

b) for every $t > c$ and $u \in [c, t)$, $h(t) > h(u)$ for all $\eta \geq \eta_0$

c) for every $t \in (0, c]$, $h(t) < h(0)$ for all $\eta < 1+2\eta_0$ $> h(0)$ for all $\eta > 1+2\eta_0$.

d) for every $t \in (0, c)$. (i) $h(t) < h(c)$ for all $\eta \geq 1+2\eta_0$

(ii) $h(t) > h(c)$ for all $\eta \leq \eta_0$

<u>Proof</u> The proof is elementary and is omitted.

Before closing this section we remark that although we concern ourselves only with combined estimators of canonical contrasts which are estimable from both intra-block and inter-block analysis similar results hold for any estimable treatment contrast in view of Theorem 4.5.2.

5.3 Special Case

In this section we assume that ϕ is non-negative so that the modification of μ suggested by Yates (1939) and Stein (1966) coincide. Note that this assumption is implicit in Shah (1971) and is satisfied in many cases. Consider

$$\hat{\mu}_* = x + \phi_*(y-x) \tag{5.3.1}$$

where $\phi_* = \min[\phi, 1/(1+\eta_0)]$ and $\eta_0 \geq 0$ is a given constant. Note that μ_* is the modification of $\hat{\mu}$, which is generally recommended [Yates (1939), Stein (1966)] if it is known that $\eta \geq \eta_0$. Under the infinite model generally used in the literature $\sigma_*^2/\sigma^2 \geq 1$ and η_0 may be taken to be β_0/α_0. We prove

<u>Theorem 5.3.1</u> Let $\phi \in \Phi$. Let $\hat{\mu}$ and $\hat{\mu}_*$ be as defined in 2.1 and 3.1 respectively. Assume that $\phi \geq 0$ a.s. and that $E\hat{\mu}$ exists.

$$V(\hat{\mu}_*) \leq V(\hat{\mu}) \quad \text{for all} \quad \eta \geq \eta_0 \tag{5.3.2}$$

with strict inequality holding unless $\hat{\mu}_* = \hat{\mu}$ a.s.

<u>Proof</u> In view of theorem 5.2.1 it suffices to show that if $\eta \geq \eta_0, h(\phi) \geq h(\phi_*)$ for every ϕ with strict inequality holding for every $\phi > 1/(1+\eta_0)$. But for $\phi \leq 1/(1+\eta_0)$, we have $\phi_* = \phi$ and by Lemma 5.2.1(b), $\phi > 1/(1+\eta_0) \Rightarrow h(\phi) > h(\phi_*)$ since $\eta \geq \eta_0$. Hence the result.

Remark 5.3.1. Note that $\eta_0 > 0$ could be arbitrary and hence the result of theorem 5.3.1 may not be dependent on a rational choice of η_0. However, suppose that the best available knowledge about η is that $\eta \geq \eta_0$. Consider the rivals of $\hat{\mu}_*$ [namely] $\hat{\mu}_1$ and $\hat{\mu}_2$ which use $\hat{\eta}$ truncated at $\eta_1 < \eta_0$ and $\eta_2 > \eta_0$, respectively. Then, $\hat{\mu}_1$ is better than $\hat{\mu}$ but is inadmissible since it would be dominated by any one which uses $\hat{\eta}$ truncated at a value $\eta_1^* \in (\eta_1, \eta_0]$.

other hand $\hat{\mu}_2$ is better than both $\hat{\mu}$ and $\hat{\mu}_*$ for all $\eta \geq \eta_2$ but may se than even $\hat{\mu}$ for some or all $\eta \in [\eta_o, \eta_2)$. For this reason, it is ended that $\hat{\mu}_*$, which uses $\hat{\eta}$, truncated at $\eta = \eta_o$ be used in preference and $\hat{\mu}_2$.

Shah (1964b) considered the truncated form of an estimator and showed it is better than the intra-block estimator for all $\eta \geq 0$ (not merely $\eta \geq \eta_o$). His proof is a bit complicated but a similar result concerning ntruncated form of his estimator is established very easily [See ion 4.6]. An interesting question which arises then is : can we t Shah's result concerning the truncated form from the similar result rning the untruncated form which is a lot more easier to deal with. generally suppose it is known that $\hat{\mu}$ is better than x for all $\eta \geq 0$. we infer from this that $\hat{\mu}_*$ is better than x for all $\eta > 0$. From rem 5.3.1 we can only infer that $\hat{\mu}_*$ is uniformly better than x for $\eta \geq \eta_o$ and the question remains unanswered. Theorem 5.3.3 gives an rmative answer to this question. We first prove,

orem 5.3.2 Let ϕ be as in theorem 3.1. Then

$$\phi \leq 1/(1+\eta_o) \text{ a.s. } \Rightarrow V(\hat{\mu}) \leq V(x) \text{ for all } \eta < 1+2\eta_o$$

h strict inequality holding unless $\hat{\mu} = x$ a.s.

of We have to show that for every $t \in (o, 1/(1+\eta_o)]$, $h(0) \geq h(t)$ for all $< 1+2\eta_o$. This holds by lemma 5.2.1(c). Hence the result.

eorem 5.3.3 below is a simple consequence of theorem 5.3.1 and theorem 3.2 once it is noted that ϕ_* satisfies the condition of theorem 5.3.2.

Theorem 5.3.3. Let $\hat{\mu}$, $\hat{\mu}_*$ be as in theorem 5.3.1. Then

(i) $V(\hat{\mu}_*) \leq V(x)$ for all $\eta < 1+2\,\eta_0$, with strict inequality holding unless $\mu_* = \hat{\mu}$ a.s.

(ii) $V(\hat{\mu}) \leq V(x)$ for all $\eta \geq 1+2\,\eta_0 \implies V(\hat{\mu}_*) \leq V(x)$ for all $\eta \geq 0$.

Remark 5.3.3 Note that $\eta_0 \geq 0$ could be arbitrary and hence the results in theorems 5.3.2 and 5.3.3 may not be dependent on a rational choice of η_0, which is no doubt desirable for reasons given in remarks 5.3.2.

5.4 General Case

In this section we drop the assumption that ϕ is non-negative. Our aim is to generalize the results of theorems 5.3.1 and 5.3.3. Note that the modification of $\hat{\mu}$ suggested by Yates (1939) do not agree with that by Stein (1966) in the general case which we consider in this section; moreover the suggestion by Stein for a non-negative ϕ can be extended in many ways. But, before we consider these alternatives we shall obtain some general results. Our first result contained in theorem 5.4.1 gives us a class of estimators for which the results of theorem 5.3.1 hold. It also tells us that any estimator which with a positive probability gives to the inter-block estimator a weight which is either negative or in excess of $1/(1+\eta_0)$ is inadmissible with respect to the restricted parameter set $[\eta \geq \eta_0]$.

Theorem 5.4.1 Let $\phi \in \Phi$ and let $\hat{\mu}$ be as defined in (5.2.1). Let $\hat{\mu}_* = x + \phi_*(y-x)$ where

$$\phi_* = \phi \quad \text{if} \quad 0 \leq \phi \leq 1/(1+\eta_0)$$

$$= \lambda(\phi) \text{ otherwise}$$

and $\lambda(\phi)$ is a measurable function of ϕ such that

(i) $\lambda(\phi) \in [\phi, |\phi|]$ for every $\phi < 0$ and

(ii) $\lambda(\phi) \in [1/(1+\eta_o), \phi]$ for every $\phi > 1/(1+\eta_o)$

Then

(i) $V(\hat{\mu}_*) \leq V(\hat{\mu})$ for all $\eta \geq \eta_o$ with strict inequality holding unless $\hat{\mu}_* = \hat{\mu}$ a.s.

(ii) If further $\phi \leq 1/(1+\eta_o)$ a.s., then $V(\hat{\mu}_*) \leq V(\hat{\mu})$ for all $\eta \geq 0$ with strict inequality holding unless $\hat{\mu}_* = \hat{\mu}$ a.s.

Proof To prove (ii) we have to show that $h(\phi) \geq h(\lambda)$ if $\phi < 0$ and $\lambda \in [\phi, |\phi|]$, with strict inequality holding if $\lambda \neq \phi$. This holds by Lemma 5.2.1(a). To prove (i) we have to show in addition that $h(\phi) \geq h(\lambda)$ if $\phi > 1/(1+\eta_o)$ and $\lambda \in [1/(1+\eta_o), \phi]$, with strict inequality holding of if $\lambda \neq \phi$. This holds by lemma 5.2.1(b). Hence the proof is complete.

Corollary 5.4.1 $\hat{\mu}$ is inadmissible with respect to the restircted parameter set $[\eta \geq \eta_o]$ if either $P(\phi < 0) > 0$ or $P[\phi > 1/(1+\eta_o] > 0$.

Taking $\eta_o = 0$, we have

Corollary 5.4.2. $\hat{\mu}$ is inadmissible with respect to the entire parameter set $[\eta \geq 0]$ if $P(0 \leq \phi \leq 1) < 1$.

The next result, contained in theorem 5.4.2, gives us a class of modified estimators for which results of theorems 5.3.1 and 5.3.3 both hold.

Theorem 5.4.2. Let $\hat{\mu}_* = x + \phi_*(y-x)$ where

$$\phi_* = \begin{cases} \text{Min}\,[\phi, 1/(1+\eta_o)] & \text{if } \phi \geq 0 \\ \lambda(\phi) & \text{otherwise} \end{cases}$$

and $\lambda(\phi)$ is a measurable function of ϕ such that

$$\lambda(\phi) \in [0, \text{Min}\{|\phi|, 1/(1+\eta_o)\}] \quad \text{for every} \quad \phi < 0$$

Then

(i) $V(\hat{\mu}_*) \leq V(\hat{\mu})$ for all $\eta \geq \eta_o$ with strict inequality holding unless $\hat{\mu}_* = \hat{\mu}$ a.s.

(ii) $V(\hat{\mu}_*) \leq V(x)$ for all $\eta < 1+2\eta_o$ with strict inequality holding unless $\hat{\mu}_* = x$ a.s.

(iii) $V(\hat{\mu}) \leq V(x)$ for all $\eta \geq 1+2\eta_o \Rightarrow V(\mu_*) \leq V(x)$ for all $\eta > 0$

<u>Proof</u> (i) follows from theorem 5.4.1 once it is noted that $\lambda(\phi)$ satisfies the condition in that theorem. Proofs of (ii) and (iii) are analogous to that of theorem 5.3.3 and follows from theorems 5.3.2 and 5.4.1.

Consider now,

$$\hat{\mu}_*^{(i)} = x + \phi_*^{(i)}(y-x), \; i = 0,1,\ldots,5.$$

where

$$\phi_*^{(0)} = \phi \quad \text{if} \quad 0 < \phi < 1/(1+\eta_o)$$
$$= 1/(1+\eta_o) \quad \text{otherwise}$$

$$\phi_*^{(1)} = \text{Min}\,[\phi, 1/(1+\eta_o)]$$

$$\phi_*^{(2)} = \text{Sgn}\,\phi \,\text{Min}\,[|\phi|, 1/(1+\eta_o)]$$

$$\phi_*^{(3)} = \phi \quad \text{if} \quad |\phi| < 1/(1+\eta_o)$$
$$= 1/(1+\eta_o) \quad \text{otherwise}$$

$$\phi_*^{(4)} \quad \text{Min}[\phi, 1/(1+\eta_o)] \text{ if } \phi > 0$$
$$= 0 \quad \text{otherwise}$$

$$\phi_*^{(5)} = \text{Min}[|\phi|, 1/(1+\eta_o)]$$

Note that $\hat{\mu}_*^{(0)}$ is the modification of $\hat{\mu}$ sugge ted by Yates (1939) and $\hat{\mu}_*^{(i)}$, $i = 1,2,\ldots,5$ are imitations of that recommended by Stein (1966) for a non-négat

Before amining other aspects let us compare these six estimators thems es. Observe that by theorem 5.3.1, each of the estimators $\hat{\mu}_*$, $\hat{\mu}_*^{(2)}$, $\hat{\mu}_*^{(3)}$ is strictly dominated by $\hat{\mu}_5$ unless it is identical with $\hat{\mu}_5$ almost sure. Hence in the following theorem, we compare the remaining three estimators.

Theorem 5. et V $(\hat{\mu}_*^{(i)})$, th

(i) V $_4$ ll $< 1+2\eta_o$

(ii) $V_o \leq V_4$ for all $\eta < 1+2\eta_o$

$\geq V_4$ for all $\eta > 1+2\eta_o$

(iii) $V_o \leq V_5$ for all $\eta \leq \eta_o$

$\geq V_5$ for all $\eta \geq 1+2\eta_o$

Furthermore, the inequality between each pair of variances as stated above holds strictly unless the corresponding estimators are identical almost sure.

Proof Let c be as in lemma 5.2.1. Then, to prove (i), we have to show that for every $t \in (0,c]$, $h(t) < h(0)$ for all $\eta < 1+2\eta_o$ and to prove (ii) we have to show that $h(c) < h(0)$ for all $\eta < 1+2\eta_o$ and $h(c) > h(0)$ for $\eta > 1+2\eta_o$. All these hold by lemma 5.2.1(c). To prove (iii) we have to show that for every $t \in (0,c)$, $h(t) > h(c)$ for all $\eta \leq \eta_o$ and $h(t) < h(c)$ for all $\eta \geq 1+2\eta_o$. This holds by lemma 5.2.1(d). Hence the proof is complete.

