GENERALISATION OF MARKOFF'S THEOREM AND TESTS OF LINEAR HYPOTHESES

By C. RADHAKRISHNA RAO
Statistical Laboratory, Calcutta

INTRODUCTION

The general theory of linear estimation, without involving the assumpions of normality or independence of variates was first given by Markoff in his book *Calculus of Probability* published in Russian. Sheppard (1912 and 1914) working independently published some results which are approximately on the same lines as that of Markoff, but of a less general character. A distinct advance was made by Aitken (1935) who removed the unnecessary limitations in both Sheppard's and Markoff's results. Recently David and Neyman (1938) have published an article giving a slight extension of Markoff's theorem.

A significant step in generalising the theory of linear estimation is due to Raj Chandra Bose (1943), who, for the first time, introduced the concept of *nonestimable parametric functions*. He distinguishes two types of linear functions of stochastic variates ; the estimating functions and the error functions. A linear function $BY' = (b_1y_1 + b_2y_2 + \ldots + b_ny_n)$ of the stochastic variates is said to belong to error if $E(BY') \equiv 0$. The totality of the independent vectors such as B constitute a vector space which is called the error space. The vector space orthogonal to this errror space is called the estimation space, and *the best unbiassed estimate of any estimable parametric function* comes out as the scalar product of the vector $Y = (y_1, y_2, \ldots, y_n)$ of the stochastic variates and a vector C of the estimation space. The present paper is the result of ideas suggested by Bose's results (1943) and his post-graduate lectures in the Calcutta University in 1943-44 on the general theory of linear estimation and the fundamental structure of the analysis of variance.

The object of the present paper is firstly to take up the most general problem in the theory of linear estimation and get suitable generalisations of the previous results and secondly to derive tests of significance connected with linear hypotheses.

2. THE GENERAL PROBLEM OF LINEAR ESTIMATION.

Let $Y = (y_1\ y_2 \ldots y_n)$ be the vector of n stochastic variates $y_1, y_2, \ldots, y_n$ and $\Theta = (\theta_1\ \theta_2 \ldots \theta_n)$ be the vector of expectations of $y_1, y_2, \ldots, y_n$. It is given that $\Theta = TA'$ where $T = (\tau_1\ \tau_2 \ldots \tau_m)$ is a row matrix of m unknown parameters and A is a known matrix with n rows and m columns. The transpose of a matrix is denoted by a dash. *Here we do not assume any equality or inequality relations between n and m or make any restrictions on the rank of A.* The rank of A is, evidently, less than or equal to the smaller of m and n. Assuming that there are no functional relationships among the y's we can get a positive definite matrix Λ of rank n as $\Lambda = E(Y-\Theta)'(Y-\Theta)$ where E stands for the mathematical expectation in which case Λ is referred to as the dispersion matrix of the stochastic variates of Y. We assume that Λ is known apart from a constant multiplier and that its elements are finite.

The set of equations $E(Y) = TA'$ are known as observational equations and a linear function such as LT' where L is a given row matrix is called a parametric function.

The problem of linear estimation is to find a linear function BY′ of the stochastic variates such that

(a) $E(BY') = LT'$ independently of T (2.1)

and (b) $V(BY')$ is minimum (2.2)

where V stands for the variance. If there exists a vector B such that (a) is satisfied then LT′ is said to be *estimable* and BY′ is called an *unbiassed estimate* of LT′. If further, (b) is satisfied then BY′ is called the best *unbiassed estimate of LT′*.

