The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 224–250

SIMULTANEOUS TEST PROCEDURES—SOME THEORY OF MULTIPLE COMPARISONS[1]

By K. R. Gabriel

University of North Carolina and Hebrew University, Jerusalem

1. Introduction and summary. When a hypothesis is tested by a significance test and is not rejected, it is generally agreed that all hypotheses implied by that hypothesis (its "components") must also be considered as non-rejected. However, when the hypothesis is rejected the question remains as to which components may also be rejected. Various writers have given attention to this question and have proposed a variety of multiple comparisons methods based either on tests of each one of the components or on simultaneous confidence bounds on parametric functions related to the various hypotheses.

An approach to such methods, apparently originally due to Tukey [27], is to test each component hypothesis by comparing its statistic with the α level critical value of the statistic for the overall hypothesis. This is called a Simultaneous Test Procedure (STP for short) as all hypotheses may be tested simultaneously and without reference to one another. An STP involves no stepwise testing of the kind employed by some other methods of multiple comparisons for means, in which subsets are tested for equality only if they are contained in sets which have already been found significant. (See [3], [4], [10], [18]).

A general formalization of STP's is attempted in this paper. Section 2 introduces the requisite concepts of families of hypotheses and the implication relations between them, as well as the monotonicity relations between the related statistics. Section 3 defines STP's and shows conditions for coherence and consonance of their decisions, these properties being that hypothesis implication relations are preserved in the decisions of the STP. Section 4 discusses comparison of various STP's for the same hypotheses and shows the advantages of the union-intersection type of statistics and of reducing the family of hypotheses tested as much as possible. Section 5 translates all these results to simultaneous confidence statements after introducing the definitions necessary to allow such translation. The analogy between simultaneous test and confidence methods is of special importance as it brings a wide spectrum of methods within this framework, most of which was originally formulated in confidence region terms. This covers the original work of Tukey [27] and Scheffé [25] and continues with that of Roy and his associates [21] and most recently Krishnaiah [12], [13]. A general discussion of this confidence approach has been given by Aitchison [1] since the first draft of the present paper. In view of the close analogies pointed out in Section 5, it is

Received 2 October 1963, revised 2 August 1967.

[1] The research was supported in part by the National Institutes of Health (Institute of General Medical Sciences) Grant No. GM-12868-03, and in part by National Center for Health Statistics Grant No. NCHS-IS-1.

not surprising that Aitchison arrives at the requirement of a constant critical value for all his tests, exactly as in an STP. In fact his approach and the present one are complementary.

The formal theory is illustrated with ANOVA examples to clarify the concepts. No essentially new techniques are presented in this paper though this approach has been used elsewhere by the author to derive a number of practically useful procedures [6], [7], [8].

2. Families of hypotheses and statistics. Let Y be a (vector) random variable having density function $f_\theta(Y)$ with respect to a σ-finite measure μ, where the parameter(s) are assumed to lie in a set ω. An *overall* hypothesis $\omega_0 (\subset \omega)$ is being considered, as well as a *family* $\Omega = \{\omega_i \mid i \varepsilon I\}$ of hypotheses ω_i implied by ω_0, i.e., $\omega_0 \subseteq \omega_i \subset \omega$ if $i \varepsilon I$, I being an index set, not necessarily denumerable. (Note that ω_i is being used to denote both a set in parameter space ω and the hypothesis $\theta \varepsilon \omega_i$ restricting the parameters to it). The hypothesis ω_0 itself is assumed to belong to the family and to form the *intersection* of its members, that is, $\omega_0 \varepsilon \Omega$ and $\omega_0 = \bigcap_I \omega_i$.

The implication relation $\omega_i \subseteq \omega_j$ for a pair of hypotheses from Ω will be referred to by saying ω_j is a *component* of ω_i. If the relation is strict, i.e., $\omega_i \subset \omega_j$, ω_j will be said to be a *proper component* of ω_i. Any such strict implication relation $\omega_i \subset \omega_j$, where $i \varepsilon I$ and $j \varepsilon I$, may be written $i \prec j$ and the set of all such relations existing between pairs of indices from I will be denoted $\mathcal{J}$. Note that $i \prec j$ is a transitive relation, but is neither reflexive nor symmetric and that $0 \prec i$ for all $i \varepsilon I - \{0\}$. Also note that this is not the only type of implication relation that may exist between members of a family of hypotheses. If, for example, μ_1, μ_2 and μ_3 are parameters one might consider the intersection hypothesis $\omega_0 : \mu_1 = \mu_2 = \mu_3$ and let the family further contain hypotheses $\omega_1 : \mu_1 = \mu_2$, $\omega_2 : \mu_1 = \mu_3$ and $\omega_3 : \mu_2 = \mu_3$. In this example the relation set $\mathcal{J}$ contains only the relations $0 \prec 1$ $0 \prec 2$, $0 \prec 3$, whereas implications $(\omega_1 \cap \omega_2 \cap \omega_3) \subset \omega_0$ or $(\omega_1 \cap \omega_2) \subset \omega_3$ are not included in $\mathcal{J}$.

A partial ordering of the hypotheses of Ω may be obtained by always assigning a higher *rank* to a hypothesis ω_i than to any of its proper components ω_j, $i \prec j$. This will assign maximum rank to the intersection hypothesis ω_0 and minimum rank to hypotheses which have no proper components in Ω. The latter hypotheses are referred to as *minimal* and their subfamily denoted $\Omega_{\min}$ and indexed by $I_{\min}$ so that $\Omega_{\min} = \{\omega_j \mid j \varepsilon I_{\min}\}$. All other hypotheses are referred to as *non-minimal*, i.e., $\Omega - \Omega_{\min} = \{\omega_i \mid i \varepsilon I - I_{\min}\}$.

Corresponding to each hypothesis ω_i of Ω let there be a real valued statistic $Z_i = Z_i(Y)$ and write the family of these statistics $\mathbf{Z} = \{Z_i \mid i \varepsilon I\}$. The collection $\{\Omega, \mathbf{Z}\}$ of hypotheses and their corresponding statistics will be called a *testing family* provided the distribution of Z_i, for every $i \varepsilon I$, is completely specified under ω_i, i.e., the same for all $\theta \varepsilon \omega_i$. If the family Ω is closed under intersection the testing family will be called *closed*. If, for any subfamily $\tilde{\Omega} = \{\omega_i \mid i \varepsilon \tilde{I}\}$, where $\tilde{I} \subseteq I$, the joint distribution of all Z_i, $i \varepsilon \tilde{I}$, is completely specified under

$\tilde{\omega}_0 = \bigcap_{\bar{I}} \omega_i$, the testing family will be called *joint.* These two categories of testing families are neither exclusive nor exhausitive.

EXAMPLE 2.1. Consider a one-way ANOVA with ω_0 as the hypothesis of equality of all $k(>3)$ means and ω_i the hypothesis of equality of a pair of means indexed by $I^P - \{0\}$ indexing all $\binom{k}{2}$ such pairs. Note that these $\binom{k}{2}$ ω_i's are minimal hypotheses, only ω_0 being non-minimal. This family is not closed under intersection for it does not contain hypotheses on equality of subsets of h means if $2 < h < k$. However, with F-ratio statistics (with a common overall estimate of error variance) being used for each hypothesis, a joint testing family $\{\Omega^P, \mathbf{Z}^F\}$ results. For the joint distribution of all the pairwise F-ratios is completely specified for any subset of $h(2 \leqq h \leqq k)$ *equal* means, irrespective of the values of the other $k - h$ means. The same obviously also holds for *augmented* F-ratios, i.e., F-ratios multiplied by their numerator df's.

EXAMPLE 2.2 In the same set-up let ω_i be a hypothesis of equality for some subset S_i of k_i means, I^S indexing all $2^k - k - 2$ such subsets. The minimal subfamily remains as in Example 2.1. If (augmented) F-ratio statistics are used for each of these hypotheses, any one of them corresponding to a true hypothesis will have a central (augmented) F distribution irrespective of any means not involved in that hypothesis. Thus one obtains a closed testing family $\{\Omega^S, \mathbf{Z}^F\}$.

EXAMPLE 2.3. Again, in the same set-up, one may start by ranking the observations of all samples together (overall ranking) and for each hypothesis use the "between" sum of squares of these ranks as a non-parametric statistic. The overall sum of squares is, apart from a constant, the Kruskal-Wallis statistic and its distribution is specified under ω_0. But for any subset of $h(<k)$ samples, the corresponding statistic depends not only on the equality of the h means involved but also on where the other $k - h$ means are relative to those h. This is evident when one considers that the ranks available for that subset of h samples depend on what ranks have been assigned to the other $k - h$ samples. Thus, such a collection of hypotheses and rank statistics does not form a testing family.

EXAMPLE 2.4. Exactly the same argument holds if instead of "between" sums of squares of ranks one uses any other statistics based on the overall ranking, such as the range of rank means or the maximum rank mean. These types of statistics do not provide testing families. (The use of such statistics and their distributions have been discussed by Nemenyi [17]).

EXAMPLE 2.5. An alternative non-parametric approach to the set-up of Example 2.2 is to use a separate ranking for the statistic of each hypothesis, including only the samples relevant to that hypothesis. One may then use the Kruskal-Wallis statistic for each hypothesis, and as its distribution is specified under the hypothesis a closed testing family is obtained.

Since for each ω_i of Ω more than one statistic can be chosen to satisfy these conditions, more than one testing family can be defined for any given family of hypotheses. Consideration of the properties of different Z_i's as test statistics for their respective hypotheses ω_i will play a part in choosing a $\mathbf{Z}$ appropriate to Ω.

In the present discussion the choice of $\mathbf{Z}$ corresponding to any given Ω is severely restricted by the following requirement of monotonicity.

A testing family $\{\Omega, \mathbf{Z}\}$ will be called *monotone* if, whenever $i \prec j$, the numerical relation

$$(2.1) \qquad Z_i(y) \geqq Z_j(y)$$

holds a.e. (almost everywhere with respect to measure μ), between the corresponding statistics. If, furthermore, for any non-minimal ω_i, i.e., $i \varepsilon I - I_{\min}$, and for any sample point y, there exists a proper component ω_j of ω_i, i.e., $i \prec j$, for which (2.1) is an equality, that is, if

$$(2.2) \qquad Z_i(y) = \max \{Z_j(y) \mid i \prec j\} \quad \text{a.e.,}$$

$\{\Omega, \mathbf{Z}\}$ will be called *strictly monotone*. Note that the actual j for which $Z_j(y) = Z_i(y)$ need not be the same for each y. Clearly, if a testing class is strictly monotone it is also monotone. Also, it is readily seen that if, and only if, (2.2) holds for every $i \varepsilon I - I_{\min}$, then, for every $i \varepsilon I - I_{\min}$

$$(2.3) \qquad Z_i(y) = \max \{Z_j(y) \mid i \prec j, j \varepsilon I_{\min}\} \quad \text{a.e.}$$

Hence (2.3) is an alternative definition of strict monotonicity. It is also clear that in a monotone testing family the statistic for ω_0, the intersection hypothesis, is

$$(2.4) \qquad Z_0(Y) = \max \{Z_i(Y) \mid i \varepsilon I\},$$

since Ω has been defined to include ω_0 itself.

EXAMPLE 2.6. In the testing family $\{\Omega^S, \mathbf{Z}^F\}$ of Example 2.2 implication relations $\prec$ exist between hypotheses on sets of means and hypotheses on proper subsets of these sets. Monotonicity holds since the "between" sum of squares, and hence the augmented F-ratio, for any set is no less than that for any of its subsets [6].

EXAMPLE 2.7. If in the subset ANOVA set-up of Examples 2.2 and 2.6, all sample sizes were equal, studentized range statistics Z^R might be used instead of (augmented) F-ratios. In that case monotonicity would be strict, for the range of all means always equals the range of some subset of the means.

EXAMPLE 2.8. In the testing family of Example 2.5 it may happen that the Kruskal-Wallis statistic for a subset exceeds that for a set containing that subset. For example, let there be three samples of four observations with the following ranks: sample 1—8, 9, 10, 11; sample 2—1, 2, 6, 7; sample 3—3, 4, 5, 12. The K-W statistic for samples 1 and 2 is 5.33, exceeding that for samples 1, 2 and 3 which is 4.77. Hence that testing family is not monotone.

EXAMPLE 2.9. In a contingency table one may consider the family of independence hypotheses for the entire table as well as for all subtables obtained by omitting or amalgamating rows and/or columns. If Pearson's chi-square statistic is used to test each hypothesis it may happen that the statistic for a table is less than that for a sub-table, though independence in the former implies independence in the latter [7]. Hence such a testing family would not be monotone.

Monotonicity holds for all testing families using likelihood ratio or union-intersection statistics, as is shown next.

LEMMA A. *If the statistics of a testing family are all likelihood ratio (LR for short) the family is monotone.*[2]

PROOF. Let the family be $\{\Omega, \mathbf{Z}\}$ and ω_i, ω_j hypotheses of Ω such that $i < j$. The LR statistic for ω_i is

$$Z_i = d(\sup_{\omega_i} f/\sup_{\omega} f)$$

where d is a monotone decreasing function, and the LR statistic Z_j for ω_j is similarly defined.

Now $i < j$ is equivalent to $\omega_i \subset \omega_j$ so that $\sup_{\omega_i} f \leqq \sup_{\omega_j} f$. Thus the ratio argument of d in Z_i is no greater than that of d in Z_j. As d is decreasing, $Z_i \geqq Z_j$, as was to be proved.

