RANK APPROACH TO THE MULTIVARIATE
TWO-POPULATION MIXTURE PROBLEM

by

Shoutir Kishore Chatterjee

University of Calcutta, India
and
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 817

MAY 1972

Rank Approach to the Multivariate Two-Population Mixture Problem*

SHOUTIR KISHORE CHATTERJEE

University of Calcutta, India
and
University of North Carolina at Chapel Hill, U.S.A.

*Work carried out at the Department of Biostatistics, U.N.C., under the Multivariate Nonparametric Project sponsored by the Aerospace Research Laboratories, Air Force Systems Command, U. S. Air Force Contract No. F 33615-71-C-1927. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

1. INTRODUCTION

We consider the following problem. Three independent random samples

$$\underset{\sim}{X}_{\alpha}^{(k)} = (X_{1\alpha}^{(k)}, X_{2\alpha}^{(k)}, \ldots, X_{p\alpha}^{(k)})', \quad \alpha=1,2,\ldots,n_k;\ k=0,1,2 \tag{1.1}$$

from three unknown p-variate populations with continuous cumulative distribution functions (cdf's) $F^{(0)}(\underset{\sim}{x})$, $F^{(1)}(\underset{\sim}{x})$, and $F^{(2)}(\underset{\sim}{x})$ are given. It is known that $F^{(1)}(\underset{\sim}{x})$ and $F^{(2)}(\underset{\sim}{x})$ are distinct and $F^{(0)}(\underset{\sim}{x})$ is a mixture of $F^{(1)}(\underset{\sim}{x})$ and $F^{(2)}(\underset{\sim}{x})$, i.e.,

$$F^{(0)}(\underset{\sim}{x}) = \theta\ F^{(1)}(\underset{\sim}{x}) + (1-\theta)F^{(2)}(\underset{\sim}{x}), \quad 0\leq\theta\leq 1 \tag{1.2}$$

The 'mixture rate' θ is unknown. To estimate θ.

For this problem we consider procedures which utilize only the observational ranks, and hence, can be used when ranks only are available.

For the i-th variate, let us rank all the $N=n_0+n_1+n_2$ observations from the three samples together, and let the rank of $X_{i\alpha}^{(k)}$ so obtained be $R_{i\alpha}^{(k)}$. In this way, from (1.1) we derive the rank vectors

$$\underset{\sim}{R}_{\alpha}^{(k)} = (R_{1\alpha}^{(k)}, R_{2\alpha}^{(k)}, \ldots, R_{p\alpha}^{(k)})', \quad \alpha=1,2,\ldots,n_k;\ k=0,1,2 \tag{1.3}$$

Now suppose a p×N score matrix (depending on N)

$$\underset{\sim}{A}_N = (a_{Ni}(\alpha))_{i=1,\ldots,p;\ \alpha=1,\ldots,N} \tag{1.4}$$

is given. With its help, we convert the ranks into rank scores

$$a_{Ni}(R_{i\alpha}^{(k)}) = a_{i\alpha}^{(k)}\ \text{(say)}, \quad \alpha=1,\ldots,n_k;\ k=0,1,2;\ i=1,\ldots,p. \tag{1.5}$$

Here $a_{i\alpha}^{(k)}$ represents the random rank score corresponding to the observation $X_{i\alpha}^{(k)}$. Let

$$\underset{\sim}{a}_{\alpha}^{(k)} = (a_{1\alpha}^{(k)}, a_{2\alpha}^{(k)}, \ldots, a_{p\alpha}^{(k)})', \qquad \alpha=1,2,\ldots,n_k;\ k=0,1,2$$

$$\bar{\underset{\sim}{a}}^{(k)} = (\bar{a}_1^{(k)}, \bar{a}_2^{(k)}, \ldots, \bar{a}_p^{(k)})' = \frac{1}{n_k} \sum_{\alpha=1}^{n_k} \underset{\sim}{a}_{\alpha}^{(k)}. \tag{1.6}$$

The relation (1.2) means that we can regard the n_0 observations from $F^{(0)}$ to have been taken in the following two steps. First, n_0 Bernoulli trials with success probability θ are performed. Then, for each trial, an observation from $F^{(1)}$ or $F^{(2)}$ is taken according as the trial shows a success or a failure. With this interpretation, it immediately follows that if the number r of successes were observable, we could disregard the rest of the data, and take $t = r/n_0$ as our estimate of θ. Since r is not observable, we use the data to find some sort of 'estimate' of t.

If r, as well as the serial numbers $\alpha_1, \alpha_2, \ldots, \alpha_r$ $(1 \le \alpha_1 < \alpha_2 < \ldots < \alpha_r \le n_0)$ of the trials resulting in success are given, then, conditionally, $\underset{\sim}{X}_{\alpha}^{(0)}$, $\alpha=\alpha_1,\alpha_2,\ldots,\alpha_r$, would have the same distribution as $\underset{\sim}{X}_{\alpha}^{(1)}$, $\alpha=1,2,\ldots,n_1$ and $\underset{\sim}{X}_{\alpha}^{(0)}$, $\alpha\neq\alpha_1,\alpha_2,\ldots,\alpha_r$, the same distribution as $\underset{\sim}{X}_{\alpha}^{(2)}$, $\alpha=1,2,\ldots,n_2$. Hence, given r and $\alpha_1,\alpha_2,\ldots,\alpha_r$, the random score vectors $\underset{\sim}{a}_{\alpha}^{(0)}$, $\alpha=\alpha_1,\alpha_2,\ldots,\alpha_r$ would be interchangeable with $\underset{\sim}{a}_{\alpha}^{(1)}$, $\alpha=1,2,\ldots,n_1$ and $\underset{\sim}{a}_{\alpha}^{(0)}$, $\alpha\neq\alpha_1,\alpha_2,\ldots,\alpha_r$ would be interchangeable with $\underset{\sim}{a}_{\alpha}^{(2)}$, $\alpha=1,2,\ldots,n_2$. Thus, we would have

$$E(\bar{\underset{\sim}{a}}^{(0)} | r,\alpha_1,\ldots,\alpha_r) = t\, E(\bar{\underset{\sim}{a}}^{(1)} | r,\alpha_1,\ldots,\alpha_r) + (1-t)E(\bar{\underset{\sim}{a}}^{(2)} | r,\alpha_1,\ldots,\alpha_r)$$

so that unconditionally also,

$$E(\bar{\underset{\sim}{a}}^{(0)} - t\,\bar{\underset{\sim}{a}}^{(1)} - (1-t)\bar{\underset{\sim}{a}}^{(2)}) = \underset{\sim}{o} \tag{1.7}$$

For any fixed non-null p-vector $\underset{\sim}{\ell}$, we have, therefore,

$$E[(\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(2)} - \underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(0)}) - t(\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(2)} - \underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(1)})] = 0 \tag{1.8}$$

This suggests that we take as a determination of t, and hence, as an estimate of θ

$$\tilde{\theta} = \tilde{\theta}(\underset{\sim}{\ell}) = (\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(2)} - \underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(0)})/(\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(2)} - \underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(1)}) \tag{1.9}$$

provided the denominator is non-zero and the ratio lies in [0,1]. As $F^{(1)}$ and $F^{(2)}$ are distinct, for a suitable score matrix $\underset{\sim}{A}_N$, and a suitably chosen $\underset{\sim}{\ell}$, these conditions are expected to be realised with high probability at least in large samples. We call $\tilde{\theta}(\underset{\sim}{\ell})$ as 'the fixed-$\underset{\sim}{\ell}$ linear rank-score estimate' of θ. In the sequel, we shall allow $\underset{\sim}{\ell}$ itself to be determined by the data so as to achieve maximum asymptotic efficiency. The corresponding estimate would be called 'optimised linear rank-score estimate'.

In the remainder of this section we consider some general results on random sequences that will be used repeatedly later.

LEMMA 1.1 <u>If</u> $g(\underset{\sim}{x})$ <u>is a real-valued function continuous over an open p-dimensional interval</u> I, $\underset{\sim}{a}_N$ <u>is a p-vector sequence such that for sufficiently large N</u>, $\underset{\sim}{a}_N \subset J \subset I$ <u>where J is a bounded closed interval and</u> $\underset{\sim}{X}_N$ <u>is a sequence of random p-vectors such that</u> $\underset{\sim}{X}_N - a_N \overset{P}{\to} \underset{\sim}{o}$, <u>then</u>, $g(\underset{\sim}{X}_N) - g(\underset{\sim}{a}_N) \overset{P}{\to} 0$.

<u>Proof</u>. Clearly we can find a closed bounded interval J', such that $J \subset J' \subset I$ such that for sufficiently large N, $a_N \varepsilon J'$ and $\underset{\sim}{x}_N \varepsilon J'$ with probability arbitrarily close to 1. As $g(\underset{\sim}{x})$ is uniformly continuous in J' the lemma follows.