3. THE GENERAL SOLUTION OF THE PROBLEM OF LINEAR ESTIMATION.

The definition of unbiassedness involved in (2.1) gives that if,

$$LT' \equiv E(BY') \equiv BE(Y') \equiv BAT' \tag{3.1}$$

then, L= BA. Also if B is such that L=BA then LT′=BAT′=E(BY′). This gives the result : *the necessary and sufficient condition that LT′ is estimable is that there exists a vector B such that L=BA.*

From the set of B's satisfying the condition L=BA we have to choose B such that V(BY′) is the least.

$$V(BY') = E[B(Y'-\theta')(Y-\theta)B'] = B\Lambda B' \quad .. \quad (3.2)$$

Introducing the vector $2C = 2(c_1 c_2 \ldots c_m)$ of Lagrangian multipliers we have to minimise $B\Lambda B' - 2C(A'B' - L')$ with respect to the elements of B and C. This leads to

$$B\Lambda = CA' \quad \text{and} \quad L = BA \tag{3.21}$$

Defining Λ^{-1} as the inverse of Λ we get

$$B\Lambda\Lambda^{-1} = CA'\Lambda^{-1} \quad \text{or} \quad B = CA'\Lambda^{-1} \quad .. \quad (3.22)$$

Using the relation L=BA we get, $L=BA=CA'\Lambda^{-1}A$ which is another form of the necessary and sufficient condition. *If C is found to satisfy $L=CA'\Lambda^{-1}A$ then $B=CA'\Lambda^{-1}$ and the best unbaissed estimate of LT′ is given by $BY'=CA'\Lambda^{-1}Y'$.*

We prove here some uniqueness results. If C_1 and C_2 are two vectors satisfying the relation $L=CA'\Lambda^{-1}A$, then $0=(C_1-C_2)A'\Lambda^{-1}A$. If B_1 and B_2 are the two vectors derived from (3.22) corresponding to C_1 and C_2 then $(B_1-B_2)=(C_1-C_2)A'\Lambda^{-1}$. Since

$$\begin{aligned} V(B_1-B_2\ Y') &= (C_1-C_2)A'\Lambda^{-1}\Lambda\Lambda^{-1}A(C_1-C_2)' \\ &= (C_1-C_2)A'\Lambda^{-1}A(C_1-C_2)' = 0 \ (C_1-C_2)' = 0 \end{aligned} \tag{3.3}$$

we get that $B_1 \equiv B_2$ for the variance of a linear function of stochastic variates cannot be zero without the coefficients being zero. This leads to the result *that the vector B or the best unbiassed estimate BY′ is unique for all C's satisfying the relation $L=CA'\Lambda^{-1}A$.*

If B is the vector derived in (3.22) and D is any other vector such that L=DA and is also a best unbiassed estimate of LT′ then

$$\begin{aligned} V(DY') = D\Lambda D' &= B\Lambda B' + (D-B)\Lambda B' + B\Lambda(D'-B') + (B-D)\Lambda(B'-D') \\ &= B\Lambda B' + (B-D)\Lambda(B'-D') \quad .. \quad (3.4) \end{aligned}$$

for $$B\Lambda(D'-B') = (D-B)\Lambda B' = (D-B)\Lambda\Lambda^{-1}AC' = (D-B)AC' = 0 \quad .. \quad (3.41)$$

$(B-D)\Lambda(B'-D')$ being positive definite cannot be negative. Hence the least variance is given when B−D=0, which leads to the result *that the calculations involved in (3.22) give a unique vector B such that V(BY′) is the least.*

4. EXTENSION OF MARKOFF'S PRINCIPLE

From Y, we construct the vector $Q=(Q_1\ Q_2 \ldots Q_m)$ composed of m linear functions of y's defined by $Q=Y\Lambda^{-1}A$. If LT' is estimable then we have $L=CA'\Lambda^{-1}A$ which gives that $LT'=CA'\Lambda^{-1}AT'=CH'$ where,

$$H=E(Q)=E(Y)\Lambda^{-1}A=TA'\Lambda^{-1}A \quad .. \quad (4.1)$$

Since the best unbiassed estimate is given by $BY'=CA'\Lambda^{-1}Y'=CQ'$ we can restate the result of section 2 as, *if there exists a vector C such that $LT'=CH'$, then LT' is estimable and the best unbiassed estimate of LT' is given by CQ' where H and Q are as defined in* (4.1) *and it is unique for all C's satisfying relation $LT'=CH'$.*

