UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

Mathematical Sciences Directorate
Air Force Office of Scientific Research
Washington 25, D. C.

AFOSR Report No.

ANTE-DEPENDENCE ANALYSIS OF AN ORDERED SET OF VARIABLES

by

K. R. Gabriel
Hebrew University, Jerusalem.

July, 1959

Contract No. AF 49(638)-213

Every variable of an ordered set depends on a number of the preceding variables. Hypotheses on the minimal number of such variables are formulated, likelihood ratio tests obtained for the normal case and these are extended to a simple decision rule. The results can be generalized to several variables at each stage of the order.

Institute of Statistics
Mimeograph Series No. 234

ANTE-DEPENDENCE ANALYSIS OF AN ORDERED SET OF VARIABLES[1]

K. R. Gabriel

University of North Carolina
and
Hebrew University, Jerusalem

1. Introduction and Summary.

For a set of variables in a given order, s-th ante-dependence will be said to obtain if each one of the variables, given at least s immediate antecedent variables in the order, is independent of all further preceding variables. If the number of variables is p, ante-dependence is of some order between 0 and p-1. 0th ante-dependence and (p-1)st-ante-dependence are equivalent to complete independence and to completely arbitrary patterns of dependence, respectively, and are defined irrespective of the ordering of the variables. 1st to (p-2)nd ante-dependence are defined in terms of a specific order only.

If $X_1, X_2, \ldots, X_p$ are multivariate normal, s-th ante-dependence is equivalent to X_i - for any $i > s$ - given its regression on $X_{i-1}, X_{i-2}, \ldots X_{i-s}, \ldots X_{i-s-z}$ -for any $z = 0,1,2,\ldots i-s-1$, being uncorrelated with $X_{i-s-z-1}, X_{i-s-z-2}, \ldots X_2, X_1$. In other words, the partial correlation of X_i and $X_{i-s-z-1}$, given all the variables $X_{i-1}, X_{i-2}, \ldots X_{i-s-z}$, is zero for all $i > s$ and $z = 0,1, \ldots i-s-1$. The hypothesis that the covariance matrix is such that all the above partial correlations vanish will be denoted by H_s $(s=0,1,\ldots p-1)$ so

1. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 49(638)-213. Reproduction in whole or in part is permitted for any purpose of the United States Government.

that for the multivariate normal H_s denotes the hypothesis of s-th ante-dependence.

It is shown that for any set of ordered variables, normal or otherwise, H_s is equivalent to the regression and correlation of any X_i on all other variables being equal to that on X_{i-s}, X_{i-s+1}, ... X_{i-1}, X_{i+1}, ... X_{i+s} only. It is also equivalent to the $(p-s)(p-s-1)/2$ elements in the upper right (and also lower left) corner of the inverse covariance matrix being zero. Indeed any null hypothesis on a set of elements of the inverse covariance matrix may be formulated, and tested, as a hypothesis H_s if the variables can be so ordered as to put the zero elements in the upper right (lower left) corner of the inverse.

Maximum likelihood estimates are derived under H_s for the normal case. Likelihood ratio tests of any one H_s aganinst any other follow immediately and may be expressed in terms of the sample partial correlations. Exact distributions are not investigated but for large sample χ^2 approximations are available. Thus a sequence of tests of H_{p-2} under H_{p-1}, H_{p-3} under H_{p-2}, ... H_o under H_1 is obtained which in effect forms a breakdown of the large sample test of independence - H_o - under the general alternative - H_{p-1}.

The anology of the model with that of an antoregressive scheme in time series is obvious. The methods of analysis suggested here are in terms of many repeated observations on the series whereas the ante-regression analysis are generally put forward for the study of single time series.

The ante-dependence models can be generalized to the case of several variables at each stage of the ordering. This would be analogous to the study of multiple time-series.

It is interesting to note that s-ante-dependent sets of variables may be generated by s successive summations of independent variables. This may be relevant for some applications of such models.

Ante-dependence models might be applicable to observations ordered in time or otherwise. Observations on growth of organisms up to each of several ages might be analyzed in such a manner. Where growth is recorded on several dimensions, e.g., height and weight, the analysis might proceed in terms of the multidimensional generalization of the model. Other possible fields of application include batteries of psychological tests increasing in complexity, and data on the successive location of travelling objects. It is hoped to publish a study of some such applications at a later date.

2. A Sequence of Null Hypotheses on Partial Correlations.

Consider a vector variable $\underline{X} = (X_1, X_2, \ldots X_p)$ with $E(\underline{X}) = \underline{\mu}$ and $Var(\underline{X}) = \Sigma$ where Σ is positive definite. Denote the (ij)th elements of Σ and Σ^{-1} by σ_{ij} and σ^{ij}, respectively. Define Δ with diagonal elements $\sigma_{11}, \sigma_{22}, \ldots \sigma_{pp}$ and all other elements zero,

$$P = \Delta^{-1/2} \Sigma \Delta^{-1/2} \text{ with (ij)th element } \rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \cdot \sigma_{jj}}}, \text{ and}$$

P^{-1} with (ij)th element ρ^{ij}.

Next define $\rho_{ij.k,\ell,\ldots m} = \dfrac{\rho_{ij.\ell,\ldots m} - \rho_{ik.\ell,\ldots m}\,\rho_{jk.\ell,\ldots m}}{\sqrt{1-\rho^2_{ik.\ell,\ldots m}}\;\sqrt{1-\rho^2_{jk.\ell,\ldots m}}}$

and $$1 - \rho^2_{i(j_1,j_2,\ldots j_s)} = \prod_{u=1}^{p} \left\{ 1 - \rho^2_{ij_u.j_1,j_2,\ldots j_{u-1}} \right\} .$$

These are the usual definitions of the population values of the following parameters,

σ_{ij} covariance of X_i and X_j,

ρ_{ij} correlation of X_i and X_j,

$\rho_{ij.k,\ell,\ldots m}$ partial correlation of X_i and X_j given $X_k, X_\ell, \ldots X_m$,

$\rho_{i(j_1,j_2,\ldots j_s)}$ multiple correlation of X_i on $X_{j_1}, X_{j_2}, \ldots X_{j_s}$.

A sequence of subhypotheses $G_o, G_1, G_2, \ldots G_{p-1}$ is defined as follows:

$$G_s : \rho_{i,i+s+1.i+1,i+2,\ldots i+s} = 0 \quad \text{for all } i=1,2,\ldots p-s-1,$$

and a sequence of hypotheses

$$H_s = \bigcap_{u=s}^{p-1} G_u .$$

Thus considering H_s under H_{s+1} is equivalent to considering G_s under the assumption that $G_{s+1}, G_{s+2}, \ldots G_{p-1}$ all hold.