I [illegible] seen that [illegible]h of $\phi_*^{(4)}$ and $\phi_*^{(5)}$ satisfies the condition of th[illegible] .2 and h[illegible] e have

<u>Theore[illegible] 4</u> Let [illegible] as in theo[illegible] 5.4.3 [illegible]en [illegible] = 4,5

(i) $V_i \leq V(\hat{\mu})$ for all $\eta \geq \eta_o$, with strict i[illegible]quality holding, unless $\hat{\mu}_*^{(i)} = \hat{\mu}$ a.s.

(ii) $V_i \leq V(x)$ for all $\eta < 1+2\,\eta_o$, with strict inequality holding , unless $\hat{\mu}^{(i)} = x$ a.s.

(iii) $V(\hat{\mu}) \leq V(x)$ for all $\eta \geq 1 + 2\,\eta_o \Rightarrow V_i \leq V(x)$ for all $\eta \geq 0$.

Finally note that $\phi_*^{(0)}$ satisfies the condition of theorem 5.3.2. Hence, we have

<u>Theorem 5.4.5</u>. Let V_i be as in theorem 5.4.3. Then, $V_5 \leq V(x)$ for all $\eta < 1+2\,\eta_o$, with strict inequality holding unless $\mu_*^{(0)} = x$ a.s.

<u>Remark 5.4.1</u> Note that $\rho_{*o} = 1$ and $\rho_* \leq 2 \Rightarrow \eta \leq 2\,\eta_o \Rightarrow \eta < 1+2\,\eta_o$ and hence theorem 5.4.5 yields the result of Shah (1964a).

CHAPTER 6

INTERVAL ESTIMATION OF COMMON MEAN OF TWO NORMAL DISTRIBUTIONS AND TREATMENT DIFFERENCES IN BLOCK DESIGNS

6.1 Introduction

The problem of point estimation of a common mean of two normal populations together with the problem of use of inter-block information for point estimation of treatment differences in incomplete block designs has received much attention in recent years. Since the probability distribution of such an estimate is not easily tractable the problem of interval estimation of these parameters has received comparatively much less attention. Useful contributions have however, been made by many authors [e.g. see Meir (1953), Cochran (1954), James (1956), Rhodes (1961), Brown and Cohen (1974), Cohen and Sackrowitz (1974), Rohatgi and Rastogi (1974), Maric and Graybill (1979a,b), Bhattacharya (1980), Khatri and Shah (1980)].

Following Brown and Cohen (1974) we consider intervals of the same width as the usual intervals based on the t-distribution which are centered around the main estimate (mean of the first sample in the two-sample problem and the intra-block estimate in the block design problem) but we center these intervals around point estimates which are hopefully more precise. In Section 2, we use numerical integration methods to compute confidence co-efficients for such intervals. Tables I and II give some illustrative computations for the two sample problem and the balanced incomplete block (BIB) design problem respectively. In Section 3, we compute constants required by Brown and Cohen (1974) for constructing confidence intervals which are uniformly better than the ones centered around the principal estimate. It turns out that for intervals with coefficient exceeding 0.9, the point estimates used are nearly identical with the corresponding principal estimates. In Section 4, we use simulation methods for computing confidence co-efficients for several intervals centered around "reasonable" point estimates of treatment differences in BIB designs. These computations indicate that some of the point estimates lead to significantly improved interval estimates.

6.2 Numerical Integration Methods

Let

$$x \sim N(\mu,\sigma_x^2),\ y \sim N(\mu,\sigma_y^2),\ U/\sigma_x^2 \sim \chi_e^2$$

and

$$V/\sigma_y^2 \sim \chi_f^2$$

We shall assume that these random variables are all independently distributed.

Let

$$w = a(U/e)/\{(U/e) + (bV/f)\}$$

We shall consider $\hat{\mu} = x+w(y-x)$. We shall consider $\hat{\mu}+t_e(\alpha)U/e$ as an interval estimate for μ where $(-t_e(\alpha),\ t_e(\alpha))$ contains a variate with 't' distribution with e degrees of freedom (d.f.) with probability $(1-\alpha)$.

Since the exact distribution of $\hat{\mu}$ is somewhat difficult to obtain we shall adopt the following method to evaluate the probability with which the above interval will contain the true parameter value μ.

We may write w as $w = af/\{f+beV/U\}$.

Let

$$\tau = \sigma_y^2/\sigma_x^2,\ D = (U/\sigma_x^2) + (V/\sigma_y^2)$$

and let

$$B = (U/\sigma_x^2)/\{(U/\sigma_x^2) + (V/\sigma_y^2)\}.$$

It is clear that B has Beta distribution with parameters $(e/2,\ f/2)$, $D \sim \chi^2_{e+f}$ and that B and D are independently distributed.

We may now write w as $w = af\,B/\{fB + be\tau(1-B)\}$. We can also express U as $U = BD\sigma_x^2$. We note that the conditional distribution of $\hat{\mu}$ given B and D is $N(\mu, (1-w)^2\sigma_x^2 + w^2\tau\sigma_x^2)$. Thus

$$P = \Pr\{|\hat{\mu}-\mu|^2 \leqq t_e^2(\alpha)U/e\}$$

$$= \underset{B,D}{E}\,\{\Pr\{z^2(\tau w^2+(1-w)^2)\sigma_x^2 \leqq t_e^2(\alpha)U/e\,|\,B,D\} \qquad (6.2.1)$$

where z has a standard normal distribution independently of B and D. Writing $U = BD\sigma_x^2$ we get

$$P = \underset{B,D}{E}\ \Pr\{z^2/D \leqq Bt_e^2(\alpha)/e(\tau w^2 + (1-w)^2)\,|\,B,D\}\ .$$

We note that w is a function of B alone and hence does not involve D. Taking expectation w.r.t. D we get

$$P = \underset{B}{E}\ I_\phi\left(\frac{1}{2}, \frac{e+f}{2}\right) \qquad (6.2.2)$$

where

$$\phi = \frac{Bt_e^2(\alpha)}{e\{\tau w^2 + (1-w)^2\} + Bt_e^2(\alpha)} \qquad (6.2.3)$$

and $I_x(p,q)$ denotes incomplete Beta function.

One may evaluate I_ϕ for appropriate values of B and use numerical integration techniques to evaluate P.

The above may be applied to evaluate confidence co-efficients for interval estimates of the common mean of two independent normal samples. Table 6.2.1 gives this for sample sizes (4,6), (4,20), (20,6) and (20,20). We take estimates with

TABLE [illegible]

Confidence co-efficients for interval estimates of the common mean from two samples

		τ	0.5			1			10			100		
		a												
m	n	1 − α	1	a_0	$2a_0$	1	a_0	$2a_0$		a_0	$2a_0$	1	a_0	$2a_0$
4	20	0.99	0.9931	0.9918	0.9930	0.9920	0.9911	0.9919	0.9903	0.9901	0.9903	0.9900	0.9900	0.9900
		0.95	0.9737	0.9650	0.9731	0.9667	0.9601	0.9661	0.9527	0.9516	0.9526	0.9503	0.9502	0.9503
		0.90	0.9523	0.9336	0.9510	0.9372	0.9233	0.9362	0.9059	0.9038	0.9058	0.9006	0.9004	0.9006
		0.80	0.9069	0.8685	0.9045	0.8744	0.8482	0.8730	0.8096	0.8080	0.8098	0.8009	0.8008	0.8009
		0.50	0.6946	0.6206	0.6930	0.6168	0.5827	0.6174	0.5075	0.5121	0.5088	0.4999	0.5012	0.5002
		0.10	0.1627	0.1381	0.1633	0.1330	0.1251	0.1337	0.1012	0.1032	0.1016	0.0998	0.1003	0.0998
4	6	0.99	0.9934	0.9907	0.9912	0.9923	0.9904	0.9908	0.9903	0.9900	0.9901	0.9900	0.9900	0.9900
		0.95	0.9748	0.9552	0.9596	0.9676	0.9536	0.9566	0.9516	0.9506	0.9511	0.9498	0.9501	0.9501
		0.90	0.9536	0.9114	0.9210	0.9379	0.9081	0.9149	0.9020	0.9015	0.9027	0.8991	0.9002	0.9003
		0.80	0.9075	0.8216	0.8410	0.8729	0.8160	0.8299	0.8004	0.8032	0.8056	0.7974	0.8004	0.8005
		0.50	0.6904	0.5315	0.5639	0.6093	0.5242	0.5476	0.4946	0.5054	0.5089	0.4941	0.5006	0.5007
		0.10	0.1606	0.1086	0.1183	0.1304	0.1067	0.1137	0.0980	0.1015	0.1025	0.0982	0.1002	0.1002
20	6	0.99	0.9997	0.9952	0.9977	0.9987	0.9941	0.9963	0.9894	0.9909	0.9914	0.9893	0.9901	0.9901
		0.95	0.9968	0.9705	0.9836	0.9862	0.9658	0.9765	0.9497	0.9534	0.9552	0.9482	0.9504	0.9504
		0.90	0.9890	0.9340	0.9591	0.9677	0.9262	0.9455	0.9003	0.9055	0.9085	0.8974	0.9006	0.9006
		0.80	0.9584	0.8506	0.8945	0.9088	0.8386	0.8705	0.8012	0.8079	0.8121	0.7967	0.8008	0.8008
		0.50	0.8316	0.5556	0.6151	0.6329	0.5419	0.5820	0.5019	0.5082	0.5125	0.4968	0.5008	0.5009
		0.10	0.1653	0.1136	0.1296	0.1341	0.1102	0.1206	0.1004	0.1019	0.1030	0.0992	0.1002	0.1002
20	20	0.99	0.9998	0.9994	0.9998	0.9989	0.9982	0.9987	0.9912	0.9920	0.9917	0.9902	[illegible]	0.9902
		0.95	0.9972	0.9946	0.9953	0.9900	0.9868	0.9861	0.9576	0.9574	0.9547	0.9508	0.9508	0.9504
		0.90	0.9904	0.9838	0.9827	0.9719	0.9658	0.9612	0.9008	0.9118	0.9063	0.9012	0.9012	0.9005
		0.80	0.9624	0.9471	0.9371	0.9173	0.9072	0.8928	0.8160	0.8167	0.8079	0.8015	0.8017	0.8004
		0.50	0.7418	0.7114	0.6737	0.6462	0.9345	0.6031	0.5157	0.5171	0.5059	0.5014	0.5017	0.5000
		0.10	0.1688	0.1593	0.1452	0.1378	0.1350	0.1254	0.1036	0.1040	0.1017	0.1002	0.1004	0.0999

Design parameters		$b = 10,\ k = 2,\ v = 5,\ \lambda = 3$			$b = 10,\ k = 3,\ v = 10,\ \lambda = 3$			$b = 30,\ k = 3,\ v = 10,\ \lambda = 2$		
	ω	1	a_0	$2a_0$	1	a_0	$2a_0$	1	a_0	$2a_0$
$1-\alpha$	δ									
0.99	1	0.9945	0.9912	0.9922	0.9917	0.9914	0.9924	0.9965	0.9963	0.9939
	2	0.9925	0.9908	0.9914	0.9895	0.9908	0.9913	0.9941	0.9940	0.9919
	4	0.9910	0.9904	0.9908	0.9888	0.9904	0.9907	0.9923	0.9923	0.9909
	8	0.9901	0.9902	0.9904	0.9889	0.9902	0.9903	0.9912	0.9912	0.9904
0.95	1	0.9715	0.9566	0.9616	0.9566	0.9557	0.9593	0.9745	0.9739	0.9635
	2	0.9611	0.9542	0.9573	0.9499	0.9532	0.9551	0.9641	0.9641	0.9562
	4	0.9537	0.9524	0.9542	0.9475	0.9517	0.9525	0.9575	0.9576	0.9526
	8	0.9499	0.9513	0.9522	0.9474	0.9509	0.9511	0.9538	0.9540	0.9511
0.90	1	0.9394	0.9123	0.9219	0.9109	0.9093	0.9153	0.9395	0.9388	0.9210
	2	0.9184	0.9078	0.9138	0.9004	0.9052	0.9082	0.9219	0.9220	0.9092
	4	0.9053	0.9046	0.9079	0.8967	0.9028	0.9040	0.9114	0.9117	0.9036
	8	0.8992	0.9025	0.9041	0.8963	0.9014	0.9018	0.9058	0.9060	0.9014
0.80	1	0.8611	0.8204	0.8371	0.8157	0.8136	0.8225	0.8570	0.8564	0.8293
	2	0.8268	0.8131	0.8231	0.8015	0.8076	0.8118	0.8303	0.8308	0.8123
	4	0.8065	0.8077	0.8130	0.7959	0.8040	0.8057	0.8154	0.8161	0.8046
	8	0.7975	0.8043	0.8068	0.7953	0.8020	0.8026	0.8077	0.8082	0.8016
0.50	1	0.5720	0.5252	0.5478	0.5164	0.5145	0.5242	0.5599	0.5598	0.5294
	2	0.5277	0.5163	0.5291	0.5015	0.5081	0.5124	0.5304	0.5313	0.5118
	4	0.5049	0.5096	0.5160	0.4993	0.5042	0.5059	0.5151	0.5159	0.5041
	8	0.4956	0.5053	0.5081	0.4954	0.5021	0.5027	0.5074	0.5080	0.5012
0.10	1	0.1181	0.1065	0.1125	0.1039	0.1035	0.1058	0.1144	0.1144	0.1069
	2	0.1066	0.1042	0.1075	0.1004	0.1019	0.1029	0.1072	0.1074	0.1027
	4	0.1010	0.1025	0.1041	0.0991	0.1010	0.1014	0.1035	0.1037	0.1009
	8	0.0987	0.1013	0.1020	0.0989	0.1005	0.1006	0.1017	0.1018	0.1002

(i) $a = 1, b = 1,$

(ii) $a = a_0 = (e-2)(f-3)/e(f+1),\ b = 1$

and

(iii) $a = 2a_0,\ b = 1.$

Here $e = m-1$ and $f = n-1$. The first is the one proposed by Graybill and Deal. The other two have the property of having uniformly smaller variance than the mean of the first sample for all values of the ratio of variances. (See Brown and Cohen (1974), Bhattacharya (1980)). We compute this when τ, the ratio of variances of the two means is 0.5, 1, 10 and 100.

