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

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

SOME RECENT RESULTS IN NORMAL MULTIVARIATE CONFIDENCE BOUNDS

by

S. N. Roy

August, 1961

Contract No. AF 49(638)-213

This paper gives a survey of some recent results, most of them available in print and the rest still unpublished, in confidence bounds on parametric functions associated with multivariate normal distributions, that have been developed with a particular end in view. The general philosophy is discussed in greater detail than heretofore, and most of the results obtained so far are exhibited in that light, the bulk of the details of the various derivations being left out. Each parametric function that is considered is interesting either by itself or as a measure of departure from some (customary) null hypothesis in the direction of some alternative that may be either standard or non-standard. Among the results discussed are those connected with one or two population dispersion matrices, multivariate linear hypotheses under a linear model, independence between two sets of variates, multiple independence among a p-set or among k sets of $p_1, p_2, \ldots, p_k$, and the vector analogues of the ratio of means and ratio of variances for a correlated bivariate normal distribution.

Institute of Statistics
Mimeograph Series No. 303

SOME RECENT RESULTS

IN NORMAL MULTIVARIATE CONFIDENCE BOUNDS*

by

S. N. Roy
Department of Statistics
University of North Carolina

1. Introduction. Confidence bounds on parametric functions associated with multivariate normal distributions, set up with a particular objective and under a particular philosophy, have been discussed in a series of publications beginning in 1953 and continuing to date. Partly for the lack of any semi-expository article that brings together most of the main results and highlights the characteristic features and the implications, and partly for other reasons, the whole point and purpose of this line of work seem to have been completely lost upon the general statistical public so far. The present paper is intended, in some measure, to some as such an expository paper and, in some measure, to offer certain very recent results not yet available in print.

1.1. Intervals based on (H_o, H). In many, though by no means in all, problems we start from a pair (H_o, H) of composite hypothesis and alternative, disjoint but not necessarily exhaustive, and seek a parametric function that might be regarded as a measure of de-

*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.

parture from H_o in the direction of H, or, alternatively, some kind of a "distance function" between the set H_o and the set H. In such problems we next seek a confidence interval for this parametric function, one-sided (one way or the other) or two-sided, depending upon the nature of the pair (H_o, H). No claim is made at this stage that the parametric function chosen or the confidence interval proposed for it is, in some sense, optimal. As to the confidence coefficient, it would no doubt be very desirable if, given any permissible $1 - \alpha$, the interval could be defined such that this coefficient was equal to $1 - \alpha$. If it does not turn out that way, the next best thing would be to have a confidence coefficient $\geq 1 - \alpha$, given any permissible α, such that the equality is attained, or, in other words, the probability of the interval covering this parametric function, for some value of this function, is equal to $1 - \alpha$. If this does not happen (or, at any rate, cannot be readily proven), then the next best thing would be, for any permissible $1 - \alpha$, to have a confidence coefficient whose greatest lower bound $\geq 1 - \alpha$ (and might, in fact, be greater than $1 - \alpha$), provided that the interval itself is not trivial, for example, $(0, \infty)$ or $(-\infty, \infty)$, etc., but is, in fact, much better than these. We shall say that such a confidence coefficient is a conservative one, or, alternatively, that the corresponding confidence region is a conservative one. For really complex problems even this may be difficult to obtain, to say nothing of intervals of the first or the second kind, and we would consider even this quite worthwhile, especially in view of the fact that we consider it far more important to estimate this "distance function" point wise or interval wise, than to test (and accept or reject) the customary null hypothesis as such. All confidence intervals discussed in this paper, that emerge as the end products for different problems (as distinct from the confidence regions that occur at the intermediate stages of the derivation) are conservative in the sense just indicated. While we still stay with (H_o, H) there are three other features that merit special notice.

1.1.1. Bounds on the "partials". In most complex problems involving (H_o, H) there are reasons, set forth partly in [6, 10] but more fully in [17, 18], why we prefer to look upon H_o as the intersection of a number of more primitive (composite) hypotheses and H as either the union or the intersection of the same number of corresponding and more primitive (composite) alternatives. In symbols this may be expressed as

$$(1) \qquad H_o = \bigcap_{i \in \Gamma} H_{oi}, \text{ and } H = \bigcup_{i \in \Gamma} H_i \text{ or } \bigcap_{i \in \Gamma} H_i \ ,$$

where (H_{oi}, H_i) form a natural pair and the subscript i belongs to an index set that might even be a part of the continuum. According to this way of looking at (H_o, H), H_{oi} and H_i (for $i \in \Gamma$) will be called the partials (or components) of H_o and H. For any such component or partial pair (H_{oi}, H_i) we would be naturally interested in a corresponding measure of departure from H_{oi} in the direction of H_i or a "distance function", with appropriate bounds on it, that goes with (H_{oi}, H_i). The question is whether it is possible to provide, with a joint conservative confidence coefficient $\geq 1 - \alpha$, simultaneous confidence bounds on the "total distance function" that goes with (H_o, H) and on the different "partial or component distance functions" that go with the different (H_{oi}, H_i)'s. For the cases discussed in this paper not only has this been possible, but the bounds on the "partial or component parametric (distance) functions" are seen to vary with the components and furthermore, for certain classes of components, the bounds are also sharper than for the "total parametric function", which is a very satisfying feature wherever it occurs. This will be amply illustrated in the following pages.

1.1.2. The nature of the interval estimation in terms of the percentage points of the distribution function used. The parametric or the socalled "distance function" set up in the following pages as also the bounds on it is, in every case, such that the percentage points used are those of a distribution function associated with the customary null hypothesis, and never those of a distribution function associated with the non null hypothesis. Not only is the latter kind of distribution functions (mostly called non-central distributions), in general, more complicated than the former type, but in multivariate problems of the types considered these are so complicated as to be practically unusable. However, parametric ("distance") functions that require for confidence interval estimation the percentage points of the "non-central distributions" are also, physically, equally meaningful. For purposes of our present work the only reason they have not been used is because of the complexity of the distribution used. Moreover, as one little illustration would indicate, they are by no means more meaningful or significant than the ones being actually used. Consider, for example, the parameter ρ (correlation coefficient) and β (the regression coefficient of, say, y or x) for a bivariate normal population. While both are equally meaningful and useful, it is well known that to put bounds on β one needs only the central distribution of r (the sample correlation coefficient) whereas, to put bounds on ρ, one needs the non-central distribution of r, which is more complicated but not intractable as in the multivariate problems discussed here. For ρ the claim is made that, unlike β, it is scale invariant, and much is made of that by many people. However, my own experience with nonparametric generalizations of multivariate analysis has more or less convinced me that in nonparametric situations the analogues of the regression line or lines are far more meaningful than the analogue of a single measure like ρ, and that is why I believe that even for the bivariate normal population β, in that it builds up the regression line, is at least as meaningful as ρ, notwithstanding the issue of scale invariance or lack of it. Given the scales in a particular situation, β builds up the appropriate regression line. The parametric functions actually used in the following pages should be looked at from this viewpoint.

1.1.3. Additional requirement beyond what is suggested by the pair (H_o, H). In problems based on a pair (H_o, H) it happens quite often that while a one-sided interval (with an upper or a lower bound) for the parametric function is indicated in a natural manner by the pair (H_o, H), there are other physical considerations that seem to call for one step further to complete the one-sided into a two-sided interval, by adjoining a lower or an upper bound to an upper or a lower bound. In other words, while a pair (H_0, H) might set in motion the search for a suitable parametric function it may not always completely determine whether the confidence interval for this function is to be one-sided or two-sided.