Given a sequence of random p-vectors $\underset{\sim}{X}_N$ and a sequence of positive definite matrices $\underset{\sim}{\Sigma}_N$, we say $\underset{\sim}{X}_N$ is asymptotically $N(\underset{\sim}{o},\underset{\sim}{\Sigma}_N)$ if for every non-null p-vector $\underset{\sim}{\ell}$, $\underset{\sim}{\ell}'\underset{\sim}{X}_N/[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_N\underset{\sim}{\ell}]^{\frac{1}{2}}$ converges in law to N(0,1).

For a symmetric matrix A we use the notations m(A) and M(A) to denote the minimum and maximum characteristic roots respectively. $\Phi(z)$ denotes the standard normal cdf.

LEMMA 1.2 Let $\underset{\sim}{X}_N$ be a sequence of random p-vectors and $\Sigma_N(p\times p)$ be a sequence of positive definite matrices such that $\underset{\sim}{X}_N$ is asymptotically $N(\underset{\sim}{o},\underset{\sim}{\Sigma}_N)$. Then, provided

$$0 < \liminf_{N\to\infty} \{m(\underset{\sim}{\Sigma}_N)/M(\underset{\sim}{\Sigma}_N)\} \tag{1.10}$$

we have, for every z

$$P\{\underset{\sim}{\ell}'\underset{\sim}{X}_N[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_N\underset{\sim}{\ell}]^{-\frac{1}{2}} \le z\} \to \Phi(z) \text{ uniformly in } \underset{\sim}{\ell}\ne\underset{\sim}{o}. \tag{1.11}$$

Proof. Let us write $m(\Sigma_N) = m_N$, $M(\Sigma_N) = M_N$. Since replacement of $\underset{\sim}{X}_N$ and Σ_N by $M_N^{-\frac{1}{2}}\underset{\sim}{X}_N$ and $M_N^{-1}\underset{\sim}{\Sigma}_N$ does not change $\underset{\sim}{\ell}'\underset{\sim}{X}_N/[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_N\underset{\sim}{\ell}]^{\frac{1}{2}}$, without any loss of generality, we may assume $M_N=1$ for all N. Then (1.10) can be restated as

$$0 < \liminf_{N\to\infty} m_N \le M_N = 1. \tag{1.12}$$

The implication of (1.12) is that given any sequence $\{N_k\}$ we can find a subsequence $\{N_k'\}$ such that $\lim_{k\to\infty} \underset{\sim}{\Sigma}_{N_k'}$ exists and is positive definite.

Suppose (1.11) does not hold i.e., there is a z' such that

$$\sup_{\underset{\sim}{\ell}\ne\underset{\sim}{o}} P\{\underset{\sim}{\ell}'\underset{\sim}{X}_N[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_N\underset{\sim}{\ell}]^{-\frac{1}{2}} \le z'\} - \Phi(z')| \not\to 0. \tag{1.13}$$

Then we can find a number $\varepsilon>0$ and a subsequence $\{N_k\}$ such that

$$\sup_{\underset{\sim}{\ell}\ne\underset{\sim}{o}} |P\{\underset{\sim}{\ell}'\underset{\sim}{X}_{N_k}[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_{N_k}\underset{\sim}{\ell}]^{-\frac{1}{2}} \le z'\} - \Phi(z')| > \varepsilon \text{ for all } k \tag{1.14}$$

and by (1.12) we can choose $\{N_k\}$ so that

$$\lim_{k\to\infty} \underset{\sim}{\Sigma}_{N_k} = \text{some positive definite matrix, say, } \underset{\sim}{\Sigma}_0. \tag{1.15}$$

Now, (1.15) together with the fact that $\ell'\underset{\sim}{X}_{N_k}/[\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_{N_k}\underset{\sim}{\ell}]^{\frac{1}{2}} \xrightarrow{\mathcal{L}} N(0,1)$ for all $\underset{\sim}{\ell}\ne\underset{\sim}{o}$, implies that $\underset{\sim}{X}_{N_k} \xrightarrow{\mathcal{L}} \underset{\sim}{Y}_o$ where we write $\underset{\sim}{Y}_o$ for a random vector distributed as $N(\underset{\sim}{o},\underset{\sim}{\Sigma}_o)$.

We now apply a general result on weak convergence due to Ranga Rao ([4] Theorem 4.2) from which we get that

$$\sup_{\underset{\sim}{\ell} \neq \underset{\sim}{o}} \left| P\{\underset{\sim}{\ell}'\underset{\sim}{X}_{N_k} \leq z'\} - P\{\underset{\sim}{\ell}'\underset{\sim}{Y}_o \leq z'\} \right| \to 0 \tag{1.16}$$

Now let us denote by $\underset{\sim}{Y}_{N_k}$ a random vector following exactly the distribution $N(\underset{\sim}{o}, \underset{\sim}{\Sigma}_{N_k})$. Then (1.15) implies $\underset{\sim}{Y}_{N_k} \xrightarrow{\ell} \underset{\sim}{Y}_o$, and again by Ranga Rao's result, we get

$$\sup_{\underset{\sim}{\ell} \neq \underset{\sim}{o}} \left| P\{\underset{\sim}{\ell}'\underset{\sim}{Y}_{N_k} \leq z'\} - P\{\underset{\sim}{\ell}'\underset{\sim}{Y}_o \leq z'\} \right| \to 0 \tag{1.17}$$

From (1.16) and (1.17) we deduce,

$$\sup_{\underset{\sim}{\ell} \neq \underset{\sim}{o}} \left| P\{\underset{\sim}{\ell}'\underset{\sim}{X}_{N_k} \leq z'\} - P\{\underset{\sim}{\ell}'\underset{\sim}{Y}_{N_k} \leq z'\} \right| \to 0$$

and hence,

$$\sup_{\underset{\sim}{\ell} \neq \underset{\sim}{o}} \left| P\{\underset{\sim}{\ell}'\underset{\sim}{X}_{N_k} [\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_{N_k}\underset{\sim}{\ell}]^{-\frac{1}{2}} \leq z'\} - P\{\underset{\sim}{\ell}'\underset{\sim}{Y}_{N_k} [\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_{N_k}\underset{\sim}{\ell}]^{-\frac{1}{2}} \leq z'\} \right| \to 0$$

i.e.,

$$\sup_{\underset{\sim}{\ell} \neq \underset{\sim}{o}} \left| P\{\underset{\sim}{\ell}'\underset{\sim}{X}_{N_k} [\underset{\sim}{\ell}'\underset{\sim}{\Sigma}_{N_k}\underset{\sim}{\ell}]^{-\frac{1}{2}} \leq z'\} - \Phi(z') \right| \to 0. \tag{1.18}$$

But this contradicts (1.14). Hence the lemma.

Note. Under the conditions of the lemma, if $\underset{\sim}{\ell}_N$ is a sequence of non-null vectors, then $\underset{\sim}{\ell}_N'\underset{\sim}{X}_N$ is asymptotically distributed as $N(0, \underset{\sim}{\ell}_N'\underset{\sim}{\Sigma}_N\underset{\sim}{\ell}_N)$.

2. SOME PRELIMINARY RESULTS

In the following sections we shall assume that for every N there is a triplet (n_0, n_1, n_2) such that $n_0 + n_1 + n_2 = N$. We write

$$\lambda_k = n_k/N, \quad k=0,1,2, \quad \lambda_0 + \lambda_1 + \lambda_2 = 1 \tag{2.1}$$

Whenever $N \to \infty$ we assume the following holds.

ASSUMPTION I There is a number λ^* $(0<\lambda^*<\frac{1}{3})$ such that

$$\lambda^* < \lambda_k < 1-\lambda^*, \quad k=0,1,2, \text{ for all } N \tag{2.2}$$

We write

$$H(\underset{\sim}{x}) = \lambda_o F^{(0)}(\underset{\sim}{x}) + \lambda_1 F^{(1)}(\underset{\sim}{x}) + \lambda_2 F^{(2)}(\underset{\sim}{x}) \tag{2.3}$$

By (1.2), this gives

$$H(\underset{\sim}{x}) = (\lambda_0\theta+\lambda_1)F^{(1)}(\underset{\sim}{x}) + (\lambda_0(1-\theta)+\lambda_2)F^{(2)}(\underset{\sim}{x}). \tag{2.4}$$

As λ_0, λ_1, λ_2 vary with N, so does $H(\underset{\sim}{x})$. (2.2) entails

$$F^{(k)}(\underset{\sim}{x}) \leq \frac{1}{\lambda^*} H(\underset{\sim}{x}), \qquad k=0,1,2. \tag{2.5}$$

We write $F^{(k)}_{[i]}(x)$, $H_{[i]}(x)$ for the marginal cdf's of $F^{(k)}(\underset{\sim}{x})$ and $H(\underset{\sim}{x})$ corresponding to the i-th coordinate and $F^{(k)}_{[i,j]}(x,y)$, $H_{[i,j]}(x,y)$ for the bivariate marginal cdf's of the same corresponding to the i-th and j-th coordinates.