EXAMPLE 2.10. In the ANOVA testing family $\{\Omega^S, \mathbf{Z}^F\}$ Example 2.2 the F-ratios are LR statistics. Lemma A confirms the monotonicity noted in Example 2.6.

EXAMPLE 2.11. In the contingency table independence testing set-up described in Example 2.9 a monotone testing family may be obtained by using the log likelihood ratio statistics instead of Pearson's chi-squares [7].

LEMMA B. *A testing family is strictly monotone if and only if its statistics are related by Roy's Union-Intersection* (UI *for short*) *principle* [20].

PROOF. Roy used the UI principle to construct a test for ω_i from tests of all its minimal components ω_j, $i < j$, $j \varepsilon I_{\min}$. If the latter are tested with critical regions $(y \mid Z_j > \zeta)$, Roy [9], [20] defined the critical region for ω_i, $i \varepsilon I - I_{\min}$, as $\bigcup_{\{j \mid i < j, j \varepsilon I_{\min}\}} (y \mid Z_j > \zeta)$, or equivalently as

$$(y \mid \max_j \{Z_j \mid i < j, j \varepsilon I_{\min}\} > \zeta).$$

Extending this definition to all possible values of ζ, Roy essentially defined the UI statistic for ω_i as $\max_j \{Z_j \mid i < j, j \varepsilon I_{\min}\}$. Use of this definition for every $i \varepsilon I$ is equivalent to using the statistics (2.3) which satisfy strict monotonicity, as was to be proved.

EXAMPLE 2.12. For the ANOVA set-up of (2.7) one notes that the studentized range Z^R for any subset of means is the maximum of the Student t statistics times $2^{\frac{1}{2}}$, i.e., studentized ranges, for all pairwise comparisons of means from that subset. Thus, the studentized range statistics generate UI relations for the testing family $\{\Omega^S, \mathbf{Z}^R\}$ as was expected by Lemma B from the fact that this family is strictly monotone.

Since the UI relation is one of hypotheses *and* statistics it is preferable not to speak of UI *statistics*, but of UI *related testing families*.

3. Simultaneous test procedures. For a testing family $\{\Omega, \mathbf{Z}\}$ and a *critical*

[2] Essentially equivalent lemmas have been proved before by Aitchison [2] and Knight [11].

value ζ a *Simultaneous Test Procedure* (STP) is defined as the family of tests of all $\omega_i \varepsilon \Omega$ which reject any ω_i, $i \varepsilon I$, if $Z_i > \zeta$, using the same constant ζ for all Z_i of $\mathbf{Z}$. Such an STP will be denoted $\{\Omega, \mathbf{Z}, \zeta\}$. The probability

$$\alpha = P_{\omega_0}(Z_0 > \zeta) \tag{3.1}$$

of falsely rejecting the intersection hypothesis is referred to as the *level* of the STP. If Z_0 is a continuous variable the same STP may alternatively be defined in terms of its testing family $\{\Omega, \mathbf{Z}\}$ and level α, the critical value ζ being then determined by (3.1).

The term *simultaneous* is used to indicate that no order or sequence is imposed on the tests of all the hypotheses of Ω and that they are made without reference to one another. It is shown below that this cannot lead to incoherent decisions.

Note that ω_0 is rejected by the STP if and only if $Z_0 > \zeta$, exactly as by a significance test of significance level α. This test of the intersection hypothesis is therefore always part of the STP. If the testing family contains only a single hypothesis ω_0 and its statistic Z_0, the STP $\{\Omega, \mathbf{Z}, \zeta\}$ is equivalent to the test of hypothesis ω_0 by means of statistic Z_0 at critical value ζ.

An essential requirement that any procedure must satisfy is that no hypothesis should be "accepted" if any hypothesis implied by it is rejected. In other words, if $i < j$ then ω_i must always be rejected if ω_j is. This requirement will be called *coherence*,[3] and an STP will be called coherent if its decisions always satisfy this.

THEOREM 1. *All* STP*'s based on testing family* $\{\Omega, \mathbf{Z}\}$ *are coherent if and only if* $\{\Omega, \mathbf{Z}\}$ *is monotone.*

PROOF. Consider any hypotheses ω_i and ω_j of Ω for which $i < j$. Monotonicity of $\{\Omega, \mathbf{Z}\}$ postulates that $Z_i \geqq Z_j$ a.e., whereas coherence of $\{\Omega, \mathbf{Z}, \zeta\}$ for any given ζ requires that $(y \mid Z_i > \zeta) \supseteq (y \mid Z_j > \zeta)$ a.e.[4] since these events determine the rejection of ω_i and ω_j, respectively

Now, the statements "$Z_i \geqq Z_j$ a.e." and "$(y \mid Z_i > \zeta) \supseteq (y \mid Z_j > \zeta)$ a.e. for all ζ" are clearly equivalent. Hence monotonicity of testing family is equivalent to coherence of all STP's based on that family, as was to be proved.

COROLLARY 1. *If the statistics of a testing family are all* LR, *all* STP*'s based on that family are coherent.*

PROOF. This follows from Lemma A and Theorem 1.

EXAMPLE 3.1. In the subset ANOVA set-up $\{\Omega^S, \mathbf{Z}^F\}$ of Examples 2.2, 2.6, 2.10 one may test equality of the h means of each subset by means of the (augmented) F-ratio against ζ^F, the upper α point of the (augmented) F distribution with $k - 1$ and n_e degrees of freedom, where n_e is the d.f. for the variance estimate. This will give an STP with coherent decisions; set S_i of k_i means will be rejected if any subset of its means is rejected [6].

---

[3] In an earlier paper [6] this properly was called transitivity. The terminology has been changed because of the different connotations of the earlier term.

[4] For set containment relations $A \subset B$ and $A \subseteq B$ the a.e. qualifications mean that $A \cap \bar{B}$ may be non-empty but of measure zero. Similarly, $A = B$ a.e. mean that both $A \cap \bar{B}$ and $\bar{A} \cap B$ may be non-empty but are of measure zero.

EXAMPLE 3.2. The non-parametric analogue of the above, using the Kruskal-Wallis statistics of Example 2.5, is not an STP since its testing family is not monotone. As was pointed out in Example 2.8, the Kruskal-Wallis statistic for a set may be less than the same statistic for a subset contained in the set. If the common critical value happens to be between these two statistics' values the subset would be rejected whilst the set was accepted. This would lead to the incoherence of asserting that all means of a set are equal whilst some of them differ from each other. Hence, such a procedure would be incoherent. (See further remarks in Section 6, below.)

EXAMPLE 3.3. In the one-way ANOVA set-up denote the expectations $\mathbf{\mu}' = (\mu_1, \mu_2, \cdots, \mu_k)$ and write $\mathbf{c}_i'\mathbf{\mu}$, $\mathbf{c}_i' \varepsilon V(C')$ for a contrast in the expectations where $V(C')$ is the k-vector space orthogonal to the vector $\mathbf{1}' = (1, 1, \cdots, 1)$. Now consider $\omega_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ and $\omega_j : \mathbf{c}_j'\mathbf{\mu} = \sum_{h=1}^k c_{jh}\mu_h = 0, j \varepsilon I^C - \{0\}$ ranging over all contrasts such that $\mathbf{c}_j' \varepsilon V(C')$. Clearly, $\omega_0 = \bigcap_{I^C} \omega_j$ as nullity of all contrasts is equivalent to equality of all expectations.

Let the sample means be written $\bar{\mathbf{x}}' = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k)$, the variance estimate s^2 and the sample sizes $n_1, n_2, \cdots, n_k$. Then LR statistics are, for ω_0, the augmented F-ratio

$$Z_0^F = \left(\textstyle\sum_{h=1}^k n_h\bar{x}_h^2 - \left(\sum_{h=1}^k n_h\bar{x}_h\right)^2 / \sum_{h=1}^k n_h\right)/s^2,$$

and for $\omega_j, j \varepsilon I^C - \{0\}$, the squared Student t

$$Z_j^F = \left(\textstyle\sum_{h=1}^k c_{jh}\bar{x}_h\right)^2 / s^2 \sum_{h=1}^k (c_{jh}^2/n_h).$$

It is readily seen that this provides a joint testing family $\{\Omega^C, \mathbf{Z}^F\}$.

An α level STP is obtained by choosing ζ^F as in Example 3.1. This STP consists of testing the overall ω_0 and each contrast ω_i by checking if $Z_0^F > \zeta^F$ and $Z_j^F > \zeta^F$, respectively. By Corollary 1 this STP $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$ is known to be coherent in the sense that rejection of any contrast hypothesis always entails rejection of the overall equality hypothesis.

It will be noted that this STP consists of the significance tests of Scheffé's method (Example 5.1 [25], [6]).

In a coherent STP the level α may be regarded as an *experimentwise* level, as follows from the next theorem.

THEOREM 2. *The probability that a coherent* STP $\{\Omega, \mathbf{Z}, \zeta\}$ *of level* α *rejects at least one true* ω_i *of* Ω *is* α *if* ω_0 *is true; it is at most* α *irrespective of the truth of* ω_0 *provided* $\{\Omega, \mathbf{Z}\}$ *is either closed or joint.*

The probability of rejecting any particular true ω_i *of* Ω *is no more than the above probability.*

PROOF. By Theorem 1, $\{\Omega, \mathbf{Z}\}$ must be monotone, so that by (2.4) $Z_0 = \max \{Z_i \mid i \varepsilon I\}$. Hence rejection of ω_0, event $(Z_0 > \zeta)$, is equivalent to rejection of any ω_i, event $U_I(Z_i > \zeta)$. Under ω_0 the probability of the former event is α, (3.1), hence so is the probability of the latter, proving the first statement.

Next, let $\tilde{\Omega} = \{\omega_i \mid i \varepsilon \tilde{I}\}$, $\tilde{I} \subseteq I$, be the subfamily of true hypotheses whose intersection $\tilde{\omega}_0 = \bigcap_{\tilde{I}} \omega_i$ is necessarily also true. Then rejection of a true ω_i occurs whenever $\max \{Z_i \mid i \varepsilon \tilde{I}\} > \zeta$ and the required probability is

$$P_{\tilde{\omega}_0}(\max \{Z_i \mid i \varepsilon \tilde{I}\} > \zeta).$$

If it can be shown that

$$P_{\tilde{\omega}_0}(\max \{Z_i \mid i \varepsilon \tilde{I}\} > \zeta) = P_{\omega_0}(\max \{Z_i \mid i \varepsilon \tilde{I}\} > \zeta) \tag{3.2}$$

then the second statement of the theorem follows, for the latter probability is clearly no greater than $P_{\omega_0}(\max \{Z_i \mid i \varepsilon I\} > \zeta) = P_{\omega_0}(Z_0 > \zeta) = \alpha$.

If $\{\Omega, \mathbf{Z}\}$ is closed $\tilde{\omega}_0 \varepsilon \Omega$, as Ω is closed under intersection. Hence $\max \{Z_i \mid i \varepsilon \tilde{I}\}$, the statistic for $\tilde{\omega}_0$, has its distribution completely specified under $\tilde{\omega}_0$, irrespective of whether ω_0 holds or not, and so (3.2) holds.[5]

If $\{\Omega, \mathbf{Z}\}$ is joint, the joint distribution of all Z_i, $i \varepsilon \tilde{I}$, is specified under $\tilde{\omega}_0$, irrespective of the truth of ω_0. Hence the same holds also for $\max \{Z_i \mid i \varepsilon \tilde{I}\}$ and (3.2) follows as before.

The last statement in the theorem is obvious.

Coherence prevents the contradiction of rejecting a hypothesis without also rejecting all other hypotheses implying it. It does not, however, preclude the *dissonance* of rejecting a hypothesis whilst not rejecting any other hypotheses implied by it. Such dissonances, though not desirable, are sometimes allowed when the alternative to rejection of a hypothesis is taken to be non-rejection rather than acceptance. They are familiar to users of tests of significance, as these merely reject a hypothesis without indicating where it fails. They occur less often when STP's are used, as they may indicate some components to be also rejected. However, the use of STP's does not eliminate all dissonances: some STP's may reject hypotheses without rejecting any of their proper components—such STP's will be called *non-consonant*. One would generally prefer procedures which are coherent as well as *consonant*. In a later section non-consonant procedures will be compared in terms of the likelihood of avoiding dissonances.

Note that consonance of an STP can equivalently be defined as the requirement that any ω_i of Ω be rejected only if some minimal component of ω_i is rejected.

EXAMPLE 3.4. In the subset ANOVA F-ratio STP $\{\Omega^S, \mathbf{Z}^F, \zeta^F\}$ of Example 3.1 the minimal hypotheses $\omega_j : \mu_{j_\alpha} = \mu_{j_\beta}$, $j \varepsilon I^S_{\min}$, are those of pairwise equality of means. Since

$$Z_0^F = \left(\textstyle\sum_{h=1}^k n_h \bar{x}_h^2 - \left(\sum_{h=1}^k n_h \bar{x}_h\right)^2 / \sum_{h=1}^k n_h\right) / s^2$$

and

$$Z_j^F = (\bar{x}_{j_\alpha} - \bar{x}_{j_\beta})^2 / s^2(1/n_{j_\alpha} + 1/n_{j_\beta}), \quad j \varepsilon I^S_{\min},$$

and

$$Z_0^F > \max \{Z_j^F \mid j \varepsilon I^S_{\min}\} \quad \text{a.e.}$$

---

[5] A similar proof for one-way ANOVA was given in [6] and a general proof for closed testing families in [11] by Knight. As a result of the requirement of closure under intersection Knight's form of the theorem cannot be applied directly to some important classes of procedures, as, for instance, Scheffé's (Example 3.3, above). Knight surmounts this difficulty by applying his lemma to an extended family which includes all intersections of sets from Ω.