1.2. Intervals not prima facie based on any (H_0, H). Starting from an appropriate test procedure for any (H_0, H) we can, under certain broad conditions, obtain a suitable confidence statement by "inversion". Such statements (on some kind of a "distance function") are almost always meaningful, though we may not always happen to be interested in such statements. However, this seems to be, to some extent, a one-way affair. In other words, there are situations where a confidence statement of a certain nature seems to be of direct physical interest in its own right and can also be obtained without any conscious attempt to associate it with a pair (H_0, H), although, with some effort, a pair (H_0, H) to tag it on to can be always obtained, which , in many cases, may not be too interesting or meaningful in itself. Sometimes the confidence statements to be looked for are suggested by an H_0 and its decomposition without explicit introduction of an H or any deliberate attempt to link the statements with a pair (H_0, H) and its decomposition. Most of the other remarks made in subsections 1.1.1. and 1.1.2. will also apply to the intervals under the heading of this section, and hence such remarks need not be repeated here.

2. Bounds related to the dispersion matrix Σ of an $N(\underset{p\times 1}{\zeta}, \underset{p\times p}{\Sigma})$ [8, 10, 11]. The case belongs to the category of section 1.2. rather than to that of section 1.1. With a sample dispersion matrix $\underset{p\times p}{S}$, let us start from the statement

$$(2) \quad \lambda_1 \leq \frac{a' S a}{a' \Sigma a} \leq \lambda_2 \text{ (for all non null } a) ,$$

with a joint probability $1 - \alpha$, such that λ_1 and λ_2 are based on the joint distribution of the largest and smallest root of $S\ \Sigma^{-1}$ (where Σ is, in fact, the population dispersion matrix). $\underset{p\times 1}{a}$ is a vector of weights to the different variates, and it is to be noted that (2) is eventually invariant under the choice of this weight system. Let us rewrite this as

$$(3) \qquad 1/\lambda_2 \ \frac{a' S a}{a' a} = \frac{a' \Sigma a}{a' a} \leq 1/\lambda_1 \ \frac{a' S a}{a' a} \text{ (for all non null } a).$$

We are, of course, interested in bounds on $(a' \Sigma a)/(a'a)$ for all non null a, that is, for all values of the weight system a. Notice that, for any a, the actual bounds on $(a' \Sigma a)/(a'a)$ and the width of the interval depend, among other things, on a as well. Among the infinity of $(a' \Sigma a)/(a'a)$ (for varying a), we are especially interested in the subset $[\underset{max}{ch} (\Sigma), \underset{min}{ch} (\Sigma)]$, $[\underset{max}{ch} (\Sigma\binom{i}{i}), \underset{min}{ch} (\Sigma\binom{i}{i})]$, $[\underset{max}{ch} (\Sigma\binom{i,i'}{i,i'}), \underset{min}{ch} (\Sigma\binom{i,i'}{i,i'})]$, and so on $(i \neq i' \neq \ldots = 1,2,\ldots,p)$; or, in other words, in the largest and smallest roots of the total matrix, and in the largest and smallest roots of the partial matrices obtained by cutting out any one variate, any two variates, and so on. With a corresponding notation for the partial sample dispersion matrices it is shown in [8, 10, 11] that (3) implies the $2^p - 1$ pairs of bounds

$$\text{(4) (i)} \quad \frac{1}{\lambda_2}\, ch_{max}(S) \leq ch_{max}(\Sigma) \leq \frac{1}{\lambda_1}\, ch_{max}(S) \;,$$

$$\frac{1}{\lambda_2}\, ch_{min}(S) \leq ch_{min}(\Sigma) \leq \frac{1}{\lambda_1}\, ch_{min}(S) \;;$$

$$\text{(ii)} \quad \frac{1}{\lambda_2}\, ch_{max}(S\{^{(i)}_{(i)}\}) \leq ch_{max}(\Sigma\{^{(i)}_{(i)}\}) \leq \frac{1}{\lambda_1}\, ch_{max}(S\{^{(i)}_{(i)}\}) \;,$$

$$\frac{1}{\lambda_2}\, ch_{min}(S\{^{(i)}_{(i)}\}) \leq ch_{min}(\Sigma\{^{(i)}_{(i)}\}) \leq \frac{1}{\lambda_1}\, ch_{min}(S\{^{(i)}_{(i)}\})$$

$$(i = 1, 2, \ldots, p-1);$$

$$\text{(iii)} \quad \frac{1}{\lambda_2}\, ch_{max}(S\{^{(i,i')}_{(i,i')}\}) \leq ch_{max}(\Sigma\{^{(i,i')}_{(i,i')}\}) \leq \frac{1}{\lambda_1}\, ch_{max}(S\{^{(i,i')}_{(i,i')}\}) \;,$$

$$\frac{1}{\lambda_2}\, ch_{min}(S\{^{(i,i')}_{(i,i')}\}) \leq ch_{min}(\Sigma\{^{(i,i')}_{(i,i')}\}) \leq \frac{1}{\lambda_1}\, ch_{min}(S\{^{(i,i')}_{(i,i')}\})$$

$$(i \neq i' = 1, 2, \ldots, p - 1);$$

and so on until one has cut out any $(p - 1)$ variates. We have thus a simultaneous confidence statement with a joint (conservative) confidence coefficient $\geq 1 - \alpha$. Notice here that the largest and smallest roots of Σ are, each, given one pair of bounds, the largest and smallest roots of any partial are, each, given another pair of tighter bounds, and the bounds get tighter and tighter as we cut out more and more variates. It is obvious that this is a desirable feature. It may be further remarked that (ii) and (iii) do not follow directly from (i) but that all of them follow from (3) by specializing the vector a. Another remark is also in order. The upper bound in the first line of (i) and the lower bound in the second line of (i) are automatically the upper and lower bounds of all roots of Σ and also of the partials. But the point is that (ii), (iii), etc. are providing much tighter bounds to the partials than these. These features will be repeated again and again in the other cases discussed in the following pages, and hence will not be discussed in as much detail as in this first case. Another set of confidence intervals (for a different set of parametric functions) based on a different principle of decomposition has been given in [5], but this will not be discussed here. As a very special case of (4) may be mentioned that of $p=1$, where we have, for $N(\zeta, \sigma^2)$, the two-sided bounds on σ^2

$$(5) \quad \lambda_1 \; s^2 \leq \sigma^2 \leq \lambda_2 \; s^2 \quad ,$$

where λ_1 and λ_2 are obtained from the percentage points of the χ^2- distribution with appropriate d.f.

3. Bounds related to the dispersion matrices (Σ_1, Σ_2) of $N(\zeta_1, \Sigma_1)$, $N(\zeta_2, \Sigma_2)$. We shall discuss the problem under two heads, namely, confidence statements in the category of section 1.2., not explicitly based on any (H_o, H), and those in the category of section 1.1, based explicitly on pairs of (H_o, H)

3.1. Intervals not based explicitly on any (H_o, H). We start from an analogue of (3) with an exact probability $1 - \alpha$ [8, 9, 10, 11], and obtain, with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, simultaneous bounds on $ch_{max}[\Sigma_1 \Sigma_2^{-1}]$, $ch_{min}[\Sigma_1 \Sigma_2^{-1}]$; $ch_{max}[\Sigma_{1(i)}^{(i)} \Sigma_{2(i)}^{(i)^{-1}}]$, $ch_{min}[\Sigma_{1(i,i')}^{(i,i')} \Sigma_{2(i,i')}^{(i,i')^{-1}}]$; etc.