These hypotheses form a chain from completely arbitrary correlations - under H_{p-1} no restrictions at all are put on the correlations - to complete null correlation among all variables - under H_o (see Theorem 1). From any H_s to H_{s-1} there are p-s further restrictions on the correlations, amounting in effect to correlation of each X_i with one less of the preceding $X_{i-1}, X_{i-2}, \ldots$ or following $X_{i+1}, X_{i+2}, \ldots$. It will be shown in Theorem 1 that H_s is equivalent to Σ^{-1} having arbitrary elements in the principal diagonal and in the first s off diagonals, and zero elements everywhere else. (The s-th off diagonals of Σ^{-1} are defined as the elements σ^{ij} for which $i-j=s$ (upper) or $j-i=s$ (lower)). In other words, the sequence $H_o, H_1, \ldots H_{p-1}$ is one in which increasingly more off diagonals become arbitrary rather than zero. For H_o all off diagonals are zero; for H_1 the first off diagonal is arbitrary, all others zero; and so forth until for H_{p-1} all elements of Σ^{-1} are arbitrary.

It should be noted that the sub-hypotheses G_s and the hypotheses H_s are defined in terms of the given order of the variables $X_1, X_2, \ldots$ to X_p. For any other order of the variables the hypotheses would have a different meaning. In fact hypotheses about certain null partial correlations or null elements in the inverse covariance matrix might be formulated as hypotheses H_s by suitable permuation of the order of the variables. Only the two extreme hypotheses, H_o and H_{p-1}, are defined irrespective of the order.

Note: In the following, sets of variables of $\underline{X}$ may be defined in a manner as to include formally variables not in $\underline{X}$, e.g., the set $X_{i-s}, X_{i-s+1}, \ldots X_i, \ldots X_{i+s}$ would have X_{i-s} undefined if $i \leq s$. In such cases the sets will be understood to include only the variables which are defined in $\underline{X}$.

LEMMA 1: Under H_s $\rho_{ij.\phi} = 0$ for all i and j where ϕ denotes any set of s or more successive variables between X_i and X_j.

PROOF: Consider first ϕ to be i+1, i+2, ... i+s, where $i+s < j$. Now

$$\rho_{ij.\phi,i+s+1} = \frac{\rho_{ij.\phi.} - \rho_{i,i+s+1.\phi}\rho_{j,i+s+1.\phi}}{\sqrt{1-\rho^2_{i,i+s+1.\phi}}\sqrt{1-\rho^2_{j,i+s+1.\phi}}} \text{ (Cramer, 1946, -23.4.4)},$$

and since under H_s $\rho_{i,i+s+1.\phi} = 0$

$$\rho_{ij.\phi,i+s+1} = 0 \iff \rho_{ij.\phi} = 0 .$$

Similarly,

$$\rho_{ij.\phi,i+s+1,i+s+2} = 0 \iff \rho_{ij.\phi,i+s+1} = 0,$$

so that by successive applications of this chain of reasoning it is found that under H_s

$$\rho_{ij.\phi,i+s+1,i+s+2,\ldots,j-1} = 0 \iff \rho_{ij.\phi} = 0.$$

Now the LHS equality holds under $H_{s+(j-i-s-1)}$ which is implied by H_s if $j-i-1 \geq s$, i.e., if s intermediate variables between X_i and X_j exist. Therefore under H_s the RHS equality also holds.

The same argument is readily extended to cases where ϕ does not include X_{i+1}.

LEMMA 2: Under H_s $\rho_{ij.\psi} = 0$ for all i and j, where ψ is any set including at least s successive variables intermediate to X_i and X_j.

PROOF: Write any s successive variables in ψ which are intermediate to X_i and X_j as ϕ and the rest of ψ as $X_{k_1}, X_{k_2}, \ldots X_{k_u}$. Now

$$\rho_{ij.\psi} = \rho_{ij.\phi,k_1,k_2,\ldots k_u} =$$

$$= \frac{\rho_{ij.\phi,k_1,\ldots k_{u-1}} - \rho_{ik_u.\phi,k_1,\ldots k_{u-1}}\,\rho_{jk_u.\phi,k_1,\ldots k_{u-1}}}{\sqrt{1-\rho^2_{ik_u.\phi,k_1,\ldots k_{u-1}}}\,\sqrt{1-\rho^2_{jk_u.\phi,k_1,\ldots k_{u-1}}}}$$

so that if the Lemma holds for all ψ with s+u-1 terms or less $\rho_{ij.\phi,k_1,\ldots k_{u-1}} = 0$ and either $\rho_{ik_u.\phi,k_1,\ldots k_{u-1}} = 0$ or $\rho_{jk_u.\phi,k_1,\ldots k_{u-1}} = 0$ whence the LHS will also be zero and the Lemma hold for all ψ with s+u terms. But by Lemma 1 it holds for u=0, and therefore it must hold for all u=0,1,2,... .

LEMMA 3: For any non-singular pxp matrix Σ the following two statements are equivalent for any i:

(a) $\sigma^{ij}=0$ for all j such that $|i-j| > s$;

(b) the i-th column of the inverse of the principal minor of the i-s,i-s+1,...,i,...,i+s rows and columns of Σ is made up of the corresponding elements of the inverse of Σ.

PROOF: By definition, for all j=1,2,...p,

$$\sum_{k=1}^{p} \sigma_{jk}\sigma^{ki} = \delta_{ij} \quad \text{(Kroeneker's delta).}$$

It can easily be checked that then

$$\sum_{k=i-s}^{i+s} \sigma_{jk}\sigma^{ki} = \delta_{ij} \quad \text{for all } j=1,2,\ldots p$$

if and only if $\sigma^{ki}=0$ for all $|k-i| > s$.

The direct statement is obvious. For the converse consider that if the above two sums hold then

$$\sum_{k=1}^{i-s-1} \sigma_{jk}\sigma^{ki} + \sum_{k=i+s+1}^{p} \sigma_{jk}\sigma^{ki} = 0 \quad \text{for } j = 1,2,\ldots p.$$

In other words, the linear combination of the $1,2,\ldots i-s-1$, and $i+s+1, i+s+2,\ldots p$ column vectors of Σ with weights

$$\sigma^{1i}, \sigma^{2i}, \ldots \sigma^{i-s-1,i}, \sigma^{i+s+1,i}, \sigma^{i+s+2,i}, \ldots \sigma^{pi}$$

is equal to the null vector. As Σ is non-singular the vectors are independent whence the combination can only be null if all these weights are zero. This establishes the converse.

