A Method for Judging all Contrasts in the Analysis of Variance

Henry Scheffe

Biometrika, Vol. 40, No. 1/2. (Jun., 1953), pp. 87-104.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195306%2940%3A1%2F2%3C87%3AAMFJAC%3E2.0.CO%3B2-D

Biometrika is currently published by Biometrika Trust.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/bio.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Mon Oct 22 15:42:09 2007

A METHOD FOR JUDGING ALL CONTRASTS IN THE ANALYSIS OF VARIANCE*

BY HENRY SCHEFFÉ, *Columbia University*

A simple answer is found for the following question which has plagued the practice of the analysis of variance: Under the usual assumptions, if the conventional F-test of the hypothesis $H: \mu_1 = \mu_2 = \ldots = \mu_k$ at the α level of significance rejects H, what further inferences are valid about the contrasts among the μ_i (beyond the inference that the values of the contrasts are not all zero)? Suppose the F-test has $k-1$ and ν degrees of freedom. For any $c_1, \ldots, c_k$ with $\sum_1^k c_i = 0$ write θ for the contrast $\sum_1^k c_i\mu_i$, and write $\hat{\theta}$ and $\hat{\sigma}^2_{\hat{\theta}}$ for the usual estimates of θ and the variance of $\hat{\theta}$. Then for the totality of contrasts, no matter what the true values of the θ's, the probability is $1-\alpha$ that they all satisfy

$$\hat{\theta} - S\hat{\sigma}_{\hat{\theta}} \leqslant \theta \leqslant \hat{\theta} + S\hat{\sigma}_{\hat{\theta}},$$

where S^2 is $(k-1)$ times the upper α point of the F-distribution with $k-1$ and ν degrees of freedom. Suppose we say that the estimated contrast $\hat{\theta}$ is 'significantly different from zero' if $|\hat{\theta}| > S\hat{\sigma}_{\hat{\theta}}$. Then the F-test rejects H if and only if some $\hat{\theta}$ are significantly different from zero, and if it does, we can say just which $\hat{\theta}$. More generally, the above inequality can be employed for all the contrasts with the obvious frequency interpretation about the proportion of experiments in which all statements are correct. Relations are considered to an earlier method of Tukey using the Studentized range tables and valid in the special case where the $\hat{\mu}_i$ all have the same variance and all pairs $\hat{\mu}_i, \hat{\mu}_j$ $(i \neq j)$ have the same covariance. Some results are obtained for the operating characteristic of the new method. The paper is organized so that the reader who wishes to learn the method and avoid the proofs may skip §§ 2 and 5.

1. STATEMENT OF THE METHOD

The general problem is that of making inferences about the contrasts among a set of 'true means' or 'true main effects' $\mu_1, \mu_2, \ldots, \mu_k$ in the analysis of variance. For example, the μ_i might be the true row effects in a two-way lay-out with possibly unequal numbers of observations per cell. The μ_i may be unrestricted or subject to a single restriction of the form

$$\sum_1^k h_i\mu_i = h, \tag{1}$$

where the h_i and h are known constants with $\sum_1^k h_i \neq 0$. *A contrast* is a linear function of the μ_i,

$$\theta = \sum_1^k c_i\mu_i, \tag{2}$$

determined by k known constants c_i satisfying the condition

$$\sum_1^k c_i = 0. \tag{3}$$

The value of the linear function for a particular set of μ_i will be called the *value* of the contrast; it will not cause any confusion in the following to use the same symbol θ both for the contrast and the value of the contrast.

We make the assumptions usual in the analysis of variance, namely, that there is at hand a set of statistics $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k$, and $\hat{\sigma}^2$, such that the $\hat{\mu}_i$ have a multivariate normal distribution and are statistically independent of $\hat{\sigma}^2$, that

$$E(\hat{\mu}_i) = \mu_i \quad (i = 1, \ldots, k),$$

and
$$\operatorname{cov}(\hat{\mu}_i, \hat{\mu}_j) = a_{ij}\sigma^2 \quad (i,j = 1, \ldots, k), \tag{4}$$

where the constants a_{ij} are known, and σ^2 is unknown. The μ_i will always be 'Model I' (non-random) effects, as discussed by Eisenhart (1947) or Mood (1950). In a pure Model I