Let the empirical cdf's based on the k-th sample be $\hat{F}^{(k)}(\underset{\sim}{x})$. Then the same based on the combined sample is

$$\hat{H}(\underset{\sim}{x}) = \lambda_0\hat{F}^{(0)}(\underset{\sim}{x}) + \lambda_1\hat{F}^{(1)}(\underset{\sim}{x}) + \lambda_2\hat{F}^{(2)}(\underset{\sim}{x}) \tag{2.6}$$

The corresponding marginal cdf's are as before denoted by $\hat{F}^{(k)}_{[i]}(x)$, $\hat{H}_{[i]}(x)$ etc. Writing

$$\begin{aligned} U(x) &= 0, \quad \text{if} \quad x < 0, \\ &= 1, \quad \text{if} \quad x \geq 0, \end{aligned} \tag{2.7}$$

we have, clearly,

$$\begin{aligned} \hat{F}^{(k)}_{[i]}(x) &= \frac{1}{n_k} \sum_{\alpha=1}^{n_k} U(x-X^{(k)}_{i\alpha}) \\ \hat{H}_{[i]}(x) &= \frac{1}{N} \sum_{k=0}^{2} \sum_{\alpha=1}^{n_k} U(x-X^{(k)}_{i\alpha}) \end{aligned} \tag{2.8}$$

Following Hájek [1] we suppose that there are p 'score functions' $\varphi_i(u)$, defined over $0<u<1$, $i=1,\ldots,p$, and that, for each i, $\varphi_i(u)$ determines the i-th row of the score matrix (1.4) by either of the following two relations:

$$a_{Ni}(\alpha) = \varphi_i\left(\frac{\alpha}{N+1}\right), \qquad \alpha=1,2,\ldots,N, \tag{2.9}$$

$$a_{Ni}(\alpha) = E\varphi_i(U_N^{(\alpha)}), \qquad \alpha=1,2,\ldots,N, \tag{2.10}$$

($U_N^{(1)} < U_N^{(2)} < \ldots < U_N^{(N)}$ are the order statistics of a sample of size N from the uniform distribution over (0,1)).

As in [1] we assume that the following condition is satisfied by the φ_i's.

ASSUMPTION II For each $i=1,2,\ldots,p$,

$$\varphi_i(u) = \varphi_{i1}(u) - \varphi_{i2}(u) \tag{2.11}$$

where $\varphi_{i1}(u)$, $\varphi_{i2}(u)$ are both nondecreasing, square integrable and absolutely continuous <u>inside</u> (0,1).

This implies that φ_i is square integrable over (0,1) and also that for any $0<a<b<1$,

$$\varphi_i(b) - \varphi_i(a) = \int_a^b \varphi_i'(u)du \tag{2.12}$$

where the derivative $\varphi_i'(u)$ exists almost everywhere in (0,1).

Let us write

$$\mu_i^{(k)} = \int_{-\infty}^{\infty} \varphi_i(H_{[i]}(x))dF_{[i]}^{(k)}(x) \tag{2.13}$$

$$\sigma_{ij}^{(k)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi_i(H_{[i]}(x))\cdot\varphi_j(H_{[i]}(y))dF_{[i,j]}^{(k)}(x,y)-\mu_i^{(k)}\cdot\mu_j^{(k)} \tag{2.14}$$

$i,j=1,2,\ldots,p;\ k=0,1,2.$

These depend on N through $H(\underset{\sim}{x})$. However, (2.5) together with the square integrability of the φ_i's, implies that $\mu_i^{(k)}$, $\sigma_{ij}^{(k)}$ are all uniformly bounded. Writing

$$\underset{\sim}{\mu}^{(k)} = (\mu_1^{(k)},\ldots,\mu_p^{(k)})', \quad \underset{\sim}{\Sigma}^{(k)} = (\sigma_{ij}^{(k)})_{i,j=1,2,\ldots,p}, \quad k=0,1,2 \tag{2.15}$$

from (1.2) we get

$$\begin{aligned} \underset{\sim}{\mu}^{(0)} &= \theta\underset{\sim}{\mu}^{(1)} + (1-\theta)\underset{\sim}{\mu}^{(2)} \\ \underset{\sim}{\Sigma}^{(0)} &= \theta\, \underset{\sim}{\Sigma}^{(1)} + (1-\theta)\underset{\sim}{\Sigma}^{(2)} + \theta(1-\theta)[\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)}][\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)}]' \end{aligned} \tag{2.16}$$

Finally, we assume that the following conditions hold as $N\to\infty$.

ASSUMPTION III

$$\begin{aligned} &\text{(a)}\quad \text{for at least one } i, \quad \liminf_{N\to\infty} (\mu_i^{(2)}-\mu_i^{(1)}) > 0, \\ &\qquad\qquad\qquad\qquad \text{or} \quad \limsup_{N\to\infty} (\mu_i^{(2)}-\mu_i^{(1)}) < 0. \end{aligned} \tag{2.17}$$

$$\text{(b)}\quad \min_{k=1,2} \{\liminf_{N\to\infty} m(\underset{\sim}{\Sigma}^{(k)})\} > 0. \tag{2.18}$$

The implication of III(a) is that, for at least one value of i, the mean value of $\varphi_i(H_{[i]}(X_i^{(1)}))$ is larger than that of $\varphi_i(H_{[i]}(X_i^{(2)}))$ (or vice versa) for all but a finite number of values of N. Since $F^{(1)}$ and $F^{(2)}$ are supposed to be distinct, for proper choices of the score functions φ_i, it would be possible to get this difference reflected in the mean value of $\varphi_i(H_{[i]}(X_i))$ for at least one i. Thus III(a) is not unduly restrictive. Assumption III(b) is a sort of 'nonsingularity assumption', which implies that for none of the two populations $F^{(1)}$, $F^{(2)}$, any of the p variables becomes ever 'useless' in the sense that its value is predictable from those of the other variables.

With the above notations and assumptions, we now state and prove some results to be used in the subsequent sections.

THEOREM 2.1 <u>Under Assumptions I and II</u>

$$\bar{a}_i^{(k)} - \mu_i^{(k)} \xrightarrow{P} 0, \qquad i=1,2,\ldots,p;\ k=0,1,2. \tag{2.19}$$

Proof. To prove (2.19), it will be sufficient to show that

$$E(\bar{a}_i^{(k)}-\mu_i^{(k)})^2 = \mathrm{Var}(\bar{a}_i^{(k)}) + \{E\bar{a}_i^{(k)}-\mu_i^{(k)}\}^2 \to 0. \tag{2.20}$$

From (2.11), writing $a_{Ni}(\alpha|\varphi_{iq})$, $a_{i\alpha}^{(k)}(\varphi_{iq})$, $\bar{a}_i^{(k)}(\varphi_{iq})$, to denote quantities obtained from the function $\varphi_{iq}(u)$ exactly as $a_{Ni}(\alpha)$, $a_{i\alpha}^{(k)}$, $\bar{a}_i^{(k)}$ are obtained from $\varphi_i(u)$, q=1,2, and letting c stand for a generic constant, we have

$$\begin{aligned}\mathrm{Var}(\bar{a}_i^{(k)}) &= \mathrm{Var}(\bar{a}_i^{(k)}(\varphi_{i1}) - \bar{a}_i^{(k)}(\varphi_{i2}))\\ &\le 2\ \mathrm{Var}(\bar{a}_i^{(k)}(\varphi_{i1})) + \mathrm{Var}(\bar{a}_i^{(k)}(\varphi_{i2}))\}\\ &\le c \cdot \frac{1}{N^2}\ \{\textstyle\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\varphi_{i1}) - \bar{a}_{Ni}(\varphi_{i1})]^2\\ &\qquad + \textstyle\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\varphi_{i2}) - \bar{a}_{Ni}(\varphi_{i2})]^2\}\end{aligned} \tag{2.21}$$

where $\bar{a}_{Ni}(\varphi_{iq}) = \frac{1}{N}\sum_{\alpha=1}^{N} a_{Ni}(\alpha|\varphi_{iq})$, q=1,2. The last inequality follows directly from Hájek's Variance Inequality ([1] Theorem 3.1), as

$$\bar{a}_i^{(k)}(\varphi_{iq}) = \frac{1}{n_k}\sum_{\alpha=1}^{n_k} a_{i\alpha}^{(k)}(\varphi_{iq}) = \frac{1}{n_k}\sum_{\alpha=1}^{n_k} a_{Ni}(R_{i\alpha}^{(k)}|\varphi_{iq})$$

and $a_{Ni}(\alpha|\varphi_{iq})$, is nondecreasing in $\alpha=1,2,\ldots,N$. Now as φ_{iq} is square integrable and $a_{Ni}(\alpha|\varphi_{iq})$ is as in (2.9) or (2.10), we have as $N\to\infty$

$$\frac{1}{N}\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\varphi_{iq}) - \bar{a}_{Ni}(\varphi_{iq})]^2 \to \int_0^1 [\varphi_{iq}(u)-\bar{\varphi}_{iq}]^2du, \qquad q=1,2 \tag{2.22}$$

where $\bar{\varphi}_{iq} = \int_0^1 \varphi_{iq}(u)du$ (see [2] pp. 158, 164). From (2.21) and (2.22) we conclude that

$$\mathrm{Var}(\bar{a}_i^{(k)}) \to 0 \quad \text{as} \quad N\to\infty . \tag{2.23}$$