(see [9]), $(Z_0^F > \zeta)$ does not necessarily imply $(Z_j^F > \zeta)$ for any $j \varepsilon I^S_{\min}$. In other words, dissonances may occur in rejecting over-all equality of means without rejecting the equality of any pair of means [6].

EXAMPLE 3.5. The contrast ANOVA F-ratio STP $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$ of Example 3.3 is consonant as well as coherent. To see this, note that the maximal squared t statistic for a contrast is equal to the overall augmented F-ratio, i.e., $Z_0^F = \max\{Z_j^F \mid j \varepsilon I^C_{\min}\}$ where $I^C_{\min} = I^C - \{0\}$. (This is established by differentiating the Z_j^F expression in Example 3.3 over all $\mathbf{c}_j$ and setting the derivative equal to zero [6]. See also [26]). Hence, $(Z_0^F > \zeta^F) = \bigcup_{I^C_{\min}} (Z_j^F > \zeta^F)$, so that whenever overall ω_0 is rejected so is ω_j for some contrast.

EXAMPLE 3.6. In the pairwise ANOVA set-up the F-ratio STP $\{\Omega^P, \mathbf{Z}^F, \zeta^F\}$ would consist of tests belonging to both $\{\Omega^S, \mathbf{Z}^F, \zeta^F\}$ of Example 3.1 and $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$ of Example 3.3, since the overall ω_0 is the same in Ω^P, Ω^S and Ω^C and the other hypotheses of Ω^P, that is, those on pairwise comparisons, are included in both Ω^S and in Ω^C.

THEOREM 3. *All* STP'*s based on testing family* $\{\Omega, \mathbf{Z}\}$ *are coherent and consonant if and only if* $\{\Omega, \mathbf{Z}\}$ *is strictly monotone.*

PROOF. STP $\{\Omega, \mathbf{Z}, \zeta\}$ is coherent and consonant, by definition, if, for every $i \varepsilon I$,

$$(y \mid Z_i > \zeta) = \bigcup_{\{i<j, j\varepsilon I_{\min}\}} (y \mid Z_j > \zeta) \quad \text{a.e.} \tag{3.3}$$

For coherence requires the right hand side event to imply the left hand side, whereas consonance requires that the left hand side imply the right hand side. All STP's based on $\{\Omega, \mathbf{Z}\}$ are coherent and consonant if and only if, (3.3) hold for all ζ, but this is equivalent to (2.3), the definition of strict monotonicity of that testing family.

COROLLARY 2. *If, and only if, a testing family is* UI *related, all* STP'*s* based on *that family are cohrent and consonant.*

PROOF. This follows from Lemma B and Theorem 3.

Corollary 2 points out the correspondence between UI statistics and consonant STP's. Roy has used the UI principle to induce a test of the intersection hypothesis from the tests of its components, whereas its equivalent is used here to resolve the test decision on the intersection hypothesis into decisions on the components. The UI test of ω_0 can be used as part of an STP. Rejection by this test is followed in the STP by further tests to resolve which components are to be rejected, whereas non-rejection obviates the need for further testing.

EXAMPLE 3.7. In the subset ANOVA set-up for equal sample sizes $n_1 = n_2 = \cdots = n_k = n$, say, one might use studentized range statistics as in Example 2.7. Thus, hypothesis $\omega_i : \mu_{i_1} = \mu_{i_2} = \cdots = \mu_{i_{k_i}}$ of equality of all k_i means of subset S_i, say, would be tested with statistic

$$Z_i^R = \max\{|\bar{x}_{i_\alpha} - \bar{x}_{i_\beta}| n^{\frac{1}{2}}/s \mid i_\alpha \neq i_\beta \varepsilon S_i\}.$$

Taking ζ^R as the upper α point of the studentized range distribution for k variables with n_e d.f., and rejecting any ω_i if $(Z_i^R > \zeta^R)$, one obtains an STP

$\{\Omega^S, \mathbf{Z}^R, \zeta^R\}$. As pointed out in Example 2.12 this testing family $\{\Omega^S, \mathbf{Z}^R\}$ is UI related and hence is strictly monotone. By Corollary 2 $\{\Omega^S, \mathbf{Z}^R, \zeta^R\}$ is coherent and consonant.

EXAMPLE 3.8. If the decisions of the STP of Example 3.7 were restricted to the overall hypothesis and all pairwise comparisons (as in Example 3.6) this would not affect the coherence and consonance properties and one would obtain an STP $\{\Omega^P, \mathbf{Z}^R, \zeta^R\}$ consisting of the significance tests of Tukey's method of allowances [27].

Coherence and consonance are relations between test decisions on hypotheses which imply one another, i.e., between some ω_i and ω_j of Ω for which $i < j$. As was noticed in Section 2, above, other implication relations may also obtain within the family ω. The STP properties of coherence and consonance do not relate to such other relations and indeed STP decisions may violate them. Thus, in a family of hypotheses on equality of sets of parameters from among $\mu_1, \mu_2, \cdots, \mu_k$, an STP may accept $\mu_1 = \mu_2$ as well as $\mu_1 = \mu_3$ but reject $\mu_2 = \mu_3$ even though it is implied *jointly* by the two accepted hypotheses. Lehmann [14], [15] has considered procedures which preserve a wider class of implication relations and has termed them *compatible*. STP's, even consonant ones, are not generally compatible in that sense.

4. Resolution of STP's. The purpose of multiple comparisons in general, and STP's in particular, is to provide *resolution* of significant test decisions on overall hypotheses into significant decisions on components, as far down as minimal components. Thus, for family $\Omega = \{\omega_i \mid i \in I\}$ rejection of intersection ω_0 is to be resolved into rejection of proper components ω_i, $i \in I$, down to minimal hypotheses ω_j, $j \in I_{\min}$.

The extent of resolution of a procedure must be defined in terms of the likelihood of rejecting minimal proper components provided the overall hypothesis is rejected. If two procedures each test overall hypothesis ω_0 at level α, the one will be said to be *no less resolvent* than the other if it always rejects any minimal hypothesis rejected by the other, and will be said to be *strictly more resolvent* if it sometimes rejects when the other does not. In other words, consider two STP's $\{\Omega, \mathbf{Z}, \zeta\}$ and $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$, such that $\omega_0 = \bigcap_I \omega_i = \bigcap_{I^*} \omega_i$, where $\Omega = \{\omega_i \mid i \in I\}$ and $\Omega^* = \{\omega_i \mid i \in I^*\}$ and $I_{\min} \subset I^*$. Then the former STP is no less resolvent than the latter if

$$P_{\omega_0}(Z_0 > \zeta) = P_{\omega_0}(Z_0^* > \zeta^*) \tag{4.1a}$$

and for all $j \in I_{\min}$

$$(Z_j > \zeta) \supseteq (Z_j^* > \zeta^*) \quad \text{a.e.} \tag{4.1b}$$

It is strictly more resolvent if the containment in (4.1b) is proper. It is clear that no meaningful comparisons can be made between STP's unless they test the same hypotheses, at least to the extent that their intersections are the same and the minimal hypotheses of one of the families are all included in the other family.

EXAMPLE 4.1. In the pairwise ANOVA set-up with equal sample sizes the studentized range STP $\{\Omega^P, \mathbf{Z}^R, \zeta^R\}$ (Tukey's method-Example 3.8) will reject any pairwise comparison of means which is rejected by the F-ratio STP $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$ (Scheffé's method-Example 3.5) provided both STP's are of the same level [9]. Note that the family of hypotheses of the latter STP is wider than that of the former as it also includes all contrasts, but that all hypotheses of the former are included in the latter family.

Resolution comparisons are made between STP's of the same level. An alternative kind of comparison can be made between STP's which give identical decisions on all minimal components. The procedure will be said to be no less (more) parsimonious if it gives these decisions at no more (less) cost in rejection of other, non-minimal, hypotheses, and specifically of the overall hypothesis. Thus $\{\Omega, \mathbf{Z}, \zeta\}$ is *no less parsimonious* than $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$, where $\Omega \subseteq \Omega^*$ and $\omega_0 = \omega_0^*$, if

(4.2a) $\quad$ for all $j \varepsilon I_{\min}$ $\quad (Z_j > \zeta) = (Z_j^* > \zeta^*)$ $\quad$ a.e.,

(4.2b) $\quad$ for all $i \varepsilon I$ $\quad (Z_i > \zeta) \subseteq (Z_i^* > \zeta^*)$ $\quad$ a.e.,

(4.2c) $\quad$ and for ω_0 $\quad P_{\omega_0}(Z_0 > \zeta) \leqq P_{\omega_0}(Z_0^* > \zeta^*)$.

If the inequality in (4.2c) is strict, $\{\Omega, \mathbf{Z}, \zeta\}$ is *strictly more parsimonious* than $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$.

A possible measure of resolution is the ratio of the rejection probabilities of minimal hypotheses to that of the overall hypothesis. Under the overall hypothesis ω_0, if all minimal hypotheses have the same probability of being rejected, i.e., if

$$\alpha_1 = P_{\omega_j}(Z_j > \zeta)$$

for all $j \varepsilon I_{\min}$, resolution may be gauged by the ratio α_1/α. Clearly, if one STP is more resolvent than another, the former will have a higher α_1/α ratio than the latter. However, a higher α_1/α ratio merely means more probable resolution, not necessarily strictly more resolution in the sense of (4.1b).

EXAMPLE 4.2. Consider an eight mean ANOVA with 40 d.f. for error [6]. The 10% level studentized range STP $\{\Omega^S, \mathbf{Z}^F, 4.10\}$ has a resolution ratio of $\alpha_1/\alpha = 0.006/0.100 = 0.06$ whereas the 10% F-ratio STP $\{\Omega^S, \mathbf{Z}^F, 13.09\}$, noted in Example 4.1 to be less resolvent, has a ratio of only $\alpha_1/\alpha = 0.001/0.100 = 0.01$.

EXAMPLE 4.3. In the subset ANOVA set-up any hypothesis ω_i on a subset S_i of $k_i(>2)$ expectations may be tested by means of the maximal pairwise F statistic

$$Z_i^M = \max_{S_j \subset S_i} Z_j^F,$$

where the index j indicates S_j is a pair. For equal sample sizes $Z_j^F = (Z_j^R)^2/2$ (Example 2.12) so that the use of Z_i^M is equivalent to the use of the studentized range (Example 3.7). For unequal samples percentage points of Z_0^M are not available. For any critical value ζ^M, the STP $\{\Omega^S, \mathbf{Z}^M, \zeta^M\}$ is clearly coherent and consonant but its level α cannot be determined exactly for unequal n_js. However, by virtue of the first Bonferroni inequality $\alpha \leqq \binom{k}{2}\alpha_1$, where α_1 is the pairwise probability of a type I error which is readily determined by noting that ζ^M is the upper α_1 point of the F distribution with 1 and n_e d.f. The resolution ratio is thus $\alpha_1/\alpha \geqq 1/\binom{k}{2}$ for the maximal pairwise F STP. For the particular case of $k = 8$, $\alpha_1/\alpha \geqq 0.036$, whereas, with $n_e = 40$, for the augmented F-ratio STP this ratio is 0.01 (as in Example 4.2). Thus, if tests are desired only for pairs and other sets, and not for all contrasts, one would have more probable resolution with this approximate STP using the maximal pairwise F-ratio than with the one using the ordinary F-ratio, and one may well be led to prefer to use the former.

A more general measure of resolution should also take into account the chances of rejecting non-minimal component hypotheses. If ranks can be assigned meaningfully to hypotheses, resolution should depend on the entire sequence

$$\alpha_1 : \alpha_2 : \cdots : \alpha_{k-1} = \alpha$$

where α_q is the probability of falsely rejecting a hypothesis of rank q and $k - 1$ is the maximum rank (as in a k mean ANOVA), i.e., the rank of the intersection hypothesis. Examples of such sequences of probabilities are given in [6] for ANOVA STP's and in [7] for contingency table STP's.

To determine which STP's are more resolvent or more parsimonious than which others, one is led to consider the following relation between monotone testing families. $\{\Omega, \mathbf{Z}\}$ will be said to be *narrower* than $\{\Omega^*, \mathbf{Z}^*\}$ if $\Omega \subseteq \Omega^*$ and $\omega_0 = \omega_0^*$ and if there exists a function g such that

(4.3a) $\quad g$ is continuous monotone increasing,

(4.3b) $\quad$ for every $j \,\varepsilon\, I_{\min} \quad Z_j^* = g(Z_j)$ a.e.,

(4.3c) $\quad$ for every $i \,\varepsilon\, I - I_{\min} \quad Z_i^* \geqq g(Z_i)$ a.e.,

and

(4.3d) $$P(Z_0^* > g(Z_0)) > 0.$$

A stronger condition which implies (4.3d) is

(4.3d′) $$Z_0^* > g(Z_0) \quad \text{a.e.}$$

If (4.3d′) holds $\{\Omega, \mathbf{Z}\}$ will be said to be *strictly narrower* than $\{\Omega^*, \mathbf{Z}^*\}$.