Results of Table 6.1 indicate that these intervals have improved confidence co-efficient when τ does not exceed 10.

The above method can also be applied to the problem of estimation of the treatment differences in incomplete block designs. Table 6.2.2 gives the computations for three balanced incomplete block designs again for the same values of a and b as in Table 6.2.1. We present the computations for $\delta = 1,2,4$ and 8 where δ denotes the ratio of inter to intra-block variances.

The computations appear to indicate substantial improvement over intra-block estimate for values of δ up to 4.

It may be noted that the class of estimates $\hat{\mu}$ considered here does not include Yates' estimate. However, these methods can be applied for some of the estimates considered in Khatri and Shah (1974) where they take $a = 1$ and use suitable values of b.

6.3. Analytical methods

Brown and Cohen (1974) considered interval estimates of μ of the type considered in Section 2 with $b = 1$. They showed that there exists a suitable value of a such that the confidence co-efficient exceeds $1-\alpha$ for all values of τ.

After integrating w.r.t. w on both sides condition (5.3) of Brown and Cohen gives

$$\frac{1-a}{2^{3/2}} \int_1^{\alpha} \frac{v^{(f-2)/2}}{(1+v)^3} \frac{\Gamma(h)}{\left[\frac{1}{2} + 2t^2 + \frac{vf}{2e}\right]^h} dv \geq a \int_0^{\alpha} \frac{\Gamma(h)\, v^{(f-6)/2}}{\left[\frac{1}{2} + \frac{fv}{2e}\right]^h} dv \qquad (6.3.1)$$

where

$$h = (e+f+1)/2 \qquad \text{and} \qquad t = t_e(\alpha)/\sqrt{e}$$

To evaluate the L.H.S. of (6.3.1) we put $k = f/e(1+4t^2)$, $x = 1/(1+kv)$. This gives

$$1/(1+v)^3 = k^3 x^3 [1+(k-1)x]^{-3}$$

Expanding $\{1+(k-1)x\}^{-3}$ in binomial series we can evaluate the L.H.S. by integrating each term of this series. The R.H.S. is easily evaluated by using the substitution $y = 1/(1+fv/e)$ and integrating w.r.t. y.

$$Ak^{3-f/2} \sum_{r=0}^{\infty} \binom{-3}{r} (k-1)^r I_{1/(1+k)}\left(\frac{e+1}{2} + r + 3, \frac{f}{2}\right) B\left(\frac{e+1}{2} + r + 3, \frac{f}{2}\right)$$

$$\geq a 2^h (e/f)^{(f-4)/2} B\left(\frac{f-4}{2}, \frac{e+5}{2}\right)\left\{1 - I_{1/(1+af/e)}\left(\frac{e+5}{2}, \frac{f-4}{2}\right)\right\} \qquad (6.3.2)$$

where

$$A = 2^{h-3/2} (1-a)(1+4t^2)^{-h}$$

We obtain the largest value of a which satisfies (6.3.2). It should be noted that the value of a depends upon α. In particular a tends to zero as α tends to zero. Table 6.3.1 gives values of a for two sample problem with m = 2,10,20 and n = 6, 10, 20. These give e = 1,9,19; f = 5,9,19. We present the results for α = 0.6, 0.4 and 0.2. We also calculated values of a for α = 0.1, 0.05 and 0.01. For these values of α the corresponding values of a turn out to be less than 10^{-8} with the result that the corresponding $\hat{\mu}$ will be virtually the same as x. Even for larger values of α as in Table 6.3.1 the values of a are rather small so that $\hat{\mu}$ will not differ much from x. It is clear that these methods can be applied in the case of incomplete block designs. The results in that case can be expected to be similar to the results of Table 6.3.1

TABLE 6.3.1

Values for 'a' for two sample problem x 10^4

	α	0.6	0.4	0.2
m	n			
2	6	0.3414	0.0856	0.0032
	10	1.3472	0.3080	0.0077
	20	2.3654	0.4936	0.0077
10	6	0.1797	0.0049	*
	10	0.9567	0.0215	*
	20	2.5126	0.0453	*
20	6	0.0057	*	*
	10	0.0340	*	*
	20	0.1162	*	*

* Indicates that $a < 10^{-8}$.

Simulation Methods

In this section we shall apply simulation methods to problem of rval estimation of treatment contrasts in balanced incomplete block) designs. When one uses the intra-block information only, one can truct the usual confidence interval centered around this estimate. estimate is based on the fact that this estimate divided by its mated standard error follows student's t distribution with appropriate rees of freedom.

For all BIB designs it is possible to obtain a combined inter-and ra-block estimate which has smaller variance than the intra-block estimate. ides, there are many good procedures to find such a combined estimate. wever, the distribution of these estimates is virtually intractable and is makes it very difficult to construct confidence intervals based on the obability distributions of these estimates. An alternative approach nsidered here is to construct intervals of the same width as the usual tervals centered around the intra-block estimates but to center the terval around a reasonable combined estimate and to hope that this interval uld contain the true parameter value with higher probability which exceeds he confidence co-efficient.

Since we do not know the probability distributions of these combined stimates we attempt to estimate these probabilities via simulation in the following manner. We consider the following six methods of estimating treatment differences: (i) Graybill and Deal (1959) (ii) Stein (1966) (iii) Yates (1940) (iv) Khatri and Shah (1974) (v) Brown and Cohen (1974) (vi) intra-block estimate. We take a treatment contrast with true value zero and construct confidence intervals of the same width as the interval for the intra-block estimate and center them around the six estimates. We mulate 50,000 samples and examine the proportion of samples for which the ntervals covered the true value (in this case zero). For estimate (vi) this portion should be the confidence co-efficient. The results of this mulation are given in Table 6.4.1 for $\delta = 1,2,4,8$ where δ is the ratio of ter to intra-block variances. In all cases, we have used truncation for $\hat{\delta}$ t unity.

Results of these simulations indicate that the estimates (iii), (iv) and (v) which are based on all components of the analysis of variance table provide good confidence intervals. Estimate (iii) provides a better interval for small values of δ while (iv) and (v) provide better intervals when δ exceeds two. As pointed out in Khatri and Shah (1974) each of the estimates (i) and (ii) ignore a component of the analysis of variance table and this results in somewhat poorer performance. Here again estimate (ii) provides better intervals than estimate (i). As expected, for estimates (i) to (v) the proportion decreases as δ increases. Results for estimate (vi) provide a check for the simulation procedure. We expect the proportions for this estimate to be fairly close to the confidence co-efficients and in this sense the results for estimate (vi) validates the rest of Table 6.4.1. It may be noted that for three of the four designs we also have the results for the first estimate in Table 6.2.1 when a = 1. Of course, these are obtained without truncation for $\hat{\delta}$. The results obtained by numerical integration and by simulation appear to be fairly close.

We also calculated the ratio of the fourth central moment to the square of the second central moment for each estimate for each design. The values were in all cases very close to 3 indicating roughly a distribution not too far from normal. This would lead one to expect that the estimates with smaller variance would have good concentration around the true parameter value.

Design: $b = 6$, $k = 2$, $v = 4$, $\lambda = 1$

Estimate $1-\alpha$	δ	Graybill–Deal	Stein	Yates	Khatri–Shah	Brown–Cohen	Intra-block
0.99	1	0.9922	0.9903	0.9914	0.9906	0.9906	0.9902
	2	0.9913	0.9903	0.9909	0.9905	0.9905	0.9902
	4	0.9907	0.9903	0.9907	0.9904	0.9904	0.9902
	8	0.9899	0.9903	0.9905	0.9903	0.9903	0.9902
0.95	1	0.9628	0.9527	0.9591	0.9542	0.9539	0.9505
	2	0.9565	0.9518	0.9556	0.9531	0.9529	0.9505
	4	0.9512	0.9511	0.9532	0.9519	0.9518	0.9505
	8	0.9466	0.9509	0.9522	0.9513	0.9512	0.9505
0.90	1	0.9253	0.9050	0.9195	0.9091	0.9079	0.8998
	2	0.9117	0.9030	0.9106	0.9053	0.9051	0.8998
	4	0.9006	0.9017	0.9049	0.9029	0.9029	0.8998
	8	0.8904	0.9008	0.9022	0.9015	0.9015	0.8998
0.80	1	0.8481	0.8132	[illegible]	0.9216	0.8179	0.8007
	2	0.8235	0.8087	[illegible]	0.8132	0.8117	0.8007
	4	0.7986	0.8050	0.8099	0.8069	0.8070	0.8007
	8	0.7805	0.8027	0.8041	0.8040	0.8041	0.8007
0.50	1	0.5734	0.5265	0.5668	[illegible]	0.5318	0.5000
	2	0.5324	0.5149	0.5346	0.5238	0.5209	0.5000
	4	0.4950	0.5076	0.5114	0.5123	0.5126	0.5000
	8	0.4766	0.5036	0.5024	0.5068	0.5071	0.5000
0.10	1	0.1231	0.1076	0.1222	0.1116	0.1081	0.0996
	2	0.1093	0.1038	0.1115	0.1066	0.1048	0.0996
	4	0.1007	0.1013	0.1019	0.1032	0.1026	0.0996
	8	0.0944	0.1019	[illegible]	0.1009	0.1007	0.0996

TABLE 6.4.5.b

Design: $b = 10, k = 2, r = 5, \lambda = 1$

Estimate		Graybill–Deal	Stein	Yates	Khatri–Shah	Brown–Cohen	Intra-block
$1-\alpha$	δ						
0.99	1	0.9948	0.9912	0.9939	0.9925	0.9923	0.9899
	2	0.9930	0.9908	0.9925	0.9916	0.9915	0.9899
	4	0.9915	0.9904	0.9914	0.9909	0.9908	0.9899
	8	0.9906	0.9901	0.9908	0.9904	0.9904	0.9899
0.95	1	0.9744	0.9600	0.9721	0.9655	0.9641	0.9513
	2	0.9653	0.9567	0.9643	0.9601	0.9597	0.9513
	4	0.9576	0.9542	0.9585	0.9560	0.9561	0.9513
	8	0.9525	0.9528	0.9546	0.9542	0.9542	0.9513
0.90	1	0.9445	0.9203	0.9404	0.9299	0.9273	0.9011
	2	0.9261	0.9133	0.9255	0.9195	0.9182	0.9011
	4	0.9121	0.9081	0.9146	0.9114	0.9111	0.9011
	8	0.9027	0.9052	0.9080	0.9068	0.9066	0.9011
0.80	1	0.8744	0.8352	0.8701	0.8508	0.8449	0.7992
	2	0.8392	0.8208	0.8391	0.8302	0.8280	0.7992
	4	0.8120	0.8112	0.8177	0.8160	0.8159	0.7992
	8	0.7990	0.8055	0.8078	0.8082	0.8082	0.7992
0.50	1	0.5911	0.5498	0.5866	0.5668	0.5594	0.4976
	2	0.5404	0.5278	0.5416	0.5370	0.5361	0.4976
	4	0.5087	0.5143	0.5162	0.5201	0.5200	0.4976
	8	0.4961	0.5069	0.5072	0.5091	0.5099	0.4976
0.10	1	0.1237	0.1134	0.1220	0.1199	0.1171	0.1008
	2	0.1117	0.1072	0.1097	0.1119	0.1102	0.1008
	4	0.1032	0.1042	0.1048	0.1056	0.1053	0.1008
	8	0.0991	0.1014	0.1020	0.1021	0.1021	0.1008