The second set of equations for $j=i-s, i-s+1, \ldots i, \ldots i+s$ define $(\sigma^{i-s,i}, \sigma^{i-s+1,i}, \ldots \sigma^{i,i}, \ldots \sigma^{i+s,i})'$ as the i-th column of the inverse of the principal minor of the $(i-s, i-s+1, \ldots i, \ldots i+s)$ rows and columns of Σ.

LEMMA 4: For any non-singular pxp matrix Σ, the condition

$$\sigma^{ij} = 0 \quad \text{for all } i,j \text{ such that } |i-j| > s$$

implies that in the inverse of any principal minor of consecutive rows and columns of Σ all elements outside the principal diagonal and the first s off diagonals are zero.

PROOF: Consider any principal minor of the first u+s rows and columns. By definition

$$\sum_{k=1}^{p} \sigma_{jk}\sigma^{ki} = \delta_{ij} \text{ for all } i \text{ and } j,$$

so that if $\sigma^{ij} = 0$ for $|i-j| > s$,

$$\sum_{k=1}^{u+s} \sigma_{jk}\sigma^{ki} = \delta_{ij} \text{ for } i = 1,2,\ldots u \text{ and } j=1,2,\ldots u+s.$$

But the latter equations define the first u columns of the inverse of the principal minor, which are therefore seen to have the same elements as the corresponding parts of the first u columns of Σ^{-1}. In particular, all elements of the minor for which $j > i+s$ must be zero

By the same reasoning about the first u rows of the inverse of that principal minor, all elements for which $i > j+s$ must be zero.

The same argument holds for principal minors of the last u+s consecutive rows. For minors of s rows and columns the argument is trivial.

Now, any principal minor of consecutive rows and columns of Σ can be obtained from Σ by first taking the principal minor whose last rows and columns are the ones concerned, and then taking the required minor as the last so and so many rows and columns of that. For example, in a matrix of four rows and columns one would obtain the principal minor of the second and third rows and columns by first striking out the fourth, and then from the remainder striking out the first. As has been shown, the property that all elements outside the first s off diagonals vanish is preserved when taking minors in this manner.

Therefore that property holds for any principal minor of consecutive rows and columns of Σ if it holds for Σ.

THEOREM 1: For a vector variable $\underline{X}$ with variance Σ and correlation P, the following three statements are equivalent:

H_s : $\rho_{i,i+u+1.i+1,\ldots i+u} = 0$ for all $u = s,s+1,s+2,\ldots(p-1)$;

H_s' : $\sigma^{ij} = 0$ for all $|i-j| > s$;

H_s'' : For any $i=1,2,\ldots p$ in the regressions of X_i on

(1) all other variables in $\underline{X}$, and

(2) all other X_j such that $|i-j| \leq s$,

the multiple correlations and the regression coefficients of X_j for $|i-j| \leq s$ are equal, and in (1) all other regression coefficients are null.

PROOF: By Lemmas 1 and 2, H_s implies that $\rho_{ij.\text{all others in } \underline{X}} = 0$ if $|i-j| > s$. But

$$\rho_{ij.\text{all others in } \underline{X}} = - \frac{\sigma^{ij}}{\sqrt{\sigma^{ii}\ \sigma^{jj}}}$$

(Cramer, 1946, -23.4.2 - gives this in terms of cofactors). Thus $H_s \longrightarrow H_s'$.

By Lemma 4, H_s' implies that, for any i, in the inverse of the principal minor of the $(i,i+1,\ldots i+u+1)$th rows and columns of Σ, the $(i,i+u+1)$th element is zero if $u \geq s$. But $\rho_{i,i+u+1.i+1,\ldots i+u}$ is a multiple of that element (by the same expression from Cramer quoted above) so it also is zero under H_s'. Thus $H_s' \longrightarrow H_s$.

The regression of any X_i on all the other variables in $\underline{X}$ has (regression coefficient of X_j) = $-\frac{\sigma^{ij}}{\sigma^{ii}}$ (Cramer, 1946, -23.2.4) and (multiple correlation) = $\sqrt{1 - \frac{1}{\sigma_{ii}\sigma^{ii}}}$ (Cramer, 1946, -23.5.2-).

Under H_s' clearly the regression coefficients of X_j for $|i-j| > s$ are zero. By Lemma 3, σ^{ij} for $|i-j| \leq s$ are equal to the corresponding elements in the ith column of the principal minor of the $(i-s, i-s+1, \ldots i, \ldots i+s)$th rows and columns of Σ. But the regression of X_i on $X_{i-s}, \ldots X_{i-1}, X_{i+1}, \ldots X_{i+s}$ involves the elements of the i-th column of that minor in the same formulae. As the elements involved are equal for both regressions, the regression coefficients and multiple correlations are necessarily also equal.

Conversely, if the regressions are the same, application of the formulae above shows the non-null elements of the i-th columns of Σ and the minor to be the same. By the converse of Lemma 3, $\sigma^{ji}=0$ for j such that $|i-j| \geq s$. Thus $H_s' \Longleftrightarrow H_s''$.

This completes the proof of the theorem.

It should be noted that H_s could equivalently have been stated in terms of $\sigma_{i,i+u+1.i+1,\ldots i+u}$, and H_s' in terms of ρ^{ij}.

The above are hypotheses of null correlation or covariance, not of independence. For a multivariate normally distributed $\underline{X}$, however, they are equivalent to independence (Roy, 1957, Ch. 3), and H_s would correspond to the definition of s-th ante-dependence given in the introduction.