EXAMPLE 4.4. Again, in the subset ANOVA set-up for equally sized samples, compare the studentized range testing family $\{\Omega^S, \mathbf{Z}^R\}$ with the augmented F-ratio testing family $\{\Omega^S, \mathbf{Z}^F\}$. The hypothesis family Ω^S and intersection $\omega_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ are the same for both. Any minimal hypothesis $j \,\varepsilon\, I_{\min}^S$ is a pairwise comparison of expectations $\omega_j : \mu_{j_\alpha} - \mu_{j_\beta} = 0$, for which $Z_j^F = (Z_j^R)^2/2$ (Example 2.12). Since Z^R can assume only positive values the function $Z^F = (Z^R)^2/2$ satisfies (4.3a) and (4.3b). Also, it has been shown ([9], see also Example 4.10, below) that for all non-minimal hypotheses ω_i on $k_i(>2)$ expectations $Z_i^F > (Z_i^R)^2/2$ a.e., so that (4.3c) and (4.3d′) also hold. This shows $\{\Omega^S, \mathbf{Z}^R\}$ to be narrower than $\{\Omega^S, \mathbf{Z}^F\}$.

EXAMPLE 4.5. For contrast hypotheses $\omega_j : \mathbf{c}_j'\boldsymbol{\mu} = \sum_{h=1}^k c_{jh}\mu_h = 0$ Tukey has generalized the studentized range statistic as

$$Z_j^R = 2n^{\frac{1}{2}}|\sum_{h=1}^k c_{jh}\bar{x}_h|/s \sum_{h=1}^k |c_{jh}|$$

and shown that $Z_0^R = \max \{Z_j^R \mid \mathbf{c}_j \,\varepsilon\, V(C')\}$ ([26], Section 3.6, [28], [16] pp. 44–46). Hence $\{\Omega^C, \mathbf{Z}^R\}$ is UI related, just as $\{\Omega^C, \mathbf{Z}^F\}$ is (Example 3.5 and Corollary 2). Thus, in the contrast ANOVA set-up for equally sized samples a generalized studentized range STP $\{\Omega^C, \mathbf{Z}^R, \zeta^R\}$ is an alternative to the (augmented) F-ratio STP $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$ of Example 3.5.

Let $g(x) = x^2/2$ for $x > 0$. This is a function which satisfies (4.3a), (4.3b) for all pairwise contrasts and (4.3c) and (4.3d) for the non-minimal hypothesis $\omega_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ (Example 4.3). However, for minimal hypotheses on contrasts involving more than two expectations (4.3b) does not hold. Hence $\{\Omega^c, \mathbf{Z}^R\}$ is not narrower than $\{\Omega^c, \mathbf{Z}^F\}$, nor vice versa.

EXAMPLE 4.6. Also compare testing families $\{\Omega^P, \mathbf{Z}^R\}$ of Example 2.1 and $\{\Omega^C, \mathbf{Z}^F\}$ of Example 3.3. Clearly, $\Omega^P \subset \Omega^C$ and ω_0 is the same for both. Furthermore, as in Example 4.3, for all $j \,\varepsilon\, I_{\min}^P$

$$Z_j^F = (Z_j^R)^2/2 \qquad \text{and} \qquad Z_0^F > (Z_0^R)^2/2 \quad \text{a.e.},$$

ω_0 being the only non-minimal hypothesis of Ω^P (and Ω^C). Hence (4.3a, b, c, d) hold so that $\{\Omega^P, \mathbf{Z}^R\}$ is found to be narrower than $\{\Omega^C, \mathbf{Z}^F\}$.

Given the narrowness relation between testing families the following two theorems show the corresponding relations between STP's based on these families.

THEOREM 4. *If* $\{\Omega, \mathbf{Z}\}$ *is narrower than* $\{\Omega^*, \mathbf{Z}^*\}$, *then for every* ζ^* *there exists* ζ *such that* $\{\Omega, \mathbf{Z}, \zeta\}$ *is no less parsimonious than* $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$. *Moreover, the former* STP *is strictly more parsimonious than the latter for some values of* ζ^*, i.e., *those for which*

$$P_{\omega_0}(Z_0^* > \zeta^* \geqq g(Z_0)) > 0, \tag{4.4}$$

where g *satisfies* (4.3a, b, c, d).

PROOF. In view of narrowness conditions, (4.3a, b, c, d) hold for some function g. For any given ζ^* choose ζ so that $\zeta^* = g(\zeta)$. Then it follows from (4.3b) that

$$\text{for all } j \,\varepsilon\, I_{\min} \quad (y \mid Z_j^*(y) > \zeta^*) = (y \mid Z_j(y) > \zeta) \quad \text{a.e.,}$$

and from (4.3c) that

$$\text{for all } i \,\varepsilon\, I \quad (y \mid Z_i^*(y) > \zeta^*) \supseteq (y \mid Z_i(y) > \zeta) \quad \text{a.e.,}$$

establishing (4.2a) and (4.2b). From (4.3c) one further obtains

$$P_{\omega_0}(Z_0^* > \zeta^*) = P_{\omega_0}(Z_0 > \zeta) + P_{\omega_0}(Z_0^* > \zeta^* \geqq g(Z_0))$$

from which (4.2c) follows—and no lesser parsimony is proved. It also follows that (4.2c) is a strict inequality—more parsimony—if (4.4) holds. That this holds at least for some values of ζ^* is ensured by (4.3d), and the theorem is proved.

EXAMPLE 4.7. In comparing the subset ANOVA studentized range STP $\{\Omega^S, \mathbf{Z}^R, \zeta^R\}$ and augmented F-ratio STP $\{\Omega^S, \mathbf{Z}^F, \zeta^F\}$, choose $\zeta^F = (\zeta^R)^2/2$. In view of the narrowness established in Example 4.4 it follows from Theorem 4 that the pairwise decisions of both STP's are identical but for all hypotheses on three or more means the F-ratio STP will reject whenever the studentized ratio STP does and more often. In particular, this will result in a higher level for the F-ratio STP—see Table 1 in [6].

THEOREM 5. *If* $\{\Omega, \mathbf{Z}\}$ *is narrower than* $\{\Omega^*, \mathbf{Z}^*\}$ *and if* $\{\Omega, \mathbf{Z}, \zeta\}$ *and* $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$ *are of the same level, then the former* STP *is no less resolvent than the latter. Moreover, it is strictly more resolvent for some values of* ζ, i.e., *those for which*

$$P_{\omega_0}(Z_0^* > g(\zeta) \geqq g(Z_0)) > 0, \tag{4.5}$$

where g *satisfies* (4.3a, b, c, d).

PROOF. According to the proof of Theorem 4, $P_{\omega_0}(Z_0^* > g(\zeta)) \geqq P_{\omega_0}(Z_0 > \zeta)$ for all ζ with strict inequality if (4.5) holds. Hence, if the levels of the two STP's are to be equal, i.e., if $P_{\omega_0}(Z_0^* > \zeta^*) = P_{\omega_0}(Z_0 > \zeta)$ (condition (4.1a)), one must have $\zeta^* > g(\zeta)$, with strict inequality conditional on (4.5). Now, minimal hypotheses ω_j, $j \varepsilon I_{\min}$, will be rejected by $\{\Omega^*, \mathbf{Z}^*, \zeta^*\}$ when $(Z_j^* > \zeta^*)$. But, in view of (4.3a)

$$(Z_j^* > \zeta^*) = (g(Z_j) > \zeta^*) \subseteq (g(Z_j) > g(\zeta)) = (Z_j > \zeta),$$

the containment being proper if (4.5) holds. Thus (4.1b) is established generally and the resolvence of $\{\Omega, \mathbf{Z}, \zeta\}$ shown to be always no less, and strictly whenever (4.5) holds. The latter occurs for at least some values of ζ, as argued in Theorem 4 for (4.4).

EXAMPLE 4.8. Again, comparing the pairwise ANOVA studentized range STP $\{\Omega^P, \mathbf{Z}^R, \zeta^R\}$ with the contrast ANOVA F-ratio STP $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$, the critical values may be chosen to ensure equal levels. In view of the narrowness comparison of the testing families in Example 4.6, it follows from Theorem 5 that $\{\Omega^P, \mathbf{Z}^R, \zeta^R\}$ is more resolvent than $\{\Omega^C, \mathbf{Z}^F, \zeta^F\}$, as stated in Example 4.1.

EXAMPLE 4.9. Numerical calculations have shown ([25], [26], Chapter 3) that neither Tukey's nor Scheffé's method is more resolvent for all contrasts (i.e., narrower confidence bounds—see Section 5, below). This accords with the finding in Example 4.5 that neither of the corresponding testing families is narrower than the other.

It is evident from these theorems that narrower testing families have the advantages of providing more[6] resolvent STP's of the same level (Theorem 5) and more[6] parsimonious STP's for the same decisions on minimal hypotheses

[6] "Providing more resolvent and more parsimonious STP's" must be understood as in Theorems 4 and 5 as strictly more for some critical values and no less for all other critical values.

(Theorem 4). In practical applications many multiple comparisons techniques have been criticised precisely for insufficient resolution or excessively frequent overall rejections for given minimal decisions. Hence, the use of narrower testing families may help to provide more practically useful techniques.

The next lemma gives a sufficient condition for one testing family to be narrower than another. In the following corollary this condition is applied to show when some STP's will be more resolvent than others.

LEMMA C. *Let* $\{\Omega, \mathbf{Z}\}$ *be a strictly monotone testing family and* $\{\Omega^*, \mathbf{Z}^*\}$ *be a monotone testing family such that* $\Omega \subseteq \Omega^*$, $\Omega_{\min} \subseteq \Omega^*_{\min}$ *and* $\omega_0 = \omega_0^*$. *A sufficient condition for* $\{\Omega, \mathbf{Z}\}$ *to be narrower than* $\{\Omega^*, \mathbf{Z}^*\}$ *is the existence of a function* g *satisfying* (4.3a, b) *and also*

$$P_{\omega_0}(Z_0^* \neq g(Z_0)) > 0. \tag{4.3e}$$

If (4.3e) *is replaced by*

$$Z_0^* \neq g(Z_0) \quad \text{a.e.}, \tag{4.3e$'$}$$

one obtains a sufficient condition for strict narrowness.

PROOF. If (4.3c) holds, (4.3e) implies (4.3d) and thus narrowness, whereas (4.3e$'$) implies (4.3d$'$) and thus strict narrowness. It therefore suffices to establish (4.3c).

Now, for any $i \,\varepsilon\, I - I_{\min}$, strict monotonicity of $\{\Omega, \mathbf{Z}\}$ gives, by (2.3), that

$$Z_i = \max\{Z_j \mid i < j, j \,\varepsilon\, I_{\min}\} \quad \text{a.e.},$$

whereas monotonicity of $\{\Omega^*, \mathbf{Z}^*\}$ requires, by (2.1) that

$$Z_i^* \geqq \max\{Z_j^* \mid i < j, j \,\varepsilon\, I^*_{\min}\} \quad \text{a.e.}$$

Under the conditions of this lemma

$$\begin{aligned}\max\{Z_j^* \mid i < j, j \,\varepsilon\, I^*_{\min}\} &\geqq \max\{Z_j^* \mid i < j, j \,\varepsilon\, I_{\min}\}\\ &= \max\{g(Z_j) \mid i < j, j \,\varepsilon\, I_{\min}\}\\ &= g(\max\{Z_j \mid i < j, j \,\varepsilon\, I_{\min}\}),\end{aligned}$$

so that, for all $i \,\varepsilon\, I - I_{\min}$, $Z_i^* \geqq g(Z_i)$ a.e., which is condition (4.3c).

EXAMPLE 4.10. In the subset ANOVA set-up for equal sample sizes the studentized range testing family $\{\Omega^S, \mathbf{Z}^F\}$ was noted (Example 2.7) to be strictly monotone and the augmented F-ratio testing family $\{\Omega^S, \mathbf{Z}^F\}$ to be monotone (Example 2.6). The family of hypotheses is the same in both cases and the monotonic increasing relation $Z_j^F = (Z_j^R)^2/2$ holds for all minimal (pairwise) hypotheses (Example 4.4). For any other hypothesis ω_i on $k_i(>2)$ means, the statistic

$$Z_i^F = \textstyle\sum_{h=1}^{k_i} n_{i_h}(\bar{x}_{i_h} - \bar{\bar{x}}_i)^2/s^2$$

where

$$\bar{\bar{x}}_i \textstyle\sum_{h=1}^{k_i} n_{i_h} = \sum_{h=1}^{k_i} n_{i_h}\bar{x}_{i_h},$$

and the statistic

$$(Z_i^R)^2/2 = \max \{|\bar{x}_{i_\alpha} - \bar{x}_{i_\beta}|n^{\frac{1}{2}}/s\, 2^{\frac{1}{2}} \mid i_\alpha, i_\beta \,\varepsilon\, S_i\}$$

are equal only if $k_i - 2$ of the means $\bar{x}_{i_h}$ are exactly equal to the average of the other 2 means, clearly an event of probability zero. Hence $Z_i^F = (Z_i^R)^2/2$ a.e., and all the conditions of Lemma C are seen to obtain. Thus the unproven statement in Example 4.3 is established and the proof that $\{\Omega^S, \mathbf{Z}^R\}$ is narrower than $\{\Omega^S, \mathbf{Z}^F\}$ is complete.