* Work sponsored by the Office of Naval Research (U.S.A.).

situation, σ^2 is the variance σ_e^2 of a single observation ('error variance'). In a mixed model situation where there exists an exact F-test of the hypothesis

$$H\colon \mu_1 = \mu_2 = \ldots = \mu_k, \tag{5}$$

σ^2 will equal σ_e^2 plus further unknown non-negative parameters. In any case, $\hat{\sigma}^2$ is an estimate of σ^2 with ν D.F. (degrees of freedom), that is, $\nu\hat{\sigma}^2/\sigma^2$ has the χ^2 distribution with ν D.F. The case where σ^2 is known can be treated by obvious modifications of the theory below, usually merely by putting $\nu = \infty$ and $\hat{\sigma}^2 = \sigma^2$ in the results. It is further assumed that if the μ_i are unrestricted the rank* of the covariance matrix with elements (4) is k, and that if the μ_i are subject to a restriction (1) then the $\hat{\mu}_i$ are subject to the same restriction (1) and the rank of the covariance matrix is $k-1$.

The hypothesis H in (5), equivalent to the statement that all the contrasts are zero, can be tested by the conventional F-statistic with $k-1$ and ν D.F. We shall refer to this test at significance level α as 'the' F-test of H. The problem of making further inferences about the contrasts, arising when the F-test rejects H, has been considered by various writers, including R. A. Fisher (1935), D. Newman (1939), J. W. Tukey (1951), and H. K. Nandi (1951). Except for Tukey's and Nandi's, the methods involve repeated tests of significance on the same data, and are hence subject to the usual objection that little is known about the joint operating characteristic. While it is often not possible in practical applications to avoid repeated tests of significance, it is possible for the particular problem we are considering.

The solution studied in this paper is based on the following probability statement about the infinite totality of contrasts:† For any contrast (2) denote its estimate by $\hat{\theta}$,

$$\hat{\theta} = \sum_1^k c_i \hat{\mu}_i,$$

the variance of $\hat{\theta}$ by $\sigma_{\hat{\theta}}^2$,

$$\sigma_{\hat{\theta}}^2 = \sum_{i=1}^k \sum_{j=1}^k a_{ij} c_i c_j \sigma^2, \tag{6}$$

and the estimate of this variance by $\hat{\sigma}_{\hat{\theta}}^2$,

$$\hat{\sigma}_{\hat{\theta}}^2 = \sum_{i=1}^k \sum_{j=1}^k a_{ij} c_i c_j \hat{\sigma}^2.$$

Define the positive constant S from

$$S^2 = (k-1) F_\alpha(k-1, \nu), \tag{7}$$

* The non-mathematical statistician may safely assume these rank conditions to be satisfied in practical applications; they are stated because they are needed later for the mathematical arguments in §§2 and 5.

† The idea of making an overall confidence statement for all the contrasts, its successful realization, and the resultant possibility of making valid tests of hypotheses suggested by the data, I first met in a lecture by Prof. J. W. Tukey on 'New methods in the analysis of variance: Range in the numerator and range in the denominator' at the Annual Princeton Conference of the American Society for Quality Control on 8 December 1951. This method was published (Tukey, 1951). The method of Tukey described in §3 was explained in 'Allowances for various types of error rates', an unpublished invited address presented before a joint meeting of the Institute of Mathematical Statistics and the Eastern North American Region of the Biometric Society on 19 March 1952 at Blacksburg, Va.; it differs from the published method (Tukey, 1951) only in the use of a root-mean-square instead of a range estimate of the error standard deviation. Recently when I communicated the method based on (8) to Prof. Tukey he wrote me that it had been familiar to him and Prof. D. B. Duncan for some time, and that they had discussed it publicly when he gave a lecture at Blacksburg in November 1951.

where $F_\alpha(k-1, \nu)$ denotes the upper α point of the F-distribution with $k-1$ and ν D.F. Then the probability is $1-\alpha$ that the values θ of *all* the contrasts simultaneously satisfy

$$\hat{\theta}-S\hat{\sigma}_{\hat{\theta}} \leqslant \theta \leqslant \hat{\theta}+S\hat{\sigma}_{\hat{\theta}}, \tag{8}$$

no matter what the values of all unknown parameters. The proof of this statement will be given in §2.