3. Likelihood Ratio Tests of Partial Independence.

For $\underline{X}$ multivariate normal, consider a sample of N, i.e., $\underline{x}_1, \underline{x}_2, \ldots \underline{x}_N$ where $\underline{x}_\nu = (x_{1\nu}, x_{2\nu}, \ldots x_{p\nu})$. Define $\bar{\underline{x}} = (\bar{x}_1, \bar{x}_2, \ldots \bar{x}_p)$ where $\bar{x}_i = \frac{1}{N} \sum_{\nu=1}^{N} x_{i\nu}$, also define $s_{ij} = \frac{1}{N} \sum_{\nu=1}^{N} (x_{i\nu} - \bar{x}_i)(x_{j\nu} - \bar{x}_j)$ for all $i,j = 1,2,\ldots p$ and let s_{ij} be the (ij)th element of S. Similarily to the population parameters, define

S^{-1} with element s^{ij},

D with diagonal $s_{11}, s_{22}, \ldots s_{pp}$ and other elements zero,

$R = D^{-1/2} S D^{-1/2}$ with element r_{ij},

R^{-1} with element r^{ij},

$$r_{ij.k,\ell,\ldots m} = \frac{r_{ij.\ell,\ldots m} - r_{ik.\ell,\ldots m}\, r_{jk.\ell,\ldots m}}{\sqrt{1 - r^2_{ik.\ell,\ldots m}}\sqrt{1 - r^2_{jk.\ell,\ldots m}}}$$

and

$$1 - r^2_{i(j_1,j_2,\ldots j_s)} = \prod_{u=1}^{s} \left\{ 1 - r^2_{ij_u.j_1,j_2,\ldots j_{u-1}} \right\} .$$

THEOREM 2: Let $\underline{X}$ have a multivariate normal distribution with expectation $\underline{\mu}$ and variance Σ, and $X = (\underline{x}_1, \underline{x}_2, \ldots \underline{x}_N)$ be a sample of N from it. Let Γ be a set of elements of Σ^{-1} including the diagonal elements, write $(i,j)\epsilon\Gamma$ if σ^{ij} belongs to Γ, and denote summation over all values of $(i,j)\epsilon\Gamma$ by $\sum_\gamma$. Consider the hypothesis H_γ that

$$\sigma^{ij} = 0 \qquad \text{if } (ij) \notin \Gamma$$

and

$$\sigma^{ij} \text{ arbitrary} \quad \text{if } (ij) \epsilon \Gamma.$$

Then under H_γ the maximum likelihood is

$$(2\pi)^{-Np/2}|\hat{\Sigma}_{(\gamma)}|^{-N/2}\exp[-Np/2]$$

where $\hat{\Sigma}_{(\gamma)}$ has elements $\hat{\sigma}^{ij} = s_{ij}$ for $(ij) \in \Gamma$,
and $\hat{\Sigma}^{-1}_{(\gamma)}$ satisfies $\hat{\sigma}^{ij} = 0$ for $(ij) \notin \Gamma$.

PROOF: Under H_γ the parameters $\underline{\mu}$ and σ_{ij} for $(ij) \in \Gamma$ specify the distribution completely. The likelihood is maximized by putting

$$\hat{\underline{\mu}} = \bar{\underline{x}}$$

and choosing $\hat{\sigma}_{ij}$ to satisfy

$$\sum_{v=1}^{p} \sum_{z=1}^{p} (\sigma_{vz} - s_{vz}) \frac{\partial \sigma^{vz}}{\partial \sigma^{ij}} = 0$$

for all $(ij) \in \Gamma$. (This is a simple extension of the well known theory. See Anderson, 1958, section 3.2). The latter is satisfied by putting

$$\hat{\sigma}_{ij} = s_{ij} \quad \text{for all } (ij) \in \Gamma.$$

The maximum likelihood can be written

$$(2\pi)^{-Np/2}|\hat{\Sigma}_{(\gamma)}|^{-N/2}\exp[-\tfrac{N}{2}\mathrm{Tr}\ \hat{\Sigma}^{-1}_{(\gamma)}\ (\bar{\underline{x}}-\hat{\underline{\mu}})\ (\bar{\underline{x}}-\hat{\underline{\mu}})' - \tfrac{N}{2}\mathrm{Tr}\ \hat{\Sigma}^{-1}_{(\gamma)}S],$$

where $\hat{\underline{\mu}}$, as above, and $\hat{\Sigma}_{(\gamma)}$ are the maximum likelihood estimates under H_γ. Clearly the first term in the exponent vanishes. The second term is

$$-\frac{N}{2}\,\mathrm{Tr}\,\hat{\Sigma}^{-1}_{(\gamma)}S = -\frac{N}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\hat{\sigma}^{ij}s_{ij}$$
$$= -\frac{N}{2}\sum_{\gamma}\hat{\sigma}^{ij}s_{ij}$$
$$= -\frac{N}{2}\sum_{\gamma}\hat{\sigma}^{ij}\hat{\sigma}_{ij}$$
$$= -\frac{N}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\hat{\sigma}^{ij}\hat{\sigma}_{ij}$$
$$= -\frac{N}{2}\,\mathrm{Tr}\,\hat{\Sigma}^{-1}_{(\gamma)}\,\hat{\Sigma}_{(\gamma)}$$
$$= -\frac{Np}{2},$$

for as $\hat{\sigma}^{ij} = 0$ for $(ij) \notin \Gamma$, $\sum_{\gamma}\hat{\sigma}^{ij}a_{ij} = \sum_{i=1}^{p}\sum_{j=1}^{p}\hat{\sigma}^{ij}a_{ij}$ for any a_{ij}.

LEMMA 5: For any correlation matrix P

$$|P| = \prod_{i=1}^{p}(1-\rho^2_{i,i+1})\prod_{i=1}^{p-1}(1-\rho^2_{i,i+2.i+1})\cdots\prod_{i=1}^{p-s}(1-\rho^2_{i,i+s.i+1,i+2,\ldots i+s-1})$$
$$\cdots(1-\rho^2_{1,1+p-1.1+1,\ldots 1+p-2}).$$

PROOF: It is well known (Cramer, 1946, -23.5.2-) that

$$\frac{|P|}{|P_{11}|} = 1-\rho^2_{1(23\ldots p)},$$

whence $\frac{|P|}{|P_{11}|} = (1-\rho^2_{12})(1-\rho^2_{13.2})\ldots(1-\rho^2_{1p.23\ldots p-1})$.

Similarly $\frac{|P_{11}|}{|P_{1122}|} = (1-\rho^2_{23})(1-\rho^2_{24.3})\ldots(1-\rho^2_{2p.34\ldots p-1})$

and so on for further principal minors of P. Multiplying these results together the above result follows.

COROLLARY 2.1 : Under H_s the maximum likelihood is

$$(2\pi)^{-Np/2} \, |\hat{\Sigma}_{(s)}|^{-N/2} \exp[-Np/2]$$

where $$|\hat{\Sigma}_{(s)}| = \prod_{i=1}^{p} s_{ii} \prod_{i=1}^{p-1} (1-r^2_{i,i+1}) \prod_{i=1}^{p-2} (1-r^2_{i,i+2.i+1}) \cdots$$

$$\cdots \prod_{i=1}^{p-s} (1-r^2_{i,i+s.i+1,i+2,\ldots i+s-1}).$$

PROOF: It was shown in Theorem 1 that H_s is a hypothesis of the type considered in Theorem 2. Hence by Theorem 2 the above expression for the likelihood follows.