Next consider the second term in (2.20). Under our Assumption I, by lemma 5.1 of Hájek [1], given any $\varepsilon>0$ we can find a decomposition

$$\varphi_i(u) = \psi_i(u) + \psi_{i1}(u) - \psi_{i2}(u) \tag{2.24}$$

such that ψ_i is a polynomial, ψ_{i1} and ψ_{i2} are nondecreasing, and

$$\int_0^1 \psi_{i1}^2(u)du + \int_0^1 \psi_{i2}^2(u)du < \varepsilon. \tag{2.25}$$

Hence, using notations similar to above,

$$\bar{a}_i^{(k)} - \mu_i^{(k)} = [\bar{a}_i^{(k)}(\psi_i) - \int_{-\infty}^{\infty}\psi_i(H_{[i]}(x))dF_{[i]}^{(k)}(x)] + [\bar{a}_i^{(k)}(\psi_{i1}) - \int_{-\infty}^{\infty}\psi_{i1}(H_{[i]}(x))dF_{[i]}^{(k)}(x)] - [\bar{a}_i^{(k)}(\psi_{i2}) - \int_{-\infty}^{\infty}\psi_{i2}(H_{[i]}(x))dF_{[i]}^{(k)}(x)].$$

Hence

$$[E\bar{a}_i^{(k)} - \mu_i^{(k)}]^2 \leq c\{[E\bar{a}_i^{(k)}(\psi_i) - \int_{-\infty}^{\infty}\psi_{i1}(H_{[i]}(x))dF_{[i]}^{(k)}(x)]^2 + [E\bar{a}_i^{(k)}(\psi_{i1})]^2 + [E\bar{a}_i^{(k)}(\psi_{i2})]^2 + [\int_{-\infty}^{\infty}\psi_{i1}(H_{[i]}(x))dF_{[i]}^{(k)}(x)]^2 + [\int_{-\infty}^{\infty}\psi_{i2}(H_{[i]}(x))dF_{[i]}^{(k)}(x)]^2\}. \tag{2.26}$$

Now ψ_i being a polynomial has bounded second derivative in (0,1), and hence, by a result of Hájek (see (4.27) in [1]), the first term on the right of (2.26) has limit 0 as $N\to\infty$. Further, by Schwartz inequality,

$$\sum_{q=1}^{2}\{[E\bar{a}_i^{(k)}(\psi_{iq})]^2 + [\int_{-\infty}^{\infty}\psi_{iq}(H_{[i]}(x))dF_{[i]}^{(k)}(x)]^2\}$$

$$\leq \sum_{q=1}^{2}\{\frac{1}{n_o}\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\psi_{iq})]^2 + \int_{-\infty}^{\infty}[\psi_{iq}(H_{[i]}(x))]^2 dF_{[i]}^{(k)}(x)\}$$

$$\leq c\sum_{q=1}^{2}\{\frac{1}{N}\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\psi_{iq})]^2 + \int_{-\infty}^{\infty}[\psi_{iq}(u)]^2 du\} \tag{2.27}$$

by (2.2) and (2.5). Now, just as in (2.22), we have

$$\frac{1}{N}\sum_{\alpha=1}^{N}[a_{Ni}(\alpha|\psi_{iq})]^2 \to \int_0^1 \psi_{iq}^2(u)\,du. \tag{2.28}$$

From (2.25), (2.27) and (2.28), it follows that the other terms on the right of (2.26) can be made less than an arbitrary quantity by choosing N large enough. Hence

$$[E\,\bar{a}_i^{(k)} - \mu_i^{(k)}]^2 \to 0 \quad \text{as} \quad N\to\infty. \tag{2.29}$$

(2.23) and (2.29) imply (2.20). Q.E.D.

We define

$$\hat{\sigma}_{ij}^{(k)} = \frac{1}{n_k}\sum_{\alpha=1}^{n_k} a_{i\alpha}^{(k)} a_{j\alpha}^{(k)} - \bar{a}_i^{(k)}\bar{a}_j^{(k)} \tag{2.30}$$

where $a_{i\alpha}^{(k)}$ is as in (1.5).

THEOREM 2.2 <u>Under assumptions I and II</u>

$$\hat{\sigma}_{ij}^{(k)} - \sigma_{ij}^{(k)} \overset{P}{\to} 0, \qquad i,j=1,2,\ldots,p;\ k=0,1,2 \tag{2.31}$$

<u>where</u> $\sigma_{ij}^{(k)}$ <u>is defined as in (2.14)</u>.

<u>Proof</u>. As $\mu_i^{(k)}$ $i=1,\ldots,p$ are uniformly bounded for all N, Theorem 2.1 together with Lemma 1.1 imply that

$$\bar{a}_i^{(k)}\bar{a}_j^{(k)} - \mu_i^{(k)}\cdot\mu_j^{(k)} \overset{P}{\to} 0.$$

Hence, to prove (2.31), it will be sufficient to show that

$$\frac{1}{n_k}\sum_{\alpha=1}^{n_k} a_{i\alpha}^{(k)} a_{j\alpha}^{(k)} - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi_i(H_{[i]}(x))\varphi_j(H_{[j]}(y))\,dF_{[i,j]}^{(k)}(x,y) \overset{P}{\to} 0. \tag{2.32}$$

The proof will be similar to that of Theorem 3.1 of Puri and Sen [3]. We only sketch the outline. First using the decomposition (2.24) we can show that the L.H.S. of (2.32) can be written as

$$\frac{1}{n_k}\sum_{\alpha=1}^{n_k} a_{i\alpha}^{(k)}(\psi_i)\cdot a_{j\alpha}^{(k)}(\psi_i) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i(H_{[i]}(x))\psi_j(H_{[j]}(y))dF_{[i,j]}^{(k)}(x,y)+R \tag{2.33}$$

where by application of Schwartz inequality and use of (2.2), (2.5) and (2.28), it can be shown that $|R| < c\cdot\varepsilon$ for sufficiently large N (c is a generic constant). Hence to prove (2.32), it will be sufficient to show that the difference between the first two terms in (2.33) converges in probability to zero. Since ψ_i is a polynomial, it is known that

$$E\ \psi_i(U_N^{(\alpha)}) - \psi_i(\tfrac{\alpha}{N+1}) = O(\tfrac{1}{N})$$

(see (2.9) and (2.10)). Hence it will be sufficient to take $a_{Ni}(\alpha|\psi_i) = \psi_i(\frac{\alpha}{N+1})$. Thus recalling (1.5) our problem reduces to showing that

$$\frac{1}{n_k}\sum_{\alpha=1}^{n_k} \psi_i(\frac{R_{i\alpha}^{(k)}}{N+1})\cdot\psi_j(\frac{R_{j\alpha}^{(k)}}{N+1}) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i(H_{[i]}(x))\psi_j(H_{[j]}(y))dF_{[i,j]}^{(k)}(x,y). \tag{2.34}$$

converges to zero. As ψ_i has bounded second derivative in (0,1), by Taylor expansion, we may write

$$\psi_i(\frac{R_{i\alpha}^{(k)}}{N+1}) = \psi_i(H_{[i]}(X_{i\alpha}^{(k)})) + (\frac{R_{i\alpha}^{(k)}}{N+1} - H_{[i]}(X_{i\alpha}^{(k)}))\psi_i'(H_{[i]}(X_{i\alpha}^{(k)}))$$
$$+ \frac{1}{2}(\frac{R_{i\alpha}^{(k)}}{N+1} - H_{[i]}(X_{i\alpha}^{(k)}))^2\cdot c, \tag{2.35}$$

$$i=1,2,\ldots,p.$$

Substituting (2.35) in the first term of (2.35) we get that (2.34) is equal to (say)

$$\frac{1}{n}\sum_{\alpha=1}^{n_k}\psi_i(H_{[i]}(X_{i\alpha}^{(k)}))\cdot\psi_j(H_{[j]}(X_{j\alpha}^{(k)}))-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_i(H_{[i]}(x))\psi_j(H_{[j]}(y))dF_{[i,j]}^{(k)}(x,y)+R^* \tag{2.36}$$

Using the notation (2.7),

$$R_{i\alpha}^{(k)} = \sum_{k'}\sum_{\alpha'} U(X_{i\alpha}^{(k)}-X_{i\alpha'}^{(k')}),$$

and hence, given $X_{i\alpha}^{(k)}$, can be considered as the sum of N Bernoulli variables. Using this fact and Schwartz inequality, it may be shown that in (2.36) R* is the sum of terms each of which converges in mean square, and hence, in probability to zero. As, in (2.36), the difference between the first two terms converges in probability to zero (by the Khinchin Law of Large Numbers), the required result follows. Q.E.D.