This result may be used for the interval estimation of all contrasts of interest, including any suggested by the way the observed means $\hat{\mu}_i$ fall out. No matter how many contrasts are estimated by the method (8), the probability that all the statements thus made will be correct will be $\geqslant 1-\alpha$.

The result may also be used to declare any estimated contrast 'significantly different from zero' or not, according as the corresponding interval (8) excludes $\theta = 0$ or not. More precisely, after selecting a set of coefficients c_i subject to (3) and thus determining a contrast we make one of the following three statements (it will be convenient also to say we make the statement *for* the contrast as well as *about* its estimate):

(i) $\hat{\theta}$ *is not significantly different from zero,*

(ii) $\hat{\theta}$ *is significantly different from zero and positive,*

(iii) $\hat{\theta}$ *is significantly different from zero and negative.*

We make statement (i) if $-S\hat{\sigma}_{\hat{\theta}} < \hat{\theta} < S\hat{\sigma}_{\hat{\theta}}$, (ii) if $\hat{\theta} \geqslant S\hat{\sigma}_{\hat{\theta}}$, (iii) if $\hat{\theta} \leqslant -S\hat{\sigma}_{\hat{\theta}}$. The operating characteristic of this method is studied in §4.

We warn the reader here that in the special case where all the $\hat{\mu}_i$ have the same variance, and all pairs $\hat{\mu}_i, \hat{\mu}_j$ $(i \neq j)$ have the same covariance, and where the *only* contrasts of interest are the $\frac{1}{2}k(k-1)$ differences $\mu_i-\mu_j$, the method of Tukey described in §3 should be used in preference to the above, because the confidence intervals will then be shorter. An example of such a case may be found in Scheffé (1952).

2. Proof of the method

The proof will be made with the aid of a linear transformation and other mathematical apparatus which will again be useful in §5 for proving results about the operating characteristic stated in §4. Although the coefficients of the transformation will be regarded as known in the mathematical discussion, they need not be computed for the practical applications.

The dimension of the space of estimated contrasts $\hat{\theta}$, regarded as linear forms in indeterminates $\hat{\mu}_1$, is $k-1$. Under the assumptions we made about the rank of the covariance matrix (4), we may find a basis for this space such that the corresponding random variables $\hat{\eta}_1, \ldots, \hat{\eta}_{k-1}$, will be statistically independent with equal variance

$$\sigma^2_{\hat{\eta}} = C^2\sigma^2,$$

where C is some chosen positive constant. The choice of C for the purposes of §4 will be discussed there; it does not matter at present. Let $\hat{\eta}_k = k^{-\frac{1}{2}} \sum_1^k \hat{\mu}_i$. Then between the $\hat{\mu}_i$ and the $\hat{\eta}_i$ there will exist a non-singular linear relationship,

$$\hat{\mu}_i = \sum_{j=1}^{k} b_{ij}\hat{\eta}_j \quad (i = 1, \ldots, k), \tag{9}$$

where the b_{ij} are constants not depending on unknown parameters. Writing $E(\hat{\eta}_j) = \eta_j$, we get from (9)

$$\mu_i = \sum_{j=1}^{k} b_{ij}\eta_j \quad (i = 1, \ldots, k).$$

Then

$$\theta = \sum_{i=1}^{k} c_i \mu_i = \sum_{i=1}^{k} \sum_{j=1}^{k} c_i b_{ij} \eta_j.$$

The coefficient of η_k must be zero for all $c_1, \ldots, c_k$ satisfying (3), so b_{ik} does not depend on i, and we may write

$$\theta = \sum_{j=1}^{k-1} d_j \eta_j, \; \hat{\theta} = \sum_{j=1}^{k-1} d_j \hat{\eta}_j,$$

where

$$d_j = \sum_{i=1}^{k} c_i b_{ij} \quad (j = 1, \ldots, k-1).$$

Now

$$\sigma^2_{\hat{\theta}} = \sum_{j=1}^{k-1} d_j^2 \sigma^2_{\hat{\eta}} = C^2\sigma^2 \sum_{j=1}^{k-1} d_j^2. \tag{10}$$

