MASTER

MULTIPLE THREE-DECISION RULES
FOR 2^k -FACTORIAL SIMPLE EFFECTSR. Bohrer, W. Chow, R. Faith,
V.M. Joshi, C.F. WuTECHNICAL REPORT NO. 8
March 22, 1977PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA There is no objection from the patent
point of view to the publication or
dissemination of the document(s)
listed in this letter.

BROOKHAVEN PATENT GROUP

1/12/78 By *ckl*Prepared For
THE U.S. ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
UNDER CONTRACT NO. EY-76-S-02-2874

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

MULTIPLE THREE-DECISION RULES FOR 2^k -FACTORIAL SIMPLE EFFECTS*

R. Bohrer, W. Chow, R. Faith, V.M. Joshi, C.F. Wu
Universities of Illinois (Urbana), California (Berkeley),
Stanford, Michigan, and California (Berkeley)

TECHNICAL REPORT NO. 8

MARCH 22, 1977

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH

PREPARED UNDER SUPPORT TO SIMS FROM
ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION (ERDA)
ROCKEFELLER FOUNDATION

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

* This paper was prepared with the joint support of the National Institutes of Health, research grant no. ES 01299-13; the U.S. Office of Naval Research, contract no. N 00014-75-C-0444/NR042-036 (approved for public release, distribution unlimited); by the auspices of the SIAM Institute for Mathematics and Society under support from the Energy Research and Development Administration and the Rockefeller Foundation; and NSF Grant MCS-77-07978.

28
DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ABSTRACT

One of the problems confronted in the investigation of environmental health and in biological research more generally is the fact that organisms often respond to combinations of "treatments" in ways which are not predictable from the way they respond to each of these taken separately. One air pollutant may, for example, disable the respiratory tract's self cleaning mechanisms and thereby increase its sensitivity to the effects of other pollutants. It is important that laboratory experiments be able to simultaneously detect the effects of the various possible combinations of treatments in a way that utilizes the data as effectively as possible. This paper deals with the optimal use of data from such experiments when the response variable can be regarded (possibly after undergoing a suitable transformation) as having a normal distribution with mean determined by the combination of treatments it has undergone.

The problem is that of making simultaneous inferences about the signs of "simple effects" in factorial experiments, as defined by Neyman [1935b]. The goal of bounding the probability of making no incorrect decisions about the parameters of concern is seen to require a three-decision rule. One such rule, $\{\phi_{\tilde{x}}^c\}$ is defined, and two fashions in which it is optimal are cited. To implement this rule in practice it is necessary to find the value of c for which the desired bound on the error probability is achieved. Existing, well-tabled methods for finding conservative bounds for c are reviewed. Use and refinement of the Bonferroni inequality are developed. Exact evaluation of c is accomplished for some simple designs, and this is used to assess the conservatism of the other bounds. Numerical comparisons of the several methods are given.

MULTIPLE THREE-DECISION RULES FOR 2^k -FACTORIAL SIMPLE EFFECTS*

R. Bohrer, W. Chow, R. Faith, V.M. Joshi, C.F. Wu
Universities of Illinois (Urbana), California (Berkeley)
Stanford, Michigan, and California (Berkeley)

1. INTRODUCTION

In many biological (and other) experiments, an observed response is a function of several different factors, each of which may be either present or absent when the response is observed. In such a "factorial" experiment, there are 2^k different possible combinations of the levels of the k factors, the i^{th} factor being at level 1 if it is present and at level 0 if it is absent. In the general situation, the expected response to a combination of j factors present is not additively predictable from

* This paper was prepared with the joint support of the National Institutes of Health, Research Grant No. ES 01299-13; the U.S. Office of Naval Research, Contract No. N 00014-75-C-0444/NR042-036 (approved for public release, distribution unlimited); and by the auspices of the SIAM Institute for Mathematics and Society under support from the Energy Research and Development Administration and the Rockefeller Foundation, and NSF Grant MCS-77-07978.

the j "simple" responses to the application of each of the factors taken separately. In such cases, the biologist says that the factors are synergistic, and the statistician says that there are interactions among the factors.

To fix ideas, consider a simple case where two chemical agents are factors which affect the response of tumor genesis. There may be cases where the two agents are catalysts for each other, i.e., where the expected tumor growth is zero when either drug is administered by itself, but where the expected growth is great when the two agents are given together. Other notable examples of synergisms are the multiplicative effect of the depressants alcohol and barbituates when taken together and the increased effect of the blood anticoagulant warfarin when administered together with chloramphenicol.

Based on the data collected from a factorial experiment, it is often important to decide whether the parameters which express the expected response values are positive or negative. For example, one may need to determine whether or not certain combinations of drugs increase or decrease the expected blood viscosity when administered together. The present work describes a rule for making such decisions, and it evaluates both the quantities necessary to implement the rule in practice and the performance characteristics of the rule.

In section 2, the basic framework, assumed throughout the paper, is defined. In section 3, alternative parametrizations are introduced which are often more natural in practice. The rule $\{\phi_x^c\}$, to be considered, is defined in section 4; and ways in which it is optimal are cited. To use

this rule in practice, the value of c must be determined so that the error probability is as low as required. Existing techniques which yield conservative values for c are cited, with numerical evaluations in section 5. These include the use of the Bonferroni inequality in this regard, together with its refinement for certain cases. For some special cases, exact values of c can be determined using arguments of geometrical and computational probability, as done in section 6. All these results are collected and compared numerically, to give practical information on the use of the rule, in section 7.

The "one parameter at a time" version of this problem was considered by Neyman [1935a,b], who evaluated the required c values and performance characteristics for that case. As described later, optimal decisions are available for "main effects" following Bohrer [1976]; but this work provides only conservative bounds for "simple effects," for which the present work provides and analyzes other conservative approximations, as well as some exact results.

2. FRAMEWORK AND BEST ESTIMATION FOR THE 2^k FACTORIAL EXPERIMENT

A "treatment" in the 2^k factorial experiment is any one of the 2^k different possible combinations of the k factors in which each factor is either present or absent. These will be denoted by vectors $\underline{v}$ consisting of k 0's and 1's, v_i being 1 if the i^{th} factor is present and 0 if not. Let V denote the set of the 2^k possible treatment vectors. For each treatment combination $\underline{v}$ in V , the experimental data consist of $n_{\underline{v}}$ observations, $Y_{\underline{v},i}$, $i=1(1)n_{\underline{v}}$. Denote the total sample size by $N = \sum_{\underline{v}} n_{\underline{v}}$. The most general parametrization considered here for the expectations of

these random variables is

$$(1) \quad E Y_{\underline{v},i} = M_{\underline{v}}.$$

Other assumptions which will be made throughout are the "usual" assumptions for fixed effects analysis of variance; see section 2.1 in Scheffé [1959]. Namely, assume that all the N observations are normally distributed, mutually independent, with the same unknown variance σ^2 .

Under these assumptions, uniformly minimum variance unbiased (UMVU) estimators for the parameters $M_{\underline{v}}$ are the cell sample averages $\hat{M}_{\underline{v}} = \sum_i Y_{\underline{v},i} / n_{\underline{v}}$, which are thus distributed independently and normally, $N(M_{\underline{v}}, \sigma^2 / n_{\underline{v}})$. Moreover, these are distributed independently from the unbiased estimator for σ^2 given by

$$(2) \quad S^2 = \sum_{\underline{v}} \sum_i (Y_{\underline{v},i} - \hat{M}_{\underline{v}})^2 / (N - 2^k);$$

and $(N - 2^k) S^2 / \sigma^2$ is $\chi^2_{N - 2^k}$.

Note, also from the standard theory, that the UMVU estimator for any linear function $\underline{x}' \underline{M}$, by which we denote $\sum_{\underline{v}} \underline{x}_{\underline{v}} \hat{M}_{\underline{v}}$, is the corresponding linear function of the estimators $\hat{M}_{\underline{v}}$, viz. $\hat{\underline{x}}' \hat{\underline{M}} = \sum_{\underline{v}} \underline{x}_{\underline{v}} \hat{M}_{\underline{v}}$, which is distributed normally, $N(\underline{x}' \underline{M}, \sigma_{\underline{x}}^2 \sigma^2)$, independently of S^2 ; here

$\sigma_{\underline{x}}^2 = \sum_{\underline{v}} \underline{x}_{\underline{v}}^2 / n_{\underline{v}}$. The covariance between any two such estimators, $\hat{\underline{x}}_1' \hat{\underline{M}}$ and $\hat{\underline{x}}_2' \hat{\underline{M}}$, is $\sum_{\underline{v}} \underline{x}_{1,\underline{v}} \underline{x}_{2,\underline{v}} \sigma^2 / n_{\underline{v}}$ and they are jointly normal.

3. ALTERNATIVE PARAMETERIZATIONS: MAIN EFFECTS, SIMPLE EFFECTS, AND INTERACTIONS

The parameters $M_{\tilde{y}}$, as noted above, define expectations of all random variables observed in a factorial experiment. The objective of such an experiment, on the other hand, is usually to assess how the various factors act separately and interact (or synergize) together. This is difficult to deduce from the values of the 2^k parameters $M_{\tilde{y}}$, whereas it is much easier to deduce from properly defined, linear functions of the $M_{\tilde{y}}$. Two such sets of linear functions, which have been found useful in this regard are 1) main effects and interactions, and 2) simple effects and simple interactions.