Now under H_s the maximum likelihood estimators of σ_{ii} are s_{ii}, and those of the non-null partial correlation coefficients are the sample partial correlations. This follows from the fact that the set of variances and correlations ρ_{ij} where $|i-j| \leq s$ are a one-to-one transformation of the set of variances and non-null partial correlations under H_s. (For similar considerations see Anderson, 1958, Corollary 3.2.1 and section 4.3.1).

Next, $\Sigma = \Delta^{1/2} P \Delta^{1/2}$, whence $|\Sigma| = |\Delta| . |P|$. Now $|\Delta| = \prod_{i=1}^{p} \sigma_{ii}$, and under H_s

$$|P| = \prod_{i=1}^{p-1} (1-\rho^2_{i,i+1}) \prod_{i=1}^{p-2} (1-\rho^2_{i,i+2.i+1}) \cdots \prod_{i=1}^{p-s} (1-\rho^2_{i,i+s.i+1,\ldots i+s-1}).$$

Introducing the maximum likelihood estimates one obtains the expression for $|\hat{\Sigma}_{(s)}|$.

COROLLARY 2.2: The α-level likelihood ratio test of H_s under H_{s+1} is

accept H_s if $-N \sum_{i=1}^{p-s-1} \log_e(1-r^2_{i,i+s+1.i+1,\ldots,i+s}) \leq \chi^2_{1-\alpha}$,

reject H_s otherwise,

where $\chi^2_{1-\alpha}$ is the $100(1-\alpha)$ percentage point of the χ^2 distribution with (p-s-1) degrees of freedom. This test is asymptotically valid.

PROOF: The ratio of the likelihood under H_s to that under H_{s+1} is - from Corollary 2.1 -

$$\lambda_s = \left\{ |\hat{\Sigma}_{(s)}| / |\hat{\Sigma}_{(s+1)}| \right\}^{-N/2} = \left\{ \prod_{i=1}^{p-s-1} (1-r^2_{i,i+s+1.i+1,\ldots,i+s}) \right\}^{-N/2}$$

so that

$$-2\log_e \lambda_s = -N \sum_{i=1}^{p-s-1} \log_e (1-r^2_{i,i+s+1.i+1,\ldots i+s}).$$

Now H_s specifies $(p-s)(p-s-1)/2$ zero elements in Σ^{-1} and H_{s+1} specifies only $(p-s-1)(p-s-2)/2$ of them as zero. Thus by the well known results on the asymptotic distribution of the likelihood criterion (Wald, 1943) $-2\log_e \lambda_s$ is asymptotically distributed as χ^2 with (p-s-1) degrees of freedom.

4. A Sequence of Tests of Partial Independence.

A large sample test of H_s - or equivalently of G_s under the model, or assumption, H_{s+1} has been given in Corollary 2.1, and will be denoted T_s. The argument holds for any s from 0 to p-2, and thus a sequence of tests is obtained as shown in the Table. These tests are asymptotically valid but nothing definite is known about the closeness of the asymptotic approximation for any finite sample size.

Model	Hypothesis	Test	Test Statistic	Degrees of Freedom of asymptotic χ^2
H_{p-1}	H_{p-2}	T_{p-2}	$-N\log_e(1-r^2_{i,i+p-1.i+1,\dots i+p-2})$	1
H_{p-2}	H_{p-3}	T_{p-3}	$-N\sum_{i=1}^{2}\log_e(1-r^2_{i,i+p-2.i+1,\dots i+p-3})$	2
. . .	. . .	. . .	. . .	. . .
H_3	H_2	T_2	$-N\sum_{i=1}^{p-3}\log_e(1-r^2_{i,i+3.i+1,i+2})$	p-3
H_2	H_1	T_1	$-N\sum_{i=1}^{p-2}\log_e(1-r^2_{i,i+2.i+1})$	p-2
H_1	H_o	T_o	$-N\sum_{i=1}^{p-1}\log_e(1-r^2_{i,i+1})$	p-1

It will be readily apparent from considerations of these tests as likelihood ratios that successive tests can be combined by adding the test statistics and the degrees of freedom of the χ^2s. Thus, for example, to test H_1 under H_3 the test statistic would be

$$-N\left\{\sum_{i=1}^{p-2}\log_e(1-r^2_{i,i+2.i+1})+\sum_{i=1}^{p-3}\log_e(1-r^2_{i,i+3.i+1,i+2})\right\}$$

distributed asymptotically as χ^2 with 2p-5 degrees of freedom.

A well known example is the test of H_o against the arbitrary alternative H_{p-1} whose test statistic is the sum of all those in the table above which becomes

$$-2\log_e\prod_{i=1}^{p-1}(1-r^2_{i,i+1})\prod_{i=1}^{p-2}(1-r^2_{i,i+2.i+1})\cdots$$
$$\cdots\prod_{i=1}^{2}(1-r^2_{i,i+p-2.i+1,\dots i+p-3})(1-r^2_{i,i+p-1.i+1,\dots i+p-2})$$
$$=-2\log_e|\hat{P}|$$

by Lemma 2, and is asymptotically distributed as χ^2 with $p(p-1)/2$ degrees of freedom. This will be recognized as the usual likelihood ratio test statistic of independence (Anderson, 1958, sections 9.2 and 9.6). Thus the sequence of tests $T_0, T_1, T_2, \ldots T_{p-2}$ may be considered as intermediate tests between independence and a completely arbitrary alternative. Just as in the case of hypotheses (section 2), all tests except that of H_o under H_{p-1} depend on the ordering of the variables. Though any ordering is possible and may be tested, the probability statements are valid only if the ordering is decided on independently of the sample data. Procedures for inferring on possible orders from the sample have not been investigated.

Any one T_s may be used to test H_s under H_{s+1}, or any combination $T_s+T_{s+1}+\ldots+T_{z-1}$ to test H_s under H_z $(z>s)$. These are strictly tests of significance with the assumption - H_{s+1} or H_z, respectively - and the hypothesis - H_s - being clearly defined and only two decisions possible.