The following lemma will be required in proving the next theorem.

LEMMA 2.1 <u>Under assumptions I and II, as $N\to\infty$</u>,

$$\sqrt{N}\,E\{\bar{a}_i^{(0)} - \theta\,\bar{a}_i^{(1)}-(1-\theta)\bar{a}_i^{(2)}\} \to 0, \quad i=1,\dots,p \tag{2.37}$$

<u>Proof</u>. We recall that, if t denotes the unobservable proportion of 'observations from $F^{(1)}$' among $\underset{\sim}{X}_{\alpha}^{(0)}$, $\alpha=1,2,\dots,n_o$, then (1.7) holds. Hence (2.37) will follow if we can show

$$E\{\sqrt{N}(t-\theta)(\bar{a}_i^{(2)}-\bar{a}_i^{(1)})\} \to 0, \quad i=1,\dots,p,$$

or equivalently, as $E(t) = \theta$,

$$E\{\sqrt{N}(t-\theta)[(\bar{a}_i^{(2)}-\bar{a}_i^{(1)})-E(\bar{a}_i^{(2)}-\bar{a}_i^{(1)})]\} \to 0, \quad i=1,\dots,p \tag{2.38}$$

As t is a binomial proportion, by Schwartz inequality, the modulus of the left hand member of (2.38) is dominated by

$$\sqrt{\{\theta(1-\theta)\cdot \mathrm{Var}(\bar{a}_i^{(2)}-\bar{a}_i^{(1)})\}} \le \sqrt{\{2\theta(1-\theta)[\mathrm{Var}(\bar{a}_i^{(1)}) + \mathrm{Var}(\bar{a}_i^{(2)})]\}}$$

which tends to zero by (2.23). Q.E.D.

THEOREM 2.3 Under assumptions I, II, and III(b), as $N\to\infty$, $\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}$ is asymptotically distributed as $N(\underset{\sim}{o}, \frac{1}{n_o}\underset{\sim}{\Sigma}^{(0)} + \frac{\theta^2}{n_1}\underset{\sim}{\Sigma}^{(1)} + \frac{(1-\theta)^2}{n_2}\underset{\sim}{\Sigma}^{(2)})$ where $\underset{\sim}{\Sigma}^{(k)}$ is given by (2.14) and (2.15).

Proof. We have to show that for any nonnull p-vector $\underset{\sim}{\ell}=(\ell_1,\ldots,\ell_p)'$,

$$\mathcal{L}\ \frac{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)})}{\{\frac{1}{n_o}\cdot\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \frac{\theta^2}{n_1}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + \frac{(1-\theta)^2}{n_2}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell}\}^{\frac{1}{2}}} \to N(0,1) \tag{2.39}$$

Now, by assumption III(b), the denominator in

$$\frac{\sqrt{N}\ E\{\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(0)} - \theta\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(1)} - (1-\theta)\underset{\sim}{\ell}'\bar{\underset{\sim}{a}}^{(2)}\}}{\{\frac{N}{n_o}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \theta^2\cdot\frac{N}{n_1}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + (1-\theta)^2\cdot\frac{N}{n_2}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell}\}^{\frac{1}{2}}} \tag{2.40}$$

is bounded away from zero. So, by lemma 2.1, the expression (2.40) converges to zero as $N\to\infty$. Hence (2.39) will follow if we can show

$$\mathcal{L}\ \frac{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}) - E\{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}}{\{\frac{1}{n_o}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \frac{\theta^2}{n_1}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + \frac{(1-\theta)^2}{n_2}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell}\}^{\frac{1}{2}}} \to N(0,1) \tag{2.41}$$

(2.41) follows from the results of Hájek.[1] in the same way as Theorem 4.1 of [3]. Specifically, one has to show first that when φ_i, i=1,2,...,p, have bounded second derivatives in (0,1), the left hand member of (2.41), has asymptotically the same distribution as

$$\{\frac{1}{n_o}\sum_{\alpha=1}^{n_o}[\sum_{i=1}^{p}\ell_i\varphi_i(H_{[i]}(X_{i\alpha}^{(0)})] - \frac{\theta}{n_1}\sum_{\alpha=1}^{n_1}[\sum_{i=1}^{p}\ell_i\varphi_i(H_{[i]}(X_{i\alpha}^{(1)}))]$$

$$- \frac{(1-\theta)}{n_2}\sum_{\alpha=1}^{n_2}[\sum_{i=1}^{p}\ell_i\varphi_i(H_{[i]}(X_{i\alpha}^{(2)}))]\}$$

$$\div \{\frac{1}{n_o}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \frac{\theta^2}{n_1}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + \frac{(1-\theta)^2}{n_2}\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell}\}^{\frac{1}{2}} \tag{2.42}$$

This follows by an application of Hájek's [1] inequality (4.26) (the inequality actually holds under both (2.9) and (2.10)) and Assumption III(b). As, by Lindeberg-Feller Central Limit Theorem, (2.42) is asymptotically distributed as N(0,1) (see (2.14)), we get that when φ_i'', $i=1,\ldots,p$ exist and are bounded, (2.41) holds. Generally, under Assumptions I, II and III(b) we use the decomposition (2.24) as in [1], to show that we can approximate the left hand member of (2.41) arbitrarily closely by another expression of exactly the same form but based on polynomial scores-generating functions ψ_i. Hence the theorem follows. (Note that because of the particular nature of our problem, we do not require a 'centering assumption' as in [3]) Q.E.D.

3. FIXED-$\underset{\sim}{\ell}$ LINEAR RANK SCORE ESTIMATE

Given a nonnull vector $\underset{\sim}{\ell}$ (p×1) in (1.9) we have proposed

$$\tilde{\theta} = \tilde{\theta}(\underset{\sim}{\ell}) = \frac{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)})}{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)})} \tag{3.1}$$

as an estimate of θ, provided the denominator is non-zero and the ratio lies in [0,1]. Let $\underset{\sim}{\ell}$ be such that

$$\liminf_{N\to\infty} \underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)} - \underset{\sim}{\mu}^{(1)}) > 0 \tag{3.2}$$

Under Assumption III(a), it is always possible to choose $\underset{\sim}{\ell}$ so that (3.2) is realised, (For instance, we make take $\ell_i>0$, if lim inf $(\mu_i^{(2)}-\mu_i^{(1)}) > 0$, $\ell_i<0$, if lim sup $(\mu_i^{(2)}-\mu_i^{(1)}) < 0$, and $\ell_i=0$, otherwise.). By Theorem 2.1, and (2.16) we have

$$\begin{aligned} &\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}) - \underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)} - \underset{\sim}{\mu}^{(1)}) \overset{P}{\to} 0 \\ &\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)}) - \theta\underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)} - \underset{\sim}{\mu}^{(1)}) \overset{P}{\to} 0 \end{aligned} \tag{3.3}$$

(3.2) and (3.3) imply that, with probability approaching 1 as $N\to\infty$, (3.1) gives a well-defined estimate of θ. In fact, as by our assumptions, the elements of $\underset{\sim}{\mu}^{(k)}$ remain bounded, by Lemma 1.1, we immediately get $\tilde{\theta}-\theta \overset{P}{\to} 0$, so that $\tilde{\theta}$ is a consistent estimate of θ. Further, by the same argument

$$\{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)})\}^{-1} - \{\underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)} - \underset{\sim}{\mu}^{(1)})\}^{-1} \overset{P}{\to} 0 \tag{3.4}$$

and by Theorem 2.3, $\sqrt{N}\,\underset{\sim}{\ell}'\{\bar{\underset{\sim}{a}}^{(0)} - \theta\bar{\underset{\sim}{a}}^{(1)} - (1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}$ is asymptotically normal with mean 0 and variance

$$\frac{N}{n_o}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \theta^2 \cdot \frac{N}{n_1}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + (1-\theta)^2 \cdot \frac{N}{n_2}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell} \tag{3.5}$$