3.1. Main effects and interactions. Perhaps the most common such parametrization is that of main effects and interactions, defined by Yates [1935]. For deciding the signs of these parameters, using the rule of this paper, full details and exact computations are given by Bohrer [1976]. For use in motivating and relating these to simple interactions and effects, defined below, definitions of these parameters are given here for the 2^2 and 2^3 cases. Note that here, as in general case, the main effect of a factor is the average of the 2^{k-1} expectations wherein a factor is present minus the average of those 2^{k-1} expectations where it is absent.

In the 2^2 case, the main effects are:

$$\text{Main}_1 = (M_{11} + M_{10} - M_{01} - M_{00})/2 = ((M_{11} - M_{01}) + (M_{10} - M_{00}))/2$$

and

$$\text{Main}_2 = (M_{11} + M_{01} - M_{10} - M_{00})/2 = ((M_{11} - M_{10}) + (M_{01} - M_{00}))/2.$$

The interaction is $I_{12} = M_{11} + M_{00} - M_{10} - M_{01}$.

For the 2^3 case, the main effects are

$$\text{Main}_1 = ((M_{111} - M_{011}) + (M_{110} - M_{010}) + (M_{101} - M_{001}) + (M_{100} - M_{000}))/4 ,$$

$$\text{Main}_2 = ((M_{111} - M_{101}) + (M_{110} - M_{100}) + (M_{011} - M_{001}) + (M_{010} - M_{000}))/4 ,$$

$$\text{Main}_3 = ((M_{111} - M_{110}) + (M_{101} - M_{100}) + (M_{011} - M_{010}) + (M_{001} - M_{000}))/4 .$$

The two-factor interactions are

$$I_{12} = (M_{111} + M_{001} - M_{101} - M_{011}) + (M_{110} + M_{000} - M_{100} - M_{010}) ,$$

$$I_{13} = (M_{111} + M_{010} - M_{110} - M_{011}) + (M_{101} + M_{000} - M_{100} - M_{001}) ,$$

$$I_{23} = (M_{111} + M_{100} - M_{110} - M_{101}) + (M_{011} + M_{000} - M_{010} - M_{001}) .$$

The three-factor interaction is defined by

$$I_{123} = (M_{111} + M_{001} - M_{101} - M_{011}) - (M_{110} + M_{000} - M_{100} - M_{010}) .$$

See Yates [1935] for these definitions. Note, however, that the parameters used are often defined as positive multiples of those given above; see, for example, Neyman and Yates [1936]. Throughout the present work, it makes no difference which such multiple is used, since inferences about these parameters are functions of estimators for the parameters divided by estimators of their standard deviations, which thus adjusts for such multipliers. Note also that the groupings of the $M_{\underline{y}}$ in the formulae above is done to facilitate comparison with the simple effects and simple interactions in section 3.4.

3.2. The cases for simple effects and simple interactions. When interactions, as defined above, are present, the main effect of a factor gives little indication of how the factor would be expected to behave in

the presence of any given other factors. For example, consider the 2^2 case where the response is lifetime after administration of treatment by neither, one, or both of two drugs (factors). Suppose the drugs are individually detrimental but positively synergistic as follows: $M_{00} = 10$, $M_{10} = M_{01} = 0$, and $M_{11} = 50$. Thus, the main effect of each factor is 20 or twice the size of the expected lifetime without treatment (M_{00}), and yet the individual treatments are real killers.

It is, however, standard statistical practice to be careful, if not silent, about interpreting main effects in the presence of interaction. In this case, different and more functions of the $M_{\underline{y}}$ are necessary to answer questions about how particular factors do in the presence of particular other factors. Such parameters are the simple effects and simple interactions, thus motivated and defined.

The case concerning parametrization in terms of main effects and interactions is even worse, however. For Neyman's [1935b] Monte Carlo study shows that there are cases when statistical tests are rather likely to accept all hypotheses of zero interaction and to indicate that exactly one factor has positive main effect, whereas in fact the indicated factor does worse than nothing if applied alone. Therefore, whether or not interactions appear to be present, one is led to define and to consider simple effects and interactions, as in the following subsections.

3.3. Simple effects and simple interactions. These are defined recursively, beginning with simple effects, which are also known as simple one-factor interactions henceforth. For a given set of factors, many simple interactions are defined, since they interact differently in the presence of different other factors.

Definition. The simple effect, or simple one-factor interaction, of factor A in the presence of other factors $B_1, \dots, B_f$, denoted $S(A; B_1, \dots, B_f)$, is the difference in expected responses when i) only factors $A, B_1, \dots, B_f$ are present, and when ii) only factors $B_1, \dots, B_f$ are present. At the risk of rigor-mortis, if $\tilde{v}_1$ is the vector in V with ones in positions $A, B_1, \dots, B_f$ and $\tilde{v}_2$ is the vector in V with ones in positions $B_1, \dots, B_f$, then $S(A; B_1, \dots, B_f) = M_{\tilde{v}_1} - M_{\tilde{v}_2}$. Note that the set of factors $B_1, \dots, B_f$ may be empty, in which case, the simple effect $S(A; \emptyset)$ is just the expected response when A is the only factor used minus the expected response from the "control", i.e., when no factors are present; this is called the isolated effect of factor A .

Definition. The simple f -factor interaction of factors $A_1, \dots, A_f$ in the presence of other factors $B_1, \dots, B_g$ is denoted $S(A_1, \dots, A_f; B_1, \dots, B_g)$. These are defined recursively by the formula

$$(3) \quad S(A_1, \dots, A_f; B_1, \dots, B_g) = S(A_1, \dots, A_{f-1}; B_1, \dots, B_g, A_f) - S(A_1, \dots, A_{f-1}; B_1, \dots, B_g).$$

Simple interactions among a given number f of factors are thus differences of simple interactions among $f-1$ factors, and they are defined differently, depending on which other factors B_i are present.

Next it is shown that the simple interaction among several factors does not depend on the order in which the factors are listed. For example, the simple interaction of A_1 and A_2 is the same as that of A_2 and A_1 . The proof of this proceeds by induction, beginning with the straightforward

verification for cases of one- and two-factor interactions. For $f \geq 3$, suppose the claim is true at all preceding stages. Let $A_1, \dots, A_f$ and $A'_1, \dots, A'_f$ denote two permutations of the same f factors, and let $B_1, \dots, B_g$ denote any other factors. If $A_f = A'_f$, then the result follows from the induction hypothesis. If $A_f \neq A'_f$, then, using the permutation of the A-factors excluding A_f and A'_f , it follows that $S(A_1, \dots, A_f; B_1, \dots, B_g) = S(A''_1, \dots, A''_{f-2}, A_f, A'_f; B_1, \dots, B_g)$ and $S(A''_1, \dots, A''_f; B_1, \dots, B_g) = S(A''_1, \dots, A''_{f-2}, A'_f, A_f; B_1, \dots, B_g)$. The result follows by noting, as in the verification for $f = 2$, that both right-hand sides are defined by the same sums and differences of simple $f-2$ factor simple interactions. It should be noted that the ordering of $B_1, \dots, B_g$ is irrelevant as well.

Since, among k factors, there are ${}_k C_f$ combinations of f factors, and since there are 2^{k-f} ways in which the remaining factors can be present or absent, this shows that there are exactly ${}_k C_f 2^{k-f}$ simple f -factor interactions. For example, in the 2^2 factorial experiment, there is only one two-factor simple interaction, which is the same as the Yates two-factor interaction in this case; and there are four simple effects (one-factor interactions), viz., those of each factor in the presence of the other factor and in its absence. In the 2^3 case, there are 12 simple effects, 6 simple two-factor interactions, and one three-factor simple interaction.

3.4. Relations among the parametrizations. In the sense of the following general result and specific examples, simple effects and simple interactions give more specific information about the performance of several factors in specific combinations than do the main effects and interactions,

which are seen to be averages or sums of the former. Thus, as in the first example in section 3.2, the latter parameters may "average out". An advantage of main effects and interactions is that they are fewer in number; for example, in the 2^2 case, there are 5 simple effects and simple interactions, but only 3 main effects and interactions, while in the 2^3 case, there are 19 of the former and only 7 of the latter. But, what profiteth a statistician if he should gain fewer parameters and lose the ability to interpret them. On the other hand, if higher-order interactions are zero, then lower-order ones, or even main effects, may have simple interpretations. However, beware the second example of section 3.2.

It is true, quite generally, that main effects are averages of simple effects. Namely, all the 2^{k-1} simple effects for a given factor average to give its main effect. For $k = 2$ or 3 , comparison of the definitions of parameters in section 3.1 and 3.3 shows that the single, order k interactions are identical, and that the 2-factor interaction of section 3.1 is, for any pair of factors, exactly the sum of two two-factor simple interactions for that pair of factors. The refinement of the simple effects, simple interactions approach is that it permits one to determine how the synergistic interplay of a set of factors $A_1, \dots, A_F$ may depend on what other factors $B_1, \dots, B_g$ are being held fixed in various combinations, rather than simply combining all these interactions into one global "average" interaction of $A_1, \dots, A_F$.