Let $\bar{T}=(t_1\ t_2 \ldots t_m)$ be a solution of the equations $Q=H$. Since $LT'\equiv CH'$ we get $L\bar{T}'=C\bar{H}'=CQ'$ where $\bar{H}$ denotes the matrix obtained from H by replacing T by $\bar{T}$. So we get the following rule of *estimating* any *estimable* parametric function. *We get a solution of the equations Q=H solved for τ's, and substitute the t's for τ's in the given parametric function. This gives the best unbiassed estimate of LT' and is unique for all solutions of τ's.* The proof of the latter part of the statement follows from the fact that whatever may be the solution, so long as $\bar{H}$ gives Q we get $L\bar{T}'=CQ'$ where $LT'\equiv CH'$.

We shall now study the effect of substituting a solution of $Q=H$ in a non-estimable parametric function NT'. If $\bar{T}_1$, and $\bar{T}_2$ are two solutions of $Q=H$, such that $\bar{T}_1\neq\bar{T}_2$, then $E(N\bar{T}'_1)\neq E(N\bar{T}'_2)$. For if $E(N\bar{T}'_1)=E(N\bar{T}'_2)$ then

$$E[N(\bar{T}'_1-\bar{T}'_2)]=0 \quad \text{or} \quad N(\bar{T}'_1-\bar{T}'_2)=0 \quad .. \quad (4.2)$$

for $(\bar{T}_1-\bar{T}_2)$ being a non-null solution of the homogeneous equations $H=0$, is independent of Q's and hence, of the stochastic variates. The result (4.2) shows that N belongs to the matrix of equations $H=0$ or $N=CA'\Lambda^{-1}A$ which means that NT' is estimable contrary to our supposition. Again, if $\bar{T}$ is a solution of $Q=H$, and if $N\bar{T}'$ is homogeneous in $y_1, y_2, \ldots, y_n$ then $E(N\bar{T}')=NT'$ when and only when NT' is estimable for otherwise it means that there exists a linear homogeneous function of Y's such that its expectation is NT'. Also the result of substitution of $\bar{T}$ in an estimable parametric function NT' leads to a homogeneous expression in y's. So we get *that the result of substitution of any solution* $\bar{T}$ of *Q=H in a parametric function LT' leads to the best unbiassed estimate of LT' if and only if*

$$\text{(i)}\quad L\bar{T}' \text{ is homogeneous in y's}$$
$$\text{(ii)}\quad E(L\bar{T}')=LT' \quad .. \quad (4.21)$$

A non-estimable parametric function will show forth *either in violation of (i) or (ii) or in giving two different expressions for two different solutions of Q=H.* We need not test for estimability before applying Markoff's principle of substitution. The above discussion shows that we can add a consistent, and a convenient and sometimes conventionally choosen set of equations in τ's not necessarily linear to $Q=H$ and get a solution for substitution.

The equations $Q=H$ are called normal equations and are readily obtained by equating the partial derivatives of

$$\sum_i \sum_j \lambda^{ij}(y_i-\theta_i)\ (y_j-\theta_j) \quad .. \quad (4.3)$$

where $\lambda^{ij}(i,j=1,2,\ldots,n)$ are elements of the matrix Λ^{-1}, with respect to $\tau_1, \tau_2, \ldots, \tau_m$, to zero. *The above equations Q=H are always solvable.* To prove this it is enough to show that if D is a vector such that $DH' \equiv 0$, then $DQ' \equiv 0$. $DH' \equiv 0$ implies that $DA'\Lambda^{-1}A = 0$ and

$$V(DQ') = V(DA'\Lambda^{-1}Y) = DA'\Lambda^{-1}\Lambda\Lambda^{-1}AD' = DA'\Lambda^{-1}AD' = 0D' = 0 \quad .. \quad (4.4)$$

which shows that $DQ' \equiv 0$ if, $DH' \equiv 0$.

We now prove some results which give the number of estimable and non estimable parametric functions.