TABLE 6.4.1c

Design: $b = 10, k = 3, v = 5, \lambda = 3$

Estimate $1-\alpha$	δ	Graybill–Deal	Stein	Yates	Khatri–Shah	Brown–Cohen	Intra-block
0.99	1	0.9939	0.9918	0.9933	0.9926	0.9926	0.9895
	2	0.9915	0.9909	0.9916	0.9911	0.9913	0.9895
	4	0.9901	0.9903	0.9906	0.9907	0.9907	0.9895
	8	0.9894	0.9900	0.9903	0.9901	0.9901	0.9895
0.95	1	0.9640	0.9580	0.9636	0.9610	0.9606	0.9493
	2	0.9558	0.9542	0.9569	0.9556	0.9554	0.9493
	4	0.9502	0.9520	0.9529	0.9528	0.9528	0.9493
	8	0.9478	0.9507	0.9512	0.9514	0.9514	0.9493
0.90	1	0.9233	0.9144	0.9225	0.9186	0.9181	0.9006
	2	0.9092	0.9083	0.9110	0.9100	0.9102	0.9006
	4	0.9010	0.9047	0.9048	0.9055	0.9054	0.9006
	8	0.8974	0.9031	0.9029	0.9035	0.9035	0.9006
0.80	1	0.8337	0.8199	0.8322	0.8274	0.8266	0.7986
	2	0.8141	0.8096	0.8151	0.8140	0.8137	0.7986
	4	0.8005	0.8054	0.8065	0.8069	0.8069	0.7986
	8	0.7965	0.8018	0.8025	0.8029	0.8030	0.7986
0.50	1	0.5347	0.5203	0.5333	0.5282	0.5262	0.4977
	2	0.5127	0.5086	0.5143	0.5130	0.5127	0.4977
	4	0.4987	0.5032	0.5036	0.5051	0.5054	0.4977
	8	0.4941	0.4998	0.5010	0.5020	0.5019	0.4977
0.10	1	0.1104	0.1053	0.1089	0.1070	0.1068	0.1007
	2	0.1039	0.1028	0.1029	0.1044	0.1036	0.1007
	4	0.0986	0.1021	0.1014	0.1025	0.1026	0.1007
	8	0.0985	0.1013	0.1018	0.1015	0.1014	0.1007

Estimate		Graybill Deal	Stein	Yates	Khatri Shah	Brown-Cohen	Intra-block
1 − α	δ						
0.99	1	0.9970	0.9955	0.9969	0.9966	0.9965	0.9899
	2	0.9945	0.9938	0.9943	0.9943	0.9943	0.9899
	4	0.9926	0.9922	0.9927	0.9925	0.9924	0.9899
	8	0.9912	0.9912	0.9912	0.9912	0.9912	0.9899
0.95	1	0.9761	0.9715	0.9759	0.9746	0.9742	0.9491
	2	0.9641	0.9624	0.9645	0.9643	0.9640	0.9491
	4	0.9568	0.9564	0.9575	0.9574	0.9575	0.9491
	8	0.9533	0.9528	0.9534	0.9531	0.9532	0.9491
0.90	1	0.9415	0.9347	0.9408	0.9389	0.9384	0.9001
	2	0.9211	0.9195	0.9222	0.9215	0.9215	0.9001
	4	0.9106	0.9106	0.9113	0.9121	0.9121	0.9001
	8	0.9057	0.9057	0.9064	0.9066	0.9066	0.9001
0.80	1	0.8587	0.8513	0.8590	0.8560	0.8558	0.7983
	2	0.8291	0.8274	0.8310	0.8306	0.8308	0.7983
	4	0.8147	0.8135	0.8158	0.8166	0.8161	0.7983
	8	0.8071	0.8064	0.8078	0.8076	0.8077	0.7983
0.50	1	0.5633	0.5527	0.5627	0.5598	0.5594	0.4977
	2	0.5296	0.5253	0.5300	0.5306	0.5305	0.4977
	4	0.5124	0.5114	0.5140	0.5143	0.5144	0.4977
	8	0.5051	0.5046	0.5058	0.5051	0.5051	0.4977
0.10	1	0.1184	0.1157	0.1177	0.1164	0.1158	0.0994
	2	0.1082	0.1092	0.1089	0.1093	0.1095	0.0994
	4	0.1042	0.1045	0.1056	0.1043	0.1046	0.0994
	8	0.1024	0.1013	0.1028	0.1032	0.1031	0.0994

CHAPTER 7

ESTIMATION OF A COMMON LOCATION

7.1 Introduction

The problem of combining two or more independent unbiased estimators arises often in practice. So far only the normal case has been studied extensively. Hogg (1960) appears to be the first to discuss unbiased estimation of a common location in the non-normal case. Cohen (1976) using Hogg's result and the techniques in Brown and Cohen (1974) obtained a combined estimator with a variance smaller than that for the first sample. He also points out situations when his combined estimator would be uniformly better than both of the individual estimators.

To use Cohen's estimator in practice one needs to know the upper limit of a constant 'a' to be used in his estimator. For this Cohen (1976) derives an upper limit [denoted by $a^*(m,n)$ where m,n are the two sample sizes] and provides a table of this for a particular situation. We shall see that this upper limit needs improvement for use in practice and provide an improved one which appears satisfactory.

A glance at the table of values of $a^*(m,n)$ in Cohen (1976) reveals that his $a^*(m,n)$ is decreasing in n once n is sufficiently large. This gives an impression, which is contrary to the fact and the intutitive feeling one should have, as explained later. In remark 2.3 of his paper Cohen (1976) considers the important problem of constructing a combined estimator which is uniformly better than both of the individual estimators. But, if we scan through his table of $a^*(m,n)$ we see that for no (m,n) such an estimator can be found following his suggestion although, as we shall show, such estimators exist, in this case, for many pairs (m,n).

With the motivation to overcome these deficiencies of Cohen's $a^*(m,n)$ derive an upper limit [denoted by $A(m,n)$] of a, which is an improvement $a^*(m,n)$. A comparison of the table of our $A(m,n)$ with that of $a^*(m,n)$ Cohen (1976) shows that the improvement is substantial in the particular se worked out in detail by Cohen (1976). Further examination shows that like Cohen's $a^*(m,n)$, $A(m,n)$ here steadily increases in n and unless m large, reaches the maximum value 2 fairly quickly. The table of $A(m,n)$ e enables one to construct combined estimator which is uniformly better an both of the individual estimators for all $n \geq 25$ if $m = n$ and for ll $n \geq 35$ if $m \leq n+5$, in contrast to the fact mentioned earlier that the ble of $a^*(m,n)$ in Cohen (1976) enables one to construct none.

In section 2, we introduce the necessary preliminary notations and ssumptions. This is followed by derivation of $A(m,n)$ and some other related sults in section 3. Finally, in section 4, we present an application to an identical situation considered by Cohen (1976) and compare our results with his.

1.2 Preliminary Notations and Assumptions

Consider two independent random samples $\underset{\sim}{x} = (x_1,\ldots,x_m)$ and $\underset{\sim}{y} = (y_1,\ldots y_n)$ of sizes m and n respectively from two distributions characterized by a common unknown location parameter θ and unknown scale parameters β_x, β_y respectively. Assume that the distributions are symmetric about θ. Let $\hat{\theta}_x$ be an odd location-scale estimator of θ and $\hat{\beta}_x$ be an even location - free scale - invariant estimator of β_x based on the first sample. That is $\hat{\theta}_x$, $\hat{\beta}_x$ satisfy

$$\hat{\theta}(ax_1+b, \ldots, ax_m+b) = a\hat{\theta}(x_1,\ldots,x_m) + b$$

$$\hat{\beta}(ax_1+b,\ldots,ax_m + b) = |a|\, \hat{\beta}(x_1,\ldots,x_m)$$

for every $a \neq 0$ and b.

Let $\hat{\theta}_y$, $\hat{\beta}_y$ be similar estimators based on the second sample. Define $T_x = (\hat{\theta}_x-\theta)/\beta_x$, $T_y = (\hat{\theta}_y-\theta)/\beta_y$, $S_x = \hat{\beta}_x/\beta_x$, $S_y = \hat{\beta}_y/\beta_y$, $v = S_y^2/S_x^2$, $h(v) = \text{Max}\ (1,1/v)$, $g(v) = \text{Min}\ (1,1/v)$, $\eta = \beta_y/\beta_x$, $\gamma = 1/(1+\eta)$, $w = 1/[\gamma+(1-\gamma)v]$. Note that the distributions of T_x, T_y, S_x, S_y do not depend on the unknown parameters and γ lies between 0 and 1. Finally assume that $E[h^2(v)\ \text{Max}(T_x^2,\ T_y^2)] < \infty$. Cohen (1976) has shown that this assumption is justified in a wide variety of situations (see the discussion preceding his lemma 2.1), provided $n \geq 6$.

7.3 Results

Consider the class of estimators of the form

$$\hat{\theta} = \hat{\theta}_x + a(\hat{\theta}_y - \hat{\theta}_x)/(1+z) \tag{7.3.1}$$

where $z = \hat{\beta}_y/\hat{\beta}_x$ and $a > 0$ is a constant to be suitably chosen. Cohen (1976) has shown (i) $\hat{\theta}$ is unbiased for θ ;

$$\text{(ii)}\quad V(\hat{\theta}) = V(\hat{\theta}_x)[1 + E\delta(v)] \tag{7.3.2}$$

where,

$$\delta(v) = (1+\eta)(1+\eta v)^{-2}(T_x^2 + \eta T_y^2) - 2a(1+\eta v)^{-1}\ T_x^2\ .$$

Since we are considering only unbiased estimators, following Cohen (1976), we shall judge the merit of the estimator $\hat{\theta}$ by its variance. It can be seen that

$$\delta(v) = a^2w^2[\gamma T_x^2 + (1-\gamma)T_y^2] - 2\,awT_x^2 \tag{7.3.3}$$

Let

$$R(\gamma) = EwT_x^2/Ew^2[\ \gamma T_x^2 + (1-\gamma)T_y^2] \tag{7.3.4}$$

Then, using (7.3.2) - (7.3.4), we have

<u>Theorem 7.3.1</u> Let $\hat{\theta}$ be as defined in (7.3.1). Then, $\hat{\theta}$ is better than $\hat{\theta}_x$, iff $a \leq 2M$, where $M = \text{Inf } R(\gamma)$.

Evaluation of M is an extremely complicated job. But, for the if-part of theorem 7.3.1 to hold it suffices to take the constant a to be less than or equal to any non-trivial lower bound for 2M. It is easy to see that

$$M \geq \text{Min } (M_1, M_2) \tag{7.3.5}$$

where $M_1 = \underset{\gamma}{\text{Inf}}\ EwT_x^2/Ew^2T_x^2$, $M_2 = \underset{\gamma}{\text{inf}}\ EwT_x^2/Ew^2T_y^2$.

Let λ denote the joint density of v and T_x^2 and let $\lambda_* = \lambda T_x^2/ET_x^2$. Then,

$M_1 = \underset{\gamma}{\text{inf}}\ E_*w/E_*w^2$. where E_* stands for expectation w.r.t. λ_*.

Note that w is of the same form as f of theorem A 2 and satisfies all conditions of that theorem. Hence, using that theorem,

$$\begin{aligned} M_1 &= \text{Min}(1, E_*v^{-1}/E_*v^{-2}) \\ &= \text{Min } (1, ET_x^2\ v^{-1}/ET_x^2\ v^{-2}) \end{aligned} \tag{7.3.6}$$

Also, it is easy to see that

$$M_2 \geq ET_x^2 g(v)/ET_y^2 h^2(v). \tag{7.3.7}$$

Let,

$$A(m,n) = \text{Min}(2, A_1, A_2) \tag{7.3.8}$$

where

$$A_1 = 2ET_x^2 v^{-1}/ET_x^2 v^{-2},\ A_2 = 2ET_x^2 g(v)/ET_y^2 h^2(v). \tag{7.3.9}$$

Then, (3.3), (3.4) and (3.5) together imply

$$A(m,n) \leq 2M \tag{7.3.10}$$

Hence,

<u>Theorem 7.3.2</u> $\hat{\theta}$ is better than $\hat{\theta}_x$, for all $a \leq A(m,n)$.

From theorem 7.3.2 and symmetry consideration, we have,

Corollary 7.3.1 If both $A(m,n) \geq 1$ and $A(n,m) \geq 1$ hold then $\hat{\theta}$ with $a = 1$ is better than both $\hat{\theta}_x$ and $\hat{\theta}_y$.

<u>Remark 7.3.1</u> It is easy to see that theorem 7.3.2 is an improvement of theorem 2.1 of Cohen (1976). For a numerical comparison refer to the next section.

<u>Remark 7.3.2</u> It can be seen that if (T_x,T_y) is independent of v, then $M_1 = \text{Min}\,(1,Ev^{-1}/Ev^{-2})$, $M_2 = M_1ET_x^2/ET_y^2$. Hence it is possible to calculate $A_*(m,n) = 2\,\text{Min}(M_1,M_2)$, which is a better upper limit of the constant a than $A(m,n)$. For the estimator $T_a(1)$ of the common mean of two normal populations considered in Brown and Cohen (1974), $A_*(m,n) = 2Ev^{-1}/Ev^{-2}$, which turns out to be best upper limit of the constant as shown in section 3.4.

<u>Remark 7.3.3</u> From a consideration of the values of $R(\gamma)$ for $\gamma = 0$ and $\gamma = 1$, it is clear that $2M \leq A^*(m,n)$ where $A^*(m,n) = \text{Min}(2,\ 2ET_x^2v^{-1}/ET_y^2v^{-2})$. Hence a necessary condition for $\hat{\theta}$ to be better than $\hat{\theta}_x$ is given by $a \leq A^*(m,n)$. For $a = 1$, this condition reduces to that in theorem 2.2 of Cohen (1976), who used a different method of proof [essentially due to Graybill and Deal (1959)], which requires the derivative of $V(\hat{\theta})$ w.r.t. $\rho = 1/\eta$ at $\rho = 0$. Our approach is obviously simpler.