4. THE RULE $\{\phi_x^C\}$

In this section, the goal of the decision rule is defined. It is seen that a three-decision rule is generally required to attain the goal,

one such rule is defined, and results concerning its performance and optimality are cited.

4.1. The goal and the necessity of three-decision rules. On the basis of the observations from a factorial experiment, it is required to decide whether each of the parameters in some subset $\mathcal{S}$ of the simple interactions is positive or negative. We arbitrarily adopt the convention here that a parameter is called positive if it is zero. The same rule would result if zero were considered negative. Subject to the requirement that, for some preassigned value of α and for all (M, σ^2) ,

$$(4) \quad P_{M, \sigma^2}^{\mathcal{S}} (\text{one or more decisions is incorrect}) \leq \alpha,$$

the expected number of correct decisions made should be large.

For example, if $\mathcal{S}$ is the set of all "simple one-factor interactions in the presence of nothing else" (that is, isolated effects), then one is trying to decide, as often as possible, whether each of the factors is as good as (positive) or worse than (negative) nothing, subject to the constraint that less than 100α percent of the time should there be even one error among all the k decisions.

However, consideration of the simplest case, of classifying a single parameter when the variance is known, shows, with the help of the Neyman Pearson lemma argument, that no rule which allows only the decisions "positive" and "negative" can achieve an arbitrary error probability bound α . In fact, the best rule is to decide that the parameter is positive if its least squares estimate is positive (and negative if the estimate is negative); and this rule has error probability bound of .5,

attained in the limit as the true parameter approaches zero. To retain our error requirement (4) thus requires adoption of three-decision rules, where the decision "no decision" is added to those of "positive" and "negative". The "no decision" is not a wrong decision in calculating (4); but neither is it a correct decision in the goal to maximize the number of correct decisions.

4.2. The rule $\{\phi_x^c\}$. A general three-decision rule of this type is defined by Bohrer [1976] and can be applied in this case. Note that the rule must make a separate decision for each of the parameters in the set $\mathcal{S}$ and that each of these parameters is a particular linear function $\tilde{x}'\tilde{M}$ of the cell means $\tilde{M}$; see section 2. More explicitly, then, the rule is $\{\phi_x^c: \text{all } \tilde{x} \text{ such that } \tilde{x}'\tilde{M} \text{ is in } \mathcal{S}\}$, and the rule for the parameter $\tilde{x}'\tilde{M}$ is given by:

$$\phi_x^c \text{ decides } \begin{cases} \text{"positive" if } \hat{\tilde{x}'\tilde{M}} \geq cS\hat{\sigma}_{\tilde{x}} \\ \text{"negative" if } \hat{\tilde{x}'\tilde{M}} \leq -cS\hat{\sigma}_{\tilde{x}} \\ \text{"no decision" otherwise.} \end{cases}$$

The rule $\{\phi_x^c\}$ thus decides the sign of a parameter only if its estimate is at least a multiple c of its sample standard deviation away from zero.

Two essential questions about these rules need to be answered. "Are the rules any good?" is answered at least mildly affirmatively in sections 4.3 and 4.4. The critical practical question, "What value of c is required to satisfy (4)?", is the subject of sections 5-8.

4.3. Probabilities of correct classification. By the definition of the decision rule ϕ_x^c for deciding about the parameter $\tilde{x}'\tilde{M}$, and by the definition of the non-central Student t distribution (see Scheffé [1959]), it

follows that, as a function of true value of $(\underline{M}, \sigma^2)$,

$$(5) \quad P_{\underline{M}, \sigma^2}(\phi_{\underline{x}}^c \text{ decides correctly about } \underline{x}'\underline{M}) \\ = P(t_{m; \delta} \geq c) ,$$

where $t_{m; \delta}$ is a Student t random variable with m degrees of freedom and non-centrality parameter δ , where $m = N - 2^k$ and $\delta = |\underline{x}'\underline{M}| / \sigma \sigma_{\underline{x}}$. These probabilities increase to unity as the parameter value $\underline{x}'\underline{M}$ departs from zero without limit in either direction. These non-central t probabilities are available in the tables of Resnikoff and Lieberman [1953] and the algorithm of Cooper [1968].

4.4. Optimality of the rule $\{\phi_{\underline{x}}^c\}$. Two general results exist in the literature which show two senses in which the rule $\{\phi_{\underline{x}}^c\}$ is optimal for this problem. However, for each of these, the error bound criterion (4) is replaced by the related criterion that, for a fixed error bound γ ,

$$(6) \quad \sum_{\mathcal{J}} P_{\underline{M}, \sigma^2}(\phi_{\underline{x}}^c \text{ decides incorrectly about } \underline{x}'\underline{M}) \leq \gamma ,$$

for all $(\underline{M}, \sigma^2)$. Note that the Bonferroni inequality (q.v.) shows that (6) is an upper bound for (4) and that this bound is approached as the chance of making two or more errors approaches zero.

The first optimality result has a minimax character as derived by Spjøtvoll [1972] and applied to this case by Bohrer [1976]. For any $c > 0$, and $\delta > 0$, let D denote the set of $(\underline{M}, \sigma^2)$ values for which each of the parameters $\underline{x}'\underline{M}$ in $\mathcal{J}$ is at a distance of at least $\delta \sigma_{\underline{x}}$ away from 0 in either direction.

Among rules which satisfy (6) and are unbiased, in the sense that their classification probabilities are at least as good when $\underline{M} \neq 0$

as when $\underline{M} = \underline{0}$, the rule of the form $\{\phi_{\underline{x}}^c\}$, with c chosen to satisfy (6), maximizes (over all such rules) the minimum (over D) expected number of correct decisions made.

The second optimality result, which follows from Lehmann's [1957a,b] work on multiple decision rules, is also concerned with the sum of the error probabilities. For each parameter $\underline{x}'\underline{M}$ in $\mathcal{S}$, suppose one incurs loss 0 if the correct decision is made, a if no decision is made, and $a+b$ if an incorrect decision is made, where a and b are positive constants. Let the overall loss $L(\underline{M}, \phi)$ for the problem be the sum over $\mathcal{S}$ of the losses for each component decision. Among rules which are unbiased in the sense that

$$E_{\underline{M}, \sigma^2}[L(\underline{M}, \phi)] \leq E_{\underline{M}', \sigma^2}[L(\underline{M}', \phi)]$$

for all $\underline{M}, \sigma^2$, and $\underline{M}'$, the rule $\{\phi_{\underline{x}}^c\}$ minimizes the expected loss uniformly for all values of $\underline{M}$ and σ^2 . The critical value c is determined by the relative sizes of a and b and by the condition of unbiasedness, which is a requirement that the rule not favor some combinations of the parameter signs at the expense of others. It follows from the foregoing that among rules which are unbiased in the present sense, for a given value of the left side of (6) the rule $\{\phi_{\underline{x}}^c\}$ maximizes the expected number of decisions made.

5. INEQUALITIES FOR CONSERVATIVE RULES

5.1. A basic inequality. The complement of the event in (4) is

$$(7) \quad (-\text{sgn } \underline{x}'\underline{M}) \underline{x}'\hat{\underline{M}} \leq c\sigma_{\underline{x}}, \quad \text{for all } \underline{x}'\underline{M} \text{ in } \mathcal{L}.$$

Thus, (4) requires that $P\{(7)\} \geq 1-\alpha$.

In succeeding subsections, the following inequalities will be useful:

$$\begin{aligned} & P_{\underline{M}, \sigma^2} \{ (-\text{sgn } \underline{x}'\underline{M}) \underline{x}'\hat{\underline{M}} \leq c\sigma_{\underline{x}}, \text{ all } \underline{x} \text{ with } \underline{x}'\underline{M} \text{ in } \mathcal{L} \} \\ &= P_{\underline{M}, \sigma^2} \{ (-\text{sgn } \underline{x}'\underline{M}) \underline{x}'(\hat{\underline{M}} - \underline{M}) \leq c\sigma_{\underline{x}} + |\underline{x}'\underline{M}|, \text{ all } \underline{x} \text{ with } \underline{x}'\underline{M} \text{ in } \mathcal{L} \} \\ (8) \quad & \geq P_{\underline{M}, \sigma^2} \{ (-\text{sgn } \underline{x}'\underline{M}) \underline{x}'(\hat{\underline{M}} - \underline{M}) \leq c\sigma_{\underline{x}}, \text{ all } \underline{x} \text{ with } \underline{x}'\underline{M} \text{ in } \mathcal{L} \} \\ (9) \quad & \geq P_{\underline{M}, \sigma^2} \{ |\underline{x}'(\hat{\underline{M}} - \underline{M})| \leq c\sigma_{\underline{x}}, \text{ all } \underline{x} \text{ with } \underline{x}'\underline{M} \text{ in } \mathcal{L} \}. \end{aligned}$$

Note that the inequalities (8) and (9) remain valid if the set " $\underline{x}$ with $\underline{x}'\underline{M}$ in $\mathcal{L}$ " is replaced by any class of vectors which contains this set.