A contrast θ for which $\sigma^2_{\hat{\theta}} = C^2\sigma^2$ will be called a *normalized* contrast and denoted by ϑ. For a normalized contrast

$$\vartheta = \sum_{1}^{k} c_i \mu_i = \sum_{1}^{k-1} d_i \eta_i, \tag{11}$$

we then have

$$\sum_{1}^{k-1} d_i^2 = 1.$$

Clearly it suffices to prove that the probability statement associated with (8) holds for the totality of normalized contrasts. We shall do this by means of some simple geometric considerations.

Let us introduce a $(k-1)$-dimensional space, which we shall call the y-space of points $y = (y_1, \ldots, y_{k-1})$, for graphing the parameter point or vector $\eta = (\eta_1, \ldots, \eta_{k-1})$, its estimate $\hat{\eta}$, and other quantities. The random variables $\sum_{1}^{k-1} (\hat{\eta}_i - \eta_i)^2/(C^2\sigma^2)$ and $\nu\hat{\sigma}^2/\sigma^2$ have independent χ^2 distributions with $k-1$ and ν D.F., respectively, and so $\nu/(k-1)$ times their quotient has the F-distribution. This yields the confidence sphere $\mathscr{S}$,

$$\sum_{i=1}^{k-1} (y_i - \hat{\eta}_i)^2 \leqslant S^2C^2\hat{\sigma}^2, \tag{12}$$

for the parameter point η, where S^2 is defined by (7). The probability is $1-\alpha$ that the parameter point η is covered by $\mathscr{S}$.

Now a normalized contrast is uniquely determined by a coefficient vector $d = (d_1, \ldots, d_{k-1})$ of unit length, and it will be convenient in this section and § 5 to identify the contrast with the vector. It is seen from (11) that the value ϑ of the contrast is the projection of η on d. Similarly, the value $\hat{\vartheta}$ of its estimate is the projection of $\hat{\eta}$ on d. The interval (8) when written for ϑ becomes

$$\hat{\vartheta} - SC\hat{\sigma} \leqslant \vartheta \leqslant \hat{\vartheta} + SC\hat{\sigma}. \tag{13}$$

If we lay this interval off on the vector d, its centre is the projection on d of the centre $\hat{\eta}$ of the sphere $\mathscr{S}$, and its half-length is the radius of $\mathscr{S}$; in other words, the interval (13) may be interpreted as the projection of $\mathscr{S}$ on d. The interval covers the true value of the contrast if and only if the projection of the point η on d lies in the projection of $\mathscr{S}$ on d. This happens for all vectors d if and only if $\mathscr{S}$ covers η, and the probability of this is $1-\alpha$.

3. COMPARISON WITH A METHOD OF TUKEY

In the special case where all the $\hat{\mu}_i$ have the same variance $a_{11}\sigma^2$ and all pairs $\hat{\mu}_i, \hat{\mu}_j$ $(i \neq j)$ have the same covariance $a_{12}\sigma^2$, the following method of Tukey* is applicable. The probability is $1-\alpha$ that the values θ of all the contrasts simultaneously satisfy

$$\hat{\theta} - T\hat{\sigma} \leqslant \theta \leqslant \hat{\theta} + T\hat{\sigma}, \tag{14}$$

where the constant T is defined as

$$T = \frac{1}{2}\sum_1^k |c_i| \, q(a_{11}-a_{12})^{\frac{1}{2}}, \tag{15}$$

and q is the upper α point of the Studentized range, for the range of a sample of k in the numerator, and ν D.F. in the denominator, that is, the upper α point of the quotient w/s, where w and s $(s > 0)$ are statistically independent, w is the range of a random sample of k standard normal deviates, and νs^2 has the χ^2 distribution with ν D.F. This has been tabled by J. M. May (1952) for $\alpha = 0{\cdot}05$ and $0{\cdot}01$, and needs to be tabled for $\alpha = 0{\cdot}10$.