COROLLARY 3. *Let* $\{\Omega, \mathbf{Z}\}$ *and* $\{\Omega^*, \mathbf{Z}^*\}$ *be monotone testing families such that* $\Omega \subseteq \Omega^*$, $\Omega_{\min} \subseteq \Omega^*_{\min}$ *and* $\omega_0 = \omega_0^*$ *and let there exist a function g satisfying* (4.3a, b). *Then each of the following are sufficient conditions that* $\{\Omega, \mathbf{Z}\}$ *provide more parsimonious and more resolvent* STP's *than* $\{\Omega^*, \mathbf{Z}^*\}$,

(I) *If* $\{\Omega, \mathbf{Z}\}$ *is* UI *related and g also satisfies* (4.3e) *or* (4.3e$'$);

(II) *if both* $\{\Omega, \mathbf{Z}\}$ *and* $\{\Omega^*, \mathbf{Z}^*\}$ *are* UI *related and*

$$P_{\omega_0}(\max \{Z_j^* \mid j \,\varepsilon\, I^*_{\min}\} > \max \{Z_j^* \mid j \,\varepsilon\, I_{\min}\} > 0. \tag{4.6}$$

PROOF. By Lemma B, both (I) and (II) require $\{\Omega, \mathbf{Z}\}$ to be strictly monotone. Hence, by Lemma C, condition (I) is sufficient for $\{\Omega, \mathbf{Z}\}$ to be narrower than $\{\Omega^*, \mathbf{Z}^*\}$. The rest follows from Theorems 4 and 5.

Under (II), both testing families are strictly monotone so that, a.e., $Z_0^* = \max \{Z_j^* \mid j \,\varepsilon\, I^*_{\min}\}$ and $Z_0 = \max \{Z_j \mid j \,\varepsilon\, I_{\min}\}$. Now by (4.3a, b) $g(Z_0) = \max \{g(Z_j) \mid j \,\varepsilon\, I_{\min}\} = \max \{Z_j^* \mid j \,\varepsilon\, I_{\min}\}$, so that (4.6) of (II) is equivalent to (4.3d) and this implies (4.3e). The rest of the argument is as for (I), above.

EXAMPLE 4.11. The conclusions of Example 4.7 regarding equal sample size ANOVA STP's for subsets would also follow from Corollary 3(I) upon noting the existence of a g function satisfying (4.3a, b, c) as in Example 4.10, and the fact that the studentized range testing family is UI related (Example 2.12).

EXAMPLE 4.12. Similarly, the conclusions of Example 4.8 regarding pairwise and contrastwise STP's follow from Corollary 3 (II). Note first, by Example 4.6, that a function $x^2/2$ exists which satisfies (4.3a, b). Next, note that both $\{\Omega^P, \mathbf{Z}^R\}$ and $\{\Omega^C, \mathbf{Z}^F\}$ are UI related (Examples 3.8 and 3.5, respectively). Finally, condition (II) is established since (4.6) holds, i.e.,

$$\max \{Z_j^R \mid \mathbf{c}_j' \,\varepsilon\, V(C')\} > \max \{Z_j^F \mid j \,\varepsilon\, I^P_{\min}\} \quad \text{a.e.,}$$

for otherwise a number of contrasts in means $\bar{x}_h$ have to be zero, an event of probability zero.

EXAMPLE 4.13. In the equal sample size ANOVA set-up let $\Omega^D = \{\omega_0, \omega_2, \omega_3, \cdots, \omega_k\}$, where $\omega_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ as before and $\omega_j : \mu_1 = \mu_j$, $j = 2, \cdots, k$ so that

$$\omega_0 = \bigcap\nolimits_{j=2}^k \omega_j$$

is the intersection of Ω^D, and ω_j, $j = 2, \cdots, k$, are minimal. Let Z_j^D, $j = 2, \cdots, k$, be the absolute value of Student's t for ω_j, so that $Z_j^D = (Z_j^F)^{\frac{1}{2}} = (Z_j^R)/2^{\frac{1}{2}}$,

$j = 2, \cdots, k$. Further, let $Z_0^D = \max \{Z_j^D \mid j = 2, \cdots, k\}$ so that $\{\Omega^D, \mathbf{Z}^D\}$ is UI related. An α level STP $\{\Omega^D, \mathbf{Z}^D, \zeta^D\}$ may be obtained by choosing ζ^D from the tables provided by Dunnett [5].

To compare Dunnett's technique with Tukey's (Example 3.8) note that $\Omega^D \subset \Omega^P$, $\Omega^D_{\min} \subset \Omega^P_{\min}$ and ω_0 is the intersection of both. The two testing families are UI related, so it remains only to check whether (4.6) holds. This is so since the event $\max \{Z_j^R \mid j \varepsilon I^P_{\min}\} > \max \{2^{\frac{1}{2}} Z_j^D \mid j = 2, \cdots, k\}$ may obviously occur with positive probability (but less than one) under ω_0.

It then follows from Corollary 3 (II) that if $\zeta^D = 2^{\frac{1}{2}} \zeta^R$ the STP $\{\Omega^D, \mathbf{Z}^D, \zeta^D\}$ will have identical decision on ω_j, $j = 2, \cdots, k$, as STP $\{\Omega^P, \mathbf{Z}^R, \zeta^R\}$, but it will be more parsimonious. On the other hand, if ζ^D and ζ^R are chosen so that both STP's are of the same level, the former will be more resolvent than the latter.

An analogous result was proved by Krishnaiah for the related confidence statements ([13] and Example 5.9, below).

Corollary 3 allows parsimony and resolution comparisons of different STP's with equivalent statistics for minimal hypotheses. Condition I of the corollary may be used to compare STP's testing the same family of hypotheses (as in Example 4.11) in which case it shows those with UI related testing classes to be preferable to any others. Condition (II) of the corollary may be used to compare STP's all of which have UI related testing classes (Examples 4.12, 4.13). In that case the most preferable STP is shown to be the one whose subfamily of minimal hypotheses is contained in that of the other STP's. This is so unless the containment is trivial in the sense that the maximum of the STP's statistics for the contained subfamily equals that for a containing subfamily with probability one. The importance of using the UI principle with as restricted a family of hypotheses as possible was already stressed by Roy and Srivastava [24]. Krishnaiah has demonstrated a confidence bound analogue of condition II (see Corollary 3'(II) below) but has not pointed out the need for non-triviality of containment ([13], Theorem 6.4).

5. Simultaneous confidence statements. If the hypotheses of a family relate to parametric functions, simultaneous confidence statements may be made regarding the values of all those functions. The correspondence between such statements for certain functions and STP's for the hypotheses regarding particular values of these functions allows the extension of the foregoing properties of STP's to corresponding properties of these confidence statements. A number of additional definitions will be introduced to clarify this correspondence.

Consider a family $\Phi = \{\phi_i \mid i \varepsilon I\}$ of parametric functions, the range of ϕ_i being denoted Λ_i. A typical value of $\phi_i(\theta)$ will be denoted λ_i, it being understood that $\lambda_i \varepsilon \Lambda_i$. Then

$$\omega_i(\lambda_i) = \{\theta \mid \phi_i(\theta) = \lambda_i\} \tag{5.1}$$

is a hypothesis regarding the value of $\phi_i(\theta)$, and the class $\{\omega_i(\lambda_i) \mid \lambda_i \varepsilon \Lambda_i\}$ of such hypotheses forms a disjoint partition of the parameter space ω.

For each value λ_i, let $Z_i(Y; \lambda_i)$ be a statistic whose distribution is the same for all $\theta \varepsilon \omega_i(\lambda_i)$ and is independent of the value of λ_i. Write $\bar{\mathbf{Z}}$ for the collection of all such statistics with $\lambda_i \varepsilon \Lambda_i$, $i \varepsilon I$. The collections of parametric functions ϕ_i and corresponding statistics $Z_i(Y; \lambda_i)$ will be denoted $[\Phi, \bar{\mathbf{Z}}]$ and referred to as an *estimation family*. For any given probability α_i let ζ be such that

$$(5.2) \qquad P_{\omega_i(\lambda_i)}(Z_i(Y;\lambda_i) \leqq \zeta) = P_\theta(Z_i(Y;\phi_i(\theta)) \leqq \zeta) = 1 - \alpha_i,$$

this probability clearly being independent of the value of λ_i or θ. Then the subset of Λ_i

$$(5.3) \qquad \Lambda_i(Y; \zeta) = \{\lambda_i \mid Z_i(Y; \lambda_i) \leqq \zeta\}$$

is a *confidence region* for $\phi_i(\theta)$ with co-efficient $1 - \alpha_i$. The *confidence statement* $\phi_i(\theta) \varepsilon \Lambda_i(Y; \zeta)$ is equivalent to the statement $\theta \varepsilon \Theta_i(Y; \zeta)$, with

$$(5.4) \qquad \Theta_i(Y; \zeta) = \{\theta \mid Z_i(Y; \phi_i(\theta)) \leqq \zeta\}$$

being a set in ω. Either of these statements is understood in the sense of acceptance of hypothesis $\omega_i(\lambda_i)$ for any value $\lambda_i \varepsilon \Lambda_i(Y; \zeta)$ by a test of level α_i.

The family of statements

$$(5.5) \qquad \{\phi_i(\theta) \varepsilon \Lambda_i(Y; \zeta) \mid i \varepsilon I\} = \{\theta \varepsilon \Theta_i(Y; \zeta) \mid i \varepsilon I\}$$

based on estimation family $[\Phi, \bar{\mathbf{Z}}]$ and critical value ζ is referred to as a *Simultaneous Confidence Statement* (SCS) and denoted $[\Phi, \bar{\mathbf{Z}}, \zeta]$. Clearly, all statements of $[\Phi, \bar{\mathbf{Z}}, \zeta]$ are true if and only if $\theta \varepsilon \bigcap_I \Theta_i(Y; \zeta)$, i.e., if the true parameters θ are in each one of the confidence regions $\Theta_i(Y; \zeta)$, $i \varepsilon I$. The probability that this event occurs, i.e.,

$$(5.6) \qquad C = P_\theta(\phi_i(\theta) \varepsilon \Lambda_i(Y; \zeta) \forall i \varepsilon I) = P_\theta(\theta \varepsilon \bigcap_I \Theta_i(Y; \zeta)),$$

is referred to as the *joint confidence co-efficient* of $[\Phi, \bar{\mathbf{Z}}, \zeta]$. By introducing (5.4) this may be written

$$(5.7) \qquad C = P_\theta(\max \{Z_i(Y; \phi_i(\theta)) \mid i \varepsilon I\} \leqq \zeta).$$

It is important to distinguish a *joint confidence region* for several parametric functions from a simultaneous confidence statement. The former is a confidence region in the sense of (5.3), where ϕ_i is a vector valued parametric function and is equivalent to a single region in parameter space—(5.4). The latter consists of a family of such regions, each one relating to either a scalar or a vector Φ_i. Note that Roy and Bose [22] have restricted consideration to the case where each ϕ_i was scalar valued and specifically to the case where each $\Lambda_i(Y; \zeta)$ could be expressed as an interval on the real line.

EXAMPLE 5.1. In the one-way ANOVA set-up a simultaneous confidence statement on all contrasts $\mathbf{c}_j'\boldsymbol{\mu}$, $\mathbf{c}_j' \varepsilon V(C')$, in expectations is

$$\{\mathbf{c}_j'\boldsymbol{\mu} \varepsilon \Lambda_{\mathbf{c}_j}(\mathbf{x}, \zeta^F) \forall \mathbf{c}_j' \varepsilon V(C')\},$$

where $\mathbf{x}$ indicates the observations, $\zeta^F = (k - 1)F_{(k-1;n_e)1-\alpha}$ and

$$\Lambda_{\mathbf{c}_j}(\mathbf{x}, \zeta^F) = \{\mathbf{c}_j'\boldsymbol{\mu} \mid \mathbf{c}_j'\bar{\mathbf{x}} - (\zeta^F s^2 \mathbf{c}_j' N^{-1}\mathbf{c}_j)^{\frac{1}{2}} \leqq \mathbf{c}_j'\boldsymbol{\mu} \leqq \mathbf{c}_j'\bar{\mathbf{x}} + (\zeta^F s^2 \mathbf{c}_j' N^{-1}\mathbf{c}_j)^{\frac{1}{2}}\},$$

with $N = \operatorname{diag}(n_1, n_2, \cdots, n_k)$, is a confidence interval for contrast $\mathbf{c}_j'\boldsymbol{\mu}$. Furthermore, since $C'\boldsymbol{\mu}$ generates all such contrasts

$$\{C'\boldsymbol{\mu} \mid \textstyle\sum_{h=1}^k n_h(\bar{x}_h - \mu_h)^2 - (\sum_{n=1}^k n_h(\bar{x}_h - \mu_h))^2/\sum_{h=1}^k n_h \leqq \zeta^F s^2\}$$

is a single joint confidence region for all contrasts in expectations. Since this region is the intersection of all the $\Lambda_{\mathbf{c}_j}(\mathbf{x}, \zeta^C)$ regions in the space of contrasts $C'\boldsymbol{\mu}$ (and in that of the parameters $\boldsymbol{\mu}, \sigma^2$) the joint region and the simultaneous statement have, by (5.6), the same confidence co-efficient $1 - \alpha$.