The right side of (8), which equals $P_{\underline{0}, \sigma^2} \{ (-\text{sgn } \underline{x}'\underline{M}) \underline{x}'\hat{\underline{M}} \leq c\sigma_{\underline{x}}, \text{ all } \underline{x} \text{ with } \underline{x}'\underline{M} \text{ in } \mathcal{L} \}$, is the minimum probability of making no errors for given values of the parameter signs $\text{sgn}(\underline{x}'\underline{M})$, taking $k_{\underline{x}} = \sigma_{\underline{x}}$. Equality in (8) is obtained in the limit as $\underline{M}$ approaches zero from any direction.

5.2. The S-Method. Let B be a smallest subset of $\mathcal{L}$ such that every combination of means in $\mathcal{L}$ is a weighted sum of mean combinations in B . Denote by $d(\mathcal{L})$ the number of parameters in B and by L the set of all linear combinations of parameters in B . Now consider inequality (9) with $k_{\tilde{x}} = \sigma_{\tilde{x}}$, $c = (d(\mathcal{L})F_{\alpha; d(\mathcal{L}), N-2^k})^{1/2}$, and, on the right-hand side, $\mathcal{L}$ replaced by L . The left-hand side is exactly $P_{M, \sigma^2} \{ (7) \}$ for the given value of c . The right-hand side is exactly what Scheffé's ([1959], section 3.5) S-method evaluates as $1-\alpha$. This says that if c is so defined, then the left-hand side of (4) is at most α . This c -value can thus be used conservatively.

5.3. Sharpening the S-Method. The inequality (8) is sharper than (9), and thus it can be used to sharpen the conservatism of the c -value from section 5.2. This is done by Bohrer [1976], where c -values are still conservative for inferring about finitely many parameters, since $\mathcal{L}$ is still replaced by the larger set L . These sharpened, conservative c -values are tabled by Hochberg and Quade [1975], whose k and v are the $d(\mathcal{L})$ and $N-2^k$ of this paper.

5.4. The T-Method and a modified rule. For balanced designs, where $n_{\tilde{v}} = 2^{-k}N = n$ (say), Tukey's T-Method (see Scheffé [1959], section 3.6) can be used with inequality (9) to define a different rule, which also satisfies (4) conservatively. This rule is like $\{\phi_{\tilde{x}}^c\}$, only it replaces $\sigma_{\tilde{x}}$ by $n^{-1/2} \sum_{\tilde{v}} |x_{\tilde{v}}|/2$; and it uses $c = q_{\alpha; 2^k, N-2^k}$, the upper α point of the Studentized range distribution with parameters 2^k , and $N-2^k$. With $\mathcal{L}$ replaced by the larger set of all "contrasts"

($\underline{x}$ whose components sum to zero), the right-hand side of (9) is seen to be the probability $1-\alpha$ evaluated by the T-method.

5.5. The Bonferroni Method. As noted and used by Dunn [1958] such multiple comparison probabilities as (4) are often well-approximated by the Bonferroni inequality:

$$(10) \quad P(\cap_i E_i) \geq 1 - \sum_i P(\bar{E}_i).$$

The present application takes E_i as an event " $(-\text{sgn } \underline{x}'\underline{M}) \underline{x}'\hat{\underline{M}} \leq c\sigma_{\underline{x}}$ " and P as P_{Q, σ^2} , so that the left-hand side of (10) is exactly the right-hand side of (8). It follows at once from (10) that (4) is satisfied conservatively if each of the $P(\bar{E}_i) = \alpha/m$ where m is the number of parameters in the set $\mathcal{A}$. This may be done uniformly for all i using a single c -value since $P(E_i)$ does not depend on the value of $\text{sgn}(\underline{x}'\underline{M})$ or of $\underline{x}$, but only on c . A conservative c -value is thus obtained as $c = t_{\alpha/m; N-2^k}$, the upper α/m point of the Student t distribution with $N-2^k$ degrees of freedom.

5.6. A refined Bonferroni inequality for the 2 factor case with known σ^2 . The method of the preceding section can be improved somewhat by utilizing the relation

$$(11) \quad P(\cap_i E_i) \geq 1 - \sum_i P(\bar{E}_i) + \sum_{i < j} P(\bar{E}_i \cap \bar{E}_j) - \sum_{i < j < k} P(\bar{E}_i \cap \bar{E}_j \cap \bar{E}_k),$$

which gives a closer bound than does (10). The evaluation of the right side is substantially more complicated than that of (10) however, even

in the simple case of the 2-factor experiment with known variance σ^2 .

For $\mathcal{L}$ consisting of all four simple effects and the simple interaction,

there are 10 terms in the double summation and 10 in the triple summation of (11). As shown below in section 6.1, there are three

distinct combinations of "sgn $\tilde{x}_i M_i$, $\tilde{x}_i$ in $\mathcal{L}$ " to be considered. Of these, case C may be disregarded since $P(\cap_i E_i)$ is maximum for this case. A lower bound for the right side of (11) is obtained for cases A and B by bounds for $P(\bar{E}_i \cap \bar{E}_j)$ and upper bounds for $P(\bar{E}_i \cap \bar{E}_j \cap \bar{E}_k)$. The bounds used are the following:

$$(12) \quad P(Z_1 > c, Z_2 > c) \geq P(Z_1 > c)P(Z_2 > c(\frac{1-\rho}{1+\rho})^{1/2})$$

if Z_1 and Z_2 are bivariate standard normal having non-negative correlation ρ , with equality if $\rho = 0$, and

$$(13) \quad P(\bar{E}_i \cap \bar{E}_j \cap \bar{E}_k) \leq P(\bar{E}_i \cap \bar{E}_j) \text{ for all } i, j, \text{ and } k.$$

Inequality (12) is obtained by writing Z_2 as $\rho Z_1 + \sqrt{1-\rho^2} Z_3$ where Z_1 and Z_3 are independent normal and then observing that

$\{Z_1 > c, Z_2 > c\} \supset \{Z_1 > c, Z_3 > c(1-\rho)\}$. Inequality (13) is trivial.

By judicious application of (13) to cancel 7 of the double and triple terms in (11) and then using (12) on the remaining terms it may be shown that in either case A or case B, (11) can be replaced by

$$(14) \quad P(\bigcap_i E_i) \geq 1 + P[Z_1 > c](-5 + (P[Z_1 > ck_1] + 2P[Z > ck_2] - 3P[Z_1 > c])P[Z_1 > c])$$

$$\text{where } k_1 = 3^{-1/2} \text{ and } k_2 = (2^{1/2} - 1).$$

Inequality (14) is somewhat less conservative than the unmodified Bonferroni bound $1 - 5P[Z_1 > c]$. For the case $\alpha = .05$ the unmodified Bonferroni upper bound obtained for c is 2.33 (see table 7.1), whereas the right side of (14) is greater than .05 for c as low as 2.30. Table 7.1 shows that the smallest value of c based on exact computation of the error probability $\alpha = .05$ is $c = 2.26$. The method of this section is not easily extendable to more complicated designs because of the extensive computations involved, but in some cases it may prove worth the effort when exact computations are unavailable and the unmodified Bonferroni method is overly conservative.

6. EXACT VALUES FOR c

The theory required to obtain exact values of the c which satisfies (4) is a matter of geometric probability and can be worked out as in the examples below. However, in the present state of the computing art, the c values can be computed only in special cases, as considered theoretically below and computationally in section 7. These cases, although special, are nonetheless important, since they permit an "absolute" comparison of the bounds of the previous section with the best possible value.

6.1. The balanced 2^2 case. For this case, inequality (8) can provide c values which satisfy (4) exactly, for the case where $\mathcal{L}$ is the class of four simple effects and the single two factor interaction. To emphasize that the dimension $d(\mathcal{L})$ in this case is three, and to simplify calculations, (8) is rewritten in terms of the three parameters $\underline{\beta}' = (\text{Main}_1, \text{Main}_2, I_{12}/2)$. In these terms, note that the five parameters of interest are exactly the $\underline{y}_i' \underline{\beta}$ in (15). Note also that the $\underline{\beta}$ elements are just the main effects and half-interaction defined for the 2^2 case in section 2.1, and their least squares estimators $\hat{\underline{\beta}}$ are unbiased and independent with variance σ^2/n , where n is the common sample size (or number of replications) from each of the four cells.

From (8) it is seen that the probability bound on the right-hand side is attained in the limit as $\underline{M}$ approaches $\underline{Q}$. However, this limiting value depends on the direction, i.e., the signs of the various true parameter values $\underline{x}' \underline{M} = \underline{y}' \underline{\beta}$, through which $\underline{Q}$ is approached. Specifically, in terms of $\hat{\underline{\beta}}$ and for $\underline{\beta} = 0$, the minimum of (8) over these approaches is

$$(15) \quad \min_{+, -} P_{\underline{Q}, \sigma^2}^2 \left(\pm \underline{y}' \hat{\underline{\beta}} \leq c S(\underline{y}' \underline{y} / n)^{1/2} \right), \quad \text{all } \underline{y} \text{ in } Y^*,$$

where $\min_{+, -}$ denotes the minimum over all possible ones of the 32 combinations of $+$ and $-$ signs for the five parameters, and where Y^* is the set of five (3×1) vectors which express the simple effects and interaction as linear functions of $\underline{\beta}$. Namely, Y^* is

the set of vectors $\underline{y}_1 = (1, 0, -1)$, $\underline{y}_2' = (1, 0, 1)$, $\underline{y}_3' = (0, 1, -1)$, $\underline{y}_4' = (0, 1, 1)$, and $\underline{y}_5' = (0, 0, 2)$, which express $(S(1;\phi), S(1;2), S(2;\phi), S(2;1), S(1,2;\phi))$ as $\beta'(\underline{y}_1, \underline{y}_2, \underline{y}_3, \underline{y}_4, \underline{y}_5)$.