Denoting the corresponding sample dispersion matrices by S_1 and S_2 (as in section 2), the simultaneous bounds on the whole set, total and partials, $2^p - 1$ in number, can be expressed in the compact form

$$(6) \quad \lambda_1 \ ch_{min} \left[S_{1(i_1,i_2,\ldots,i_{p_1})}^{(i_1,i_2,\ldots,i_{p_1})} \ S_{2(i_1,i_2,\ldots,i_{p_1})}^{(i_1,i_2,\ldots,i_{p_1})^{-1}} \right]$$

$$\leq ch_{min} \left[\Sigma_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})} \ \Sigma_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}} \right]$$

$$\leq ch_{max} \left[\Sigma_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})} \ \Sigma_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}} \right]$$

$$\leq \lambda_2 \ ch_{max} \left[S_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})} \ S_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}} \right] ,$$

where $i_1 < i_2 < \ldots < i_{p_1}$ $(0 \leq p_1 \leq p - 1)$ is any ordered subset from $(1, 2, \ldots, p)$ and the right hand upper and lower subscripts indicate that the corresponding rows and columns have been cut out. Notice that the bounds as also the width of the interval depend upon $(i_1, i_2, \ldots, i_{p_1})$, getting tighter and tighter as we increase p_1. Another point worth mention is that, unlike the case of section 2, here we don't have separate pairs of bounds for the largest and the smallest root. As a very special case of (6), putting $p = 1$, we obtain for $N(\xi_i, \sigma_i^2)$ $(i = 1, 2)$, the confidence interval

$$(7) \quad \lambda_1 \ s_1^2/s_2^2 \leq \sigma_1^2/\sigma_2^2 \leq \lambda_2 \ s_1^2/s_2^2 ,$$

where λ_1 and λ_2 are based on the percentage points of the (central) F-distribution with appropriate d.f. For another set of confidence statements (on a different set of parametric functions) based on a different principle of decomposition, the reader is referred to [5].

3.2. Intervals based explicitly on pairs of (H_o, H) [18]. Consider an H_o: $\Sigma_1 = \Sigma_2$ as expressed in the form of a decomposition

$$(8) \quad H_o: (\Sigma_1 = \Sigma_2) \iff \bigcap_a (a'\Sigma_1 a = a'\Sigma_2 a)$$

$$\iff \text{all } \operatorname{ch}(\Sigma_1 \Sigma_2^{-1}) = 1,$$

and three different alternatives, given respectively by

$$(9) \quad \text{(i)} \ H^{(1)}: \ \bigcap_a (a'\Sigma_1 a > a'\Sigma_2 a) \iff \operatorname{ch}_{min}(\Sigma_1 \Sigma_2^{-1}) > 1$$

$$\text{(ii)} \ H^{(2)}: \ \bigcap_a (a'\Sigma_1 a < a'\Sigma_2 a) \iff \operatorname{ch}_{max}(\Sigma_1 \Sigma_2^{-1}) < 1, \text{ and}$$

$$\text{(iii)} \ H^{(3)}: \ H^{(1)} \qquad H^{(2)} = \operatorname{ch}_{max}(\Sigma_1 \Sigma_2^{-1}) < 1 \text{ or } \operatorname{ch}_{min}(\Sigma_1\Sigma_2^{-1}) > 1.$$

Starting from suitable analogues of (3) (different for the three different pairs $(H_o, H^{(i)})$ $i = 1, 2, 3$)), with an exact probability $1 - \alpha$, we have, with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, the following three sets of confidence intervals for the three cases [18].

$$(10) \quad \operatorname{ch}_{min}\left[\Sigma_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})} \ \Sigma_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}}\right]$$

$$\geq \lambda_1 \operatorname{ch}_{min}\left[S_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})} \ S_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}}\right],$$

$$(11)\quad ch_{max}\left[\Sigma_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})}\ \Sigma_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}}\right]$$

$$\leq \lambda_2\ ch_{max}\left[S_{1(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})}\ S_{2(i_1,\ldots,i_{p_1})}^{(i_1,\ldots,i_{p_1})^{-1}}\right],$$

(12) (10) $\bigcup$ (11), with different values for λ_1 and λ_2, say λ_3 and λ_4 ($\lambda_3 > \lambda_4$).

The right hand side upper and lower subscripts have the same interpretation as in subsection 3.1. For other interesting alternatives $H^{(i)}$ the reader is referred to [18]. Two other comments are in order. First, for the very special case of $p = 1$, we easily obtain analogues of (7) appropriate to (10), (11), and (12). Second, applying the principle of section 1.1.3., we can extend (10) or (11) and obtain a set of two-sided bounds of the form (6).

4. Linear hypothesis under a multivariate linear model. This belongs to the category of section 1.1. For simplicity, the model we consider here is called model I of ANOVA and MANOVA. This, together with the associated tests and confidence bounds is discussed in [10, 11, 14, 15]. For Model II and mixed models for MANOVA (with the tests and the associated confidence intervals) the reader is referred to [15, 16]. The model I for MANOVA is the following.

$\underset{p\times n}{X} \equiv [\underset{1}{x_1}\ \underset{1}{x_2}\ \ldots\ \underset{1}{x_n}]\ p$ consists of n independently distributed vectors such that

$$(13)\quad \text{(i)}\quad x_i:\ N\left[\mathscr{E}(x_i),\ \underset{p\times p}{\Sigma}\right]\ (i = 1, 2, \ldots, n),$$

$$\text{(ii)}\quad \mathscr{E}(\underset{n\times p}{X'}) = \underset{n\times m}{A}\ \underset{m\times r}{\zeta},$$

where $\mathscr{E}$ stands for expectation, ζ is a matrix of unknown parameters, A is a matrix of given constants (that might include some observed concomitant variables as well) depending upon the design and the model, of rank $r \leq m \leq n - p$. Under this model we set up

$$(14)\quad H_0:\ \underset{s\times m}{C}\ \underset{m\times p}{\zeta}\ \underset{p\times u}{U} = \underset{s\times u}{0} \quad \text{against}\quad H:\ C\ \zeta\ U \neq 0\ (=\underset{s\times u}{\eta},\ \text{say}),$$

where C and U are given matrices of constants of ranks $s \leq r$ and $u \leq p$ and where it is furthermore assumed that

$$\operatorname{rank}\left[\underset{n\times m}{A}\right] = \operatorname{rank}\begin{array}{c}\begin{bmatrix} A \\ C \end{bmatrix}\begin{array}{l} n \\ s \end{array} \\ m\end{array},$$

which is characterized in [14, 19] as the condition for testability of H_0 in the strong sense. Let us denote by Q_{H_0}, Q_{error} and $Q^*_{H_0}$ the symmetric positive semi-definite matrices given in terms of (A, C), A and (A, C) in [10, 11, 14, 15]. They are of course available in other forms as well from the extensive literature of ANOVA, and they might be conveniently called the matrices due to the hypothesis, the error and the alternative. As distinct from these let us define

$$(15)\quad \underset{u\times u}{S_{H_0}} = \frac{1}{s}\left[\underset{u\times p}{U'}\ \underset{p\times n}{X}\ \underset{n\times n}{Q_{H_0}}\ \underset{n\times p}{X'}\ \underset{p\times u}{U}\right],$$

$$\underset{u\times u}{S_{\text{error}}} = \frac{1}{n-r}\left[\underset{u\times p}{U'}\ \underset{p\times n}{X}\ \underset{n\times n}{Q_{\text{error}}}\ \underset{n\times p}{X'}\ \underset{p\times u}{U}\right],$$

where the left side matrices are called respectively the dispersion matrices due to the hypothesis and to the error. Starting from a statement analogous to (3) (except that it is one-sided) with an exact probability $1 - \alpha$, it is possible [9, 10, 14, 15] to put, with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, simultaneous confidence bounds on $ch_{max}^{1/2} [\eta' \; Q^{*-1}_{H_o} \; \eta]$ and on partials obtained by cutting out (i) one or more columns of η (i.e., one or more rows of η') and (ii) one or more rows of η (i.e., one or more columns of η') together with the corresponding rows (and the same columns of the symmetric matrix $Q^{*}_{H_o}$, which, as can be checked from [10, 11, 14, 15] means cutting out one or more rows of C). If, as in many problems, C stands for a set of treatment contrasts, then this means cutting out one or more contrasts. Also, cutting out one or more columns of η (i.e., rows of η') means cutting out one or more columns of U, i.e., certain linear combinations out of the total set of linear combinations of the variates in which we are interested, or , in case U is an identity matrix, cutting out just one or more variates. We can also combine (i) at any level of cutting with (ii) at any level of cutting, and thus obtain $(2^u - 1) \times (2^s - 1)$ simultaneous confidence intervals altogether. $ch_{max}^{1/2} [\eta' \; Q^{*-1}_{H_o} \; \eta]$ might be regarded as a measure of deviation from the total H_o in the direction of the total alternative H which is the complement of H_o, while any partial of