We propose to compare the efficiency of the two methods by use of the ratio R of the squared lengths of the confidence intervals (8) and (14),

$$R = (S^2\hat{\sigma}^2_{\hat{\theta}})/(T^2\hat{\sigma}^2).$$

The motivation for using the squared lengths is that if for a particular contrast the value of R thus defined is R_0, we may say that for large samples the method (8) requires R_0 times as many measurements as (14) to give the same accuracy on this contrast.

After noting that
$$\sigma^2_{\hat{\theta}} = (a_{11}-a_{12})\,\sigma^2 \sum_1^k c_i^2$$

in the present case, we may express the ratio R as

$$R = \frac{S^2}{q^2}\left(\sum_1^k c_i^2\right)\bigg/\left(\frac{1}{2}\sum_1^k |c_i|\right)^2. \tag{16}$$

The contrasts usually of the greatest practical interest are perhaps those consisting of the difference between the average of m of the μ_i and the average of r of the other μ_i $(m+r \leqslant k)$; we shall symbolize this type of contrast by $\{m, r\}$. For instance, a contrast of the type $\{2, 3\}$ is

$$\tfrac{1}{2}(\mu_2+\mu_7) - \tfrac{1}{3}(\mu_1+\mu_4+\mu_8). \tag{17}$$

It may be shown that R attains its maximum value $R_{\text{max.}}$ for a contrast of the type $\{1, 1\}$, that is, a difference of two μ_i, and its minimum value $R_{\text{min.}}$ for one of the type $\{\frac{1}{2}k, \frac{1}{2}k\}$ if k is even, $\{\frac{1}{2}(k-1), \frac{1}{2}(k+1)\}$ if k is odd, and hence

$$R_{\text{max.}} = 2(S^2/q^2),$$
$$R_{\text{min.}} = \begin{cases} 4k^{-1}(S^2/q^2) & \text{for } k \text{ even}, \\ 4k(k^2-1)^{-1}(S^2/q^2) & \text{for } k \text{ odd}. \end{cases}$$

Table 1 shows how the relative efficiency of the two methods varies with k, the number of means, in the case $\nu = \infty$. The rows headed $1/R_{\text{max.}}$ show the efficiency of method (8) relative to method (14) on contrasts which are differences of two μ_i, $R_{\text{min.}}$ shows the efficiency of method (14) relative to method (8) on some other contrasts. The value of S^2/q^2 is also tabled

* See footnote on p. 88.

for use with (16) for the calculation of R. Table 2 shows that the value of R is not very sensitive to ν. Of course any increase of S^2/q^2 above its value for $\nu = \infty$ listed in Table 1 favours method (14).

Table 1. *Variation of relative efficiency R of two methods for different contrasts* ($\nu = \infty$)

α	Value of	For $k=$ 2	3	4	5	6	8	10	13	16	20
0·10	S^2/q^2	0·50	0·55	0·60	0·64	0·69	0·78	0·86	0·98	1·09	1·23
	$1/R_{\text{max.}}$	1·00	0·91	0·84	0·78	0·73	0·64	0·58	0·51	0·46	0·40
	$R_{\text{min.}}$	1·00	0·82	0·60	0·54	0·46	0·39	0·34	0·30	0·27	0·25
0·05	S^2/q^2	0·50	0·54	0·59	0·64	0·68	0·76	0·85	0·96	1·07	1·20
	$1/R_{\text{max.}}$	1·00	0·92	0·84	0·79	0·73	0·65	0·59	0·52	0·47	0·42
	$R_{\text{min.}}$	1·00	0·82	0·59	0·53	0·45	0·38	0·34	0·30	0·27	0·24
0·01	S^2/q^2	0·50	0·54	0·59	0·63	0·67	0·74	0·81	0·92	1·01	1·14
	$1/R_{\text{max.}}$	1·00	0·92	0·85	0·80	0·75	0·67	0·61	0·55	0·49	0·44
	$R_{\text{min.}}$	1·00	0·81	0·59	0·52	0·44	0·37	0·33	0·28	0·25	0·23