A multiple decision rule may be adapted from these tests in the following manner: if $T_{p-2}, T_{p-3}, \ldots T_s$ are not significant, but T_{s-1} is, made decision D_s, i.e., accept $H_{p-1}, H_{p-2}, \ldots H_s$ and reject $H_{s-1}, H_{s-2}, \ldots H_0$. The levels of significance of the tests will be denoted by α_z for test T_z.

Some properties of such a decision procedure are straightforward. If H_s holds, $T_{p-2}, T_{p-3}, \ldots T_s$ are distributed asymptotically as χ^2 with $1, 2, \ldots (p-s-1)$ degrees of freedom, respectively; $T_{p-2}+T_{p-3}+\ldots+T_s$ is distributed asymptotically as χ^2 with $(p-s)(p-s-1)/2$ degrees of freedom.

Hence $T_{p-2}, T_{p-3}, \ldots T_s$ are independent under H_s. So the probability of none of them being signifcant is $(1-\alpha_{p-2})(1-\alpha_{p-3})\ldots(1-\alpha_s)$.

In terms of the suggested decision rule this means that if H_s holds but H_{s-1} does not, the probability of decision D_z is

$$\Pr\left\{ D_z \mid H_s \right\} = (1-\alpha_{p-2})(1-\alpha_{p-3})\ldots(1-\alpha_z)\alpha_{z-1} \qquad \text{if } z > s,$$

$$\Pr\left\{ D_z \mid H_s \right\} \leq (1-\alpha_{p-2})(1-\alpha_{p-3})\ldots(1-\alpha_s) \qquad \text{if } z \leq s$$

and

$$\Pr\left\{ \bigcup_{z=0}^{s} D_z \mid H_s \right\} = (1-\alpha_{p-2})(1-\alpha_{p-3})\ldots(1-\alpha_s).$$

Thus the probability of making a decision consonant with H_s, i.e., D_s or D_{s-1} or ... or D_0, when H_s is true and H_{s-1} is not, decreases with the complexity of the hypothesis (i.e., as s decreases). Some arguments might be advanced for the desirability of such a procedure, but for the present these properties are simply stated for consideration. (For a somewhat analogous problem in analysis of variance see Duncan, 1955, section 4.4).

A simple form of this procedure would be to require that

$$\Pr\left\{ \bigcup_{z=0}^{s} D_z \mid H_s \right\} = (1-\alpha)^{p-s-1} \text{ for all } s = 0,1,\ldots(p-1),$$

and this might be made to include a β-level significance test of independence against the most general alternative by choosing α such that $(1-\alpha)^{p-1} = 1-\beta$.

The suggested rules have analogies in polynomial regression and in factor analysis. In the former a similar procedure may be used for deciding the degree of polynomial required, starting from some high degree polynomial and testing successively the regression coefficients of the highest remaining powers of the independent variable. In factor analysis one would start from the maximal number of factors q and test whether only q-1 may be required, then q-2 under q-1, etc. An important difference between these problems and that of ante-dependence is that in polynomial regression and presumably also in factor analysis the successive test statistics are not independent. This would introduce further complications.

In any of these problems it would seem logically unsatisfactory to start from the most restrictive hypothesis and test it under the next less restrictive one, etc., until non significant tests are reached. Under such an inverse procedure the tests might be meaningless for the model need not hold, e.g., testing H_0 under H_1 if H_1 did not hold. The more satisfactory procedure seems to be to start testing the least restrictive hypothesis under the most general assumptions and then to add further restrictions and test them successively.

5. Some Aspects of Null Partial Correlation.

The hypotheses, or models, $H_0, H_1, \ldots H_{p-1}$ may be considered from several different aspects, and the present discussion will be concerned with regression on preceding variables in the sequence. Under H_s each variable X_i of $\underline{X}$ may be regarded as linearly correlated with the preceding s variables $X_{i-1}, X_{i-2}, \ldots X_{i-s}$ and, given those, uncorrelated with all further preceding, i.e., with $X_{i-s-1}, X_{i-s-2}, \ldots X_2, X_1$.

This can be expressed in the usual way by regression of X_i on $X_{i-1}, X_{i-2}, \ldots X_{i-s}$, and the residual is uncorrelated with $X_1, X_2, \ldots X_{i-1}$. This model is somewhat analogous to that of an auto-regressive scheme of order s (Kendall, 1948, Ch. 30), though the latter is usually treated with the assumption of equal regression functions at each stage, which is a special case of the present model. Moreover, the statistical treatment of autoregression is usually in terms of a single observation on the whole series, whereas the present discussion is in terms of repeated samples of the sequence.

It will be interesting to consider a particular way in which any H_s may be generated from a set of uncorrelated variables $Z_1, Z_2, \ldots Z_p$.

Define

$$X_i^{(s)} = \sum_{k_1=1}^{i} \sum_{k_2=1}^{k_1} \cdots \sum_{k_s=1}^{k_{s-1}} Z_{k_s}$$

for every j=0,1,2,...p-1 and i=1,2,...p.

Thus

$$X_i^{(0)} = Z_i \ ,$$

$$X_i^{(1)} = \sum_{k=1}^{i} Z_k,$$

$$X_i^{(2)} = \sum_{k=1}^{i} (i+1-k)Z_k,$$

and in general,

$$X_i^{(s)} = \sum_{k=1}^{i} \binom{i-k+s-1}{s-1} Z_k$$

(Riordan, 1958, p.24).

THEOREM 3: $X_i^{(s)}$ can be expressed as $Z_i+\sum_{k=1}^{s} (-1)^{k-1}\binom{s}{k}X_{i-k}^{(s)}$, and H_s holds for the sequence $X_1^{(s)},X_2^{(s)},\ldots X_p^{(s)}$.