Actually, not all 32 of the + and - combinations are possible. For example, (+--+) is impossible, since if both simple effects in the presence of nothing else are positive and if both simple effects in the presence of the other factor are negative, then the interaction, which is the difference of the former from the latter, is necessarily negative. The 18 possible sign combinations are expressed conveniently in the following four subsets:

- A) (+ + + + +), (+ + + + -), (+ + - - +), (+ + - - -) ;
- B) (+ + + - -), (+ + - + +), (+ - + + -), (+ - - - -) ;
- C) (- + - + +) ;
- D) any vector of signs in A, B, or C with + replaced by - and - by + throughout.

The desired, exact c-value is that for which (15) attains the desired α in (4), i.e., so that $P\{(7)\} = 1-\alpha$. This can be found by calculating these probabilities exactly for given values of c until the desired α is attained. For a given c, and for any one of the 18 possible sign combinations, the probability (15) is the chance that a multivariate Student t random vector falls in a given set.

The following geometrical argument simplifies the evaluation of the 18 probabilities, by reducing these to only three distinct cases. For any possible combination of + and - signs, the event to be

measured in (15) is the event that the trivariate Student t random vector $\underline{T} = n^{1/2} \hat{\underline{\beta}}/S$ falls on the $\underline{Q}$ -side of the plane of support of the three-dimensional sphere, with center at $\underline{Q}$ and radius c , and perpendicular to each vector $\pm \underline{y}_i$, for $\underline{y}_i$ ($i=1(1)5$) in Y^* . Geometrically, the probabilities of these 18 events are seen to depend only on the orientation of the vectors $\pm \underline{y}_i$, for $\underline{y}_i$ in Y^* , with respect to each other. For any of the sign combinations in the set D above, this orientation is the same as for the corresponding vector in the set A , B , or C . Also, within each of the sets A , B , and C , each of the sign combinations gives the same orientation of vectors $\pm \underline{y}_i$. For example, $(++++)$ and $(+++-)$ are seen to give the same angles, by verifying that cosines of angles between the (permuted) vectors are equal, verified most easily by interchanging the roles of $\underline{y}_1$ and $\underline{y}_2$ with $\underline{y}_3$ and $\underline{y}_4$, respectively, in the second case. Thus the 18 calculations are reduced to only three, e.g., $(++++)$, $(++-++)$, and $(-+-++)$.

For case A , $(++++)$ for example, the probability to be evaluated is the chance that $\underline{T}$ falls in the set

$$\{(T_1, T_2, T_3): T_3 \leq c; T_i \leq c2^{1/2} - |T_3|, i=1,2\}.$$

For case B , the corresponding probability is

$$P(T_1 \leq c2^{1/2} - |T_3|, |T_2| \leq c2^{1/2} - T_3, T_3 \leq c),$$

while that for case C is

$$P(|T_1| \leq c2^{1/2} - T_3, |T_2| \leq c2^{1/2} - T_3, T_3 \leq c).$$

To compare these probabilities of cases A through D, the result of Slepian [1962] shows that the case C gives the maximizing, rather than minimizing, value, since in case C all the estimators $\pm \hat{y}'\hat{\beta}$ have non-negative correlations. The trivariate t probabilities in cases B and C are trivariate normal probabilities when conditioned on the sample standard deviation, and they are thus, unconditionally, positively weighted combinations of these normal probabilities. As a function of c , the normal probabilities in A and B are both $1/2^4$ at $c = 0$ and both 1 in the limit as c increases. The derivative of the difference of the case B probability from that of case A at $c = 0$ can be shown to be $(2\pi)^{-1/2} (\text{Arctan } 2^{1/2}/\pi - .25) > 0$, so that case A exceeds case B for c near zero. On the other hand, for most positive c values, case B exceeds case A.

Computations show that case A exceeds case B for the normal integrals only for c on the interval $(0, .33)$, on which the difference never exceeds 0.00156; while on the entire interval $(.39, 2.73)$ case B exceeds case A by at least 0.00156. This, together with the fact that the density of S increases on the interval $(0, (df=2))$ yields an inequality which covers most cases of practical interest; namely, if $c \geq 1$ and $df \geq 3$ then case B exceeds case A.

6.2. Simple effects $S(i;\phi)$ in the 2^k experiment. If, in a 2^k factorial experiment, interest is confined to the parameters $S(i;\phi)$, i.e., the simple effects of the k factors in the presence of nothing else, then observations are required only on the control, $\underline{y} = 0$, and for each of the one factor at a time cases, where $\underline{y} = \underline{y}_i$ has a 1 in the i -th place only; other observations are wasted. It is natural to take an equal number of observations on each factor. However, since the estimator of M_0 appears in each of the k simple effect estimators, it may be worth while to spend more observations on this class. For example, with a total of N observations, if n is the sample size for each of the individual factors and m that for the "control" ($\underline{y} = 0$), so that $N = kn + m$, the variance of each simple effect estimator is $\sigma^2(\frac{1}{m} + \frac{1}{n})$. To a continuous approximation the minimum for this variance is $\sigma^2(k^{1/2}+1)^2/N$, attained when $n = N/(k^{1/2}+k)$ and $m = k^{1/2}n$. This contrasts with the case where equal numbers, $N/(k+1)$, are observed with each vector and with the control; then each simple effect estimator has variance $2\sigma^2(k+1)/N$. The correlations of these estimators are all $\rho = (k^{1/2}+1)^{-1}$ in the minimum variance case, while for the "equal sampling" case, the common correlation is $\rho = \frac{1}{2}$.

For this case of simple effects, the dimension $d(\underline{\phi}) = k$ is that of the k parameters $S(i;\phi) = M_{\underline{y}_i} - M_0 = \theta_i$ (say). The rule classifies θ_i as positive or negative according as its least squares estimator $\hat{\theta}_i = \hat{M}_{\underline{y}_i} - \hat{M}_0$ exceeds $c\sigma_{\hat{\theta}}$, or is less than $-c\sigma_{\hat{\theta}}$, where $\sigma_{\hat{\theta}} = (\frac{1}{m} + \frac{1}{n})^{1/2}$. To satisfy (4) requires that

$$1-\alpha = \min_{+, -} P_{\theta, \sigma^2}(\hat{\theta}_i - \theta_i \leq c S \sigma_{\hat{\theta}}^2, \quad i = 1(1)k)$$

$$= \min_{+, -} P(+ T_i \leq c, \quad i = 1(1)k).$$

Here, "min" is as in section 6.1, only now all 2^k sign combinations are possible; and $T \quad (k \times 1) = \underline{Z}/S$, where $\underline{Z}$ is a k -variate normal random vector with all means 0, variances σ^2 , correlations ρ , given in the preceding paragraph, and independent of S .

The maximum, rather than minimum, of these probabilities is attained when either all signs are positive or all signs are negative. Also, when all signs are the same, the percentile c -values are exactly those for the one-sided case tabled by Dunnett [1955], which values thus provide liberal (i.e., too small) lower bounds for the required c . The required minimum probability is attained when half (or $(k-1)/2$ if k is odd) of the signs are negative, as is shown next.

Both the "maximum" and "minimum" statements of the preceding paragraph are established by proving that

$$P_r = P(Z_i - Z \leq c, \quad i=1(1)r, \quad \text{and} \quad -(Z_i - Z) \leq c, \quad i=r+1(1)k)$$

is symmetric about $k/2$ and decreases from 0 to $k/2$. The only properties of the Z_i and Z required to establish this are that they be mutually independent, that Z (with density g , say) have distribution symmetric about 0, and that the Z_i have common

distribution F (say) which is symmetric about 0 . Since

$$\begin{aligned} P_r &= \int_{-\infty}^{\infty} F^r(c+z) (1-F(-c+z))^{k-r} g(z) dz \\ &= \int_{-\infty}^{\infty} F^r(c+z) F^{k-r}(c-z) g(z) dz \end{aligned}$$

it follows that

$$P_r - P_{r+1} = \int_{-\infty}^{\infty} F^r(c+z) F^{k-r-1}(c-z) [F(c-z) - F(c+z)] g(z) dz.$$

For the term in square brackets, its sign is the opposite of the sign of z , and its absolute value is the same for z and $-z$. Therefore

$$\begin{aligned} P_r - P_{r+1} &= \int_0^{\infty} [F^r(c+z) F^{k-r-1}(c-z) - F^r(c-z) F^{k-r-1}(c+z)] \\ &\quad [F(c-z) - F(c+z)] g(z) dz \end{aligned}$$

and the result follows since the first term in square brackets is of the form $a^r b^n = a^n b^r$, which is non-positive whenever $(a/b)^{r-n} \leq 1$, as is the case when $r \leq (k-1)/2$. By replacing c in the preceding argument by cS and then integrating $P_r - P_{r+1}$ with respect to the density of S , we obtain the desired result for the case of unknown variance as well.