As, by our assumptions, (3.5) is bounded, $\sqrt{N}\,\underset{\sim}{\ell}'\{\bar{\underset{\sim}{a}}^{(0)}- \bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}$ is bounded in probability. Hence, by well-known results,

$$\sqrt{N}(\tilde{\theta}-\theta) = \sqrt{N}\,\underset{\sim}{\ell}'\{\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}\{\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})\}^{-1}$$

has asymptotically the same distribution as

$$\sqrt{N}\,\underset{\sim}{\ell}'\{\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}\{\underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)})\}^{-1}$$

So, $\sqrt{N}\,(\tilde{\theta}-\theta)$ is asymptotically normal with mean 0 and variance

$$\{\frac{N}{n_o}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell} + \theta^2 \cdot \frac{N}{n_1}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell} + (1-\theta)^2 \cdot \frac{N}{n_2}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}^{(2)}\underset{\sim}{\ell}\}\{\underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)})\}^{-2} \tag{3.6}$$

By the second relation of (2.16), (3.6) can be written as

$$\theta(1-\theta) \cdot \frac{N}{n_o} + \{\underset{\sim}{\ell}'(\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)})\}^{-2}\underset{\sim}{\ell}'\{\theta(\frac{N}{n_o} + \theta\frac{N}{n_1})\underset{\sim}{\Sigma}^{(1)} + (1-\theta)(\frac{N}{n_o} + (1-\theta)\frac{N}{n_2})\underset{\sim}{\Sigma}^{(2)}\}\underset{\sim}{\ell} \tag{3.7}$$

It is to be noted that the first term in (3.7) represents the asymptotic variance of $\sqrt{N}(t-\theta)$.

So far we have considered $\underset{\sim}{\ell}$ as a fixed non-null vector. Before concluding this section, we note that, if in (3.1) we replace $\underset{\sim}{\ell}$ by $\underset{\sim}{\ell}_N$, where $\underset{\sim}{\ell}_N$ is a sequence of non-null vectors, then provided (i) the elements of $\underset{\sim}{\ell}_N$ are uniformly bounded and (ii) $\underset{\sim}{\ell}_N$ satisfies (3.2), the above conclusions will hold true for $\tilde{\theta}(\underset{\sim}{\ell}_N)$ as well. This is because, (3.3) and (3.4) obviously apply to $\underset{\sim}{\ell}_N$, and asymptotic normality of $\sqrt{N}\,\underset{\sim}{\ell}_N'\{\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)}\}$ with mean 0 and variance

$$\frac{N}{n_o}\,\underset{\sim}{\ell}_N'\underset{\sim}{\Sigma}^{(0)}\underset{\sim}{\ell}_N + \theta^2\frac{N}{n_1}\,\underset{\sim}{\ell}_N'\underset{\sim}{\Sigma}^{(1)}\underset{\sim}{\ell}_N + (1-\theta)^2 \cdot \frac{N}{n_2}\,\underset{\sim}{\ell}'\underset{\sim}{\Sigma}\quad\underset{\sim}{\ell}$$

follows by lemma 1.2, since by our assumptions the latent roots of $\sum_{k=0}^{2} \frac{N}{n_k}\,\underset{\sim}{\Sigma}^{(k)}$ are bounded away from both zero and ∞.

4. OPTIMISED LINEAR RANK SCORE ESTIMATE

In the previous section we have seen that the second term in the expression (3.7) for the asymptotic variance of $\sqrt{N}\,(\tilde{\theta}(\underset{\sim}{\ell})-\theta)$ depends on $\underset{\sim}{\ell}$. In this section we investigate what will be its minimum value and whether it is possible, by any means, to attain this minimum. The problem is involved because the matrix in the second term of (3.7) involves the unknown θ but is still tractable provided we are prepared to solve higher degree polynomial equations. We first state the following well-known algebraic lemma.

LEMMA 4.1 For a positive definite matrix $G(p\times p)$ and non-null vector $\underset{\sim}{\delta}(p\times 1)$,

$$\min_{\underset{\sim}{\ell}\neq \underset{\sim}{o}} \frac{\underset{\sim}{\ell}'\underset{\sim}{G}\underset{\sim}{\ell}}{(\underset{\sim}{\ell}'\underset{\sim}{\delta})^2} = \{\underset{\sim}{\delta}'\underset{\sim}{G}^{-1}\underset{\sim}{\delta}\}^{-1}$$

and the minimum is attained for $G\underset{\sim}{\ell} = g\cdot\underset{\sim}{\delta}$ where g is an arbitrary non-zero number.

By our Assumptions III(a) and (b), for sufficiently large N, $\underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)}=\underset{\sim}{\delta}$ (say) $\neq \underset{\sim}{o}$, and

$$\theta(\frac{N}{n_o} + \theta\frac{N}{n_1})\underset{\sim}{\Sigma}^{(1)} + (1-\theta)(\frac{N}{n_o} + (1-\theta)\frac{N}{n_2})\underset{\sim}{\Sigma}^{(2)} = \underset{\sim}{G}(\theta) \text{ (say)} \tag{4.1}$$

is positive definite. Hence, by the above lemma, the minimum value of the asymptotic variance (3.7) of $\sqrt{N}(\tilde{\theta}(\underset{\sim}{\ell})-\theta)$, over all possible choices of $\underset{\sim}{\ell}$, is

$$\theta(1-\theta) \cdot \frac{N}{n_o} + [\underset{\sim}{\delta}'\underset{\sim}{G}^{-1}(\theta)\underset{\sim}{\delta}]^{-1} \tag{4.2}$$

If θ were known, to attain this minimum in $\tilde{\theta}(\underset{\sim}{\ell})$ we could take $\underset{\sim}{\ell}$ so that

$$\underset{\sim}{G}(\theta)\underset{\sim}{\ell} = \underset{\sim}{\delta} \tag{4.3}$$

Of course, the solution of (4.3), which we denote by $\underset{\sim}{\ell}_N$ would depend on N. But as noted at the end of section 3, the results of that section would still remain true ($\underset{\sim}{G}(\theta)$ and $\underset{\sim}{\delta}$ being bounded, so are the elements of $\underset{\sim}{\ell}_N$).

These considerations suggest that to estimate θ in an optimal manner we may proceed as follows.

Let $\hat{\underset{\sim}{\Sigma}}^{(k)} = (\hat{\sigma}_{ij}^{(k)})$ where $\hat{\sigma}_{ij}^{(k)}$ is given by (2.30). For any s ($0\leq s\leq 1$), we set up

$$\hat{\underset{\sim}{G}}(s) = s(\frac{N}{n_o} + s\frac{N}{n_1})\hat{\underset{\sim}{\Sigma}}^{(1)} + (1-s)(\frac{N}{n_o} + (1-s)\frac{N}{n_2})\hat{\underset{\sim}{\Sigma}}^{(2)}. \tag{4.4}$$

Consider the set of (p+1) simultaneous equations

$$\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)}) - s\cdot\underset{\sim}{\ell}'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}) = 0 \tag{4.5}$$

$$\hat{G}(s)\underset{\sim}{\ell} = \bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)} \tag{4.6}$$

in the (p+1) unknowns s and $(\ell_1,\ldots,\ell_p)=\underset{\sim}{\ell}'$. By Theorem 2.1 we have

$$\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)} - \underset{\sim}{\delta} \overset{P}{\to} \underset{\sim}{o} \tag{4.7}$$

Also by Assumption I, Theorem 2.2, and expression (4.4)

$$\hat{\underset{\sim}{G}}(s) - \underset{\sim}{G}(s) \overset{P}{\to} \underset{\sim}{O} \text{ uniformly in } 0\leq s\leq 1. \tag{4.8}$$

Hence, the equations (4.5) and (4.6) are suggested naturally from (3.1) and (4.3). (We write s instead of θ, to avoid confusion with the true value). If $s=\hat{\theta}_N$, $\underset{\sim}{\ell}=\hat{\underset{\sim}{\ell}}_N$ is a real solution of (4.5) and (4.6), such that $0\leq\hat{\theta}_N\leq 1$, we propose $\hat{\theta}_N$ as an estimate of θ. The following two theorems describe its properties.