PROOF: Assume that for some $m-1 < s$,

$$(*)\quad X_i^{(s)}=Z_i+\sum_{k=1}^{m-1} (-1)^{k-1}\binom{s}{k}X_{i-k}^{(s)}+(-1)^{m-1} \sum_{k=1}^{i-m} \frac{(i-k+s-m)!}{(i-k-m)!(s-m)!(m-1)!(i-k)} Z_k.$$

Now

$$(-1)^{m-1} \sum_{k=1}^{i-m} \frac{(i-k+s-m)!}{(i-k-m)!(s-m)!(m-1)!(i-k)} Z_k - (-1)^{m-1}\binom{s}{m}X_{i-m}^{(s)}$$

$$= (-1)^{m-1} \sum_{k=1}^{i-m} \left\{ \frac{(i-k+s-m)!}{(i-k-m)!(s-m)!(m-1)!(i-k)}Z_k-\frac{s!}{(s-m)!m!}\frac{(i-k+s-m-1)!}{(i-k-m)!(s-1)!}Z_k\right\}$$

$$= (-1)^{m} \sum_{k=1}^{i-m} \frac{(i-k+s-m-1)!}{(i-k-m)!(s-m)!m!} \left(s - \frac{(i-k+s-m)m}{i-k}\right)Z_k$$

$$= (-1)^{m} \sum_{k=1}^{i-m-1} \frac{(i-k+s-m-1)!}{(i-k-m)!(s-m!m!}\left(s - \frac{(i-k+s-m)m}{i-k}\right)Z_k$$

$$=\begin{cases} = 0 & \text{if } m = s, \\ = (-1)^{m} \sum_{k=1}^{i-m-1} \frac{(i-k+s-m-1)!}{(i-k-m-1)!(s-m-1)!m!(i-k)}Z_k & \text{if } m < s. \end{cases}$$

Thus, if $m < s$, (*) holds for m; and if m=s,

$$X_i^{(s)} = Z_i +\sum_{k=1}^{s} (-1)^{k-1} \binom{s}{k} X_{i-k}^{(s)}.$$

Checking the assumption of (*) holding for m=1, one notes that

$$(*)\qquad X_i^{(s)} = Z_i + \sum_{k=1}^{i-1} \frac{(i-k+s-1)!}{(i-k-1)!(s-1)!(i-k)} Z_k$$

$$= \sum_{k=1}^{i} \frac{(i-k+s-1)!}{(i-k)!(s-1)!}Z_k$$

which is the definition of $X_i^{(s)}$. Then by mathematical induction (*) holds for all m = 1,2,...s-1, and for m = s the equation of (*) holds with the last sum omitted.

$X_i^{(s)}$ can thus be expressed as a linear combination of Z_i and $X_{i-1}^{(s)}, X_{i-2}^{(s)}, \ldots X_{i-s}^{(s)}$; by their definitions $X_{i-s-1}^{(s)}, X_{i-s-2}^{(s)}, \ldots X_1^{(s)}$ are sums of $Z_1, Z_2, \ldots Z_{i-s-1}$ and are therefore uncorrelated with Z_i. It is clear, then, that $X_i^{(s)}$, given $X_{i-1}^{(s)}, X_{i-2}^{(s)}, \ldots X_{i-s}^{(s)}$, is uncorrelated with $X_{i-s-1}^{(s)}, X_{i-s-2}^{(s)}, \ldots X_2^{(s)}, X_1^{(s)}$. In other words, the variables $X_1^{(s)}, X_2^{(s)}, \ldots X_p^{(s)}$ have property H_s, which proves the theorem.

Thus s successive cumulative summations of p uncorrelated variables yield p variables for which the hypothesis H_s holds. The converse is not true, so that it cannot be inferred from H_s that the variables $X_1, X_2, \ldots X_p$ are generated in such a manner.

The case s = 1 is known as Guttman's Perfect Simplex and has been proposed as a model for batteries of psychological tests increasing in complexity (Guttman, 1954, pp. 309-311 and 1955). In that case the Z_is are assumed to be uncorrelated "elementary components" of the tests, and each successive test is made up of all the components of the preceding tests plus one more component (the model holds only if the relative contributions of the earlier components remain unchanged as more components are added on).

For s > 1 the interpretation of this scheme does not seem as simple. It may be noted, however, that with increasing s the relative weight of the earlier Z_is becomes greater. Perhaps this might describe some pattern in which uncorrelated increments are of decreasing importance as one moves up in the order of the sequence of variables.

6. Generalization to Several Dimensions.

A multidimensional generalization of the foregoing will be to consider p sets of k variables each

$$\underline{X_1} = X_{1_1}, X_{1_2}, \ldots X_{1_k},$$
$$\underline{X_2} = X_{2_1}, X_{2_2}, \ldots X_{2_k},$$
$$\cdot\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$\underline{X_p} = X_{p_1}, X_{p_2}, \ldots X_{p_k}.$$

The corresponding hypotheses can be written

$$G_s^{(k)} : \quad \rho_{i_\alpha j_\beta \cdot \underline{i+1}, \underline{i+2}, \ldots \underline{j-1}} = 0$$

for all $i = 1,2,\ldots p-s-1$ and all $\alpha,\beta = 1,2,\ldots k$, where the subscript $\underline{i}$ denotes all k variables of the set $X_{\underline{i}}$. As before

$$H_s^{(k)} : \bigcap_{u=s}^{p-1} G_u^{(k)} .$$

Extension of the results of previous sections to this more general case will be indicated by reformulating the Lemmas and Theorems - the extended versions will be denoted by a prime. The proofs are mostly parallel to those in the earlier sections and will be sketched only if there is anything substantially different in them.

Further generalization to the case of unequal numbers of variables in the p sets would follow lines essentially similar to those of this section. The relevant expressions could be written down as an extension of those given here.

For either equal or unequal numbers of variables in each set the hypothesis relates only to the order between the sets and is invariant under permutations of the variables within any or all sets.

LEMMA 1': Under $H_s^{(k)}$, $\rho_{i_\alpha j_\beta . \phi} = 0$ for all i, j, α and β where ϕ denotes any class of s or more successive sets of variables between the sets $X_{\underline{i}}$ and $X_{\underline{j}}$.

LEMMA 2': Under $H_s^{(k)}$, $\rho_{i_\alpha j_\beta . \psi} = 0$ for all i, j, α and β where ψ is any set of variables including at least s successive sets of variables intermediate to the sets $X_{\underline{i}}$ and $X_{\underline{j}}$.

LEMMA 3': For any non-singular kpxkp matrix Σ the following two statements are equivalent for any i:

(a') $\sigma^{i_\alpha j_\beta} = 0$ for all α, β and all j such that $|i-j| > s$;

(b') the $\underline{i}$-th (i.e., $i_1, i_2, \ldots i_k$-th) columns of the inverse of the principal minor of the $\underline{i-s}, \underline{i-s+1}, \ldots \underline{i}, \ldots \underline{i+s}$ rows and columns of Σ are made up of the corresponding elements of the inverse of Σ.