7. COMPUTATIONS, COMPARISONS, AND CONCLUSIONS

To compare the rules $\{\phi_{\tilde{x}}^c\}$, using the various exact and conservative c values derived in the previous two sections, two simple, qualitative results are immediate. First, conservatism increases as c increases away from the exact value. That is, the true chance of making at least one error, to be denoted α_T , is less than the desired value of α , which is attained by the exact value. Also, α_T approaches zero monotonically as c increases without limit. Second, the non-central t probability of making a correct decision decreases as c increases. An interesting, practical measure of correct decision chances is the size E_P , measured in units of the standard deviation σ , which a single parameter $\tilde{x}'\tilde{M}$ must exceed if it is to be classified correctly with probability $1-P$. For example a parameter of size $\sigma E_{.5}$ would be classified correctly half of the time. These parameters, which were introduced by Neyman [1935a], constitute one way of assessing the power of each test.

A meaningful, quantitative comparison of the conservative rules thus requires evaluation of the exact value of c , as well as of values of α_T and E_P for interesting P values. Of some particular interest are $E_{.5}$, which gives half a chance of a correct classification, and E_{α} , which makes the chance of correct classification as large as the error probability is small.

For the examples of sections 6.1 and 6.2, Cooper's [1968] non-central t algorithm is used to calculate the E_P values in Tables 7.1 and 7.2.

For the exact c value, as well as all the α_T values, the required multivariate t probabilities can be reduced to bivariate integrals, as described below, and calculated using modifications of the Milton-Tavernini algorithm, as described by Milton [1972].

7.1. Simple effects and interaction in the 2^2 design. Since always in this case there are at least four degrees of freedom for error and c exceeds unity, the α_T values are the probabilities for case A in section 6.1. Namely,

$$\alpha_T = 1 - \int_0^\infty \int_{-\infty}^{cs} \Phi^2(cs2^{1/2} = |z|) \phi(z) f(s) dz ds,$$

where Φ and ϕ are, respectively, the standard normal distribution and density functions, and f is the density for the "sample standard deviation" S . Milton's [1972] algorithm can be used to calculate this after replacing his $X(1)$, $X(2)$, and U by S , z , and the integrand above. For the conservative methods, the values for c are obtained from existing tables, as detailed in section 5; while the exact (to two decimal places) value for c is obtained through a computer search.

Table 7.1 exhibits comparisons of four conservative methods (Bonferroni, Tukey, Sharpened-Scheffé, and Scheffé) with each other and with the exact method, as measured by each of the four values: c , α_T , $E_{.5}$, and E_α . Cases considered include $\alpha = .05$ and $.01$, and cases where replications are small ($n = 2, df = 4$), moderate,

($n = 6$, $df = 20$), and large ($n = 16$, $df = 60$). For each of these cases, both variance known and variance unknown comparisons are given.

(Note to editor: Insert Table 7.1 near here.)

Table 7.1 shows that when the variance σ^2 is known, the conservatism in c and α_T is independent of the number n of replications; however, the sensitivity of the rule increases with n , since the parameters E_P decrease to zero like $n^{-1/2}$. The Bonferroni bound is by far the sharpest, having $E_{.5}$ only a few percent greater than for the sharp bound and with E_α even better. The Tukey bound is slightly sharper than the sharpened Scheffe bound, and both are substantially better than the Scheffé bound; but these are about 13, 18, and 23 percent longer than the sharp bound. For all the bounds, the conservatism, as measured by the sensitivity of E_P , is less when α is smaller and/or when P is smaller.

When the variance σ^2 is unknown, the conservatism is worse yet, especially when the number n of replications is small. The Bonferroni bound is still clearly the best among the conservative methods, although its $E_{.5}$ value is as much as 11 percent larger than the sharp value when $n = 2$. Compared to their performance with known variance, the other three conservative bounds behave qualitatively the same, but their E_P conservatism can be up to half-again as bad. The conservatism reflected in these E_P is again best when α and/or P is small and when n is large. In general, the computations for $n = 16$ replications do not differ much from the corresponding, limiting values for variance known, found in part A of Table 7.1.

TABLE 7.1. Operating Characteristics for $\{\phi_x^c\}$ for the Five Simple Interactions in the 2^2 Factorial Experiment

α is the desired global error bound; n is the number of replications; α_T is the true, conservative global error bound; and σE_p is the effect-size detected with probability $1-P$. Note that, for the Tukey method: a) the table applies for simple (first-order) effects and $I/2$; and b). $c = 2^{-1/2} q_{\alpha; 4, 4(n-1)}$ so that the "decision points" are $\pm c\sigma_{\tilde{x}}$, comparable to the other methods.

A. Variance known

α	c	Method	α_T	$n^{1/2} E_{.5}$	$n^{1/2} E_{\alpha}$
0.5	2.26	Exact	.050	3.20	5.53
	2.33	Bonferroni	.042	3.30	5.61
	2.57	Tukey	.022	3.63	5.95
	2.66	Sharpened-Scheffé	.017	3.76	6.10
	2.79	Scheffé	.012	3.95	6.28
.01	2.85	Exact	.010	4.03	7.31
	2.88	Bonferroni	.009	4.07	7.37
	3.11	Tukey	.004	4.40	7.69
	3.24	Sharpened-Scheffé	.003	4.58	7.88
	3.37	Scheffé	.002	4.77	8.06

TABLE 7.1 Continued

B. Variance Unknown

α	n	c	Method	α_T	$E_{.5}$	E_α
0.05	2	3.37	Exact	.050	3.12	5.75
		3.75	Bonferroni	.036	3.47	6.29
		4.07	Tukey	.028	3.76	6.75
		4.12	Sharpened-Scheffé	.026	3.81	6.82
		4.45	Scheffé	.021	4.11	7.30
	6	2.43	Exact	.050	1.38	2.40
		2.53	Bonferroni	.041	1.44	2.47
		2.80	Tukey	.023	1.60	2.64
		2.88	Sharpened-Scheffé	.020	1.64	2.69
		3.05	Scheffé	.014	1.74	2.79
	16	2.32	Exact	.050	0.82	1.41
		2.39	Bonferroni	.042	0.84	1.44
		2.64	Tukey	.023	0.93	1.53
		2.73	Sharpened-Scheffé	.018	0.96	1.56
		2.88	Scheffé	.012	1.01	1.62
0.01	2	5.43	Exact	.010	5.00	10.39
		5.95	Bonferroni	.007	5.48	11.30
		6.49	Tukey	.005	5.97	12.25
		6.60	Sharpened-Scheffé	.005	6.07	12.44
		7.08	Scheffé	.004	6.51	13.29
	6	3.19	Exact	.010	1.82	3.33
		3.25	Bonferroni	.009	1.85	3.37
		3.55	Tukey	.004	2.02	3.57
		3.68	Sharpened-Scheffé	.003	2.10	3.66
		3.85	Scheffé	.002	2.19	3.78
	16	2.95	Exact	.010	1.04	1.89
		2.99	Bonferroni	.009	1.05	1.91
		3.25	Tukey	.004	1.14	2.00
		3.38	Sharpened-Scheffé	.003	1.19	2.05
		3.52	Scheffé	.002	1.24	2.10

7.2. Simple effects $S(i;\phi)$ in the 2^k experiment. It is shown in section 6.2 that the probability of no errors is smallest when r of the parameter signs are negative and $k-r$ are positive where r is $k/2$ for k even and $(k-1)/2$ for k odd. Hence for a given value of c , the true minimum probability of making no errors is

$$\alpha_T = 1 - \int_0^\infty \int_{-\infty}^\infty \phi^r \left(c \left(1 + \frac{n}{m} \right)^{\frac{1}{2}} s + \left(\frac{n}{m} \right)^{\frac{1}{2}} z \right) \phi^{k-r} \left(c \left(1 + \frac{n}{m} \right)^{\frac{1}{2}} s - \left(\frac{n}{m} \right)^{\frac{1}{2}} z \right) \phi(z) f(s) dz ds ,$$

where m is the number of replications in the control and n is the number in each of the k single factor groups. As in the preceding section this integral may be evaluated numerically for each c value, to obtain the exact c (to two decimal places) for which the worst case error probability α_T attains the desired value α . The conservative values for c are obtained from existing tables using the methods outlined in section 5.

Table 7.2 exhibits comparisons of the four conservative methods (Bonferroni, Tukey, Sharpened Scheffé, and Scheffé) with the exact method, as measures by each of the four values: c , α_T , $E_{.5}$, and E_α . Cases considered are for error probability $\alpha = .01$ and $.05$, number of factors $k = 2, 3, 4$, and 5 , total number of observations N moderate (24, 30, 36) and large (96, 120, 144), and the design balanced (all cells of equal size) and optimal (m observations in the control cell and n observations in each of the remaining cells, as indicated in section 6.2). For each of these cases the variance known and variance unknown comparisons are given.

(Note to editor: Insert table 7.2 near here.)