(a) If $NQ' \equiv 0$, then the parametric function NT' is not estimable. For, if NT' is estimable then $N = CA'\Lambda^{-1}A$ and $NN' = NA'\Lambda^{-1}AC' = 0$ for $NQ' \equiv 0$ implies that $NH' \equiv 0$ and this $(NN' = 0)$ is impossible unless $N = 0$.

(b) There cannot be any non-vanishing linear function of Q's whose expectation is zero for it can be easily shown that the variance of such a linear function, if it exists, should be zero.

(c) The rank of $A'\Lambda^{-1}A$ is the same as the number of Q's which are linearly independent, and also of the number of linear parametric functions that are estimable. This follows from (a) and (b) given above. Since the rank of $A'\Lambda^{-1}A$ is the same as that of A it follows that the number of estimable parametric functions is equal to 's' the rank of A which is less than or equal to the smaller of m and n. The number of non-estimable parametric functions is therefore, $(m-s)$. If the rank of A is m, then all parametric functions are estimable.

5. INTRINSIC PROPERTIES OF NORMAL EQUATIONS.

The normal equations are $Q = TA'\Lambda^{-1}A$. The dispersion matrix of $Q_1, Q_2, \ldots, Q_m$ is

$$E[(Q'-H')(Q-H)] = E[A'\Lambda^{-1}(Y'-\Theta')(Y-\Theta)\Lambda^{-1}A]$$
$$= A'\Lambda^{-1}\Lambda\Lambda^{-1}A = A'\Lambda^{-1}A \quad .. \quad (5.1)$$

which shows that variance (Q_i) = coefficient of τ_i in the i-th normal equation and covariance $(Q_i\,Q_j)$ = coefficient of τ_j in the i-th normal equation.

If $L_k\,T'$ is estimated by $C_k\,Q', (k = 1, 2, \ldots, p)$ then

$$V(C_iQ') = C_iA'\Lambda^{-1}AC'_i = L_iC_i'$$
$$\text{Cov}\,[(C_iQ')\,(C_jQ')] = C_iA'\Lambda^{-1}AC'_j = L_iC'_j = L_jC'_i \quad .. \quad (5.2)$$

Let us eliminate τ_1 from the equations $Q = TA'\Lambda^{-1}A$ by the usual method of sweep-out. We divide both sides of the first equation by the coefficient of τ_1 and subtract it, after multiplying it by the coefficient of τ_1 in the i-th equation from the i-th equation. We now get $(m-1)$ equations which can be represented by $Q_1 = T_1B'$. It is easy to show that these equations also satisfy the properties of normal equations given in (5.1) and (5.2) above. The dispersion matrix of Q_1's is B and if $L_{i1}\,T'_1$ is estimated by $C_{k1}\,Q'_1$ then

$$V(C_{i1}\,Q'_1) = L_{i1}\,C'_{i1}$$
$$\text{Cov}\,[(C_{i1}\,Q'_1)\,(C_{j1}\,Q'_1)] = L_{i1}\,C'_{j1} = L_{j1}\,C'_{i1} \quad .. \quad (5.3)$$

These results hold good even if we eliminate more than one τ, the resulting equations satisfy these intrinsic properties of normal equations.

6. THE NATURE OF LINEAR HYPOTHESES.

The hypothesis involved in the theory of linear estimation is the assignment of the values of a single parametric function or a set of parametric functions. The necessary statistics and their distributions when the stochastic variates form a multivariate normal system are discussed below. The problem of distributions for other types of populations will be discussed in a future communication.