THEOREM 4.1 _With probability approaching 1 as $N\to\infty$, the equations (4.5)-(4.6) possess a real solution. If $\hat{\theta}_N$, $\hat{\underset{\sim}{\ell}}_N$ is any such solution, and $\underset{\sim}{\ell}_N$ is the unique solution of (4.3), then, as $N\to\infty$_

$$\hat{\theta}_N \overset{P}{\to} \theta, \qquad \hat{\underset{\sim}{\ell}}_N - \underset{\sim}{\ell}_N \overset{P}{\to} \underset{\sim}{o} \tag{4.9}$$

Proof. Because of our assumptions I, II, and III(b) we can find numbers m_o and M_o such that

$$0 < m_o \leq \min_{0\leq s\leq 1} m[\underset{\sim}{G}(s)] \leq \max_{0\leq s\leq 1} M[\underset{\sim}{G}(s)] \leq M_o < \infty \tag{4.10}$$

for all sufficiently large values of N. Hence, by (4.8), we can find numbers m_o' and M_o' such that, as $N\to\infty$.

$$\text{Prob}\{0 < m_o' \leq \min_{0\leq s\leq 1} m[\hat{\underset{\sim}{G}}(s)] \leq \max_{0\leq s\leq 1} M[\hat{\underset{\sim}{G}}(s)] \leq M_o' < \infty\}\to 1. \tag{4.11}$$

If $\hat{G}(s)$ is positive definite, from (4.6),

$$\underset{\sim}{\ell} = [\hat{\underset{\sim}{G}}(s)]^{-1}(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}), \tag{4.12}$$

which, on substitution in (4.5), gives

$$(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})'[\hat{\underset{\sim}{G}}(s)]^{-1}(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(0)})-s(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})'[\hat{G}(s)]^{-1}(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})=0 \tag{4.13}$$

Since each cofactor of $\hat{G}(s)$ is a (2p-2)-th degree polynomial in s, (4.13) represents a (2p-1)-th degree polynomial equation in s. Clearly, if with probability approaching 1, (4.13) has a real solution for s, (4.5)-(4.6) will also have a real solution for s, $\underset{\sim}{\ell}$.

Now, by Theorem 2.1 and the first relation in (2.16)

$$\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)} - \theta\underset{\sim}{\delta} \overset{P}{\rightarrow} \underset{\sim}{o} \tag{4.14}$$

By (4.7), (4.8), (4.14) and Lemma 1.1, in view of (4.10) and the uniform boundedness of $\underset{\sim}{\delta}$, the polynomial in (4.13) has a stochastically vanishing difference with

$$(\theta-s)\underset{\sim}{\delta}'[\underset{\sim}{G}(s)]^{-1}\underset{\sim}{\delta} \tag{4.15}$$

By Assumption III(a), $\underset{\sim}{\delta}$ is non-null for all sufficiently large N. Hence the values of (4.15), and therefore, with probability approaching 1, those of the left hand member of (4.13), at $s = \theta\pm\varepsilon$ are of opposite signs. Thus with probability approaching 1 as $N\to\infty$, (4.13) would have a root between $\theta\pm\varepsilon$, whatever ε. This proves the first part of the theorem.

Again, if $\hat{\theta}_N$, $\underset{\sim}{\ell}_N$ is any solution of (4.5)-(4.6), we can write

$$\hat{\ell}_N = [\hat{G}(\hat{\theta}_N)]^{-1}(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}), \tag{4.16}$$

$$\hat{\theta}_N = \frac{\hat{\ell}_N'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)})}{\hat{\ell}_N'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)})}, \tag{4.17}$$

provided $\hat{G}(\hat{\theta}_N)$ is positive definite and the denominator in (4.17) is positive. By (4.11), $\hat{G}(\hat{\theta}_N)$ is positive definite in probability. Also by (4.7) and (4.11), as by Assumption III(a), $\underset{\sim}{\delta}'\underset{\sim}{\delta}$ is bounded away from zero, we can find an $\varepsilon>0$ such that

$$\text{Prob}\{\hat{\underset{\sim}{\ell}}_N'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}) > \varepsilon\} \to 1. \tag{4.18}$$

Now, from (4.17),

$$\hat{\theta}_N - \theta = \frac{\hat{\ell}_N'\{(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)}) - \theta(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)})\}}{\hat{\ell}_N'(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)})}. \tag{4.19}$$

By (4.7) and (4.14), $\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(0)} - \theta(\bar{\underset{\sim}{a}}^{(2)} - \bar{\underset{\sim}{a}}^{(1)}) \overset{P}{\to} \underset{\sim}{0}$. By (4.7) and (4.11) the elements of $\hat{\ell}_N$ are bounded in probability. These facts together with (4.18), imply from (4.19), that $\hat{\theta}_N - \theta \overset{P}{\to} 0$. Hence we get that

$$\hat{G}(\hat{\theta}_N) - G(\theta) \overset{P}{\to} \underset{\sim}{0}. \tag{4.20}$$

(4.7), (4.16), and (4.20) imply $\hat{\ell}_N - \ell_N \overset{P}{\to} \underset{\sim}{0}$ by Lemma 1.1. Q.E.D.

<u>Note</u>: The above theorem shows that $\hat{\theta}_N$ is a consistent estimate of θ. If $0<\theta<1$, then for large N, with high probability $\hat{\theta}_N$ lies in [0,1]. Trouble arises if $\theta=0$ or 1. If we conventionally take $\hat{\theta}_N$ to be 0(1) whenever s-solution of (4.5)-(4.6) is negative (exceeds 1), then, whatever θ, $\hat{\theta}_N$ always lies in [0,1] and is a consistent estimate of θ. The cases $\theta=0$ or 1 are of importance in the context of the problem of classification, where the 0-th sample may have come exclusively from $F^{(1)}$ or $F^{(2)}$. It is intended to discuss this problem in a subsequent communication.

THEOREM 4.2 <u>If $\hat{\theta}_N$, $\hat{\underset{\sim}{\ell}}_N$ is any real solution of (4.5)-(4.6), as $N\to\infty$,</u> $\sqrt{N}\,(\hat{\theta}_N-\theta)$ <u>is asymptotically normally distributed with mean 0 and variance</u>

$$\theta(1-\theta)\,\frac{N}{n_o} + [\underset{\sim}{\delta}'\underset{\sim}{G}^{-1}(\theta)\underset{\sim}{\delta}]^{-1} \tag{4.21}$$

<u>Proof</u>. As earlier, with probability approaching 1, we can represent $\hat{\theta}_N$ by (4.17), so that

$$\sqrt{N}(\hat{\theta}_N-\theta) = \frac{\sqrt{N}\,\underset{\sim}{\ell}_N'\{(1-\theta)\bar{\underset{\sim}{a}}^{(2)}+\theta\bar{\underset{\sim}{a}}^{(1)}-\bar{\underset{\sim}{a}}^{(0)}\}}{\hat{\underset{\sim}{\ell}}_N'(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})} + \frac{\sqrt{N}(\hat{\underset{\sim}{\ell}}_N-\underset{\sim}{\ell}_N)'\{(1-\theta)\bar{\underset{\sim}{a}}^{(2)}+\theta\bar{\underset{\sim}{a}}^{(1)}-\bar{\underset{\sim}{a}}^{(0)}\}}{\hat{\underset{\sim}{\ell}}_N'(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})}, \tag{4.22}$$

where $\underset{\sim}{\ell}_N = \underset{\sim}{G}(\theta)^{-1}\underset{\sim}{\delta}$ is the solution of (4.3).

Now by theorem (2.3), $\sqrt{N}\{(1-\theta)\bar{\underset{\sim}{a}}^{(2)} + \theta\bar{\underset{\sim}{a}}^{(1)} - \bar{\underset{\sim}{a}}^{(0)}\}$ is asymptotically normal with $\underset{\sim}{o}$ mean vector, and a dispersion matrix which by (2.16), and (4.1) can be written as

$$\theta(1-\theta)\cdot\frac{N}{n_o}\,\underset{\sim}{\delta}\underset{\sim}{\delta}' + \underset{\sim}{G}(\theta) \tag{4.23}$$

By our assumptions (4.23) is uniformly bounded. Hence the elements of $\sqrt{n}\,\{(1-\theta)\bar{\underset{\sim}{a}}^{(2)} + \theta\bar{\underset{\sim}{a}}^{(1)} - \bar{\underset{\sim}{a}}^{(0)}\}$ are stochastically bounded. Therefore, the second relation in (4.9), and (4.19) imply that the second term on the right of (4.22) converges in probability to zero.

Again, by our assumptions, the roots of (4.23) are bounded away from both 0 and ∞. Hence, as noted at the end of section 3, by Lemma 1.2, $\sqrt{N}\,\underset{\sim}{\ell}_N'\{(1-\theta)\bar{\underset{\sim}{a}}^{(2)} + \theta\bar{\underset{\sim}{a}}^{(1)} - \bar{\underset{\sim}{a}}^{(0)}\}$ is asymptotically normally distributed with mean 0 and variance

$$\theta(1-\theta)\,\frac{N}{n_o}\,[\underset{\sim}{\delta}'\underset{\sim}{G}^{-1}(\theta)\underset{\sim}{\delta}]^2 + \underset{\sim}{\delta}'\underset{\sim}{G}^{-1}(\theta)\underset{\sim}{\delta}\,.$$

Also, by (4.7) and (4.9),

$$\hat{\ell}_N'(\bar{a}^{(2)} - \bar{a}^{(1)}) - \delta' G^{-1}(\theta)\delta \xrightarrow{P} 0.$$

Combining these facts, we get that the first term on the right of (4.22) is asymptotically normal with mean 0 and variance given by (4.21). This completes the proof of the theorem. Q.E.D.