$ch_{max}^{1/2} [\eta' \; Q^{*-1}_{H_o} \; \eta]$ (even one that combines (i) and (ii)) might be interpreted as a measure of deviation from the corresponding partial of H_o in the direction of its alternative which is its complement. To indicate the nature of these simultaneous confidence bounds we first note that the exact probability statement analogous to (3) which is our starting point in this case implies

$$(16)\quad ch_{max}^{1/2}\ (S_{H_o}) - \lambda\ ch_{max}^{1/2}\ (S_{error}) \le ch_{max}^{1/2}\ [\eta'\ Q_{H_o}^{*}\ \eta]$$

$$\le\ ch_{max}^{1/2}\ (S_{H_o}) + \lambda\ ch_{max}^{1/2}\ (S_{error})\ ,$$

where λ is obtained from the percentage point of the distribution of the central largest root [4].

But the analog of (3) also implies all the partials, and the whole set of partials together with (16), $(2^u - 1 \times 2^s - 1)$ in number, can be expressed compactly as

$$(17)\quad ch_{max}^{1/2}\ \left[(i_1, i_2, \ldots, i_{s_1})\ S_{H_o(j_1, j_2, \ldots, j_{u_1})}^{(j_1, j_2, \ldots, j_{u_1})}\right] - \lambda\ ch_{max}^{1/2}\ \left[S_{error(j_1, j_2, \ldots, j_{u_1})}^{(j_1, j_2, \ldots, j_{u_1})}\right]$$

$$\le\ ch_{max}^{1/2}\ \left[\eta_{(j_1, j_2, \ldots, j_{u_1})}^{(i_1, i_2, \ldots, i_{s_1})'}\ Q_{H_o(i_1, i_2, \ldots, i_{s_1})}^{*\,(i_1, \ldots,\, i_{s_1})^{-1}}\right.$$

$$\left.\times\ \eta_{(j_1, j_2, \ldots, j_{u_1})}^{(i_1, i_2, \ldots, i_{s_1})}\right]$$

$$\leq \; ch_{\max}^{1/2} \left[{}_{(i_1,\ldots,i_{s1})} S_{H_o(j_1,j_2,\ldots,j_{u_1})}^{(j_1,j_2,\ldots,j_{u_1})} \right]$$

$$+ \quad \lambda \; ch_{\max}^{1/2} \left[S_{error(j_1,j_2,\ldots,j_{u_1})}^{(j_1,j_2,\ldots,j_{u_1})} \right] \qquad ,$$

with the following explanation for the symbols. $j_1 < j_2 < \ldots < j_{u_1}$ $(0 \leq u_1 \leq u - 1)$ is any ordered subset from $(1, 2, \ldots, u)$ and $i_1 < i_2 < \ldots < i_{s_1}$ $(0 \leq s_1 \leq s - 1)$ is any ordered subset from $(1, 2, \ldots, s)$. For the symmetric matrices S_{Ho}, S_{error} and Q^*_{Ho}, the right hand super and subscripts indicate that the corresponding rows (and the same columns) have been cut out. For the matrix η (in general rectangular and asymmetric) the right hand super and subscripts indicate that the corresponding rows and columns (not of course now matching the rows) have been cut out. So far as η and $Q^*_{H_o}$ are concerned the relationship of the partials to the columns of U and rows of C has been already explained. So far as S_{H_o} and S_{error} are concerned the right hand super and subscripts go with just the corresponding columns of U. For the symmetric matrix S_{H_o}, the left hand subscripts indicate that the corresponding rows of C have been cut out, resulting in a change in the structure (or composition) of Q_{H_o} that does not alter its order and finally in a change in the structure of S_{H_o} that also does not alter its order. For the actual forms of these partials (and the total) involved in (9), especially the forms of the parametric "distance" functions that occur in the middle, in the case of treatment contrasts under BIB and PBIB designs the reader is referred to [2]. It is clear from (9) that the bounds depend upon both sets $(i_1, i_2, \ldots, i_{s_1})$ and $(j_1, j_2, \ldots, j_{u_1})$, while the width of the bound depends upon $(j_1, j_2, \ldots, j_{u_1})$ alone. A set of bounds for which even the width should depend upon $(i_1, i_2, \ldots, i_{s_1})$ as well have also been investigated in [17], but will not be discussed here. It is also obvious that for the partials we are getting tighter and tighter bounds, especially for the partials with respect to $(j_1, j_2, \ldots, j_{u_1})$. A further remark is called for here. The pair (H_o, H) or the partial pairs would seem to suggest a lower bound rather than the two-sided bounds. However, we still seek the two-sided bounds on the principle of section 1.1.3. For a set of confidence statements different from (9) and based on a different principle of decomposition of H_o the reader is referred to [5].

Three special cases of (17) deserve special mention here.

4.1. <u>The case $u = 1$.</u> This is the univariate or the ANOVA situation. Here S_{H_o}, S_{error} are scalar, usually referred to the mean S.S. due to the hypothesis and the error respectively, while η' is an $1 \times s$ row vector. This means that we can drop the "$ch_{max}^{1/2}$", take $u_1 = 0$, and express (17) in the form

$$(18) \quad [{}_{(i_1,\ldots,i_{s_1})} s_{H_o}]^{\frac{1}{2}} - \lambda [s_{error}]^{\frac{1}{2}} \leq$$

$$\leq \left[\eta^{(i_1,\ldots,i_{s_1})'} \, Q^{*\,(i_1,\ldots,i_{s_1})^{-1}}_{H_o(i_1,\ldots,i_{s_1})} \, \eta^{(i_1,\ldots,i_{s_1})} \right]^{1/2}$$

$$\leq [{}_{(i_1,\ldots,i_{s_1})} s_{H_o}]^{\frac{1}{2}} + \lambda [s_{error}]^{\frac{1}{2}},$$

where λ is based on the percentage point of F-distribution with suitable d.f.

4.2. <u>The problem of two mean vectors in $N(\underset{p\times 1}{\zeta_i}, \underset{p\times p}{\Sigma})$ $(i = 1, 2)$ [7, 10, 11].</u> Here H_0: $\zeta_1 = \zeta_2$ against H: $\zeta_1 \neq \zeta_2$, and this case is easily handled by a proper choice of A, C and U. Notice here that $U = I$. To fix A, we assume that two random samples of sizes n_1 and n_2 have been drawn from the two populations. Denote by $\underset{p\times 1}{\bar{x}_1}$, $\underset{p\times 1}{\bar{x}_2}$ the two sample mean vectors and by $\underset{p\times p}{S_1}$, $\underset{p\times p}{S_2}$, the two sample dispersion matrices and set $S = (n_1S_1 + n_2S_2)/(n_1 + n_2)$. Then the set (17) reduces to

$$(19)\quad \left[(\bar{x}_1 - \bar{x}_2)^{(j_1,\ldots,j_{p_1})'}(\bar{x}_1 - \bar{x}_2)^{(j_1,\ldots,j_{p_1})}\right]^{1/2}$$

$$- \lambda\, ch^{1/2}_{max}\left[S^{(j_1,\ldots,j_{p_1})}_{(j_1,\ldots,j_{p_1})}\right]$$

$$\leq \left[(\zeta_1 - \zeta_2)^{(j_1,\ldots,j_{p_1})'}(\zeta_1 - \zeta_2)^{(j_1,\ldots,j_{p_1})}\right]^{1/2}$$

$$\leq \left[(\bar{x}_1 - \bar{x}_2)^{(j_1,\ldots,j_{p_1})'}(\bar{x}_1 - \bar{x}_2)^{(j_1,\ldots,j_{p_1})}\right]^{1/2}$$

$$+ \lambda\, ch^{1/2}_{max}\left[S^{(j_1,\ldots,j_{p_1})}_{(j_1,\ldots,j_{p_1})}\right],$$

where $j_1 < j_2 < \ldots < j_{p_1}$ $(0 \leq p_1 \leq p - 1)$ is any ordered subset from $(1, 2, \ldots, p)$ and λ is based on the percentage point of the distribution of Hotelling's T^2 with appropriate d.f., i.e. of the F-distribution with proper d.f. The number of intervals in (19) is $2^p - 1$.