Section 7.2 σ^2 known

α	k	N	Design n m		Exact				Bonferroni				Tukey				Sharpened-Scheffé				Scheffé			
					c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α
.05	2	24	7 10		1.96	.950	0.97	1.78	1.96	.950	0.97	1.78	2.34	.981	1.15	1.97	2.27	.977	1.12	1.93	2.45	.986	1.21	2.02
			8 8		1.96	.950	0.98	1.80	1.96	.950	0.98	1.80	2.34	.981	1.17	1.99	2.27	.977	1.13	1.96	2.45	.986	1.22	2.05
		96	28 40		1.96	.950	0.48	0.89	1.96	.950	0.48	0.89	2.34	.981	0.58	0.98	2.27	.977	0.56	0.96	2.45	.986	0.60	1.01
			32 32		1.96	.950	0.49	0.90	1.96	.950	0.49	0.90	2.34	.981	0.59	1.00	2.27	.977	0.57	0.98	2.45	.986	0.61	1.02
	3	24	5 9		2.11	.950	1.18	2.10	2.13	.952	1.19	2.10	2.57	.985	1.43	2.35	2.66	.988	1.48	2.40	2.80	.992	1.56	2.48
			6 6		2.11	.950	1.22	2.17	2.13	.953	1.23	2.18	2.57	.985	1.48	2.43	2.66	.989	1.54	2.49	2.80	.992	1.61	2.56
		96	20 36		2.11	.950	0.59	1.05	2.13	.952	0.59	1.05	2.57	.985	0.72	1.18	2.66	.988	0.74	1.20	2.80	.992	0.78	1.24
			24 24		2.11	.950	0.61	1.09	2.13	.953	0.61	1.09	2.57	.985	0.74	1.22	2.66	.989	0.77	1.24	2.80	.992	0.81	1.28
	4	30	5 10		2.23	.950	1.22	2.12	2.24	.952	1.23	2.13	2.73	.988	1.49	2.40	2.96	.994	1.62	2.52	3.08	.996	1.69	2.59
			6 6		2.21	.950	1.28	2.23	2.24	.954	1.29	2.24	2.73	.988	1.58	2.52	2.96	.994	1.71	2.66	3.08	.996	1.78	2.73
		120	20 40		2.23	.950	0.61	1.06	2.24	.952	0.61	1.06	2.73	.988	0.75	1.20	2.96	.994	0.81	1.26	3.08	.996	0.84	1.29
			24 24		2.21	.950	0.64	1.11	2.24	.954	0.65	1.12	2.73	.988	0.79	1.26	2.96	.994	0.85	1.33	3.08	.996	0.89	1.36
	5	36	5 11		2.31	.950	1.25	2.13	2.33	.953	1.26	2.14	2.85	.989	1.54	2.42	3.22	.997	1.74	2.62	3.33	.998	1.79	2.68
			6 6		2.29	.950	1.32	2.27	2.33	.955	1.34	2.29	2.85	.990	1.65	2.59	3.22	.997	1.86	2.81	3.33	.998	1.92	2.87
		144	20 44		2.31	.950	0.63	1.06	2.33	.953	0.63	1.07	2.85	.989	0.77	1.21	3.22	.997	0.87	1.31	3.33	.998	0.90	1.34
			24 24		2.29	.950	0.66	1.14	2.33	.955	0.67	1.15	2.85	.990	0.82	1.30	3.22	.997	0.93	1.40	3.33	.998	0.96	1.44
.01	2	24	7 10		2.58	.990	1.27	2.42	2.58	.990	1.27	2.42	2.91	.996	1.44	2.58	2.88	.996	1.42	2.57	3.03	.998	1.50	2.64
			8 8		2.58	.990	1.29	2.45	2.58	.990	1.29	2.45	2.91	.996	1.46	2.62	2.88	.996	1.44	2.60	3.03	.998	1.52	2.68
		96	28 40		2.58	.990	0.63	1.21	2.58	.990	0.63	1.21	2.91	.996	0.72	1.29	2.88	.996	0.71	1.28	3.03	.998	0.75	1.32
			32 32		2.58	.990	0.64	1.23	2.58	.990	0.64	1.23	2.91	.996	0.73	1.31	2.88	.996	0.72	1.30	3.03	.998	0.76	1.34
	3	24	5 9		2.71	.990	1.51	2.81	2.71	.990	1.51	2.81	3.11	.997	1.74	3.03	3.24	.998	1.81	3.10	3.37	.999	1.88	3.18
			6 6		2.71	.990	1.57	2.91	2.71	.990	1.57	2.91	3.11	.997	1.80	3.14	3.24	.998	1.87	3.21	3.37	.999	1.94	3.29
		96	20 36		2.71	.990	0.76	1.41	2.71	.990	0.76	1.41	3.11	.997	0.87	1.52	3.24	.998	0.90	1.55	3.37	.999	0.94	1.59
			24 24		2.71	.990	0.78	1.45	2.71	.990	0.78	1.45	3.11	.997	0.90	1.57	3.24	.998	0.94	1.61	3.37	.999	0.97	1.64
		30	5 10		2.80	.990	1.54	2.81	2.81	.990	1.54	2.81	3.25	.998	1.78	3.06	3.54	.999	1.94	3.21	3.64	1.000	2.00	3.27
			6 6		2.80	.990	1.62	2.96	2.81	.990	1.62	2.96	3.25	.998	1.88	3.22	3.54	.999	2.04	3.39	3.64	1.000	2.10	3.45
		120	20 40		2.80	.990	0.77	1.41	2.81	.990	0.77	1.41	3.25	.998	0.89	1.53	3.54	.999	0.97	1.61	3.64	1.000	1.00	1.63
			24 24		2.80	.990	0.81	1.48	2.81	.990	0.81	1.48	3.25	.998	0.94	1.61	3.54	.999	1.02	1.69	3.64	1.000	1.05	1.72
	5	36	5 11		2.87	.990	1.55	2.81	2.88	.990	1.55	2.81	3.36	.998	1.81	3.07	3.79	1.000	2.04	3.30	3.88	1.000	2.09	3.35
			6 6		2.86	.990	1.65	3.00	2.88	.991	1.66	3.01	3.36	.998	1.94	3.29	3.79	1.000	2.19	3.53	3.88	1.000	2.24	3.59
		144	20 44		2.87	.990	0.78	1.40	2.88	.990	0.78	1.40	3.36	.998	0.91	1.53	3.79	1.000	1.02	1.65	3.88	1.000	1.05	1.67
			24 24		2.86	.990	0.83	1.50	2.88	.991	0.83	1.50	3.36	.998	0.97	1.64	3.79	1.000	1.09	1.77	3.88	1.000	1.12	1.79