This was first pointed out by Scheffé [25] (see also Example 3.3 for the related STP on null contrasts).

An estimation family is said to be *related* if the functions of Φ have a transitive relation $\mathcal{G}$ which is neither symmetric nor reflexive, such that if $i \prec j$ then $\omega_i(\lambda_i)$ implies $\omega_j(\lambda_j)$ for a particular value λ_j, and $0 \prec i$ for all $i \varepsilon I - \{0\}$. For $i \prec j$ the relation $\mathcal{G}$ defines a function r_{ji} such that $\lambda_j = r_{ji}(\lambda_i)$. Since $\mathcal{G}$ is transitive, $i \prec j$ and $j \prec g$ require $i \prec g$ so that if $\omega_i(\lambda_i) \subset \omega_j(\lambda_j) \subset \omega_g(\lambda_g)$, $\lambda_g = r_{gi}(\lambda_i) = r_{gj}(\lambda_j) = r_{gj}(r_{ji}(\lambda_i))$, which imposes the relation $r_{gi} = r_{gj}(r_{ji})$ on these functions. From these definitions it follows that for $i \prec j$

$$\omega_j(\lambda_j) = \bigcup_{\{\lambda_i \mid r_{ji}(\lambda_i) = \lambda_j\}} \omega_i(\lambda_i). \tag{5.8}$$

The functions ϕ_j for which there exists no $\phi_g \varepsilon \Phi$ such that $j \prec g$ will be indexed by $j \varepsilon I_{\min}$. The adjective *minimal* will be applied to these functions and all other concepts depending on these functions, as, for instance, the confidence regions $\{\Lambda_j(Y; \zeta) \mid j \varepsilon I_{\min}\}$. Similarly, the adjective *related* will be applied to SCSs based on related estimation families.

Example 5.2. In the subset ANOVA set-up one may be concerned with estimation of all contrasts in the expectations of each subset. For each subset S_i, $i \varepsilon I^S$, of k_i expectations let C_i be a $k_i \times (k_i - 1)$ matrix such that $C_i'\boldsymbol{\mu}$ generates all contrasts of S_i. It is well known that whenever $S_j \subset S_i$ there exists a B_{ji} such that $C_j' = B_{ji}'C_i'$. Hence, if $C_i'\boldsymbol{\mu} = \boldsymbol{\lambda}_i$ this implies that $C_j'\boldsymbol{\mu} = \boldsymbol{\lambda}_j$ where $\boldsymbol{\lambda}_j = B_{ji}'\boldsymbol{\lambda}_i$. Thus, the functions $\phi_i(\boldsymbol{\mu}) = C_i\boldsymbol{\mu}$, are defined for each $i \varepsilon I^S$ and a relation $\phi_j(\boldsymbol{\mu}) = B_{ji}'\phi_i(\boldsymbol{\mu})$ exists whenever $S_j \subset S_i$. An estimation family for these contrast sets will thus be related.

A simultaneous 1-α confidence statement on all these linear sets of contrasts is

$$\{C_i'\boldsymbol{\mu} \,\varepsilon\, \Lambda_i(\mathbf{x}, \zeta^F) \mid i \varepsilon I^S\}$$

where, with the same definitions as before,

$$\begin{aligned}\Lambda_i(\mathbf{x}, \zeta^F) &= \{C_i'\boldsymbol{\mu} \mid (\bar{\mathbf{x}} - \boldsymbol{\mu})'C_i(C_i'N^{-1}C_i)^{-1}C_i'(\bar{\mathbf{x}} - \boldsymbol{\mu}) \leqq \zeta^F s^2\} \\ &= \{C_i'\boldsymbol{\mu} \mid \textstyle\sum_{h\varepsilon S_i} n_h(\bar{x}_h - \mu_h)^2 - (\sum_{h\varepsilon S_i} n_h(\bar{x}_h - \mu_h))^2/\sum_{h\varepsilon S_i} n_h \leqq \zeta^F s^2\}\end{aligned}$$

is the confidence region for $C_i'\boldsymbol{\mu}$, using the augmented F-ratio.

The corresponding STP for hypotheses of null contrast sets was given in Examples 3.1 and 3.4.

EXAMPLE 5.3. In the same set-up as that of Example 5.2 Roy and Gnanadesikan [23] have proposed putting confidence bounds on non-centrality parameter

$$\phi_i(\boldsymbol{\mu}) = \{\boldsymbol{\mu}' C_i [C_i' N^{-1} C_i]^{-1} C_i' \boldsymbol{\mu}\}^{\frac{1}{2}}$$

for each $i \,\varepsilon\, I^S$. It is readily checked that if $S_i \subset S_j$, then for any λ_i, $\phi_i(\boldsymbol{\mu}) = \lambda_i$ does not generally imply a particular single value λ_j for $\phi_j(\boldsymbol{\mu})$, (except that if $\lambda_i = 0$ then also $\lambda_j = 0$). Hence these non-centrality parameter functions do not allow of a related estimation family.

Defining statistics

$$Y_i = \bar{\mathbf{x}}' C_i [C_i' N^{-1} C_i]^{-1} C_i' \bar{\mathbf{x}}$$

for each $i \,\varepsilon\, I^S$, the Roy-Gnanadesikan confidence statement is $\{\phi_i(\boldsymbol{\mu}) \,\varepsilon\, \Lambda_i^{RG} \cdot (\mathbf{x}, \zeta^F) \mid i \,\varepsilon\, I^S\}$, with the confidence regions

$$\Lambda_i^{RG}(\mathbf{x}, \zeta^F) = \{\lambda_i \mid Y_i^{\frac{1}{2}} - (\zeta^F s^2)^{\frac{1}{2}} \leqq \lambda_i \leqq Y_i^{\frac{1}{2}} - (\zeta^F s^2)^{\frac{1}{2}}\}.$$

The sense in which this SCS is non-related is illustrated in Example 5.5, below.

For a related estimation family $[\Phi, \bar{\mathbf{Z}}]$ consider, for any $\lambda_0 \,\varepsilon\, \Lambda_0$, the family of hypotheses

$$\Omega_{\lambda_0} = \{\omega_i(\lambda_i) \mid \lambda_i = r_{i0}(\lambda_0)\}, \tag{5.9}$$

and the family of statistics

$$\mathbf{Z}_{\lambda_0} = \{Z_i(Y; \lambda_i) \mid \lambda_i = r_{i0}(\lambda_0)\}. \tag{5.10}$$

It is easily checked that $\mathcal{J}$ is an implication relation for Ω_{λ_0} and that $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}\}$ satisfies the requirements of a testing family (Section 2, above). Hence $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}, \zeta\}$ is an STP of level

$$\alpha = P_{\omega_0(\lambda_0)}(Z_0(Y; \lambda_0) > \zeta). \tag{5.11}$$

This STP accepts $\omega_i(\lambda_i)$, where $\lambda_i = r_{i0}(\lambda_0)$, if $Z_i(Y; \lambda_i) \leqq \zeta$. Thus, acceptance of $\omega_i(\lambda_i)$ by STP $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}, \zeta\}$, where $\lambda_i = r_{i0}(\lambda_0)$, is determined by the same event as inclusion of the parametric value λ_i in SCS $[\Phi, \bar{\mathbf{Z}}, \zeta]$, i.e., $\lambda_i \,\varepsilon\, \Lambda_i(Y; \zeta)$ as in (5.3). This brings out the correspondence between the STP's $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}, \zeta\}$ for all $\lambda_0 \,\varepsilon\, \Lambda_0$ and the related SCS $[\Phi, \bar{\mathbf{Z}}, \zeta]$. The experimentwise error rate of $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}, \zeta\}$ is, provided $Z_0(Y; \lambda_0) = \max\{Z_i(Y; \lambda_i) \mid \lambda_i = r_{i0}(\lambda_0)\}$ (see (5.16) below),

$$\alpha = P_{\omega_0(\lambda_0)}(\max\{Z_i(Y; \lambda_i) \mid \lambda_i = r_{i0}(\lambda_0)\} > \zeta)$$
$$= P_\theta(\max\{Z_i(Y; \phi_i(\theta)) \mid i \,\varepsilon\, I\} > \zeta)$$

which is seen to be $1 - C$, for joint confidence co-efficient C.

Implication relations have been defined for a related estimation family;

hence the terms coherence and consonance may also be defined for a related SCS. For non-related SCS's these terms have no clear meaning. Coherence was defined in Section 3 as the requirement that if $\omega_i(\lambda_i) \subset \omega_j(\lambda_j)$, acceptance of $\omega_i(\lambda_i)$ would entail acceptance of $\omega_j(\lambda_j)$. Consonance was defined as the requirement that $\omega_i(\lambda_i)$ be accepted if all $\omega_j(\lambda_j)$ were accepted for which $\omega_i(\lambda_i) \subset \omega_j(\lambda_j)$. Correspondingly, related SCS $[\Phi, \bar{\mathbf{Z}}, \zeta]$ is said to be *coherent* if, when $\lambda_j = r_{ji}(\lambda_i)$ then $\lambda_i \,\varepsilon\, \Lambda_i(y; \zeta)$ ensures $\lambda_j \,\varepsilon\, \Lambda_j(y; \zeta)$. It is said to be *consonant* if $\lambda_j \,\varepsilon\, \Lambda_j(y; \zeta) \forall \lambda_j = r_{ji}(\lambda_i)$ ensures $\lambda_i \,\varepsilon\, \Lambda_i(y; \zeta)$. Thus $[\Phi, \bar{\mathbf{Z}}, \zeta]$ is coherent and consonant when the statements $\lambda_i \,\varepsilon\, \Lambda_i(y; \zeta)$ and $\lambda_j \,\varepsilon\, \Lambda_j(y; \zeta)$ for all j such that $\lambda_j = r_{ji}(\lambda_i)$ must occur together for all $i \,\varepsilon\, I$.

In terms of the θ set formulation the coherence requirement is that for all $i \,\varepsilon\, I$

$$(5.12) \qquad \Theta_i(y; \zeta) \subset \bigcap_{\{j \mid i < j\}} \Theta_j(y; \zeta)$$

and the consonance requirement is that for all $i \,\varepsilon\, I$

$$(5.13) \qquad \Theta_i(y; \zeta) \supset \bigcap_{\{j \mid i < j\}} \Theta_j(y; \zeta).$$

Thus $[\Phi, \bar{\mathbf{Z}}, \zeta]$ is coherent and consonant if

$$(5.14a) \qquad \Theta_i(y; \zeta) = \bigcap_{\{j \mid i < j\}} \Theta_j(y; \zeta)$$

for every $i \,\varepsilon\, I$.

In view of the analogous definitions of coherence and consonance for STP's and SCS's, it is immediately seen that a SCS is coherent and/or consonant, if and only if every one of the corresponding STP's is coherent and/or consonant, respectively.

Finally, to complete the analogy between $[\Phi, \bar{\mathbf{Z}}, \zeta]$ and $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}, \zeta\}$ for all $\lambda_0 \,\varepsilon\, \Lambda_0$, the estimation family $[\Phi, \bar{\mathbf{Z}}]$ will be called *monotone* if

$$(5.14b) \qquad Z_i(y; \lambda_i) \geqq Z_j(y; \lambda_j) \quad \text{a.e.}$$

whenever $\lambda_j = r_{ji}(\lambda_i)$, and *strictly monotone* if

$$(5.15) \qquad Z_i(y; \lambda_i) = \max\,\{Z_j(y; \lambda_j) \mid \lambda_j = r_{ji}(\lambda_i)\} \quad \text{a.e.}$$

for all $i \,\varepsilon\, I$. In particular, for any monotone $[\Phi, \bar{\mathbf{Z}}]$

$$(5.16) \qquad Z_0(Y; \lambda_0) = \max\,\{Z_i(Y; \lambda_i) \mid \lambda_i = r_{i0}(\lambda_0)\}.$$

These definitions correspond, for any $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}\}$ to (2.1), (2.2) and (2.4), so that (strict) monotonicity of $[\Phi, \bar{\mathbf{Z}}]$ corresponds to the same property of $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}\}$ for each $\lambda_0 \,\varepsilon\, \Lambda_0$. Also, estimation family $[\Phi, \bar{\mathbf{Z}}]$ will be said to be UI related if testing family $\{\Omega_{\lambda_0}, \mathbf{Z}_{\lambda_0}\}$ is UI related for each λ_0.

With these analogous definitions and properties of STP's and SCS's it is clear that the theorems and corollaries that have been proved for STP's will apply, *mutatis mutandis,* to SCS's. They will be stated below without further proof; the number of each statement being the same as for STP's except for a prime.

THEOREM 1$'$. *All* SCS's *based on* $[\Phi, \bar{\mathbf{Z}}]$ *are coherent if and only if* $[\Phi, \bar{\mathbf{Z}}]$ *is monotone.*

COROLLARY 1$'$. *If* $Z_i(Y; \lambda_i)$ *is a* LR *statistic for* $\omega_i(\lambda_i)$ *for all* $Z_i(Y; \lambda_i) \varepsilon \bar{\mathbf{Z}}$, *then all* SCS's *based on* $[\Phi, \bar{\mathbf{Z}}]$ *are coherent.*

THEOREM 2$'$. *The joint confidence co-efficient of* coherent $[\Phi, \bar{\mathbf{Z}}, \zeta]$ *is* $C = 1 - \alpha$, *for* α *of* (5.11). *The confidence co-efficient of any* ϕ_i *is at least* $1 - \alpha$.