The following reduction shows the structure of the equations (4.5)-(4.6) more clearly. With probability approaching 1, in large samples, $\hat{\Sigma}^{(1)}$ and $\hat{\Sigma}^{(2)}$ are both positive definite, and therefore, it is possible to find a non-singular matrix C (p×p) such that

$$C'\Sigma^{(1)}C = I, \qquad C'\Sigma^{(2)}C = \text{Diag}(\gamma_1,\gamma_2,\ldots,\gamma_p), \tag{4.24}$$

where $\gamma_1,\gamma_2,\ldots,\gamma_p$ are the latent roots of $\{\hat{\Sigma}^{(1)}\}^{-1}\hat{\Sigma}^{(2)}$. Let

$$C'(\bar{a}^{(2)}-\bar{a}^{(0)}) = d=(d_1,\ldots,d_p)'$$

$$C'(\bar{a}^{(2)}-\bar{a}^{(1)}) = b=(b_1,\ldots,b_p)'.$$

Then, putting $\ell=Ch$, $h=(h_1,\ldots,h_p)'$, and premultiplying (4.6) by C', from (4.5)-(4.6) we derive the (p+1) equations in s and h:

$$h'd - s\cdot h'b = 0$$

$$C'\hat{G}(s)Ch = b$$

By (4.4) and (4.24), these can be explicitly written as

$$\sum_{i=1}^{p} h_i d_i - s\sum_{i=1}^{p} h_i b_i = 0 \tag{4.25}$$

$$\{s(\frac{N}{n_o} + s\,\frac{N}{n_1}) + (1-s)(\frac{N}{n_o} + (1-s)\,\frac{N}{n_2})\gamma_i\}h_i = b_i,\ i=1,\ldots,p \tag{4.26}$$

Here d_i, b_i, γ_i, i=1,...,p are all known. To solve these we may substitute for h_i in (4.25), using (4.26), and then solve the resulting polynomial equation in s. Alternatively, we may put arbitrary values of $h_1,\ldots,h_p$ in (4.25) to get a rough estimate of s and use (4.26) to solve for $h_1,\ldots,h_p$ and then proceed by iteration.

As for the sampling variance of the estimate $\hat{\theta}_N$, in large samples, we may use

$$\frac{\hat{\theta}_N(1-\hat{\theta}_N)}{n_o} + \frac{1}{N}\,[(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})'\;\hat{G}(\hat{\theta}_N)^{-1}(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})]^{-1} \tag{4.27}$$

where $\hat{G}(\hat{\theta}_N)$ is obtained by putting $\hat{\theta}_N$ for s in (4.4). As $\hat{\theta}_N$ is a consistent estimator of θ, (4.27) also is consistent in the sense that the ratio of (4.27) to the true asymptotic variance of $\hat{\theta}_N$ converges in probability to 1.

5. DISCUSSION

From Theorem 4.2, we get that the asymptotic variance of the optimised linear rank score estimate $\hat{\theta}_N$ is given by

$$\frac{1}{n_o}\,\theta(1-\theta) + \frac{1}{N}\,[\underset{\sim}{\delta}'\{G(\theta)\}^{-1}\underset{\sim}{\delta}]^{-1} \tag{5.1}$$

where $\underset{\sim}{\delta} = \underset{\sim}{\mu}^{(2)}-\underset{\sim}{\mu}^{(1)}$ and $G(\theta)$ is given by (4.1). Of this the first term is the variance of $t = r/n_o$, the standard estimate that we would use if r (the number of 'first-population observations' in 0-th sample) were observable. The second term represents the inflation in the variance due to the unobservability of r. Clearly, larger the value of $\underset{\sim}{\delta}'\{G(\theta)\}^{-1}\underset{\sim}{\delta}$, more accurate the estimate. We know that if $\underset{\sim}{G}(p\times p)$ is positive definite, then for the partitioning

$$\underset{\sim}{\delta} = \begin{pmatrix}\underset{\sim}{\delta}_1\\ \underset{\sim}{\delta}_2\end{pmatrix}, \qquad \underset{\sim}{G} = \begin{pmatrix}\underset{\sim}{G}_{11} & \underset{\sim}{G}_{12}\\ \underset{\sim}{G}_{21} & \underset{\sim}{G}_{22}\end{pmatrix},$$

where δ_1 is a q-vector and G_{11} is of order q×q, we have

$$\delta' G^{-1} \delta \geq \delta_1' G_{11}^{-1} \delta_1.$$

The equality holds if and only if $\delta_2 = G_{21}G_{11}^{-1}\delta_1$. Hence, so long as this special condition is not met and the matrix $G(\theta)$ remains positive definite, consideration of a larger number of variables would increase the accuracy of $\hat{\theta}_N$.

The choice of the score functions $\varphi_1(u),\ldots,\varphi_p(u)$ is, of course, of great importance. In choosing these, we should use any knowledge that we may possess about the way $F^{(1)}$ and $F^{(2)}$ differ from each other. Again, the general principle would be to choose the scores so that $\delta'\{G(\theta)\}^{-1}\delta$ is as large as possible. We propose to deal with these aspects of the problem in a later communication. Incidentally, we remark that the application of the procedures developed in this paper require, only knowledge of the observational ranks. However, when the observational values (1.1) are themselves available, we may use these in the same way as rank scores, to get an estimate of θ based on the sample means provided (i) all second order moments exist for $F^{(1)}$, $F^{(2)}$ and (ii) $F^{(1)}$, $F^{(2)}$ differ in location (i.e., in mean vector). Under appropriate conditions, the asymptotic variance of this estimate would have the same form as (5.1), with $\mu^{(k)}$, $\Sigma^{(k)}$ now standing for the mean vector and dispersion matrix of $F^{(k)}$, k=1,2. The procedures considered in this paper encompass a wider variety of problems, since the score functions are free to be chosen to take account of any kind of divergence between $F^{(1)}$ and $F^{(2)}$.

Another point of practical importance is the effect of the relative values of n_0, n_1, and n_2 on the accuracy of the estimate. While specific recommendations here must depend on the choice of score functions, some general observations can be made. Generally, it would be profitable to have a large n_0 not only because that reduces the first term in the asymptotic variance (5.1) but also because in (4.1) (at least for θ near around $\frac{1}{2}$) the dominant terms being made small, that reduces the second term in (5.1) as well. Similarly, it would be preferable to have n_1 large relative to n_2, or vice versa, according as θ is expected to be close to 1 or 0.

Finally, we remark that, although we have followed the 'best linear combination' approach to obtain the estimate of section 4, we could have obtained an estimate by an alternative approach. Thus starting from the vector $\bar{\underset{\sim}{a}}^{(0)} - \theta\, \bar{\underset{\sim}{a}}^{(1)} - (1-\theta)\, \bar{\underset{\sim}{a}}^{(2)}$, instead of equating a linear combination of its elements to zero, we could minimize a positive definite quadratic form in its elements to get the estimate. Minimization of

$$(\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)})'\underset{\sim}{A}(\bar{\underset{\sim}{a}}^{(0)}-\theta\bar{\underset{\sim}{a}}^{(1)}-(1-\theta)\bar{\underset{\sim}{a}}^{(2)})$$

(where $\underset{\sim}{A}$ is a positive definite matrix) with respect to θ, gives us an estimate in the form

$$(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})'\underset{\sim}{A}(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(0)})/(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)})'\underset{\sim}{A}(\bar{\underset{\sim}{a}}^{(2)}-\bar{\underset{\sim}{a}}^{(1)}).$$

Optimisation with respect to the choice of A then leads us practically to the same solution as before.

ACKNOWLEDGMENT

The author wishes to thank Professor P. K. Sen for helpful discussions.

REFERENCES

[1] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist. **39** 325-346.

[2] Hájek, J. and Šidák, Z. (1967). Theory of Rank Tests. Academic Press, New York and London, Academia Publishing House, Prague.

[3] Puri, M. L. and Sen, P. K. (1969). A class of rank order tests for a general linear hypothesis. Ann. Math. Statist. **40** 1325-1343.

[4] Ranga Rao, R. (1962). Relations between weak and uniform convergence of measures with applications. Ann. Math. Statist. **33** 659-680.