Table 2. *Values of S^2/q^2, showing dependence of relative efficiency on ν*

α	ν \ k	3	6	10	13	16	20
0·05	5	0·56	0·70	0·89	1·03	1·15	1·31
	7	0·55	0·69	0·87	1·00	1·13	1·28
	10	0·55	0·69	0·87	0·99	1·11	1·26
	20	0·55	0·68	0·86	0·98	1·09	1·25
	40	0·55	0·68	0·85	0·97	1·08	1·23
	∞	0·54	0·68	0·85	0·96	1·07	1·20
0·01	5	0·57	0·72	0·91	1·05	1·18	1·34
	7	0·55	0·69	0·87	1·00	1·12	1·27
	10	0·55	0·68	0·85	0·98	1·09	1·24
	20	0·55	0·68	0·84	0·96	1·07	1·21
	40	0·54	0·67	0·83	0·94	1·04	1·18
	∞	0·54	0·67	0·81	0·92	1·01	1·14

The difference between the two methods is seen to increase with k; as k increases (8) gets relatively worse on the differences $\mu_i - \mu_j$, while (14) gets relatively worse on some other contrasts (and gets worse faster). It is clear that the choice between the two methods in the special case where (14) is applicable depends on which kind of contrast we think will be of interest.* (It is, of course, not permissible to use both on the same data and then choose the

* For the reader who has not skipped §2, a further comparison of the two methods is contained in the following loose but perhaps helpful statements: In using method (8) the user 'pays' for more than he will use. If he knew beforehand just which contrasts were going to interest him, the corresponding tangent planes to the confidence sphere $\mathscr{S}$ of §2, normal to the coefficient vectors d of the contrasts, could be constructed, and these would bound a circumscribed polyhedron $\mathscr{H}$. The user pays for the information that the parameter point is in $\mathscr{S}$ but he uses only the information that it is in $\mathscr{H}$. In

one with the results we like better—unless we are willing to settle for an overall confidence coefficient known only to be $\geqslant 1-2\alpha$, in which case we may choose for each contrast the shorter of the two intervals.) If we are interested exclusively in the differences $\mu_i - \mu_j$, we should choose (14); if we are interested in many types of contrasts, and in investigating contrasts suggested by the data, (8) would seem superior. Tables 3a and 3b show the relative efficiencies in the cases $k = 4$ and $k = 6$ on most contrasts likely to be of practical

Table 3a. *Relative efficiency of two methods when* $k = 4$ $(\alpha = 0{\cdot}05, \nu = \infty)$

Type of contrast	$1/R$	R
{1, 1}	0·84	
{1, 2}		0·89
{1, 3}		0·79
{2, 2}, quadratic		0·59
Linear, cubic		0·74

Table 3b. *Relative efficiency when* $k = 6$ $(\alpha = 0{\cdot}05, \nu = \infty)$

Type of contrast	$1/R$	R
{1, 1}	0·73	
{1, 2}	0·98	
{1, 3}		0·91
{1, 4}		0·85
{1, 5}		0·82
{2, 2}		0·68
{2, 3}		0·57
{2, 4}		0·51
{3, 3}		0·45
Linear		0·59
Quadratic		0·57
Cubic		0·48
Quartic		0·53
Quintic		0·67

interest.* The type $\{m, r\}$ is described above (17). The contrasts for linear, quadratic, etc., effects are the contrasts used in fitting orthogonal polynomials when the μ_i correspond to equal steps of some independent variable. The coefficients c_i for this type of contrast are listed in Fisher & Yates's tables (1943). The relative efficiency of method (14) is given in the column headed R, of (8) in the column $1/R$. It is seen that for $k = 4$, method (14) is superior only on the type {1, 1}, for $k = 6$ also the type {1, 2}. These tables are for $\nu = \infty$ and require some correction in favour of (14) for small values of ν; the correction factor is the ratio of entries in Table 2 for the given ν and for $\nu = \infty$ at $\alpha = 0{\cdot}05$.

principle it is possible to calculate an $\mathscr{H}'$ obtained by a uniform contraction of $\mathscr{H}$ about the centre of $\mathscr{S}$ and to pay just for this; in practice this would be a hopelessly complicated calculation for most cases. In the special case where the $\hat{\mu}_i$ have equal variances and equal covariances, and the contrasts of interest are the $\frac{1}{2}k(k-1)$ differences $\mu_i - \mu_j$, Tukey's method does calculate the $\mathscr{H}'$. If $\mathscr{H}'$ is available, we can of course infer about other contrasts by projecting $\mathscr{H}'$ on their vectors d, as we projected $\mathscr{S}$. This is equivalent to (14) for Tukey's method, and we will see below that this does not work as well for many contrasts of interest. We suggest that it will often be worthwhile to the customer to pay for more than he will use if he does not know beforehand just what he will need and can buy a good package.