THEOREM 3$'$. *All* SCS's *based on* $[\Phi, \bar{\mathbf{Z}}]$ *are coherent and consonant if and only if* $[\Phi, \bar{\mathbf{Z}}]$ *is strictly monotone.*

COROLLARY 2$'$. *All* SCS's *based on* $[\Phi, \bar{\mathbf{Z}}]$ *are coherent and consonant if, and only if,* $[\Phi, \bar{\mathbf{Z}}]$ *is* UI *related.*

EXAMPLE 5.4. To show the monotonicity of the estimation family of Example 5.2 and the coherence of the resulting SCS's one may argue as follows. It follows from a well known theorem (applying 1f.1.1 of [19] with $\mathbf{c}' = \mathbf{x}'C_i'$), that

$$\max_{\mathbf{c}'\varepsilon V(C_i')} \{\mathbf{c}'(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)\}^2/s^2\mathbf{c}'N^{-1}\mathbf{c} = (C_i'\bar{\mathbf{x}} - C_i'\boldsymbol{\mu}_0)'(C_i'N^{-1}C_i)^{-1}(C_i'\bar{\mathbf{x}} - C_i'\boldsymbol{\mu}_0)/s^2,$$

where the right hand side is $Z_i(Y; C_i'\boldsymbol{\mu}_0)$, the statistic for ω_i in Example 5.2. Now, hypotheses $\omega_i(\lambda_i) : C_i'\boldsymbol{\mu} = \boldsymbol{\lambda}_i$ have implication relations $\omega_i(\lambda_i) \subset \omega_j(\lambda_j)$ for all $S_i \supset S_j$ provided these $\boldsymbol{\lambda}_i$s, $i \varepsilon I^S$ are such that there exists a $\boldsymbol{\mu}_0$ for which $C_i'\boldsymbol{\mu}_0 = \lambda_i \forall i \varepsilon I^S$. Further, $S_i \supset S_j$ is equivalent to $V(C_i') \supset V(C_j')$ so that

$$\max_{\mathbf{c}'\varepsilon V(C_i')} [\mathbf{c}'(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)]^2/s^2\mathbf{c}'N^{-1}\mathbf{c} \geqq \max_{\mathbf{c}'\varepsilon V(C_j')} [\mathbf{c}'(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)]^2/s^2\mathbf{c}'N^{-1}\mathbf{c},$$

i.e., $Z_i(Y; \boldsymbol{\lambda}_i) \geqq Z_j(Y; \boldsymbol{\lambda}_j)$, which proves monotonicity.

Coherence of the SCS of Example 5.2 follows by a similar argument. Let $S_i \supset S_j$, let $\boldsymbol{\lambda}_i$ be a value of $C_i'\boldsymbol{\mu}$ and write $\boldsymbol{\mu}_0$ for any value such that $\boldsymbol{\lambda}_i = C_i'\boldsymbol{\mu}_0$. Now $\boldsymbol{\lambda}_i \varepsilon \Lambda_i(\mathbf{x}, \zeta)$ is equivalent to saying that $Z_i(\mathbf{x}; C_i'\boldsymbol{\mu}_0) \leqq \zeta$. Hence, by the inequality above, also $Z_j(\mathbf{x}; C_j'\boldsymbol{\mu}_0) \leqq \zeta$, which is equivalent to saying that $\boldsymbol{\lambda}_j \varepsilon \Lambda_j(\mathbf{x}, \zeta)$, where $\boldsymbol{\lambda}_j = C_j'\boldsymbol{\mu}_0$, the value corresponding to $\boldsymbol{\lambda}_i = C_i'\boldsymbol{\mu}_0$. Inclusion of $\boldsymbol{\lambda}_i$ in the SCS implies inclusion of $\boldsymbol{\lambda}_j$, establishing coherence.

EXAMPLE 5.5. In the Roy-Gnanadesikan set-up of Example 5.3 one notes, arguing as in Example 5.4, that if $S_i \supset S_j$ then for any $\boldsymbol{\mu}$, $\phi_i(\boldsymbol{\mu}) \geqq \phi_j(\boldsymbol{\mu})$ and $Y_i \geqq Y_j$. Now, for any given $\boldsymbol{\mu}$ the random event

$$Y_i^{\frac{1}{2}} + (\zeta^F s^2)^{\frac{1}{2}} \geqq \phi_i(\boldsymbol{\mu}) \geqq \phi_j(\boldsymbol{\mu}) \geqq Y_j^{\frac{1}{2}} + (\zeta^F s^2)^{\frac{1}{2}},$$

may occur. In that case $\phi_i(\boldsymbol{\mu}) \varepsilon \Lambda_i^{RG}(\mathbf{x}, \zeta^F)$ but $\phi_j(\boldsymbol{\mu}) \not\varepsilon \Lambda_j(\mathbf{x}, \zeta^F)$. Thus for a given set of expectations $\boldsymbol{\mu}$, the non-centrality parameter $\phi_i(\boldsymbol{\mu})$ may be included in the Roy-Gnanadesikan SCS whereas the other non-centrality parameter $\phi_j(\boldsymbol{\mu})$ is not. Note also that in the even that

$$\phi_i(\boldsymbol{\mu}) \geqq Y_i^{\frac{1}{2}} + (\zeta s^2)^{\frac{1}{2}} \geqq Y_j^{\frac{1}{2}} + (\zeta s^2)^{\frac{1}{2}} \geqq \phi_j(\boldsymbol{\mu})$$

the converse would occur, $\phi_j(\boldsymbol{\mu})$ being included whereas $\phi_i(\boldsymbol{\mu})$ was excluded. Evidently there is no coherence in the inclusions within this SCS.

EXAMPLE 5.6. An alternative approach may be taken in the subset ANOVA set-up if sample sizes are equal. For any subset S_i of k_i expectations $\mu_{i_1}, \mu_{i_2}, \cdots, \mu_{i_{k_i}}$ one may define $\phi_i^S(\mathbf{\mu})$ as the set of $\binom{k_i}{2}$ differences $\mu_{i_\alpha} - \mu_{i_\beta}$, $i_\alpha \neq i_\beta$. For any values $\boldsymbol{\lambda}_i$ of the differences in set $\phi_i^S(\mathbf{\mu})$ one may define statistic

$$Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i) = n^{\frac{1}{2}} \max \{[\bar{x}_{i_\alpha} - \bar{x}_{i_\beta} - (\mu_{i_\alpha} - \mu_{i_\beta})]/s \mid \mu_{i_\alpha}, \mu_{i_\beta} \varepsilon S_i\}.$$

The distribution of $Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i)$ under $\omega_i(\boldsymbol{\lambda}_i) = \{\mathbf{\mu} \mid \phi_i^S(\mathbf{\mu}) = \boldsymbol{\lambda}_i\}$ is that of a studentized range of k_i means and n_e d.f. irrespective of the value of $\boldsymbol{\lambda}_i$. Hence $[\Phi^S, \bar{\mathbf{Z}}^R]$ is an estimation family.

Defining, for any i, j such that $S_i \supset S_j$, r_{ji} as the function picking out the subset $\phi_j^S(\mathbf{\mu})$ from the set $\phi_i^S(\mathbf{\mu})$ of differences, $\phi_j^S(\mathbf{\mu}) = r_{ji}(\phi_i^S(\mathbf{\mu}))$ whenever $S_i \supset S_j$. Now, if $\phi_i^S(\mathbf{\mu}) = \boldsymbol{\lambda}_i$, clearly $\phi_j^S(\mathbf{\mu}) = r_{ji}(\boldsymbol{\lambda}_i)$, so that $\omega(\boldsymbol{\lambda}_i) \subset \omega_j(\boldsymbol{\lambda}_j)$ if $\boldsymbol{\lambda}_j = r_{ji}(\boldsymbol{\lambda}_i)$. Hence the estimation family $[\Phi^S, \mathbf{Z}^R]$ is related, and so is the SCS.

The resulting SCS of confidence co-efficient $1 - \alpha$ is $\{\phi_i^S(\mathbf{\mu}) \varepsilon \Lambda_i^R(\mathbf{x}, \zeta^R) \mid i \varepsilon I^S\}$ where ζ^R is the upper α point of the k mean studentized range distribution with n_e d.f. and

$$\Lambda_i^R(\mathbf{x}, \zeta^R) = \{\lambda_i \mid Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i) \leqq \zeta^R\}.$$

Note, in particular, that for S_j consisting of just two expectations $\phi_j^S(\mathbf{\mu}) = \mu_{j_\alpha} - \mu_{j_\beta}$, and the confidence region becomes the interval

$$\Lambda_j^R(\mathbf{x}, \zeta^R) = \{\lambda_j \mid \bar{x}_{j_\alpha} - \bar{x}_{j_\beta} - \zeta^R s/n^{\frac{1}{2}} \leqq \mu_{j_\alpha} - \mu_{j_\beta} \leqq \bar{x}_{j_\alpha} - \bar{x}_{j_\beta} + \zeta^R s/n^{\frac{1}{2}}\}$$

on the value of the difference $\lambda_j = \mu_{j_\alpha} - \mu_{j_\beta}$. This type of SCS is due to Tukey [28], being referred to as the method of allowances. The corresponding STP for null differences was discussed in Example 3.7.

Since $\phi_i^S(\mathbf{\mu}) \supset \phi_j^S(\mathbf{\mu})$ for $S_i \supset S_j$ it is clear that $Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i) \geqq Z_j^R(\mathbf{x}, \boldsymbol{\lambda}_j)$ provided $\boldsymbol{\lambda}_j = r_{ji}(\boldsymbol{\lambda}_i)$. Moreover, it is readily apparent that

$$Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i) = \max \{Z_j^R(\mathbf{x}, \boldsymbol{\lambda}_j) \mid \boldsymbol{\lambda}_j = r_{ji}(\boldsymbol{\lambda}_i), S_j \subset S_i\}$$

so that $[\Phi^S, \bar{\mathbf{Z}}^R]$ is strictly monotone.

Coherence and consonance of SCS $[\Phi^S, \bar{\mathbf{Z}}^R, \zeta^R]$, which follows from Theorem 3′, may be checked as follows. From the above relation between $Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i)$ for S_i and the statistics $Z_j^R(\mathbf{x}, \boldsymbol{\lambda}_j)$ for all pairs of expectations $S_j(\subset S_i)$, it follows that $(Z_i^R(\mathbf{x}, \boldsymbol{\lambda}_i) \leqq \zeta^R)$ is equivalent to $(Z_j^R(\mathbf{x}, \boldsymbol{\lambda}_j) \leqq \zeta^R$, where $\boldsymbol{\lambda}_j = r_{ji}(\boldsymbol{\lambda}_i)$, $\forall$ pairs $S_j(\subset S_i))$. Thus the statement $\boldsymbol{\lambda}_i \varepsilon \Lambda_i^R(\mathbf{x}, \zeta^R)$ is equivalent to the statement that $\boldsymbol{\lambda}_j \varepsilon \Lambda_j^R(\mathbf{x}, \zeta^R)$ where $\boldsymbol{\lambda}_j = r_{ji}(\boldsymbol{\lambda}_i)$, for all pairs $S_j(\subset S_i)$, and this establishes coherence and consonance.

Resolution may be defined and compared for SCS's analogously to the way it was for STP's in Section 4, above. However, a comparison of SCS's $[\Phi, \bar{\mathbf{Z}}, \zeta]$ and $[\Phi^*, \bar{\mathbf{Z}}^*, \zeta^*]$ has meaning only if the hypotheses on the functions Φ are the same as those on some of the functions of Φ^*. Thus, for example, for each value λ_0 of $\phi_0 \varepsilon \Phi$ there must exist another value λ_0^*, say, of $\phi^* \varepsilon \Phi^*$ such that $\omega_0(\lambda_0) = \omega_0^*(\lambda_0^*)$. If that is so and if $I_{\min} \subseteq I^*$ then the conditions

$$\text{(5.17a)} \qquad P_{\omega_0(\lambda_0)}(Z_0(Y; \lambda_0) \leqq \zeta) = P_{\omega_0^*(\lambda_0^*)}(Z_0^*(Y; \lambda_0^*) \leqq \zeta^*)$$

equally for all $\lambda_0 \varepsilon \Lambda_0$, and

$$\Lambda_j(y; \zeta) \subseteq \Lambda_j^*(y; \zeta^*) \quad \text{a.e.,} \tag{5.17b}$$

define $[\Phi, \bar{Z}, \zeta]$ to be *no less resolvent* than $[\Phi^*, \bar{Z}^*, \zeta^*]$. If the containment in (5.17b) is proper the former SCS is said to be *strictly more resolvent* than the latter. Thus, if two SCS's which can be compared have the same joint confidence co-efficient, but the minimal confidence regions of the one are contained in the corresponding regions of the other, the former is said to be more resolvent.