If LT′ is an estimable parametric function with a specified value ξ and its best unbiassed estimate is CQ′ then to test the hypothesis LT′$=\xi$ we construct the statistic

$$v=(CQ'-\xi)/\sqrt{CL'} \quad .. \quad (6.1)$$

which becomes a normal deviate with unit variance. We can extend this result to test the composite hypothesis $L_i\ T'=\xi_i (i=1, 2, .. k<s)$. Let $P=(P_1\ P_2 \ldots P_k)$ be the vector giving the estimates of $L_1T'-\xi_1$, $L_2T'-\xi_2$, ... $L_kT'-\xi_k$ with the dispersion matrix of P's as D. Following Fisher's (1939) technique we take GP′ a linear compound of P's, where $G=(g_1\ g_2 \ldots g_k)$ with its variance GDG′ and maximise the statistic

$$v^2=(GP')^2/GDG' \quad .. \quad (6.2)$$

If we denote by V the maximum value of v^2 we get V as the root of the determinental equation $|P'P-VD|=0$. This gives $V=\Sigma\Sigma d^{ij}\ P_i\ P_j$ where d^{ij} are the elements of matrix D^{-1}, the inverse of D. This may be called the *generalised variance statistic*. This statistic is invariant for any set of relations derived by linear combinations of the relations $L_i\ T'-\xi_i$. For in every case, we will be led to maximise the expression

$$V=(GP')^2/\bar{G}D\bar{G}' \quad .. \quad (6.21)$$

where the elements of $\bar{G}$ are linear combinations of the elements of G with which we started.

Under the normality assumption we get the distribution of P's as

$$\text{Const.}\quad e^{-\frac{1}{2}\psi} . \pi . dP_i \quad .. \quad (6.3)$$

where $$\psi=\Sigma\Sigma d^{ij}\ P_i\ P_j \quad .. \quad (6.31)$$

from which it immediately follows that the distribution of V is

$$\text{Const.}\quad [exp\ (V/2)]\ V^{(k-2)/2}\ dV \quad .. \quad (6.32)$$

which is the distribution of χ^2 with k degrees of freedom and it supplies the distribution of the V-statistic on the null hypothesis.

If the quantities $\phi_i=L_i\ T'-\xi_i\neq 0$ for at least a single i, then there is departure from the null-hypothesis. We define the quantity

$$\phi=\Sigma\Sigma d^{ij}\ \phi_i\ \phi_j \quad .. \quad (6.4)$$

which is also invariant for any linear combinations of ϕ's. If we transform the estimating functions P's to R's by linear combinations to make them independent and having unit variances the distribution (6.3) transforms to

$$\text{Const.}\quad e-\tfrac{1}{2}\psi\ dR_1\ dR_2 \ldots dR_k \quad . \quad (6.41)$$

where $$\psi=(R_1-\rho_1)^2+(R_2-\rho_2)^2+\ldots+(R_k-\rho_k)^2 \quad .. \quad (6.42)$$

and ρ's are linear combinations of ϕ's. V and ϕ transform to $R=\Sigma R^2$ and $\rho=\Sigma\rho_i^2$. By an argument similar to Bose (1935) we get the distribution of R or V as

$$\text{Const.}\quad [exp(V/2)]\ V^{(k-2)/2}\ I_{k-2}(\sqrt{V\phi})dV \quad .. \quad (6.43)$$

where I is the Bessel function of the second kind. This is the distribution of the generalised variance statistic V on the non null hypothesis and is dependent on k and ϕ only.

We consider here a class of linear hypotheses which are of special interest in tests of significance connected with biometrics and field experimentation. These are no doubt, special cases of the general composite hypothesis discussed above, but are amenable to a logical deduction of the necessary statistic. The hypothesis consists in specifying the equality of a number of parametric functions such as

$$L_1 T' = L_2 T' = \dots = L_k T' \quad \text{.. (6.5)}$$

If $P_1, P_2, \dots, P_k$ are the estimates of these linear functions then we take the linear compound

$$g_1 P_1 + g_2 P_2 + \dots + g_k P_k \quad \text{.. (6.51)}$$

such that $g_1 + g_2 + \dots + g_k = 0$ and maximise the expression

$$v^2 = (GP')^2/GDG' \quad \text{.. (6.52)}$$

This comes to testing for a suitable contrast of the parametric functions $L_i T'$. The maximisation leads to the determinental equation