* Interactions may also be regarded as contrasts. For example, if μ_{ij} is the true mean for the (i, j) cell in a two-way lay-out, the two-factor interaction in the (1, 2) cell is $\mu_{12} - \mu_{1.} - \mu_{.2} + \mu_{..}$, where the dots have their usual connotation, and this is a contrast among the μ_{ij}. General superiority of (8) over (14) for the interaction contrasts is indicated by some numerical calculations.

4. Operating characteristic of the method

For many purposes it may suffice to decide whether an experiment has adequate sensitivity by considering the lengths of the confidence intervals (8). In this section we consider the effects of the method from a viewpoint close to the Neyman-Pearson (1933) concept of the two kinds of error possible in hypothesis-testing. We preface this rather long development with a few remarks.

The power of the method will turn out to seem low to one accustomed to power calculations for a single t-test.* This has its counterpart in estimation, in that for $k > 2$ our confidence interval (8) for a contrast will usually be much longer than the $100(1-\alpha)$ % confidence interval for this contrast alone based on the t-distribution (which it is, of course, not valid

Table 4. *Values of* t^2/S^2

α	ν \ k	2	3	4	5	6	8	10	13	16	20
0·10	5	1·00	0·54	0·37	0·29	0·24	0·17	0·14	0·10	0·08	0·07
	7	1·00	0·55	0·39	0·30	0·25	0·18	0·15	0·11	0·09	0·07
	10	1·00	0·56	0·40	0·32	0·26	0·19	0·16	0·12	0·10	0·08
	20	1·00	0·57	0·42	0·33	0·28	0·21	0·17	0·13	0·11	0·09
	40	1·00	0·58	0·42	0·34	0·28	0·22	0·18	0·14	0·11	0·09
	∞	1·00	0·59	0·43	0·35	0·29	0·23	0·18	0·15	0·12	0·10
0·05	5	1·00	0·57	0·41	0·32	0·26	0·19	0·15	0·12	0·10	0·08
	7	1·00	0·59	0·43	0·34	0·28	0·21	0·17	0·13	0·11	0·09
	10	1·00	0·61	0·45	0·36	0·30	0·23	0·18	0·14	0·12	0·09
	20	1·00	0·62	0·47	0·38	0·32	0·25	0·20	0·16	0·13	0·11
	40	1·00	0·63	0·48	0·39	0·33	0·26	0·21	0·17	0·14	0·12
	∞	1·00	0·64	0·49	0·40	0·35	0·27	0·23	0·18	0·15	0·13
0·01	5	1·00	0·61	0·45	0·36	0·30	0·22	0·18	0·14	0·11	0·09
	7	1·00	0·64	0·48	0·39	0·33	0·25	0·20	0·16	0·13	0·10
	10	1·00	0·66	0·51	0·42	0·36	0·28	0·23	0·18	0·15	0·12
	20	1·00	0·69	0·55	0·46	0·39	0·31	0·26	0·21	0·18	0·14
	40	1·00	0·71	0·57	0·48	0·42	0·33	0·28	0·23	0·19	0·16
	∞	1·00	0·72	0·58	0·50	0·44	0·36	0·31	0·25	0·22	0·18

to apply if the choice of the contrast has been suggested by the configuration of the observed means). The ratio of the squared lengths of the confidence intervals, namely, t^2/S^2, where t is the upper $\frac{1}{2}\alpha$ point of the t-distribution with ν D.F. and S^2 is given by (7), is listed in Table 4. While these ratios depart far from unity for $k > 2$, the writer was somewhat surprised that the departure was not greater. For $k = 6$, for example, it requires only about three times as many measurements to get confidence intervals for as many contrasts as we please, including any suggested by the data, with 95 % confidence that all are correct, as for a 95 % confidence interval for a single one of the contrasts selected before the data are examined, the confidence interval in both cases having the same expected length.