7.4 Application

Assume that the density for x is

$$f(x;\ \theta,\beta_x) = 1/\beta_x, \quad \text{if } |x-\theta| < \beta_x/2$$
$$= 0, \qquad \text{otherwise.}$$

Let $\hat{\theta}_x = (x_{(m)} + x_{(1)})/2, \hat{\beta}_x = x_{(m)} - x_{(1)}$, where $x_{(1)}, x_{(2)}, \ldots, x_{(m)}$ are the order statistics from the x-population. Let $\hat{\theta}_y$, $\hat{\beta}_y$ be defined similarly. The purpose of this section is to evaluate A(m,n) and to compare it with a*(m,n) of Cohen (1976).

Let $L_x = 2T_x$, $L_y = 2T_y$ and note that the joint density of L_x, S_x, L_y, S_y is

$$f_*(L_x,S_x,L_y,S_y) = c\, S_x^{m-2}\, S_y^{n-2} \text{ if } |L_x| < 1-S_x,\ 0 < S_x < 1,\ |L_y| < 1-S_y, 0<S_y<1$$
$$= 0 \quad \text{otherwise}$$

where $c = m(m-1)\, n(n-1)/4$. Then,

$$ET_x^2/v = c\int_0^1 \int_0^1 \int_0^{1-S_x}\int_0^{1-S_y} L_x^2\, S_x^m\, S_y^{n-4}\, dL_y\, dL_x\, dS_y\, dS_x$$

$$ET_x^2/v^2 = c\int_0^1 \int_0^1 \int_0^{1-S_x}\int_0^{1-S_y} L_x^2\, S_x^{m+2}\, S_y^{n-6}\, dL_y\, dL_x\, dS_y\, dS_x$$

$$ET_x^2 g(v) = c[\int_0^1 \int_0^{S_x} \int_0^{1-S_x} \int_0^{1-S_y} L_x^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_y\, dS_x$$
$$+ \int_0^1 \int_{S_x}^1 \int_0^{1-S_x} \int_0^{1-S_y} L_x^2\, S_x^m\, S_y^{n-4}\, dL_y\, dL_x\, dS_y\, dS_x]$$

$$ET_y^2 h^2(v) = c[\int_0^1 \int_{S_y}^1 \int_0^{1-S_x} \int_0^{1-S_y} L_y^2\, S_x^{m+2}\, S_y^{n-6}\, dL_y\, dL_x\, dS_x\, dS_y$$
$$+ \int_0^1 \int_0^{S_y} \int_0^{1-S_x} \int_0^{1-S_y} L_y^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_x\, dS_y]$$

For convenience, define

$$G(m,n) = (\tfrac{1}{4}) \int_0^1 \int_0^1 \int_0^{1-S_x} \int_0^{1-S_y} L_x^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_y\, dS_x$$

$$H(m,n) = (\tfrac{1}{4}) \int_0^1 \int_0^{S_x} \int_0^{1-S_x} \int_0^{1-S_y} L_x^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_y\, dS_x$$

$$C(m,n) = (\tfrac{1}{2}) \int_0^1 \int_0^1 \int_0^{1-S_x} \int_0^{1-S_y} L_y^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_x\, dS_y$$

$$D(m,n) = (\tfrac{1}{2}) \int_0^1 \int_0^{S_y} \int_0^{1-S_x} \int_0^{1-S_y} L_y^2\, S_x^{m-2}\, S_y^{n-2}\, dL_y\, dL_x\, dS_x\, dS_y$$

Then, it is easy to see that

$$ET_x^2/v = 2c\, G(m+2,\ n-2);\ RT_x^2/v^2 = 2\, c\, G(m+4,\ n-4)$$

$$ET_x^2 g(v) = 2c[H(m,n) + G(m+2,\ n-2) - H(m+2,\ n-2)] \qquad (7.4.1$$

$$ET_y^2 h^2(v) = 2c[C(m+4,\ n-4) - D(m+4,n-4) + D(m,n)]$$

It is also easy to see that

$$G(m,n) = (\tfrac{1}{6}) \int_0^1 \int_0^1 (1-S_x)^3\ (1-S_y) S_x^{m-2}\ S_y^{n-2}\ dS_y\ dS_x$$

$$(7.4.2)$$

$$H(m,n) = (\tfrac{1}{6}) \int_0^1 \int_0^{S_x} (1-S_x)^3\ (1-S_y) S_x^{m-2}\ S_y^{n-2}\ dS_y\ dS_x$$

and that

$$C(m,n) = (\tfrac{1}{6}) \int_0^1 \int_0^1 (1-S_x)(1-S_y)^3\ S_x^{m-2}\ S_y^{n-2}\ dS_x\ dS_y = G(n,m)$$

$$(7.4.3)$$

$$D(m,n) = (\tfrac{1}{6}) \int_0^1 \int_0^{S_y} (1-S_x)(1-S_y)^3\ S_x^{m-2}\ S_y^{n-2}\ dS_x\ dS_y = H(n,m)$$

Using (7.4.1) and (7.4.3), A_1 and A_2 defined by (7.3.7) are given by

$$A_1 = 2\ G(m+2,\ n-2)/G(m+4,\ n-4)\ ; \qquad (7.4.4)$$

$$A_2 = \frac{2[G(m+2,n-2) + H(m,n) - H(m+2,n-2)]}{G(n-4,m+4)+H(n,m)-H(n-4,m+4)}$$

from (7.4.2)

$$G(m,n) = (\tfrac{1}{6})\ (\frac{1}{n-1} - \frac{1}{n}) \int_0^1 (1-S_x)^3\ S_x^{m-2}\ dS_x\ ;$$

$$H(m,n) = (\tfrac{1}{6}) \int_0^1 (1-S_x)^3 \, S_x^{m-2} \left(\frac{S_x^{n-1}}{n-1} - \frac{S_x^n}{n} \right) dS_x$$

$$= (\tfrac{1}{6}) \int_0^1 (1-S_x)^3 \left(\frac{S_x^{p-3}}{n-1} - \frac{S_x^{p-2}}{n} \right) dS_x$$

where $p = m+n$. Note that

$$\int_0^1 (1-S_x)^3 \, S_x^{m-2} = \frac{1}{m-1} - \frac{3}{m} + \frac{3}{m+1} - \frac{1}{m+2}$$

$$= \left(\frac{1}{m-1} - \frac{1}{m-2}\right) - 3\left(\frac{1}{m} - \frac{1}{m+1}\right)$$

$$= \frac{3}{(m-1)(m+2)} - \frac{3}{m(m+1)} = \frac{6}{(m-1)\;m(m+1)(m+2)}$$

Hence

$$G(m,n) = 1/[n^{(2)}(m+2)^{(4)}]$$

$$H(m,n) = 1/[(n-1)(p+1)^{(4)}] - 1/[n(p+2)^{(4)}] = 1/[n^{(2)}(p+2)^{(5)}] \quad (7.4.5)$$

using (7.4.5) and after some manipulation we arrive from (7.4.4) at the following formulae for practical computation of A_1 and A_2 :

$$A_1 = 2(n-4)^{(2)}(m+6)^{(2)}/[(n-2)^{(2)}(m+2)^{(2)}]$$

$$A_2 = \frac{(n-4)^{(2)}\left[.0625 - \frac{(m+4)^{(4)}}{(p+2)^{(4)}}\left\{Q + \frac{n(.25-Q) - .375}{n^{(2)}}\right\}\right]}{(m+2)^{(2)}\left[.03125 + \frac{(n-2)^{(4)}}{(p+2)^{(4)}}\left\{Q + \frac{m(.25 + 5Q)+.375}{m^{(2)}}\right\}\right]} \quad (7.4.6)$$

where $Q = .5/(p-2)$.

Using (7.3.9) and (7.4.6), the values of $A(m,n)$ are compiled in table 7.4.1 for values of (m,n) as in Cohen (1976). Unlike in Cohen (1976), we give the actual value of $A(m,n)$ even when it is greater than or equal to 1 and a blank entry in the table here means $A(m,n) = 2$, which is the best upper limit of a. Note also that the arrangement of the values of $A(m,n)$ here is different from that of $a^*(m,n)$ in Cohen (1976). The entry in cell (i,j) here should be compared with that in cell (j,i) in Cohen (1976).

It can be seen that the entries in the table are consistent with the necessary condition stated in Remark 7.3.3, which in the model of this section reduces to

$$a \leq \text{Min}\,[2, 2(n-4)(n-5)/(m+2)(m+1)]$$

A comparison between the table of $A(m,n)$ here with that of $a^*(m,n)$ in Cohen (1976) leads to the following conclusions:

1) Each entry in the table here shows improvement over the corresponding entry in Cohen (1976) and the improvement is remarkable for each m provided n is not too small.

2) As a natural consequence, the table here reveals many pairs (m,n), in contrast to none in Cohen (1976), for which both $A(m,n) \geq 1$ and $A(n,m) \geq 1$ hold. In view of corollary 7.3.1, $\hat{\theta}$ with $a = 1$ is readily seen to be better than both $\hat{\theta}_x$ and $\hat{\theta}_y$ for all $n \geq 25$ if $m = n$ and for all $n \geq 35$ if $m \leq n+5$.

TABLE 7.4.1

Values of A(m,n)

m	n = 6	7	8	9	10	11	12	13	14	15	20	25	30	35	40	45	50
2	.2569	.6570	1.0738	1.4721	1.8557												
3	.1610	.4794	.8977	1.3673	1.8621												
4	.1050	.3310	.6575	1.0568	1.5068	1.9928											
5	.0725	.2356	.4828	.8002	1.1733	1.5905											
6	.0525	.1739	.3633	.6138	.9164	1.2631	1.6473										
7	.0395	.1325	.2806	.4802	.7260	1.0125	1.3349	1.6869									
8	.0306	.1038	.2219	.3832	.5847	.8225	1.0932	1.3936	1.7212								
9	.0243	.0831	.1790	.3115	.4784	.6775	.9060	1.1618	1.4426	1.7468							
10	.0198	.0679	.1471	.2573	.3973	.5654	.7599	.9789	1.2[illegible]	1.4842							
11	.0163	.0564	.1226	.2155	.3342	.4777	.6445	.8333	1.0429	1.2721							
12	.0137	.0474	.1036	.1828	.2845	.4080	.5522	.7162	.8989	1.0995							
13	.0116	.0404	.0886	.1567	.2447	.3519	.4775	.6209	.7813	.9578							
14	.0100	.0348	.0765	.1357	.2124	.3061	.4164	.5427	.6842	.8405	1.8249						
15	.0086	.0302	.0667	.1185	.1859	.2685	.3659	.4777	.6034	.7425	1.6239						
20	.0047	.0167	.0372	.0667	.1053	.1532	.2102	.2761	.3508	.4342	.9738	1.7038					
25	.0029	.0105	.0235	.0423	.0671	.0980	.1350	.1779	.2269	.2817	.6411	1.1352	1.7556				
30	.0020	.0072	.0161	.0291	.0463	.0677	.0935	.1235	.1578	.1964	.4511	.8052	1.2534	1.7920			
35	.0014	.0052	.0117	.0211	.0337	.0494	.0683	.0904	.1157	.1442	.3334	.5984	.9359	1.3435	1.8189		
40	.0011	.0039	.0088	.0160	.0256	.0376	.0520	.0689	.0883	.1102	.2559	.4610	.7236	1.0419	1.4143	1.8397	
45	.0008	.0031	.0069	.0126	.0201	.0295	.0409	.0542	.0695	.0868	.2023	.3655	.5752	.8301	1.1291	1.4714	1.8561
50	.0007	.0024	.0056	.0101	.0162	.0238	.0329	.0437	.0561	.0700	.1637	.2965	.4676	.6760	.9211	1.2020	1.5183

(3) Cohen's $a^*(m,n)$ is at first increasing and then decreasing in n, whereas $A(m,n)$ here steadily increases in n and unless m is large, reaches the maximum value 2 fairly quickly. It is reasonable to expect $\hat{\theta}$ to be better than $\hat{\theta}_x$ for all $n \geq n_0$ if it is so for $n = n_0$ provided m remains fixed. Our table of $A(m,n)$ supports this and helps to correct an impression to the contrary one gets from the table of Cohen's $a^*(m,n)$.

BIBLIOGRAPHY

1. Barnard,G.A. (1963). Some aspects of the fiducial argument. J. Roy. Stat. Soc. B, 25 111-114.

2. Bartlett,M.S. (1936). The information available in Small Samples, Proc. Camb. Phil. Soc. 32, 560-566.

3. Bartlett,M.S. (1937). Properties of Sufficiency and Statistical Tests, Proc. R. Soc. A 160, 268-282.

4. Bement,T.R. and Williams,J.S. (1963). Variance of weighted regression estimators when sampling errors are independent and heterosecedastic J. Amer. Statisk. Ass. 64 1369-1382.

5. Berk,R.H. (1967). A special group structure and equivariant estimation. Ann. Math. Statist. 38 1436-1445.

6. Bhattacharya,C.G. (1978). Yates type estimators of a common mean Ann. Inst. Statist. Math. A 30 407-414.

7. Bhattacharya,C.G. (1979). A Note an estimating the common mean of K-normal populations, Sankhyā B 40, 272-275.

8. Bhattacharya,C.G. (1980). Estimation of a common mean and recovery of inter-block information, Ann. Statist. 8, 205-211.