$$|(P_i - \bar{P})(P_j - \bar{P}) - VD| = 0 \quad \text{.. (6.53)}$$

where

$$\bar{P} = \sum_i P_i (\sum_j d^{ij}) / \sum_i \sum_j d^{ij} \quad \text{.. (6.54)}$$

which gives

$$V = \sum\sum d^{ij}(P_i - \bar{P})(P_j - \bar{P}) \quad \text{.. (6.55)}$$

The distribution of this follows directly from previous considerations for it is designed to test a hypothesis concerning $(k-1)$ independent parametric functions. Hence V follows the distribution

$$\text{Const. } [exp(V/2)]\, V^{(k-3)/2}\, dV \quad \text{.. (6.56)}$$

on the null hypothesis and

$$\text{Const. } [exp(V/2)]\, V^{(k-3)/2}\, I_{(k-3)/2}(\sqrt{V}\phi)\, dV \quad \text{.. (6.57)}$$

on the non-null hypothesis, ϕ being defined as

$$\phi^2 = \sum\sum d^{ij}(\theta_i - \bar{\theta})(\theta_j - \bar{\theta}) \quad \text{.. (6.58)}$$

where θ_i stands for $L_i T'$ and $\bar{\theta}$ is the same functions of θ_i's as $\bar{P}$ is of P_i's.

In the case of the multinomial distribution with the proportions $\pi_1, \pi_2, \dots \pi_k$ in the k classes the linear observational equations corresponding to a sample of $n_1, n_2, \dots n_k$ $(\Sigma n_i = N)$ in the k classes are $E(n_i) = N\pi_i (i = 1, 2, \dots k)$. We want to test the hypothesis whether the sample is a reasonable one from the population with the assigned proportions $\pi_1, \pi_2, \dots \pi_k$. The observational equations $E(n_i) = N\pi_i (i = 1, 2, \dots k-1)$ which may be taken to be independent will themselves be estimating equations. Using the properties

$$V(n_i) = \pi_i(1 - \pi_i)N \quad \text{.. (6.6)}$$

$$\text{Cov}(n_i\, n_j) = -\pi_i \pi_j N \quad \text{.. (6.61)}$$

we try to maximum the statistic

$$v = \sum_{i=1}^{k-1} g_i(n_i - N\pi_i) / [\Sigma g_i^2 \pi_i (1 - \pi_i) N - \Sigma\Sigma g_i g_j \pi_i \pi_j N]^{\frac{1}{2}} \quad \text{.. (6.62)}$$

The maximum value comes out as

$$V = \sum_{i=1}^{k} (n_i - N\pi_i)^2 / N\pi_i \quad \text{.. (6.63)}$$

which is the same as the χ^2-statistic of Karl Pearson.

The discussion in the above paragraphs form the basis of the perimeter test derived by the author and discussed in another paper (1944). There are t, p-variate populations with means given by the matrix $((m_{ij}))$ $i=1, 2, \ldots, p$ referring to the variates of a population $j=1, 2, \ldots, t$ referring to the populations, the hypothesis to be tested is that

$$m_{i1}=m_{i2}\ldots=m_{it}, \ (i=1, 2, \ldots : p) \qquad \ldots (6.7)$$

There are $p(t-1)$ independent parametric functions to be tested. Making use of previous results we get the generalised variance statistic as

$$V=\sum_{j=1}^{t}\sum_{r=1}^{p}\sum_{s=1}^{p} {}_j\alpha^{rs}(\bar{x}_{rj}-\bar{x}_r)\ (\bar{x}_{sj}-\bar{x}_s)n_j \qquad \ldots (6.71)$$

where ${}_j\alpha^{rs}$ are the elements reciprocal to the variance and covariance matrix for the j-th p-variate population from which a sample of size n_j is drawn and $\bar{x}_{rj}$ is the sample mean for the r-th character in the j-th population and $\bar{x}_r$ is defined by