4.3. The problem of two comparable mean vectors in a

$$N\left[\begin{array}{cc} \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} & \begin{array}{c} p \\ p \end{array} \\ 1 & \end{array}, \begin{array}{cc} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} & \begin{array}{c} p \\ p \end{array} \\ \begin{array}{cc} p & p \end{array} & \end{array}\right] \quad [10, 11].$$

Here we have a random sample of size n, which fixes the A. We have H_0: $\zeta_1 = \zeta_2$ against H: $\zeta_1 \neq \zeta_2$, which easily fixes C and sets

$$(20) \qquad \underset{2p\times p}{U} = \begin{array}{cc} \begin{bmatrix} I \\ -I \end{bmatrix} & \begin{array}{c} p \\ p \end{array} \\ p & \end{array} .$$

As in the preceding section, let $\bar{x}_1$ and $\bar{x}_2$ stand for the sample mean vectors for the two different p-sets and, unlike the preceding section, set now $\underset{p\times p}{S^*} = S_{11} + S_{22} - S_{12} - S_{12}'$. Then, with $u = p$, (17) now reduces to a form which looks exactly like (20), except that S^* is now differently constituted and λ is different and is based on the percentage point of the distribution of a Hotelling's T^2 with different d.f.

5. Independence between two sets of variates [8, 10, 11]. This belongs to the category of section 1.1. We have here a population

$$N\left[\begin{array}{cc} \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} & \begin{array}{c} p \\ q \end{array} \\ 1 & \end{array} \quad \begin{array}{ccc} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} & \begin{array}{c} p \\ q \end{array} \\ \begin{array}{cc} p & \qquad q \end{array} & \end{array}\right] \qquad (p \leq q),$$

and a sample of size $n + 1$ ($\geq p + q$), with a sample dispersion matrix denoted by

$$\begin{array}{cc} \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} & \begin{array}{c} p \\ q \end{array} \\ \begin{array}{cc} p & \qquad q \end{array} & \end{array} \quad .$$

Under this model we consider

(21) $$H_o: \underset{p\times q}{\Sigma_{12}} = \underset{p\times q}{0} \quad \text{against} \quad H: \Sigma_{12} \neq 0 \quad .$$

Let us now set

(22) $$\underset{p\times q}{\beta} = \underset{p\times q}{\Sigma_{12}} \underset{q\times q}{\Sigma_{22}^{-1}}, \quad \underset{p\times q}{B} = \underset{p\times q}{S_{12}} \underset{q\times q}{S_{22}^{-1}} \quad \text{and}$$

$$\underset{p\times p}{S_{1.2}} = \underset{p\times p}{S_{11}} - \underset{p\times q}{S_{12}} \underset{q\times q}{S_{22}^{-1}} \underset{q\times p}{S_{12}'} \quad .$$

Notice that β and B are the population and the sample "regression" matrices. Then it has been explained in [10, 11] that $ch^{1/2}_{max}(\beta \beta')$ can be regarded as a measure of departure from H_o in the direction of H, and similarly the partials of β obtained by cutting out the rows and columns of β (i.e., cutting out from the p-set and the q-set) can be also regarded as measures of departures from the corresponding partial hypothesis in the direction of the appropriate (complementary) alternatives. As in section 3, starting from the one sided analogue of (3) with an exact probability $1 - \alpha$, we can put, with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, simultaneous bounds on $ch_{max}^{1/2} (\beta \beta')$ and the partials, the whole set, total and the partials, being expressed, as in section 3, in the compact form

$$(23) \quad ch_{max}^{1/2} \left[B_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})} \; B_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})'} \right]$$

$$- \lambda \, ch_{max}^{1/2} \left[S_{1.2(i_1, i_2, \ldots, i_{p_1})}^{(i_1, i_2, \ldots, i_{p_1})} \right] ch_{max}^{1/2} (S_{22}^{-1})$$

$$\leq ch_{max} \left[\beta_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})} \; \beta_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})'} \right]$$

$$\leq ch_{max}^{1/2} \left[B_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})} \; B_{(j_1, j_2, \ldots, j_{q_1})}^{(i_1, i_2, \ldots, i_{p_1})'} \right]$$

$$- \lambda \; ch_{max}^{1/2} \left[S_{1.2(i_1, i_2, \ldots, i_{p_1})}^{(i_1, i_2, \ldots, i_{p_1})} \right] ch_{max}^{1/2} \; (S_{22}^{-1}) \quad ,$$

where the symbols have the same general explanation as for (9). It will be instructive here to drop the superscripts and subscripts and write the leading interval (involving just the "total") in a much simpler form. However, certain additional remarks are called for. Here $i_1 < i_2 < \ldots < i_{p_1}$ $(0 \leq p_1 \leq p - 1)$ is any ordered subset from (1, 2, ..., p) and $j_1 < j_2 < \ldots < j_{q_1}$ $(p - p_1 \leq q - q_1)$ is an ordered subset from (1, 2, ..., q). Furthermore, in keeping with $p \leq q$, we also arrange that $p - p_1 \leq q - q_1$ which means that the total number of confidence statements is less than $(2^p - 1) \times (2^q - 1)$, and can be easily calculated. In conclusion, as in section 3, while (H_0, H) and the partials would seem to suggest just the lower bounds we still seek the upper bounds as well, i.e., the two-sided bounds on the principle of section 1.1.3. As in section 3 so also here, the bounds depend on both sets $(i_1, i_2, \ldots, i_{p_1})$ and $(j_1, j_2, \ldots, j_{q_1})$, but the width depends only on $(i_1, i_2, \ldots, i_{p_1})$. If we require the width also to depend on $(j_1, j_2, \ldots, j_{q_1})$ as well then we have to use a different procedure altogether which will not be discussed here. Bounds tighter than those in (12) but more or less on the same principles have been offered by other workers, but these will not be discussed here. A special case of (23) deserves particular mention.

5.1. <u>The case $p = 1$, i.e., of independence between one variate and a q-set.</u> Here $p = 1$ and $p_1 = 0$. Furthermore, (Σ_{11}, S_{11}) become scalars, (Σ_{12}, S_{12}), (1xq) row vectors, (Σ_{22}, S_{22}) stay (qxq) matrices, and, in consequence, β and B become two (1xq) row vectors while $S_{1.2}$ becomes scalar. The $\underset{1xq}{\beta}$ and $\underset{1xq}{B}$ are usually called the population and the sample regression vector of one variate on the q-set, while $S^2_{1.2}$ is called the residual variance. (23) now reduces to

$$(24) \quad \left[\underset{1\times q}{b}(j_1,j_2,\ldots,j_{q_1}) \underset{q\times 1}{b'}(j_1,j_2,\ldots,j_{q_1}) \right]^{1/2}$$

$$- \lambda \; s_{1.2}^{1/2} \; ch_{max}^{1/2} \; (S_{22}^{-1})$$

$$\leq \left[\underset{1\times q}{\beta'}(j_1,\ldots,j_{q_1}) \quad \underset{q\times 1}{\beta}(j_1,\ldots,j_{q_1}) \right]^{1/2}$$

$$\leq \left[\underset{1\times q}{b}(j_1,\ldots,j_{q_1}) \quad \underset{q\times 1}{b'}(j_1,\ldots,j_{q_1}) \right]^{1/2}$$

$$+ \quad \lambda \; s_{1.2}^{1/2} \; ch_{max}^{1/2} \; (S_{22}^{-1})$$

where λ is based on the percentage point of an F-distribution with appropriate d.f. The well known interval for the regression coefficient in a bivariate normal distribution ($p = q = 1$) is, of course, a special case of (24).