9. Blyth,C. (1951). On minimax statistical decision procedures and their admissibility. Ann. Math. Statist. 22, 22-42.

10. Bose,R.C. (1944). The fundamental Theorem of Linear Estimation (Abstract). Proc. 31st Indian Science Congress 4 2-3.

11. Bose,R.C. (1947). Presidential address. Proceedings of the 34th Indian Science Congress P. 12.

12. Bose,R.C., Clatworthy,W.H. and Shrikhande,S.S. (1954). Tables of partially balanced designs with two associate classes. North Carolina Agriculture Experimental Station, Technical Bulletin 107.

13. Box,G.E.P. and Tiao,G.C. (1973). Bayesian Inference in Statistical Analysis. Addison Wesley, Reading, Mass.

14. Brown,L.D. and Cohen,A., (1974). Point and confidence estimation of a common mean and recovery of inter-block information. Ann. Statist. 2 963-976.

15. Chakravarty,M.C. (1962), Mathematics of design and analysis of experiments. Asia Publishing House, Bombay.

17. Chakravarty,M.C. (1963). On the C-matrix in design of experiments, J.I.S.A. I, 8-23.

18. Cochran,W.G. (1937). Problems arising in the analysis of a series of similar experiments. J. R. Stat. Soc. Supp. 4, 102-118.

19. Cochran,W.G. (1954). Combination of estimators from different experiments. Biometrics 10 , 101-137.

20. Cochran,W.G. and Caroll,S.P. (1953). A sampling investigation of the efficiency of weighting inversely as the estimated variance. Biometrics 9 447-459.

21. Cohen,A. (1975). Personal communication.

22. Cohen,A. (1976). Combining estimates of location. J. Amer. Statist. Assoc. 71 172-175.

23. Cohen,A. and Sackrowitz,H.B. (1974). On estimating the common mean of two normal distributions. Ann. Statist. 2 1274-1282.

24. Cox,D.R. (1975). A note on partially Bayes inference and the linear model. Biometrika 62 651-654.

25. Cunningham,E.P. and Henderson,C.R. (1968). An iterative procedure for estimating fixed effects and variance components in mixed model situations. Biometrics 24 13-25.

26. Fisher,R.A. (1935). Design of experiments. Oliver and Boyd, Edinburg.

27. Fisher,R.A. and Yates,F. (1963). Statistical tables for biological, Agricultural + Medical Research, Sixth edition, Oliver and Boyd. Edinburg.

28. Fraser,D.A.S.(1957). A note on combining of inter-block and inter-block estimates. Ann. Math. Stat. 28 814-846.

29. Fraser,D.A.S. (1968). The structure of inference, Wiley, New York.

30. Graybill,F.A. and Deal,R.B. (1959). Combining unbiased estimators. Biometrics 15 543-550.

31. Graybill,F.A. and Seshadri,V. (1960). On the unbiasedness of Yates method of estimation using inter-block information. Ann. Math. Statist. 3 786-787.

32. Graybill,F.A. and Weeks,D.L. (1959). Combining inter-block and intra-block information in balanced incomplete blocks. Ann. Math. Statist. 30 799-805.

33. Gurland,J. and Mehta,J.S. (1969). Combination of unbiased estimators of the mean which consider inequality of unknown variances. J. Amer. Statist. Ass. 64 1042-1055.

34. Hardy,G.H., Littlewod,J.E. and Polya,G. (1952). Inequalities. 2nd edition, Cambridge University Press, London.

35. Hartley,H.O. and Rao,J.N.K. (1967). Maximum likelihood estimation for the mixed analysis of various models. Biometrika 54 93-108.

36. Hinkley,D.V. (1979). A note on the weighted mean problem. Scand. J. Statist. 6 37-40.

37. Hogg,R.V. (1960). On conditional expectations of location statistics. J. Amer. Statist. Ass. 55 714-717.

38. Hsu,P.L. (1938). On the best unbiased quadratic estimate of variance. Statist. Res. Mem. 2 91-104.

39. James,G.S. (1956). On accuracy of weighted means and ratios. Biometrics 43 304-321.

40. Kalbfleisch,J.D. and Sprott ,D.A. (1970). Application of likelihood methods to models involving large numbers of parameters (with discussion). J.R. Statist. Soc. B 32 175-208.

41. Khatri,C.G. and Shah,K.R. (1974). Estimation of location parameters from two linear models under normality. Comm. Statist. 8, 647-669.

42. Khatri,C.G. and Shah,K.R. (1975). Exact variance of combined inter and intra-block estimates in incomplete block designs. J. Amer. Statist. Assoc. 70 402-406.

43. Khatri,C.G. and Shah,K.R. (1980). Improved confidence bounds for a common mean. Comm. Statist. B (To appear).

44. Kimball,B.F. (1951). On dependent tests of significance in the analysis of variance, Ann. Math. Statist. 22

45. Lebedev,N.N. (1972). Special functions and their applications. Rev. enl. ed. Translated and edited by Richard A. Silverman, Dover, New York.

46. Lehmann,E.L. (1950). Testing statisticals hypotheses. Wiley, New York.

47. Lehman,E.L. and Scheffe,H. (1950) Completenes similar regions and unbiased estimation - Part I, Sankhyā 10 305-340.

48. Levy,P. (1970). Combining independent estimators - an empirical study. Technometrics 12 162-165.

49. Maric',N. and Graybill,F.A. (1979). Evaluation of a method for setting confidence intervals on the common mean of two normal populations. Comm. Statist. Simula, Compute B.8 53-60

50. Marie,N. and Graybill,F.A. (1979). Small samples confidence intervals on common means of two normal distributions with unequal variances. Comm. Statist. Theor. Math. A 8 1255-1269.

Meir,P. (1953). Variance of a weighted mean. Biometrics 9 59-73.

Nair,K.R. (1944). The recovery of inter-block information in incomplete block designs. Sankhyā 6 353-390.

Nelder,J.A. (1968). The combination of information in generally balanced designs. J.R. Statist. Soc. B 30 303-311.

Neyman,J. and Scott,E.L. (1948). Consistent estimators based on partially consistent observations. Econometrika 16 1-32.

55. Norwood,T.E. and Hinkelmann,K. (1977). Estimating the common mean of several normal populations. Ann. Statist. 5 1047-1050.

56. Ogawa,J. (1974). Statistical theory of the analysis of experimental designs. Marcel decker, New York.

57. Olkin,J. and Pratt,J (1958). *Unbiased estimation of certain correlation* co-efficients. Ann. Math. Statist. 29 201-211.

58. Patterson,H.D. and Thompson,R. (1971). Recovery of inter-block information when block sizes are unequal. Biometrika 58 545-554.

59. Rao,C.R. (1947). General methods of analysis for incomplete block design. J. Amer. Statist. Ass. 42 541-561.

60. Rao,C.R. (1952). Some theorems on minimum variance unbiased estimation. Sankhyā 12 27-42.

61. Rao,C.R. (1970). Estimation of heteroscedastic variances in linear models. J. Amer. Statist. Ass. 65 161-172.

62. Rao,C.R. (1971). Minimum variance quadratic unbiased estimation of variance components. J. Multivariate Analysis. 1 445-456.

63. Rao,C.R. (1973). Linear Statistical Inference and its Applications. Second edition, Wiley, New York.

64. Rao,C.R. and Mitra,S.K. (1971). Further contributions to the theory of generalized inverse of matrices and its applications. Sankhyā A 33 289-300.

65. Rao,J.N.K. and Subrahmaniam,K. (1971). Combining independent estimators and estimation in linear regression with unequal variances. Biometrics 27 971-990.

66. Rhodes,B.T. (1961). The use of combined information in Internal estimation. Ph.D. thesis submitted to Oklahuma State University.

67. Rohatgi,V.K. and Rastogi,S.C. (1974). On unbiased estimation of the common mean of two independent normal distributions based on samples of unequal size. Biom. Z. Bd. 16 445-450.

68. Roy,J. and Laha,R.G. (1956). Classification and analysis of linked block designs. Sankhyā 17 115-132.

69. Roy,J. and Shah,K.R. (1962). Recovery of inter-block information. Sankhya A 24 269-280.

70. Ruben,H. (1962) Probability content of regions under spherical normal distributions ; IV: the distribution of homogeneous and non-homogeneous quadratic forms of normal variables. Ann Math. Statist. 33, 542-570.

71. Seshadri,V. (1963a). Constructing uniformly better estimators. J. Amer. Stat. Assoc. 58 172-175.

72. Seshadri,V. (1963b). *Combining unbiased estimators* Biometrics 19 163-170.

73. Shaarawi,A.E. Prentice,R.L. and Shah,K.R. (1975). Marginal procedures for mixed models with reference to block design. Sankhyā B 37 91-99.

74. Shah,K.R. (1964a). On a local property of combined inter and intra-block estimators. Sankhyā 26 87-90.

75. Shah,K.R. (1964b). Use of inter-block information to obtain uniformly better estimators. Ann. Math. Statist. 35 1064-1078.

76. Shah,K.R. (1970). On the loss of information in combined inter and intra-block estimation. J. Amer, Statist. Asso. 65 1562-1564.

77. Shah,K.R. (1971). Use of truncated estimator of variance ratio in recovery of inter-block information. Ann. Math. Statist. 42 816-819.

78. Shah,K.R. (1975). Analysis of block designs. Gujrat Statistical Review 2 1-11.

79. Shah,K.R. and Puri,S.C. (1976). Application of MINQUE procedures to block designs. Comm. Statist. 5 191-196.

80. Shinozaki,N. (1978). A note on estimating the common mean of K-normal distributions and stein problem Comm. Statist. Theor Math. A7 1421-1432.

81. Sprott,D.A. (1956). A note on combined inter-block and intra-block estimation in incomplete block design. Ann. Math. Stat. 27 633-641 [Erratum in 28 (1957) 269]

82. Stein,C. (1950). Unbiased estimates with minimum variance, Ann. Math. Statist. 21 406-425.

83. Stein,C. (1966). An approach to the recovery of inter-block information in balanced incomplete block designs. Reserch papers in Statistics ed. F.N. David, Wiley, New York.

Thompson,R. (1968). Iterative estimation of variance components for *non-orthogonal* data. Biometrics 25 767-773.

Tocher,K.D. (1952). The design and analysis of block experiments with discussion. J.R. Statist. Soc. B 14 45-100.

Wijsman,R.A. (1967). Cross sections and orbits and their applications to densities of maximal invariant. Proc. fifth Berkeley Symp. Math. Statist. Prob. 1, 389-400.

Williams,J.S. (1967). The variance of weighted regression estimators. J. Amer. Statist. Ass. 62 1290-1301.

Williams,J.S. (1975). Lower bounds on convergence rates of weighted least squares to best linear unbiased estimators. A Survey of Statistical Design and Linear Models Ed. J.N. Shrivastava. North-Holland, Amsterdam.

89. Yates,F. (1939a). An apparent inconsistency arising from tests of significance based on fiducial distributions of unknown parameters. Proc. Camb. Phil. Soc. 35 579-591.

90. Yates,F. (1939b). The recovery of inter-block information in varietal trials arranged in three dimensional lattices. Ann. Eugenics 9 135-156.

91. Yates,F. (1940). The recovery of inter-block information in incomplete block designs. Ann. Eugenics 10 325.

92. Yates,F. and Cochran,W.G. (1938). The analysis of groups of experiments Journ. of agricultural science 28 556-580

93. *Youden,W.J. (1951). Linked blocks : a new class of incomplete block* designs (Abstract), Biometrics 7 124.

94. Zacks,S. (1966). Unbiased estimation of the commom mean of two normal distributions based on small samples of equal size. J. Amer. Statist. Ass. 61 467-476.

95. Zacks,S. (1970). Bayes and fiducial equivariant estimators of the common mean of two normal distributions. Ann. Math. Statist. 41 59-67.

Addendum

1. Blackwell (1947). Conditional expectation and unbiased sequential estimation, Ann. Math. Statist. 18 105-110.

2. Kakwani,N.C. (1967). The unbiasedness of Zellner's seemingly unrelated regression equation estimators. J.Amer. Statist. Ass. 62 141-142.

3. *Rao,C.R. (1945). Information and accuracy attainable in the estimation* of statistical parameters. Bull. Cal. Math. Soc. 37 81-91

4. Rao,C.R. (1971a). Estimation of variance and covariance components - MINQUE theory. J.Multivariate Anal. 1 257-275.

5. Sinha,B.K. (1978). Is the maximum likelihood estimate of the common mean of several normal populations admissible? Sankhyā 40 B 192-196.