Section 7.2. σ^2 unknown

α	k	N	Design n m	Exact				Bonferroni				Tukey				Sharpened-Scheffé				Scheffé			
				c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α	c	α_T	$E_{.5}$	E_α
.05	2	24	7 10	2.08	.950	1.01	1.86	2.08	.950	1.01	1.86	2.52	.980	1.23	2.10	2.43	.976	1.18	2.05	2.63	.984	1.28	2.16
			8 8	2.08	.950	1.03	1.89	2.08	.950	1.03	1.89	2.52	.980	1.24	2.13	2.43	.976	1.20	2.08	2.63	.984	1.30	2.19
		96	28 40	1.99	.950	0.49	0.90	1.99	.950	0.49	0.90	2.39	.979	0.59	1.00	2.31	.975	0.57	0.98	2.52	.985	0.62	1.03
			32 32	1.99	.950	0.50	0.91	1.99	.950	0.50	0.91	2.39	.979	0.60	1.01	2.31	.975	0.58	0.99	2.52	.985	0.63	1.05
	3	24	5 9	2.27	.950	1.25	2.23	2.29	.952	1.26	2.24	2.80	.984	1.54	2.55	2.88	.986	1.59	2.60	3.05	.991	1.68	2.70
			6 6	2.25	.950	1.27	2.28	2.29	.953	1.31	2.32	2.80	.984	1.60	2.64	2.88	.987	1.64	2.69	3.05	.991	1.74	2.79
		96	20 36	2.15	.950	0.60	1.06	2.16	.952	0.60	1.07	2.62	.985	0.73	1.20	2.71	.988	0.75	1.22	2.86	.992	0.80	1.26
			24 24	2.14	.950	0.61	1.09	2.16	.953	0.62	1.10	2.62	.985	0.75	1.24	2.71	.988	0.78	1.26	2.86	.992	0.82	1.31
	4	30	5 10	2.36	.950	1.28	2.23	2.38	.952	1.29	2.24	2.94	.986	1.59	2.57	3.20	.993	1.73	2.73	3.32	.995	1.80	2.80
			6 6	2.35	.950	1.34	2.34	2.38	.953	1.36	2.36	2.94	.987	1.68	2.71	3.20	.993	1.83	2.87	3.32	.995	1.90	2.95
		120	20 40	2.26	.950	0.62	1.07	2.27	.952	0.62	1.08	2.77	.987	0.76	1.21	3.00	.993	0.82	1.28	3.13	.996	0.86	1.32
			24 24	2.24	.950	0.64	1.12	2.27	.954	0.65	1.13	2.77	.987	0.80	1.28	3.00	.994	.086	1.35	3.13	.996	0.90	1.39
	5	36	5 11	2.43	.950	1.30	2.23	2.46	.953	1.32	2.25	3.04	.988	1.63	2.58	3.46	.995	1.85	2.82	3.56	.996	1.90	2.88
			6 6	2.41	.950	1.37	2.37	2.46	.956	1.41	2.41	3.04	.988	1.74	2.76	3.46	.995	1.98	3.02	3.56	.996	2.04	3.08
		144	20 44	2.33	.950	0.63	1.08	2.36	.953	0.64	1.08	2.89	.989	0.78	1.23	3.26	.996	0.88	1.33	3.36	.997	0.90	1.36
			24 24	2.31	.950	0.66	1.14	2.36	.955	0.68	1.16	2.89	.989	0.83	1.31	3.26	.996	0.94	1.42	3.36	.997	0.97	1.45
.01	2	24	7 10	2.83	.990	1.38	2.63	2.83	.990	1.38	2.63	3.26	.996	1.59	2.88	3.20	.995	1.56	2.84	3.40	.997	1.65	2.96
			8 8	2.83	.990	1.40	2.67	2.83	.990	1.40	2.67	3.26	.996	1.61	2.92	3.20	.995	1.58	2.88	3.40	.997	1.68	3.00
		96	28 40	2.65	.990	0.65	1.24	2.65	.990	0.65	1.24	3.01	.995	0.74	1.33	2.97	.994	0.73	1.32	3.14	.996	0.77	1.36
			32 32	2.65	.990	0.66	1.25	2.65	.990	0.66	1.25	3.01	.995	0.75	1.35	2.97	.994	0.74	1.34	3.14	.996	0.78	1.38
	3	24	5 9	3.03	.990	1.67	3.11	3.03	.990	1.67	3.11	3.55	.997	1.95	3.45	3.68	.998	2.02	3.54	3.85	.998	2.12	3.65
			6 6	3.03	.990	1.73	3.22	3.03	.990	1.73	3.22	3.55	.997	2.02	3.57	3.68	.998	2.10	3.66	3.85	.998	2.19	3.78
		96	20 36	2.78	.990	0.77	1.44	2.78	.990	0.77	1.44	3.21	.997	0.89	1.56	3.34	.998	0.93	1.60	3.48	.999	0.97	1.64
			24 24	2.78	.990	0.80	1.49	2.78	.990	0.80	1.49	3.21	.997	0.92	1.61	3.34	.998	0.96	1.65	3.48	.999	1.00	1.70
	4	30	5 10	3.08	.990	1.67	3.06	3.08	.990	1.67	3.06	3.64	.998	1.97	3.41	3.99	.999	2.16	3.63	4.09	.999	2.22	3.70
			6 6	3.08	.990	1.76	3.23	3.08	.990	1.76	3.23	3.64	.998	2.08	3.60	3.99	.999	2.28	3.83	4.09	.999	2.34	3.90
		120	20 40	2.86	.990	0.78	1.43	2.86	.990	0.78	1.43	3.32	.998	0.91	1.56	3.60	.999	0.98	1.64	3.72	.999	1.02	1.67
			24 24	2.86	.990	0.82	1.51	2.86	.990	0.82	1.51	3.32	.998	0.96	1.64	3.60	.999	1.04	1.73	3.72	.999	1.07	1.76
	5	36	5 11	3.12	.990	1.67	3.02	3.12	.990	1.67	3.02	3.71	.997	1.98	3.38	4.22	.999	2.26	3.69	4.30	.999	2.30	3.74
			6 6	3.12	.990	1.79	3.24	3.12	.990	1.79	3.24	3.71	.997	2.12	3.62	4.22	.999	2.41	3.95	4.30	.999	2.46	4.00
		144	20 44	2.93	.990	0.79	1.43	2.93	.990	0.79	1.43	3.42	.998	0.92	1.56	3.85	.999	1.04	1.68	3.95	.999	1.06	1.71
			24 24	2.93	.990	0.84	1.53	2.93	.990	0.84	1.53	3.42	.998	0.99	1.67	3.85	.999	1.11	1.80	3.95	.999	1.14	1.83

Table 7.2 shows that the Bonferroni bound for c is only slightly conservative, less than 2% larger than the exact value of c in all cases considered. This nearness is reflected in the values of α_T , $E_{.5}$, and E_α , which are all very close to the corresponding values for the exact test, and frequently in agreement with them to the accuracy of table. The conservatism of the Bonferroni bound is greatest in the case where σ^2 is unknown, $\alpha = .05$, the number of factors is large, and the design is balanced, but even here the difference is minor.

The other three conservative bounds do not fare nearly so well. As might be expected, the Scheffé bound for c is uniformly the most conservative one, ranging between 17% and 48% larger than exact value in the cases tabled. As for the other two bounds, the Tukey is the least conservative of these when there are three or more factors in the experiment, whereas the sharpened Scheffé bound is the smaller of the two when there are just two factors. If the least conservative of the two bounds is used it ranges between 11% and 26% larger than the exact value of c , which is substantially greater than the $\leq 2\%$ conservatism of the Bonferroni bound. For each of the bounds the conservatism increases with the number of factors and is greatest in the unknown variance case.

The known variance part of table 7.2 is redundant in the sense that for a given number of factors and or given sample design, the values of c , α_T , $N^{1/2}E_{.5}$, and $N^{1/2}E_\alpha$ do not depend on N , the total sample size. It follows that the relative conservatism of the various bounds shown by the table will be the same for values of N other than those tabled. On the other hand, in the unknown variance case this is not so, and the

behavior of the methods may differ somewhat for other values of N . In particular it may be conjectured that, by analogy with table 7.1, the conservatism of the Bonferroni method will be increased if the sample size is reduced to, say, two observations per cell.

7.3. Conclusions. In this paper the rule $\{\phi_x^c\}$ has been proposed as an optimal one which is easy to apply in a wide class of problems involving the signs of simple effects and interactions in 2^k factor experiments. The only substantive difficulty in the application of the rule is the determination of the value of c for which the rule has a specified probability α of making no errors, or when such a c is not readily computable, a close upper bound for c . The preceding sections give the exact value of c for a number of cases of interest, and indicate algorithms for the computation of this value for other situations of the same type. However the primary practical purpose of these sections is to compare various upper bounds for c that are readily obtainable even in more complex situations. The unequivocal evidence of tables 7.1 and 7.2 is that the bound which is based on the Bonferroni inequality is superior to the others and is in fact remarkably close to the true value of c . This is in agreement with the findings of other investigators (e.g. Dunn [1958], Miller [1966]) in the multiple comparisons context. It therefore would appear to recommend itself as a preferred method in more complex situations as well.

For any given problem, all four bounds can be computed and the smallest of them used, but a deeper question that is being asked is whether it would be worthwhile to try to overcome the formidable obstacles to the exact computation of c in the more complex situations, since even the best of these upper bounds may then be grossly conservative.

The near optimality of the Bonferroni method for the simple designs considered is encouraging in this respect. It would be even more encouraging if it could be proven that it remains near optimal in at least one comparatively complex situation, such as the case of all 19 simple effects and interactions in the 2^3 factorial experiment.

References

- Bohrer, R. (1976). Multiple three-decision rules for parametric signs.
Preprint.
- Cooper, B.E. (1968). Algorithm AS5: The integral of the non-central
t-distribution. *Appl. Statist.*, 17, 193-194.
- Dunn, O.J. (1958). Estimation of the means of dependent variables. *Ann
Math. Statist.*, 29, 1095-1111.
- Dunnett, C.W. (1955). A multiple comparison procedure for comparing several
treatments with a control. *J. Amer. Statist. Assoc.*, 50, 1096-1121.
- Hochberg, Y. and Quade, D. (1975). One-sided simultaneous confidence
bounds on regression surfaces with intercepts. *J. Amer. Statist.
Assoc.*, 70, 889-891.
- Lehmann, E.L. (1957a). A theory of some multiple decision problems, I.
Ann. Math. Statist., 28, 1-25.
- Lehmann, E.L. (1957b). A theory of some multiple decision problems,
II. *Ann. Math. Statist.*, 28, 547-572.
- Miller, R.G. Jr. (1966). *Simultaneous Statistical Inference*. McGraw-Hill,
New York.
- Milton, R.C. (1972). Computer evaluation of the multivariate normal
integral. *Technometrics*, 14, 881-889.
- Neyman, J. (1935a). Statistical problems in agricultural experimentation.
J. Roy. Statist. Soc. Suppl., 2, 107-180.
- Neyman, J. (1935b) Complex experiments. Contribution to the discussions
of the paper by F. Yates. *J. Roy. Statist. Soc. Suppl.*, 2, 235-242.
- Neyman, J. and Yates, F. (1936). Correspondence. Complex experiments.
J. Roy. Statist. Soc. Suppl., 2, 83-85.

- Resnikoff, G.J. and Lieberman, G.J. (1957). Tables of the non-central t-distribution: density function, cumulative distribution function, and percentage points. Stanford University Press.
- Scheffé, H. (1959). The Analysis of Variance. John Wiley and Sons, New York.
- Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. Bell System Technical Journal, 41, 463-501.
- Spjøtvoll, E. (1972). On the optimality of some multiple comparison procedures. Ann. Math. Statist., 43, 398-411.
- Yates, F. (1935). Complex experiments. J. Roy. Statist. Soc. Suppl., 2, 181-223.