6. Multiple independence between a p-set or between sets of p_1, $p_2,\ldots,p_k$ [12]. We shall consider the two cases separately, although the first one is a special case of the second.

6.1. Multiple independence between a p-set. For an $N[\underset{p\times 1}{\zeta}, \underset{p\times p}{\Sigma}]$ we express H_o: $\Sigma = D_{\sigma_{ii}}$ (a diagonal matrix) as the intersection of the independence of the p-th variate with variates $(1,2,\ldots,p-1)$, of the $(p-1)$-th variate with variates $(1,2,\ldots,p-2)$ and so on till we get to variate 2 and its independence with 1. The associated alternatives are the respective complements. This is called a step-down procedure and the results that we ultimately obtain are not of course invariant under the order in which we take the variates. This is supposed to be a blemish on the technique, but I feel rather differently for two reasons. First, there are problems and situations when the experimenter has some order in mind in which he is more interested. The second, and by far the more important reason is the following. My own experience in the search for appropriate nonparametric analogues of MANOVA and association has more or less persuaded me that characterizing a nonparametric multivariate distribution as the product of a chain of conditional probabilities (in the sense of a step-down procedure) right from the beginning may, in most situations, be the only feasible line of attack. Be that as it may, let us come back to the concrete problem here [ ]. With

$$(26) \qquad \underset{p\times p}{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}, \quad \underset{p\times p}{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} & \cdots & s_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

let

$$(27) \qquad \underset{1\times(i-1)}{\beta'_{i.12\ldots i-1}} = [\sigma_{1i}\ \sigma_{2i} \cdots \sigma_{i-1,i}] \times \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,i-1} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{1,i-1} & \sigma_{2,i-1} & \cdots & \sigma_{i-1,i-1} \end{bmatrix}^{-1},$$

$$\underset{1\times(i-1)}{b'_{i.12\ldots i-1}} = [s_{1i} \quad s_{2i} \cdots s_{i-1,i}]$$

$$\times \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1,i-1} \\ s_{12} & s_{22} & \cdots & s_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ s_{1,i-1} & s_{2,i-1} & \cdots & s_{i-1,i-1} \end{bmatrix}^{-1}$$

(for $i = p, p-1, \ldots, 2$) denote respectively the population and the sample regression vectors of (i) on $(1,2,\ldots,i-1)$, and also let

(28) $$\underset{1\times(i-1)}{s'_i} = [s_{1i}, s_{2i}, \ldots, s_{i-1,i}],$$

$$\underset{(i-1)\times(i-1)}{S_{i-1}} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1,i-1} \\ s_{12} & s_{22} & \cdots & s_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ s_{1,i-1} & s_{2,i-1} & \cdots & s_{i-1,i-1} \end{bmatrix},$$

(for $i = p, p-1, \ldots, 2$).

Then, we have with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, the simultaneous set of intervals

(29) $$\left[b_i^{(j_1,j_2,\ldots,j_{i_1})'} \; b_i^{(j_1,j_2,\ldots,j_{i_1})}\right]^{1/2}$$

$$- \lambda \left[s_{ii} - s'_i \; S_{i-1}^{-1} \; s_i\right]^{1/2} \; ch_{max}^{1/2} \; (S_{i-1}^{-1})$$

$$\leq \left[\beta_i^{(j_1,\ldots,j_{i_1})'} \; \beta_i^{(j_1,\ldots,j_{i_1})}\right]$$

$$\leq \left[b_i^{(j_1,\ldots,j_{i_1})'} \; b_i^{(j_1,\ldots,j_{i_1})} \right]^{1/2}$$

$$+ \lambda \left[s_{ii} - s_i' \; S_{i-1}^{-1} \; s_i \right]^{1/2} \mathrm{ch}_{\max}^{1/2} \, (S_{i-1}^{-1}) \, ,$$

where $j_1 < j_2 < \ldots < j_{i_1}$ $(0 \leq i_1 \leq i - 2;\ i = p,\ p - 1, \ldots, 2)$ is any ordered subset from $(1,2,\ldots,i-1)$ and where $\lambda = + \left[\mu/(1-\mu) \right]^{1/2}$ and μ is obtained from the following equation

$$(30) \quad \prod_{i=2}^{p} P \left[r_i^2 \leq \mu \,|\, \rho = 0 \right] = 1 - \alpha \, ,$$

r_i^2 having an incomplete β-distribution with d.f. $(i-1)/2$ and $(n-i)/2$, for $i = p,\ p-1,\ \ldots,\ 2$ and n = sample size -2. It is to be noticed that the total number of confidence intervals is $\sum_{i=2}^{p} (2^{i-1} - 1)$. Also, for any i, the bounds depend i and on the set $(j_1, j_2, \ldots, j_{i_1})$, but the width depends only on i. The leading interval in the set (29) is obtained by putting $i_1 = 0$, whence we have

$$(31) \quad (b_i' \; b_i)^{1/2} - \lambda \left[s_{ii} - s_i' \; S_{i-1}^{-1} \; s_i \right]^{1/2} \mathrm{ch}_{\max}^{1/2} \, (S_{i-1}^{-1})$$

$$\leq (\beta_i' \; \beta_i)^{1/2} \leq (b_i' \; b_i)^{1/2} + \lambda \left[s_{ii} - s_i' \; S_{i-1}^{-1} s_i \right]^{1/2}$$

$$\times \mathrm{ch}_{\max}^{1/2} \, (S_{i-1}^{-1})$$

6.2. <u>Multiple independence between sets of $p_1, p_2, \ldots, p_k$.</u> We have here

$$\mathcal{N}\left[\begin{array}{cc} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \cdot \\ \zeta_k \end{bmatrix} & \begin{array}{c} p_1 \\ p_2 \\ \\ p_k \end{array} \\ 1 & \end{array} , \begin{array}{cc} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma'_{12} & \Sigma_{22} & \cdots & \Sigma_{2k} \\ \cdot & \cdot & \cdots & \cdot \\ \Sigma'_{1k} & \Sigma'_{2k} & \cdots & \Sigma_{kk} \end{bmatrix} & \begin{array}{c} p_1 \\ p_2 \\ \\ p_k \end{array} \\ \begin{array}{cccc} p_1 & p_2 & & p_k \end{array} & \end{array}\right] ,$$

and we assume, without any loss of generality that $p_i \leq \sum_{j=1}^{i-1} p_j$.