Again, let estimation families $[\Phi, \bar{Z}]$ and $[\Phi^*, \bar{Z}^*]$ have related hypotheses, and for each $\lambda_0 \varepsilon \Lambda_0$ and the corresponding λ_0^* such that $\omega_0(\lambda_0) = \omega_0^*(\lambda_0^*)$, let $\Omega_{\lambda_0} \subseteq \Omega^*_{\lambda_0^*}$ also hold. Then $[\Phi, \bar{Z}, \zeta]$ will be said to be *no less parsimonious* than $[\Phi^*, \bar{Z}^*, \zeta^*]$ if

$$\text{for all } j \varepsilon I_{\min} \quad \Lambda_j(y; \zeta) = \Lambda_j^*(y; \zeta^*) \quad \text{a.e.,} \tag{5.18a}$$

$$\text{for all } i \varepsilon I \quad \Lambda_i(y; \zeta) \subseteq \Lambda_i^*(y; \zeta^*) \quad \text{a.e.,} \tag{5.18b}$$

$$\text{and for } \omega_0(\lambda_0) = \omega_0^*(\lambda_0^*) \quad C \geqq C^*. \tag{5.18c}$$

Again, if the latter inequality is strict the former SCS is said to be *strictly more parsimonious* than the latter. In other words, for equal minimal confidence regions the more parsimonious SCS has all non-minimal confidence regions contained in the corresponding regions of the less parsimonious SCS, and its joint confidence co-efficient is larger. Parsimony must here be understood in the sense of increasing joint confidence, and that is equivalent to reducing the experimentwise error rate.

Further, estimation families with related hypotheses may allow narrowness comparisons. Thus $[\Phi, \bar{Z}]$ is said to be *narrower* than $[\Phi^*, \bar{Z}^*]$ if, for each λ_0 and the corresponding λ_0^* (i.e., $\omega_0(\lambda_0) = \omega_0^*(\lambda_0^*)$), $\{\Omega_{\lambda_0}, Z_{\lambda_0}\}$ is narrower than $\{\Omega^*_{\lambda_0^*}, Z^*_{\lambda_0^*}\}$. The requirements for this are that for each λ_0 there exist a function g_{λ_0} satisfying (4.3a, b, c, d) for the corresponding hypotheses. This function may or may not be the same for all $\lambda_0 \varepsilon \Lambda_0$. Thus, narrowness of estimation families is defined simply as narrowness of the testing families which they generate.

Again, the following theorem and corollaries may be stated in view of these further analogies between properties of STP's and SCS's.

THEOREM 4′. *If $[\Phi, \bar{Z}]$ is narrower than $[\Phi^*, \bar{Z}^*]$, then for every ζ^* there exists ζ such that $[\Phi, \bar{Z}, \zeta]$ is no less parsimonious than $[\Phi^*, \bar{Z}^*, \zeta^*]$. Moreover, the former* SCS *is strictly more parsimonious than the latter for some values of ζ^*,* i.e., *those for which*

$$P_{\omega_0(\lambda_0)}(Z_0^*(Y; \lambda_0) > \zeta^* \geqq g_{\lambda_0}(Z_0(Y; \lambda_0))) > 0,$$

where g_{λ_0} satisfies (4.3a, b, c, d) *for all $\lambda_0 \varepsilon \Lambda_0$.*

THEOREM 5′. *If $[\Phi, \bar{Z}]$ is narrower than $[\Phi^*, \bar{Z}^*]$ and if $[\Phi, \bar{Z}, \zeta]$ and $[\Phi^*, \bar{Z}^*, \zeta^*]$ have the same co-efficient, then the former* SCS *is no less resolvent than the latter. Moreover, it is strictly more resolvent for some values of ζ,* i.e., *those for which*

$$P_{\omega_0(\lambda_0)}(Z_0^*(Y; \lambda_0) > g_{\lambda_0}(\zeta) \geqq g_{\lambda_0}(Z_0(Y; \lambda_0))) > 0$$

where g_{λ_0} satisfies (4.3a, b, c, d) *for all $\lambda_0 \varepsilon \Lambda_0$.*

Clearly, narrower estimation families have the advantage of providing more resolvent SCS's with the same joint confidence co-efficient and more parsimonious SCS's for identical minimal confidence regions.

EXAMPLE 5.7. In a subset ANOVA set-up with equal sample sizes, compare the estimation families $[\Phi^S, \bar{Z}^R]$ of Example 5.6 and $[\Phi^S, \bar{Z}^F]$ of Example 5.2. The relation $Z^F = g(Z^R) = (Z^R)^2/2$ holds for all $Z_j^F(\mathbf{x}, \lambda_j)$ and $Z_j^R(\mathbf{x}, \lambda_j)$ if S_j is a pair. As in Example 4.4, above, so also here $Z_i^F(\mathbf{x}, \lambda_i) > [Z^R(\mathbf{x}, \lambda_i)]^2/2$ if S_i contains three or more expectations and these inequalities are strict a.e. Hence $[\Phi^S, \bar{Z}^R]$ is narrower than $[\Phi^S, \bar{Z}^F]$ and it follows from Theorem 4′ and 5′ that for all subsets the SCS's due to Tukey are more parsimonious or more resolvent than the corresponding subset SCS's based on Scheffé's approach.

EXAMPLE 5.8. The comparison in Example 5.7 between Tukey and Scheffé confidence statements for subsets of expectations also follows, by virtue of Corollary 3′, condition (I), from the UI and LR relations of the respective estimation families, and the fact that $Z_j^F(\mathbf{x}, \lambda_j) = [Z_j^R(\mathbf{x}, \lambda_j)]^2/2$ for all pairs S_j.

COROLLARY 3′. *Let* $[\Phi, \bar{Z}]$ *and* $[\Phi^*, \bar{Z}^*]$ *be monotone estimation families such that* $\Phi \subseteq \Phi^*$, $\Phi_{\min} \subseteq \Phi^*_{\min}$ *and* $\omega_0(\lambda_0) = \omega_0^*(\lambda_0^*)$ *for all* $\lambda_0 \varepsilon \Lambda_0$ *and let there exist for each* $\lambda_0 \varepsilon \Lambda_0$ *a function* g_{λ_0} *satisfying* (4.3a, b) *for the corresponding hypotheses. Then each of the following are sufficient conditions that* $[\Phi, \bar{Z}]$ *provide more parsimonious and more resolvent* SCS's *than* $[\Phi^*, \bar{Z}]$,

(I) *if* $[\Phi, \bar{Z}]$ *is* UI *related and* g_{λ_0} *also satisfy* (4.3e) *or* (4.3e′) *for each* λ_0,

(II) *if both* $[\Phi, \bar{Z}]$ *and* $[\Phi^*, \bar{Z}^*]$ *are* UI *related and*

$$P_{\omega_0(\lambda_0)}(\max \{Z_j^*(Y; r_{j0}(\lambda_0)) \mid j \varepsilon I_{\min}\} > \max \{Z_j^*(Y; r_{j0}(\lambda_0)) \mid j \varepsilon I_{\min}\}) > 0 \tag{5.19}$$

for all $\lambda_0 \varepsilon \Lambda_0$.

This corollary covers some results of Krishnaiah for ANOVA and MANOVA set-ups [13]. He uses UI related families throughout and his Theorem 6.2, 6.3 and 6.4 establish that for an identity function g and finite $I_{\min} \subset I^*_{\min}$, the minimal confidence regions of $[\Phi, \mathbf{Z}, \zeta]$ will be properly contained in those of $[\Phi^*, \bar{Z}^*, \zeta^*]$ where both SCS's have the same co-efficient. Apart from cases where (5.19) might be zero his conditions are seen to correspond to (II) of Corollary 3′, and his containment conclusion corresponds, again apart from possible sets of probability zero, to the greater resolution conclusion of Corollary 3′. He does not explicitly discuss anything corresponding to parsimony.

The following are two examples of Krishnaiah's results:

EXAMPLE 5.9. Dunnett's simultaneous $1 - \alpha$ confidence bounds on comparisons of the expectations under $(k - l)$ treatments with that under a control [5] are more resolvent than Tukey's simultaneous $1 - \alpha$ confidence bounds on all pairwise comparisons of these k expectations [27]. (Krishnaiah, [13], Corollary 6.3).

The corresponding comparison of STP's was given in Example 4.13.

EXAMPLE 5.10. Simultaneous SANOVA $1 - \alpha$ confidence bounds on a finite subset of contrasts in k expectations are more resolvent than Scheffé-type simultaneous $1 - \alpha$ confidence bounds on all contrasts (Krishnaiah [13], Theorem 6.2).

6. Other methods of multiple comparisons. The previous discussion relates almost entirely to STP's and SCS's which are based on monotone families, and thus provide coherent decisions. Some examples have been mentioned of procedures for non-monotone families in which case the decisions are not always coherent and the properties in the theorems do not apply. In particular, we may mention the use of the Kruskal-Wallis statistic for nonparametric ANOVA (Example 3.2) proposed by Nemenyi [17] (see also the Miller's discussion [16], Section 4.6). One would presume that incoherences, though possible with this technique, are pretty rare, especially with large samples. In that case such a technique might be said to be asymptotically coherent. This way warrant further investigation.

Other techniques have a step-wise procedure built in which ensures coherence but destroys simultaneity. Thus, Newman [18] and Keuls [10], in the equal sample size ANOVA, decide on each subset of means according to whether the studentized range for that subset *or any other set containing it* is α significant in the distribution of its range. Duncan [3] proceeds in a more intricate way, choosing a different percentage point for each one of the ranges. He also extended the method to F-statistics [4].

In these methods the decision on a particular ω_i depends not only on the statistic Z_i but on all Z_j such that $j < i$. Clearly, this does not provide a testing family in the sense of Section 2 and the STP results do not apply. These methods have the disadvantage of making the decision on ω_i depend in part on statistics Z_j whose distribution depends not only on ω_i but also on $\omega_j - \omega_i$ which is irrelevant to ω_i. On the other hand, these methods offer better resolution—in the sense of Section 4—than STP's and are therefore preferred by some workers. (For a detailed discussion see [6], Section 9.)

Acknowledgment. I am indebted to my colleagues in Jerusalem and Chapel Hill with whom I have discussed these techniques. I am particularly happy to acknowledge the help of J. Putter who critically read an early version of the paper and that of W. J. Hall who made many useful suggestions in the last revisions of this paper.

REFERENCES

[1] AITCHISON, J. (1964). Confidence-region tests. *J. Roy. Statist. Soc. Ser. B* **26** 462–476.

[2] AITCHISON, J. (1965). Likelihood-ratio and confidence region tests. *J. Roy. Statist. Soc. Ser. B* **27** 245–250.

[3] DUNCAN, D. B. (1951). A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.* **2** 171–189.

[4] DUNCAN, D. B. (1955). Multiple range and multiple F-tests. *Biometrics* **13** 164–176.

[5] DUNNETT, CHARLES W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50** 1096–1121.

[6] Gabriel, K. R. (1964). A procedure for testing the homogeneity of all sets of means in analysis of variance. *Biometrics* **20** 459–477.

[7] Gabriel, K. R. (1966) Simultaneous test procedures for multiple comparisons on categorical data. *J. Amer. Statist. Assoc.* **61** 1081–1096.

[8] Gabriel, K. R. (1968). Simultaneous test procedures in multivariate analysis of variance. *Biometrika* **55**.

[9] Gabriel, K. R. (1968). On the relation between union intersection and likelihood ratio tests. To appear in *S. N. Roy memorial volume* (R. C. Bose, ed.).

[10] Keuls, M. (1952). The use of the 'studentized range' in connection with an analysis of variance. *Euphytica* **1** 112–122.

[11] Knight, William (1965). A lemma for multiple inference. *Ann. Math. Statist.* **36** 1873–1874.

[12] Krishnaiah, P. R. (1964). Multiple comparison tests in multivariate case. U.S.A.F. Office of Aerospace Research, ARL 64–124.

[13] Krishnaiah, P. R. (1965). On the simultaneous Anova and Manova tests. *Ann. Inst. Statist. Math.* **17** 35–53.

[14] Lehmann, E. L. (1957). A theory of some multiple decision problems, I. *Ann. Math. Statist.* **28** 1–25.

[15] Lehmann, E. L. (1957). A theory of some multiple decision problems, II. *Ann. Math. Statist.* **28** 547–572.

[16] Miller, Rupert G. (1966). *Simultaneous Statistical Inference.* McGraw-Hill, New York.

[17] Nemenyi, Peter (1963). *Distribution-free Multiple Comparisons.* (mimeographed).

[18] Newman, D. (1939). The distribution of the range in samples from the normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika* **31** 20–30.

[19] Rao, C. Radhakrishna (1965). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[20] Roy, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

[21] Roy, S. N. (1961). A survey of some recent results in normal multivariate confidence bounds. *Proc. of the 33rd Ses. of the Interna. Statist. Inst., Paris.*

[22] Roy, S. N. and Bose, R. C. (1953). Simultaneous confidence interval estimation. *Ann. Math. Statist.* **24** 513–536.

[23] Roy, S. N. and Gnanadesikan, R. (1959). Some contributions to Anova in one or more dimensions: I. *Ann. Math. Statist.* **30** 304–317.

[24] Roy, S. N. and Shrivastava, J. N. (1961). Inference on treatment effects and design of experiments in relation to such inference. Univ. of North Carolina, Inst. of Statist., Mimeo Series, No. 274.

[25] Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika* **40** 87–104.

[26] Scheffé, Henry (1959). *The Analysis of Variance.* Wiley, New York.

[27] Tukey, J. W. (1951). Quick and dirty methods in statistics. Part II. Simple analyses for standard designs. *Proc. Fifth Annual Conven. Amer. Soc. Quality Control,* 189–197.

[28] Tukey, J. W. (1953). The problem of multiple comparisons. Unpublished manuscript.