APPENDIX

SOME INEQUALITIES

We present here the derivations of some inequalities which were used in the text. Suppose f and g are functions of k random variables $x_1,\dots,x_k$. We shall use the symbol $E_n f$ where $n \leq k$ to denote the conditional expectation of f given $(x_1,\dots,x_n)$, we shall also use the abbreviations: $f \uparrow x_i$ for the statement f is non-decreasing in x_i's; $f \downarrow x_i$ for the statement f is non-increasing in x_i'. The abbreviation $f\ SDg|x_i$ would mean that f and g are monotonic in the same direction with respect to x_i in the sense that either $f \uparrow x_i$ and $g \uparrow x_i$ or $f \downarrow x_i$ and $g \downarrow x_i$. Similarly, $f\ ODg|x_i$ would mean that f and g are monotonic in opposite directions with respect to x_i in the sense that either $f \uparrow x_i$ and $g \downarrow x_i$ or $f \downarrow x_i$ and $g \uparrow x_i$. We now prove,

<u>Theorem A.1</u> Let u,v,t be a.s. positive functions of k random variables $x_1,\dots,x_k$. Assume that v has a finite expectation. Let, $g_n = E_n(tv)/E_nv$; $h_n = E_n u/E_nv$. Then

$$\text{(i)}\quad g_n\ SD\ h_n|x_n\ \forall\ n \leq k \implies \text{(ii)}\ E(tu)/E(tv) \geq Eu/Ev \tag{A.1}$$

$$\text{(iii)}\ g_n\ OD\ h_n|x_n\ \forall\ n \leq k \implies \text{(iv)}\ E(tu/E(tv) \leq Eu/Ev \tag{A.2}$$

<u>Proof</u>

$$\text{(ii)} \iff E(tu)/Ev \geq [(Eu/Ev)\ E(tv)/Ev] \iff E^*(tw) \geq E^*w\ E^*t \tag{A.3}$$

where $w = u/v$; E^* stands for expectation with respect to the density $f^* = vf/Ev$; f is the joint density of $(x_1,\dots,x_k)$. Note that, $E_n^*(t) = g_n$; $E_n^*(w) = h_n$. Hence

$$\text{(i)} \iff E_n^*(t)\ SD\ E_n^*(w)|x_n\quad \forall\ n \leq k \implies t\ SD\ w|x_k \tag{A.4}$$

is well-known that [see e.g. Hardy, Littlewood and Polya (1952), p.43 Kimball (1951), p. 600],

$$f \ \text{SD}\ g|x \implies E(fg) \geq Ef\, Eg \qquad (A.5)$$

$$f \ \text{OD}\ g|x \implies E(fg) \leq Ef\, Eg \qquad (A.6)$$

view of (A.4), (A.5) and (A.3)

$$\implies E^*(tw) = E^*E^*_{k-1}(tw) \geq E^*[E^*_{k-1}(t)\cdot E^*_{k-1}(w)]$$

$$= E^*E^*_{k-2}[E^*_{k-1}(t)E^*_{k-1}(w)] \geq E^*[E^*_{k-2}(t)E^*_{k-2}(w)] =$$

$$\cdots \geq E^*[E^*_1(t)E^*_1(w)] \geq E^*(t)\ E^*(w) \implies \text{(ii)}$$

is proves (A.1). *The proof of (A.2) is similar. In this case we use* (A.6) instead of (A.5).

Theorem A.2 Let $x_1,\ldots,x_k$ be mutually independent a.s. positive random variables and let

$$f = 1/\sum_{i=1}^{k} p_i x_i \ ; \quad 0 \leq p_i \leq 1; \ \sum_{i=1}^{k} p_i = 1.$$

Assume that Ex_i^{-2} is finite for every i. Then

$$Ef/Ef^2 \geq \min_{1 \leq i \leq k} Ex_i^{-1}/Ex_i^{-2}$$

Proof The theorem is trivial for k = 1. *It will be proved for k = 2 from* which the extensions to higher value of k will be obvious. To avoid subscripts, let p,x,y stand for p_1, x_1, x_2 respectively and let $m = \text{Min}(Ex^{-1}/Ex^{-2},\ Ey^{-1}/Ey^{-2})$. Then

$$m \leq pEx^{-1}/Ex^{-2} + (1-p)Ey^{-1}/Ey^{-2} = [pEx^{-1}Ey^{-2}+(1-p)Ey^{-1}Ex^{-2}]/(Ex^{-2}Ey^{-2})$$

$$= E[px^{-1}y^{-2} + (1-p)y^{-1}\,x^{-2}]/E(x^{-2}\,y^{-2}) \qquad (A.7)$$

since x and y are independent. Now, define $g = p/y+(1-p)/x$ and note that $fg = \bar{x}^{1}\ \bar{y}^{1}$. Then (A.7) can be written as:

$$m \leq E(tu)/E(tv) \qquad (A.8)$$

where $t = g^2$; $u = f$; $v = f^2$. Obviously, $t \downarrow y$ and $u/v \uparrow y$. Also, $E(tv|x)/E(v|x) = E(f^2g^2|x)/E(f^2|x) = E(x^{-2}\ y^{-2}|x)/E(x^{-2}\ y^{-2}\ t^{-1}|x)$ $= Ey^{-2}|x)/E(y^{-2}t^{-1}|x) \downarrow x$, since $t \downarrow x$; and, $E(u|x)/E(v|x) =$ $E(f|x)/E(f^2|x) \uparrow x$, since its derivative with respect to x is: $p[E(f|x)E(f^3|x) - E^2(f^2|x)]/E^2(f^2|x) \geq 0$, in view of an well-known inequality concerning absolute moments. Hence, using Theorem A.1, (A.8) gives:

$$m \leq Eu/Ev = Ef/Ef^2$$

This completes the proof for $k = 2$. When $k > 2$, $\sum_{i=1}^{k} p_i\ x_i$ can be written in the form: $q_k x + (1-q_k)y$, where $x = \sum_{i=1}^{k-1} q_i x_i$; $y = x_k$; $0 \leq q_i \leq 1$; $\sum_{i=1}^{k-1} q_i = 1$. Hence, it is easy to see that the result follows by induction.

We shall derive two more inequalities, for which we need the following lemmas:

<u>Lemma A.1</u> Let, $f = 1/(\gamma+h)$; $\bar{f} = 1/h$, where $h = p-q\gamma$, γ is a real variable and p,q are constants. Let primes denote derivation with respect to γ. Then

(i) provided $\gamma \neq 0$,

$$f' = (pf^2-f)/\gamma \quad ; \quad \bar{f}' = (p\bar{f}^2 - \bar{f})/\gamma$$

(ii) provided $h \neq 0$

$$f' = (qf - pf^2)/h$$

f Elementary and hence omitted

a A.2 Let f be a function of (y,γ), where y is a random variable and a constant which can assume values in a specified range. Let

$$g \quad Ef/Ef^2 \tag{A.9}$$

ume that (i) the distribution of y does not depend on γ; and for every (γ_1, γ_2), (ii) $f > 0$ a.s. (iii) $Ef^2 < \infty$ (iv) f is differentiable with spect to γ and,

$$f' = qf - pf^2 \tag{A.10}$$

ere p is a function of (y,γ); q is a function of γ only and the prime ands for derivation with respect to γ(v) p is measurable in y and we have ither $p > 0$ a.s. or, $p < 0$ a.s. Then, for any given $\gamma \in (\gamma_1,\gamma_2)$,

A) $Ef' \leq 0 \implies g' \geq 0$ if (vi): f SD pf|y

B) $Ef' \geq 0 \implies g' \leq 0$ if (vii) f OD pf|y

roof First assume that $p > 0$ a.s. Note that (ii) and (iii) imply $0 < Ef^2 < \infty$, $0 < Ef < \infty$. Hence in view of (i), (A.9) gives ;

$$g' = (Ef' \, Ef^2 - 2Ef \, Eff')/E^2f^2$$

using (A.10). This expression reduces to

$$g' = (2Ef \, Epf^3 - Epf^2Ef^2 - qEf \, Ef^2)/E^2f^2 \tag{A.11}$$

Also from (A.10)

$$qEf \leq Epf^2 \quad \text{if} \quad Ef' \leq 0; \quad qEf \geq Epf^2 \quad \text{if} \quad Ef' \geq 0$$

Using this (A.11) gives :

$$g' \geq C \quad \text{if} \quad Ef' \leq 0; \quad g' \leq C \quad \text{if} \quad Ef' \geq 0 \tag{A.12}$$

here

$$C = 2(Ef\ Epf^3 - Ef\ Epf^2)/E^2f^2 \tag{A.13}$$

ly Theorem A.1

$$\begin{aligned} &(vi) \Rightarrow Epf^3/Epf^2 \geq Ef/Ef^2 \Rightarrow C \geq 0 \\ &(vii) \Rightarrow Epf^3/Epf^2 \leq Ef/Ef^2 \Rightarrow C \leq 0 \end{aligned} \tag{A.14}$$

The desired results follow from (A.12) and (A.14). If $p < 0$ a.s. we can write (A.13) as:

$$C = -2(Ef Ep_*f^3 - Ef^2Ep_*f^2)/E^2f^2$$

where $p_* = -p > 0$ a.s. Note that, f SD (OD) $pf|y \Longleftrightarrow f$ OD (SD) $p_*f|y$.

Hence, by Theorem A.1,

$$\begin{aligned} &(vi) \Rightarrow Ep_*f^3/Ep_*f^2 \leq Ef/Ef^2 \Rightarrow C \geq 0 \\ &(vii) \Rightarrow Ep_*f^3/Ep_*f^2 \geq Ef/Ef^2 \Rightarrow C \leq 0 \end{aligned} \tag{A.15}$$

The desired results now follow from (A.12) and (A.15).

<u>Lemma A.3</u> Let f be a function of (y,γ), where y is a random variable and γ is a constant. Let $\bar{f}$ be a measurable function of y only. Let

$$g = Ef/E(f\bar{f}) \tag{A.16}$$

Assume that (i) the distribution of y does not depend on γ; and for every $\gamma \in (\gamma_1, \gamma_2)$, (ii) $f > 0$ a.s. (iii) $\bar{f} > 0$ a.s. (iv) $Ef < \infty$ (v) $E(f\bar{f}) < \infty$ (vi) f is differentiable with respect to γ and,

$$f' = qf - pf^2 \tag{A.17}$$

where p is a function of (y,γ); q is a function of γ only and the prime stands for derivation with respect to γ (vii) p is measurable in y and we have either $p > 0$ a.s. or $p < 0$ a.s. Then, for any given $\gamma \in (\gamma_1,\gamma_2)$, α) $\bar{g}' \geq 0$ if (viii) $\bar{f}$ SD $pf|y$

β) $\bar{g}' \leq 0$ if (ix) $\bar{f}$ OD $pf|y$

Proof Note that (ii), (iii) and (v) imply : $0 < E(f\bar{f}) < \infty$; $0 < Ef < \infty$. Hence in view of (i), (A.16) gives:

$$\bar{g}' = [Ef'E(f\bar{f}) - EfE(f'\bar{f})]/E^2(f\bar{f})$$

using (A.17), this expression becomes,

$$\bar{g}' = [EfE(pf^2\bar{f}) - E(f\bar{f})E(pf^2)]/E^2(f\bar{f}) \qquad (A.18)$$

First assume that $p > 0$ a.s. Then by Theorem A.1,

$$\begin{aligned} (vii) &\Rightarrow E(pf^2\bar{f})/E(pf^2) \geq E(f\bar{f})/Ef \\ (ix) &\Rightarrow E(pf^2\bar{f})/E(pf^2) \leq E(f\bar{f})/Ef \end{aligned} \qquad (A.19)$$

The desired results follow from (A.18) and (A.19). If $p < 0$ a.s., a modification of the above arguments yield the desired results as in the proof of the previous lemma.

Lemma A.4 Let f and $\bar{f}$ be measurable functions of a random variable y such that (i) $f > 0$ a.s. (ii) $f \leq \bar{f}$ a.s. (iii) $Ef < \infty$ (iv) $\bar{f}$ SD $\bar{f}/f|y$. Then

$$Ef/Ef^2 \geq E\bar{f}/E\bar{f}^2$$

Proof In view of (i) - (iii),

$$Ef/Ef^2 \geq Ef/E(\bar{f}f) \qquad (A.20)$$

Also, in view of (i) - (iv), theorem A.1 gives

$$E\bar{f}^2/E(\bar{f}f) \geq E\bar{f}/Ef \qquad (A.21)$$

The desired result follows from (A.20) and (A.21).