The null hypothesis that the population dispersion matrix has all the off diagonal submatrices zero is expressed as the intersection of the independence of the i-set with sets $(1,2,\ldots,i-1)$ (for $i = k, k-1, \ldots, 2$), with the associative alternatives as the respective complements. With a corresponding partition of the total sample dispersion matrix, let

$$(33) \qquad \underset{p_i \times (p_{i-1}+\cdots+p_1)}{\beta_i} = \left[\Sigma'_{1i} \;\; \Sigma'_{2i} \;\cdots\; \Sigma'_{i-1,i}\right] \times \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1,i-1} \\ \Sigma_{12} & \Sigma_{22} & \cdots & \Sigma_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ \Sigma'_{1,i-1} & \Sigma'_{2,i-1} & \cdots & \Sigma_{i-1,i-1} \end{bmatrix}^{-1} ,$$

$$\underset{p_i \times (p_{i-1}+\cdots+p_1)}{B_i} = [S_{1i}' \quad S_{2i}' \cdots S_{i-1,i}']$$

$$\times \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1,i-1} \\ S_{12}' & S_{22} & \cdots & S_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ S_{1,i-1}' & S_{2,i-1}' & \cdots & S_{i-1,i-1} \end{bmatrix}^{-1},$$

and

$$(34) \qquad \underset{p_i \times (p_{i-1}+\cdots+p_1)}{S^{(i)'}} = [\underset{p_1}{S_{1i}'} \quad \underset{p_2}{S_{2i}'} \quad \cdots \quad \underset{p_{i-1}}{S_{i-1,i}'}]p_i \;,$$

$$\underset{(p_1+\cdots+p_{i-1}) \times (p_1+\cdots+p_{i-1})}{S_{i-1}} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1,i-1} \\ S_{12}' & S_{22} & \cdots & S_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ S_{1,i-1}' & S_{2,i-1}' & \cdots & S_{i-1,i-1} \end{bmatrix}.$$

Then, wtth a joint (conservative) confidence coefficient $\geq 1 - \alpha$, we have the sets of simultaneous confidence intervals, of which the leading set is

$$(35) \quad ch_{max}^{1/2} [B_i \;\; B_i'] - \lambda \;\; ch_{max}^{1/2} [S_{ii} - S^{(i)'} S_{i-1}^{-1} S^{(i)}] \; ch_{max}^{1/2}(S_{i-1}^{-1})$$

$$\leq ch_{max}^{1/2} [\beta_i \;\; \beta_i']$$

$$\leq ch_{max}^{1/2} [B_i B_i'] + \lambda\, ch_{max}^{1/2} [S_{ii} - S^{(i)'} S_{i-1}^{-1} S^{(i)}]\, ch_{max}^{1/2}(S_{i-1}^{-1})$$

$$(i=k,\ k-1,\ \ldots,\ 2),$$

and the whole set of sets can be expressed in a compact form by replacing

$$(36)\quad \beta_i \text{ by } \beta_{i(\ell_1,\ell_2,\ldots,\ell_{v_i})}^{(j_1,j_2,\ldots,j_{u_i})},\quad B_i \text{ by } B_{i(\ell_1,\ell_2,\ldots,\ell_{v_i})}^{(j_1,j_2,\ldots,j_{u_i})},$$
$$S_{ii} \text{ by } S_{ii(j_1,j_2,\ldots,j_{u_i})}^{(j_1,j_2,\ldots\, j_{u_i})} \quad\text{and}\quad S^{(i)} \text{ by } S^{(i)'(j_1,j_2,\ldots,j_{u_i})},$$

where $j_1 < j_2 < \ldots < j_{u_i} (0 \leq u_i \leq p_i - 1)$ is any ordered subset from the set $(1, 2, \ldots, p_i)$ and $\ell_1 < \ell_2 < \ldots < \ell_{v_i} (0 \leq v_i \leq p_1+p_2+ \ldots p_{i-1}-p_i+u_i$ is any ordered subset from the set $(1, 2, \ldots, p_1+p_2 + \ldots + p_{i-1})$. $\lambda = +[\mu/(1-\mu)]^{1/2}$ and μ is obtained from the equation

$$(37)\quad \prod_{i=2}^{k} P[ch_{max}(S_{ii}^{-1}\ S^{(i)'}\ S_{i-1}^{-1}\ S^{(i)}) \leq \mu \mid \text{null hypothesis}]$$
$$= 1 - \alpha .$$

7. The problem of vector analogue of the ratio of the two variances and two means for a bivariate normal distribution [13]. We start from the same model as in subsection 4.3, with two comparable p-vectors following a 2p-variate normal distribution. We shall consider the two cases separately [13].

7.1. The vector analogue of the ratio of two variances. To fix our ideas we shall first consider the case of the ratio of two variances σ_1^2/σ_2^2 for a (correlated) bivariate normal distribution. It is easy to see that if x_1 and x_2 has a bivariate normal distribution, then $x_1 - (\sigma_1/\sigma_2)\, x_2$ and $x_1 + (\sigma_1/\sigma_2)\, x_2$ are uncorrelated in the population (and hence independent in the case of a normal distribution). It is possible by utilizing this to use the t-table and obtain for σ_1^2/σ_2^2, with a confidence coefficient $1 - \alpha$, the confidence interval

$$(38) \quad \frac{s_1^2}{s_2^2}\left(1 + \frac{2}{n-2}\, t_{\alpha/2}^2\,(1-r^2)\right) - \left[\left(1 + \frac{2}{n-2}\, t_{\alpha/2}^2(1-r^2)\right)^2 - 1\right]^{1/2} \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$\le \frac{s_1^2}{s_2^2}\left(1 + \frac{2}{n-2}\, t_{\alpha/2}^2\,(1-r^2)\right) + \left[\left(1 + \frac{2}{n-2}\, t_{\alpha/2}^2(1-r^2)\right)^2 - 1\right]^{1/2},$$

where n is the sample size and $t_{\alpha/2}$ is the upper $100\,\alpha/2$ °/o point of the t-distribution. We notice that $\sigma_1/\sigma_2 = 1$ if and only if the customary null hypothesis is true.

For two vectors $(\underset{p\times 1}{x_1}, \underset{p\times 1}{x_2})$ having the 2p-variate normal distribution it can be checked that the set $(x_1 - \lambda\, x_2)$ and the set $(x_1 + \lambda\, x_2)$ are uncorrelated in the population (and hence also independent because of the normal distribution), where $\lambda = \mu\, \nu^{-1}$ and μ and ν are given by

$$(39) \quad \Sigma_{11} = \mu\, \mu', \quad \Sigma_{22} = \nu\, \nu', \quad \Sigma_{12} = \mu\, D_\gamma\, \nu',$$

where D_γ is a diagonal matrix with diagonal elements $\gamma_1, \gamma_2, \ldots, \gamma_p$ and the γ's are the roots of $\Sigma_{11}^{-1}\, \Sigma_{12}\, \Sigma_{22}^{-1}\, \Sigma_{12}'$. By utilizing this fact and using the distribution of the square of the largest canonical correlation coefficient it is possible to obtain confidence bounds on $ch^{1/2}_{max}(\lambda\lambda')$ and $ch^{1/2}_{min}(\lambda\lambda')$ and on the "partials". Before we go on to that, some remarks are in order as to the physical role of the matrix λ. For $p = 2$, i.e., for the bivariate case λ becomes σ_1/σ_2. Furthermore, it can be easily shown that $\lambda\lambda' = I$, i.e., λ is an orthogonal matrix, if and only if Σ_{11} and Σ_{22} are similar matrices, i.e., $\Sigma_{11} = \lambda\Sigma_{22}\lambda'$, where λ is an orthogonal matrix and $\Sigma_{12} = \lambda\Sigma'_{12}\lambda$ (with the same orthogonal λ), or equivalently $\Sigma_{12} = \lambda$ x a symmetric matrix. This orthogonal $\lambda = I$, if and only if $\Sigma_{11} = \Sigma_{22}$ and $\Sigma_{12} = \Sigma'_{12}$. With a joint (conservative) confidence coefficient $\geq 1 - \alpha$, a set of simultaneous confidence interval pairs (including the "partials") of which the leading one (involving the "total") is

$$(40)\quad \frac{1}{e_\alpha}\, ch_{min}(S_{11} - S_{12}S_{22}^{-1}S'_{12})/ch_{max}(S_{22}) \leq ch_{min}(\lambda\lambda')$$