Now define the $\underline{i}, \underline{j}$-th k-minor of Σ as the minor of the $(i_1, i_2, \ldots i_k)$-th rows and the $(j_1, j_2, \ldots j_k)$-th columns of Σ. Also denote the partitioning of a kpxkp matrix into p^2 k-minors as the k-partitioned matrix and define diagonals and off diagonals as the k-partitioned matrix in terms of k-minors. It is then clear that the statement:

the $\underline{i}, \underline{j}$-th k-minor of Σ^{-1} is zero for all $|i-j| > s$, is equivalent to: $\sigma^{i_\alpha j_\beta} = 0$ for all α, β and i, j such that $|i-j| > s$.

LEMMA 4': For any non-singular kpxkp matrix Σ the condition:

$\sigma^{i_\alpha j_\beta} = 0$ for all α, β and i, j such that $|i-j| > s$,

implies that in the k-partitioned inverse of any principal minor of consecutive rows and columns of k-minors, all k-minors outside the principal diagonal and first s off diagonals are zero.

THEOREM 1': For a vector variable $\underline{X} = (X_{1_1}, \ldots X_{1_k}, X_{2_1} \ldots X_{p_k})$ with variance Σ and correlation P, the following three statements are equivalent:

$H_s^{(k)}$: $\rho_{i_\alpha (i+u+1)_\beta . \underline{i+1}, \underline{i+2}, \ldots \underline{i+u}} = 0$ for all $u = s, s+1, \ldots p-1$ and $\alpha, \beta = 1, 2, \ldots k$;

$H_s^{(k)'}$: $\sigma^{i_\alpha j_\beta} = 0$ for all α, β and all i, j such that $|i-j| > s$;

$H_s^{(k)''}$: for any i and any α in the regressions of X_{i_α} on:

(1) all other variables in $\underline{X}$ apart from those in $X_{\underline{i}}$,

(2) all other X_{j_β} with any β and j such that $0 < |i-j| \leq s$,

the multiple correlations and the regression coefficients of X_{j_β} are equal, and in (1) all other regression coefficients are zero.

PROOF: The proof for $H_s^{(k)} \longleftrightarrow H_s^{(k)'}$ is a straightforward extension of that in Theorem 1. $H_s^{(k)''}$, however, differs from H_s'' in omitting all X_{i_γ}, $\gamma \neq \alpha$ from both the regressions mentioned.

With this omission X_{i_α}, given all $X_{\underline{i-s}}, X_{\underline{i-s+1}}, \ldots X_{\underline{i-1}}, X_{\underline{i+1}}, \ldots \ldots X_{\underline{i+s}}$, is uncorrelated with any other X_{j_β} such that $|i-j| > s$. Using the matrix with these rows and columns omitted, the proof of $H_s^{(k)} \longleftrightarrow H_s^{(k)''}$ is along the same lines as that of $H_s \longleftrightarrow H_s''$.

Theorem 2 is not stated specifically in terms of H_s and therefore stands as it is also for the generalized case.

Denote the partial correlation of X_{i_α} and X_{j_β}, given $X_{i_{\alpha+1}}, X_{i_{\alpha+2}}, \ldots X_{j_{\beta-1}}$, as $\rho_{i_\alpha j_\beta . int}$ (if $j \leq i$ and/or $\beta \leq \alpha$, the definition would be adjusted suitably). Define next, Γ_s the class of all pairs (i_α, j_β) such that $j-i = s$, α,β arbitrary. Then note that Lemma 5 can be reformulated as

LEMMA 5': $$|P| = \prod_{u=0}^{p-1} \prod_{(i_\alpha, j_\beta)\epsilon\Gamma_s} (1 - \rho^2_{i_\alpha j_\beta . int}).$$

It follows immediately that

COROLLARY 2.1': Under $H_s^{(k)}$ the maximum likelihood is

$$(2\pi)^{-Nkp/2} \left| \hat{\Sigma}_{(s,k)} \right|^{-N/2} \exp \left[-Nkp/2 \right]$$

where

$$\left| \hat{\Sigma}_{(s,k)} \right| = \prod_{i=1}^{kp} s_{ii} \prod_{u=0}^{s-1} \prod_{(i_\alpha, j_\beta)\epsilon\Gamma_s} (1 - r^2_{i_\alpha j_\beta . int}) .$$

COROLLARY 2.2': The α-level likelihood ratio test of $H_s^{(k)}$ under $H_{s+1}^{(k)}$ is:

accept $H_s^{(k)}$ if $-N \sum_{(i_\alpha, j_\beta)\epsilon\Gamma_s} \log_e (1-r^2_{i_\alpha j_\beta . int}) \leq \chi^2_{1-\alpha}$,

reject $H_s^{(k)}$ otherwise,

where $\chi^2_{1-\alpha}$ is the $100(1-\alpha)$ percentage point of the χ^2 distribution with $k^2(p-s-1)$ degrees of freedom.

This provides a sequence of tests $T_s^{(k)}$, $s = p-2, p-3, \ldots 2,1,0$, corresponding to those in section 4. The hypotheses are analogous to those of multiple time series with autoregressive schemes.

A model corresponding to that of section 5 would presumably be in terms of k separate cumulative summations based on k sets of variates $Z_{11}, Z_{12}, \ldots Z_{1p}$; $Z_{21}, Z_{22}, \ldots Z_{2p}$; . . . ; $Z_{k1}, Z_{k2}, \ldots Z_{kp}$, such that any Z_{i_α} and Z_{j_β} are uncorrelated if $i \neq j$.

Acknowledgement: I am greatly indebted to my teacher Professor S. N. Roy for his stimulating lectures and discussions and for his suggestions on the presentation of this work. I am also indebted to Professor L. Guttmann who first set me thinking along the lines developed in this paper.

REFERENCES

Anderson, T. W., 1958, An Introduction to Multivariate Statistical Analysis. New York, Wiley.

Cramer, H., 1946, Mathematical Methods of Statistics. Princeton University Press.

Duncan, D. B., 1955,"Multiple range and multiple F tests", Biometrics, 11, pp. 1-42.

Guttman, L., 1954, "A New Approach to Factor Analysis: The Radex." In Mathematical Thinking in the Social Sciences, Ed. P.F. Lazarsfeld. Glencoe, The Free Press.

Guttman, L., 1955, "A Generalized Simplex for Factor Analysis," Psychometrika, 20, pp. 173-192.

Kendall, M. G., 1948, The Advanced Theory of Statistics, Vol. 2. London, Griffin.

Riordan, J., 1958, An Introduction to Combinatorial Analysis. New York, Wiley.

Roy, S. N., 1957, Some Aspects of Multivariate Analysis. New York, Wiley.

Wald, A., 1943, "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large", Trans. Amer. Math. Soc., 54, pp. 426-482.