<u>Lemma A.5</u> Let,

$$f = 1/[\gamma + dh(\gamma)y] \tag{A.22}$$

where $h(\gamma) = p - q\gamma$, y is a random variable, γ is a parameter and q, d are given constants. Assume that (i) $d > 0$ (ii) $p > \max(0, q)$ (iii) $y > 0$ a.s. (iv) $Ey^{-2} < \infty$ (v) the distribution of y does not depend on γ. Let, $g(\gamma) = Ef/Ef^2$; $\pi = \inf_{\gamma\epsilon(0,1)} g(\gamma)$; $\pi_* = \inf_{\gamma\epsilon(0,\gamma_0)} g(\gamma)$, where $\gamma_0 \in (0,1)$; $a_0 = Ey^{-1}/Ey^{-2}$; $\delta_1 = da_0p$; $\delta_2 = p/q$; $\delta_3 = \max[\delta_2, da_0h(1)]$; $\delta_4 = \max[\delta_2, da_0h(\gamma_0)]$. Then,

(A) $\pi = \delta_1$; $\pi_* = \delta_1$ if $q \le 0$

(B) $\pi \ge \text{Min}(\delta_1, \delta_3)$; $\pi_* \ge \text{Min}(\delta_1, \delta_4)$ if $q > 0$ (A.23)

<u>Proof</u> Assume throughout that $\gamma \in (0,1)$. Then assumption (ii) implies: $h > 0$. Hence by lemma A.1

$$f' = (qf - pf^2)/h \tag{A.24}$$

Assumptions (i), (ii) and (iii) imply: $0 < Ef^2 < \infty$; $0 < Ef < \infty$. It is now easy to see that f, y satisfy all conditions of part A of lemma A.2. Hence, by that lemma,

$$Ef' \le 0 \Rightarrow g' \ge 0 \tag{A.25}$$

From (A.22), it is easy to see that $f'' \ge 0$. This means f' and hence $Ef' \uparrow \gamma$. Hence

$$Ef' \le 0 \text{ for some } \gamma = \gamma_1 \in (0,1) \Rightarrow Ef' \le 0 \ \forall\ \gamma \in (0, \gamma_1) \tag{A.26}$$

(A.25) and (A.26) together imply that $g \uparrow \gamma\ \forall\ \gamma \in (0, \gamma_1)$ if $Ef' \le 0$ for some $\gamma = \gamma_1 \in (0,1)$. Hence

$$g \ge g(0) = da_0p \text{ if } Ef' \le 0 \tag{A.27}$$

Now, using (A.24),

$$q \leq 0 \Longrightarrow f' \leq 0 \Rightarrow Ef' \leq 0, \forall \gamma \epsilon (0,1) \qquad (A.28)$$

(A.27) and (A.28) proves the part A. From (A.28), note that $Ef' > 0 \Rightarrow q > 0$. Hence, it follows from (A.24) that,

$$g > p/q \quad \text{if} \quad Ef' > 0 \qquad (A.29)$$

(A.27) and (A.29) imply

$$\pi \geq \delta_o \;;\; \pi_* \geq \delta_o \quad \text{if} \quad q > 0 \qquad (A.30)$$

where $\delta_o = \min(\delta_1, \delta_2)$. Now, let, $\bar{f} = 1/[dh(\gamma)y]$. Then, by lemma A.4,

$$g = Ef/Ef^2 \geq E\bar{f}/E\bar{f}^2 = da_o h(\gamma), \quad \forall \gamma \epsilon (0,1) \qquad (A.31)$$

Note that $q \neq 0$ implies that $h(\gamma) \neq \gamma$, $\forall \gamma \epsilon (0,1)$. Hence, (A.31) gives:

$$\pi \geq da_o h(1) \;;\; \pi_* \geq da_o h(\gamma_o) \quad \text{if} \quad q > 0 \qquad (A.32)$$

(A.30) and (A.32) imply

$$\pi \geq \max[da_o h(1), \delta_o)];\; \pi_* \geq \max\,[da_o h(\gamma_o), \delta_o] \text{ if } q > 0 \qquad (A.33)$$

Note that $q > 0 \Longrightarrow da_o h(1) < da_o h(\gamma) < da_o h(0) = \delta_1$, $\forall \gamma \epsilon (0,1)$. Hence, it is easy to see that (A.33) is equivalent to (A.23).

<u>Lemma A.6</u> Let f_1 and f_2 be functions of a pair of random variables x and y. Let S denote the support of x and let the symbol E_t stand for the conditional expectation with respect to y given x = t. Then $Ef_1/Ef_2 \geq \inf_{t \epsilon S} E_t f_1/E_t f_2$.

Proof The proof is simple and hence omitted. We now prove

Theorem A.3 Let

$$f = 1/h(x;\gamma) \tag{A.34}$$

where x is a random variable; γ is a parameter; $h(x;\gamma) = p(x)-q(x)\gamma$; and $p(x)$, $q(x)$ are measurable functions of x, which do not depend on γ. Assume that $p(x) > \max[0,q(x)]$ a.e. (ii) $Ef^2 < \infty$, $\forall\ \gamma \in [0,1]$ (iii) The distribution of x does not depend on γ. Let, $\bar{f} = 1/\min[h(x;0),h(x;\gamma_0)]$; $\tilde{f} = 1/\min[h(x;\gamma_0), h(x;1)]$, where $\gamma_0 \in (0,1)$. Let $g(\gamma) = Ef/Ef^2$; $\bar{g}(\gamma) = Ef/E(f\bar{f})$; $\tilde{g}(\gamma) = Ef/E(f\tilde{f})$; $\pi = \inf_{\gamma\in(0,1)} g(\gamma)$; $\pi_* = \inf_{\gamma\in(0,\gamma_0)} g(\gamma)$. Then,

(A) $\pi \geq \min[\bar{g}(0),\tilde{g}(1)]$; $\pi_* \geq \bar{g}(0)$ if (iv) f OD $pf|x$, $\forall\ \gamma \in (0,\gamma_0)$; f SD $pf|x$, $\forall\ \gamma \in (\gamma_0,1)$.

(B) $\pi \geq \min[\bar{g}(0), g(1)]$ if in addition to (iv), we have (v):
(v): p SD $f|x$, $\forall\ \gamma \in (\gamma_0,1)$.

Proof Assume throughout that $\gamma \in (0,1)$. Then by (i), $p, f, \bar{f}, \tilde{f}$ and hence pf are all a.s. positive. This together with (ii) imply that Ef, $Ef\bar{f}$ and $Ef\tilde{f}$ are all positive and finite along with Ef^2. Let primes denote derivation with respect to γ. Then using lemma A.1, (A.34) gives

$$f' = [p(x)f^2-f]/\gamma = q_*f - p_*f^2 \tag{A.35}$$

where $p_* = -p(x)/\gamma$; $q_* = -1/\gamma$. Clearly, we have : f SD $\bar{f}|x$, $\forall\gamma \in (0,\gamma_0)$; f SD $\tilde{f}|x$, $\forall\ \gamma \in (\gamma_0,1)$. Hence, (iv) implies : $\bar{f}$ SD $p_*f|x$, $\forall\ \gamma \in (0,\gamma_0)$; $\tilde{f}$ OD $p_*f|x$, $\forall\ \gamma \in (\gamma_0,1)$. Hence, by lemma A.2, we have:

$$\bar{g}(\gamma) \uparrow \gamma \quad \forall\ \gamma \in (0,\gamma_0);\ \tilde{g}(\gamma) \downarrow \gamma \in (\gamma_0,1). \quad \text{Hence,}$$

$$\bar{g}(\gamma) \geq \bar{g}(0), \quad \forall\ \gamma \in (0,\gamma_0) \tag{A.36}$$

$$\tilde{g}(\gamma) \geq \tilde{g}(1), \quad \forall\ \gamma \in (\gamma_0,1) \tag{A.37}$$

t is easy to see that: $f \leq \hat{f}$, $\forall\ \gamma \in (0,\gamma_o)$; $f \leq \tilde{f}$, $\forall\ \gamma \in (\gamma_o,1)$. Hence, $(\gamma) \geq \bar{g}(\gamma)$, $\forall\ \gamma \in (0,\gamma_o)$; $g(\gamma) \geq \tilde{g}(\gamma)$, $\forall\ \gamma \in (\gamma_o,1)$. This proves part (A) In view of (A.36) and (A.37). The additional assumption in part (B) implies: $E[p(x)f^2] \geq E[p(x)]\, Ef^2$, $\forall\ \gamma \in (\gamma_o,1)$. Hence using (A.35), we have: $Ef' \leq 0 \Rightarrow g(\gamma) \geq Ep(x)$, $\forall\ \gamma \in (\gamma_o,1)$. On the other hand by lemma A.2, $Ef' \geq 0 \Rightarrow g(\gamma) \geq g(1)$, $\forall\ \gamma \in (\gamma_o,1)$. Hence,

$$g(\gamma) \geq \min[Ep(x),\ g(1)], \quad \forall\ \gamma \in (\gamma_o,1) \tag{A.38}$$

(A.36) and (A.38) give: $\pi \geq \min\ [\bar{g}(0),\ Ep(x), g(1)]$. This proves part (B) since,

$$\bar{g}(0) \leq g(0) = E[p(x)]^{-1}/E[p(x)]^{-2} \leq Ep(x)$$

<u>Theorem A.4</u> Let, $f = 1/[\gamma + dh(x;\gamma)y]$, where x,y are random variables; γ, d are constants which can assume values in specified ranges; $h(x;\gamma) = p(x) - q(x)\gamma$; and p,q are measurable functions of x, which do not depend on γ. Assume that (i) $d > 0$ (ii) $p(x) > \max[0,q(x)]$ a.s. (iii) $y > 0$ a.s. (iv) $Ey^{-2} < \infty$ (v) x and y are independent (vi), the distribution of (x,y) does not depend on γ. Let $f_* = 1/h(x;\gamma)$; $\hat{f}_* = 1/\min[h(x;0), h(x;\gamma_o)]$; $\tilde{f}_* = 1/\min[h(x;\gamma_o),\ h(x;1)]$; $g(\gamma) = Ef/Ef^2$; $g_*(\gamma) = Ef_*/Ef_*^2$; $\bar{g}_*(\gamma) = Ef_*/E(f_*\hat{f}_*)$; $\tilde{g}(\gamma) = Ef_*/E(f_*\tilde{f}_*)$; $\pi = \inf\limits_{\gamma\in(0,1)} g(\gamma)$; $\pi_* = \inf\limits_{\gamma\in(0,\gamma_o)} g(\gamma)$, where $\gamma_o \in (0,1)$. Let S be the support of x and let, $S_* = \{t | t \in S,\ q(t) > 0\}$. Let, $a_o = Ey^{-1}/Ey^{-2}$; $\delta_1 = \inf\limits_{t \in S} p(t)$; $\delta_2 = \inf\limits_{t \in S_*} p(t)/q(t)$; $\delta_3 = \inf\limits_{t \in S_*} h(t;1)$; $\delta_4 = \inf\limits_{t \in S_*} h(t;\gamma_o)$; $\delta_5 = \max(\delta_2,\ da_o\ \delta_3)$;

$\delta_6 = \max(\delta_2,\ da_o\ \delta_4)$; $\delta_7 = \min[\bar{g}_*(0),\ \bar{g}_*(1)]$; $\delta_8 = \min[\bar{q}_*(0), q_*(1)]$;

$\pi_1 = \min(da_o\delta_1, \delta_5)$; $\pi_2 = \min(da_o\ \delta_1,\ \delta_6)$. Then,

(A)
$$\pi \geq \pi \quad ; \quad \pi_* \geq \pi_2 \tag{A.39}$$

(B) $\pi \geq \max(\pi_1, da_o\delta_7)$; $\pi_* \geq \max[\pi_2, da_o\bar{g}_*(0)]$ if (vii):
f DD pf|x, $\forall \gamma \in (0,\gamma_o)$; f SD pf|x, $\forall \gamma \in (\gamma_o,1)$

(C) $\pi \geq \max(\pi_1, da_o\delta_8)$, if in addition to (vii), we have (viii): p SD f|x, $\forall \gamma \in (\gamma_o,1)$.

Proof Let, $\delta_1(t) = da_o p(t)$; $\delta_2(t) = p(t)/q(t)$;

$\delta_3(t) = \max[\delta_2(t), da_o h(t;1)]$; $\delta_4(t) = \max[\delta_2(t), da_o h(t;\gamma_o)]$;

$$\pi_1(t) \quad \min[\delta_1(t), \delta_3(t)] \text{ if } t \in S_*$$

$$= \delta_1(t) \text{ if } t \in S-S_*;$$

$$\pi_2(t) = \min[\delta_1(t), \delta_4(t)] \text{ if } t \in S_*$$

$$= \delta_1(t) \text{ if } t \in S-S_* \;;$$

Let, $g_t = E_t f/E_t f^2$, where E_t stands for conditional expectation given $x = t$; $\pi(t) = \inf_{\gamma\varepsilon(0,1)} g_t$; $\pi_*(t) = \inf_{\gamma\varepsilon(0,\gamma_o)} g_t$. By lemma A.5,

$$\pi(t) \geq \pi_1(t); \quad \pi_*(t) \geq \pi_2(t) \tag{A.40}$$

It is easy to see that,

$$\pi_1(t) \geq \pi_1 \;; \quad \pi_2(t) \geq \pi_2, \quad \forall t \in S \tag{A.41}$$

By lemma A.6,

$$g \geq \inf_{t \in S} g_t$$

Using (A.42), (A.40) and (A.41),

$$\pi \geq \inf_{\gamma\varepsilon(0,1)} \inf_{t \in S} g_t = \inf_{t \in S} \pi(t) \geq \inf_{t \in S} \pi_1(t) \geq \pi_1.$$

Similarly, $\pi_* \geq \pi_2$. This proves part A. Now, let $\bar{f} = 1/[dh(x;\gamma)y]$. Then, using lemma A.4,

$$g \geq E\bar{f}/E\bar{f}^2 = da_o g_* \tag{A.43}$$

since y is independent of x. Using theorem A.3, Part (B) and part (C) follow from (A.39) and (A.43).

LIST OF AUTHOR'S PUBLICATIONS RELATED TO THIS THESIS

1. Bhattacharya,C.G. and Shah,K.R. (1978). Interval estimation of treatment differences in block designs. Journal of Statistical Computation and Simulation 6 243-255.

2. Bhattacharya,C.G. (1978). Yates type estimators of a common mean. Annals of the Institute of Statistical Mathematics 30 A 407-414.

3. Bhattacharya,C.G. (1979). A note on estimating the common mean of k-normal populations. Sankhyā 40 B 272-275.

4. Bhattacharya,C.G. (1980). Estimation of a common mean and recovery of inter-block information. Annals of Statistics 8 205-211.

5. Bhattacharya,C.G. (1981). Estimation of a common location. To appear in communications in statistics - theory and methods A 10 No. 10.

INDIAN STATISTICAL INSTITUTE LIBRARY
7095
13 FEB 1985
Date
CALCUTTA