$$\leq e_\alpha \min\left[ch_{min}(S_{11})/ch_{min}(S_{22} - S'_{12}S_{11}^{-1}S_{12}),\right.$$

$$\left. ch_{max}(S_{11})/ch_{max}(S_{22} - S'_{12}S_{11}^{-1}S_{12})\right],$$

$$\frac{1}{e_\alpha}\max\left[ch_{min}(S_{11} - S_{12}S_{22}^{-1}S'_{12})/ch_{min}(S_{22}),\right.$$

$$\left. ch_{max}(S_{11} - S_{12}S_{22}^{-1}S'_{12})/ch_{max}(S_{22})\right]$$

$$\leq ch_{max}(\lambda\lambda') \leq e_\alpha ch_{max}(S_{11})/ch_{min}(S_{22} - S'_{12}S_{11}^{-1}S_{12}),$$

where $e_\alpha = 4/(1-c_\alpha)$, and c_α is the upper $100\,\alpha$ % point of the distribution of the square largest canonical correlation coefficient with d.f. $(p, p, n-1)$. The partials obtained by cutting out one or more variate pairs are not displayed; the form of such partials is by now clear. (38) of course follows as a very special case of (40).

7.2. <u>The vector analogue of the ratio of two means</u>. Let $\underset{1xp}{\zeta_1'} = (\zeta_{1i})$ and $\zeta_2' = (\zeta_{2i})$ $(i = 1,2,\ldots,p)$ and $q_i = \zeta_{1i}/\zeta_{2i}$. Also let D_q stand for a diagonal matrix with diagonal elements $q_1, q_2, \ldots, q_p$. Then $x_1 - D_q x_2$ is distributed as a p-variate normal with zero mean vector and $\Sigma^* = \Sigma_{11} - D_q \Sigma_{12}' - \Sigma_{12} D_q + D_q \Sigma_{22} D_q$. We utilize this and use the distribution of Hotelling's T^2 to obtain, with a joint (conservative) confidence coefficient $\geq 1 - \alpha$, a set of simultaneous confidence intervals (including the "partials") of which the leading interval (involving the "total") is

$$(41) \quad \frac{1}{n} T_\alpha^2 \, ch_{max}(S) \max \left[1 + q_1^2, 1 + q_2^2, \ldots, 1 + q_p^2 \right] + 2 \sum_{i=1}^{p} |\bar{x}_{1i} \bar{x}_{2i}| \max \left[q_1^2, q_2^2, \ldots, q_p^2 \right]^{1/2} - \bar{x}_1' \bar{x}_1 - \bar{x}_2' \bar{x}_2 \min (q_1^2, \ldots, q_p^2) \geq 0;$$

where T_α^2 is of course the 100α °/o point of Hotelling's T^2 with appropriate d.f., S is the total $2p \times 2p$ sample dispersion matrix (partitioned in the same manner as the total population matrix), $\underset{1xp}{x_1'} (= x_{1i})$, $\underset{1xp}{x_2'} (= x_{2i})$ $(i = 1, 2, \ldots, p)$ are the sample mean vectors of the two p-sets. The partials (whose forms are by now obvious) are not separately displayed. For $p = 1$, i.e., for the bivariate case, this reduces, with an exact coefficient $1 - \alpha$, to

$$(42) \quad (\bar{x}_2^2 - k^2 s_2^2)\, q^2 - 2(\bar{x}_1 \bar{x}_2 - k\, s_1 s_2\, r)q + (\bar{x}_1^2 - k\ \ s_1^2) \leq 0,$$

where $\bar{x}_1$, $\bar{x}_2$, s_1^2, s_2^2, r are the sample means, variances and correlation coefficient and $k = t^2_{\alpha/2}/n$, $t_{\alpha/2}$ is the upper $100\ \alpha/2$ % point of the t-distribution with proper d.f. and n is the sample size.

8. Concluding remarks.

It has been already stated that we are excluding the problems, treated elsewhere, involving model II and mixed models of MANOVA. We are also excluding problems involving other patterns on Σ [1], including, in particular, the problem of factor analysis and also problems involving different Σ's for more than two populations [3]. Nor are we including some other interesting results, somewhat on the same lines as the ones discussed here, that have been published in recent years by workers from groups other than the one I have worked with.

REFERENCES

[1] Bargmann, R., "A study of independence and dependence in multivariate normal analysis," Institute of Statistics, University of North Carolina Miemograph Series No. 186.

[2] Bhapkar, V., "Confidence bounds connected with ANOVA and MANOVA for balanced and partially balanced incomplete block designs," Ann. Math. Stat., Vol. 3 (1960) pp. 741-748.

[3] Gnanadesikan, R., "Equality of more than two variances and of more than two dispersion matrices against certain alternatives," Ann. Math. Stat., Vol. 30 (1959) pp. 177-184.

[4] Heck, David, "Charts of some upper percentage points of the distribution of the largest characteristic root," Ann. Math. Stat., Vol. 31 (1960) pp. 625-642.

[5] Roy, J., "Step-down procedures in multivariate analysis," Ann. Math. Stat., Vol. 29 (1958) pp. 1177-1187.

[6] Roy, S. N., "On a heuristic method of test construction and its use in multivariate analysis," Ann. Math. Stat., Vol. 20 (1953)pp.220-238.

[7] Roy, S. N. and Bose, R. C., "Simultaneous confidence interval estimation," Ann. Math. Stat., Vol. 24 (1953) pp. 513-536.

[8] Roy, S. N., "Some further results in simultaneous confidence interval estimation," Ann. Math. Stat., Vol. 25 (1954) pp. 752-761.

[9] Roy, S. N., "A note on some further results in simultaneous confidence interval estimation," Ann. Math. Stat., Vol. 27 (1956) pp.856-858.

[10] Roy, S. N., Some Aspects of Multivariate Analysis, New York: John Wiley and Sons, Inc., 1957.

[11] Roy, S. N. and Gnanadesikan, R., "Further contributions to multivariate confidence bounds," Biometrika, Vol. XLIV (1957) pp.399-410.

[12] Roy, S. N. and Bargmann, R., "Tests of multiple independence and the associated confidence bounds," Ann. Math. Stat., Vol. 29 (1958) pp. 491-503.

[13] Roy, S. N. and Potthoff, R. F., "Confidence bounds on vector analogues of the 'ratio of means' and 'ratio of variances' for two correlated normal variates and some associated tests," Ann. Math. Stat., Vol. 29 (1958) pp. 829-841.

[14] Roy, S. N. and Roy, J. , "Analysis of variance with univariate or multivariate, fixed or mixed classical models," Institute of Statistics, University of North Carolina Mimeograph Series No. 208 (1952).

[15] Roy, S. N. and Gnanadesikan, R., "Some contributions to ANOVA in one or more dimensions," Ann. Math. Stat., Vol. 30 (1959) pp. 304-340.

[16] Roy, S. N. and Cobb, Whitfield, "Mixed model variance analysis with normal error and possibly non-normal other effects. Univariate and multivariate cases," Ann. Math. Stat., Vol. 31 (1960) pp.939-968.

[17] Roy, S. N. and Srivastava, J. N., "Inference on treatment effects and design of experiments in relation to such inference," Institute of Statistics, University of North Carolina Mimeograph Series No.274 (1961).

[18] Roy, S. N. and Gnanadesikan, R., "Equality of two dispersion matrices against alternatives of intermediate specificity," Institute of Statistics, University of North Carolina Mimeograph Series No.282 (1961).

[19] Roy, S. N. and Roy, J., "On testability in normal ANOVA and MANOVA with all 'fixed effects'," Calcutta Mathematical Society Golden Jubilee Commemoration Volume (1958-59) (issued 1961).