$$(-1)^r\,\bar{x}_r\begin{vmatrix} c_{11} & c_{12} & \ldots & c_{1p} \\ c_{21} & c_{22} & \ldots & c_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ c_{p1} & c_{p2} & \ldots & c_{pp} \end{vmatrix} = \begin{vmatrix} c_{11} & \ldots & c_{1r-1} & d_1 & c_{1\,r+1} & \ldots & c_{1p} \\ c_{21} & \ldots & c_{2r-1} & d_2 & c_{2\,r+1} & \ldots & c_{2p} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{p1} & \ldots & c_{pr-1} & d_p & c_{p\,r+1} & \ldots & c_{pp} \end{vmatrix} \qquad \ldots (6.72)$$

where $$d_s=\sum_{j=1}^{t}\sum_{i=1}^{p} {}_j\alpha^{is}\,\bar{x}_{ij}\,n_j \quad \text{and} \quad c_{lm}=\sum_{j=1}^{t} n_j\,{}_j\alpha^{lm} \qquad \ldots (6.73)$$

The distribution of this statistic on the null and non-null hypotheses are obtained directly from (6.32) and (6.43) by putting $k=p(t-1)$.

An important problem in the classification of three p variate populations π_1, π_2 and π_3 is to test whether π_1 is nearer to π_2 or π_3 which are known to be different. If m_{ij} is the mean of the j-th character in the i-th population we have to test the composite hypothesis

$$2m_{1j}=m_{2j}+m_{3j} \quad j=(1, 2, \ldots p) \qquad \ldots (6.8)$$

If $\bar{x}_{ij}$ is the sample mean for the j-th character of the i-th population and n_i is the sample size for the i-th population then the necessary statistic is

$$V=\Sigma\,\Sigma\alpha^{ij}(2\bar{x}_{1j}-\bar{x}_{2j}-\bar{x}_{3j})\ (2\bar{x}_{1i}-\bar{x}_{2i}-\bar{x}_{3i}) \qquad \ldots (6.81)$$

where α^{ij} are the elements of the matrix reciprocal to

$$\left(\left(4\frac{\alpha_{ij}^{(1)}}{n_1}+\frac{\alpha_{ij}^{(2)}}{n_2}+\frac{\alpha_{ij}^{(3)}}{n_3}\right)\right) \qquad \ldots (6.82)$$

where α^k_{ij} is the covariance of the i-th and the j-th characters in the k-th population. The distribution of V on the null and non-null hypotheses are obtained from (6.32) and (6.43) by putting $k=p$.

The necessary statistics when the variances and the covariances are not known are obtained by studentising the above statistics and their distributions in the studentised forms have been derived and are being used in the Statistical Laboratory at Calcutta for various problems. A detailed discussion will be attempted in another paper.

REFERENCES

1. Aitken, A. C. (1935) On the least square and linear combination of observations. *Proc. Roy. Soc Edin.* **55**, 42-48.

2. Bose, R.C. (1944) The fundamental theorem of linear estimation. *Proc. Ind. Sc. Congress.* 4(3).

3. David, F. N. and Neyman, J. (1939) Extension of Markoff's theorem on least squares. *Biom. Research Memoirs.* **2**, 103-116.

4. Fisher, R. A. (1939) The sampling distribution of some statistics obtained from nonlinear regression. *Annals of Eugenics.* **9**, 238-249.

5. Markoff, A. A. (1924) Calculus of Probability. Russian Edition.

6. Rao, C. R. (1944) The generalised variance of populations. *Proc. Ind. Sc. Congress.* 4 (3)

7. Sheppard, W. F. (1912) Reduction of errors by negligible differences. *Proc. Fifth Int. Congr. Math (Cambridge)* **2**, 348-84.

8. Sheppard, W. F. (1914 a) Fitting of polynomials by method of least squares. *Proc. Lond. Math. Soc.* **13** (3), 97-108.

9. Sheppard, W. F. (1914 b) Graduation by a reduction of mean square of error. *Jour. Inst. Actuaries.* **48**, 171-185, 390-412. and **49**, 148-157.

Paper received : 20 January, 1944