* Those accustomed in applied statistics to making repeated t-tests at the 5 % significance level might consider choosing $\alpha = 10$ % rather than 5 % with the new method (8): The user of the repeated 5 % tests is working at some 'overall' significance level that is unknown but greater than 5 %; perhaps he would be glad to settle for a guaranteed 10 %. How bad the repeated t method may get to be from the 'overall' point of view is indicated below for confidence intervals in connexion with Tables 5 and 6.

Instead of considering the present method from the point of view of one accustomed to the method of repeated t-tests, or repeated confidence intervals based on the same data, it is instructive also to do the reverse. The theory of the method (8) permits the calculation of a lower bound for the overall confidence coefficient implied by the use of repeated t confidence intervals for the contrasts, calculated from the same data. Thus, for repeated 95 % confidence intervals based on t we may calculate the bound by substituting in place of S in (7) the two-tailed 5 % point of t with ν D.F., and solving the resulting equation for $1-\alpha$. The values of this bound shown in Table 5 may startle the reader when he considers that by increasing the number of repeated t confidence intervals, the overall probability

Table 5. *Lower bound for probability that all repeated t confidence intervals for contrasts will be correct* ($\nu = \infty$)

Conf. coeff. \ k	2	3	4	5	6	8	10	13	16	20
0·90	0·90	0·74	0·56	0·39	0·25	0·09	0·03	0·003	$0{\cdot}0^{3}2$	$<0{\cdot}0^{4}1$
0·95	0·95	0·85	0·72	0·57	0·43	0·20	0·08	0·01	0·002	$0{\cdot}0^{4}8$
0·99	0·99	0·96	0·92	0·84	0·75	0·53	0·32	0·12	0·03	0·004

Table 6. *Lower bound in certain cases where repeated t confidence intervals are used only on differences* $\mu_i - \mu_j$ ($\nu = \infty$)

Conf. coeff. \ k	2	3	4	5	6	8	10	13	16	20
0·90	0·90	0·77	0·65	0·53	0·43	0·28	0·17	0·08	0·04	0·01
0·95	0·95	0·88	0·80	0·71	0·63	0·49	0·37	0·24	0·15	0·08
0·99	0·99	0·97	0·95	0·93	0·90	0·84	0·77	0·67	0·58	0·48

may be brought arbitrarily close to this bound. If repeated t confidence intervals are used only on the differences $\mu_i - \mu_j$, and if the $\hat{\mu}_i$ have equal variances and equal covariances, then the theory of Tukey's method (14) leads to a better bound, found by solving for $1-\alpha$ the equation $q = 2^{\frac{1}{2}} t_{0{\cdot}05}$, where q is the same function of α, k, ν as in (15). This bound is attained if all $\frac{1}{2}k(k-1)$ statements are made about the differences $\mu_i - \mu_j$. Values of the bound are shown in Table 6. We may try to take comfort from the thought that if we always made all $\frac{1}{2}k(k-1)$ statements based on 5 % t, we should in the long run of experiments have 95 % of all the statements we made correct, but we should remember that of the statements to which we are likely to pay the most attention more than 5 % tend to be wrong; for example, statements associated with the larger observed differences.

The sensitivity of the new method is exactly the same as that of the F-test of the hypothesis H in (5) at significance level α in the following sense: If the F-test accepts the hypothesis H that all the contrasts are zero, the new method will make statement (i) for all contrasts, that is, say no contrast is significantly different from zero; if the F-test rejects the hypothesis, the new method will make statements (ii) and (iii) for some contrasts, that is, say they are significantly different from zero (and positive or negative, respectively). This may readily be seen* from the geometrical picture introduced in § 2.

To formulate the problem of the power of the method, we consider the probabilities of making each of the statements (i), (ii) and (iii). We might think at first that what we would like is to make statement (i) for a contrast whose true value is zero, (ii) for one whose true value is greater than some assigned θ_0, (iii) for one whose true value is less than $-\theta_0$, and that what we want are the probabilities of the desired statements in each case. This formulation is at once seen to be nonsensical, because for any contrast θ_1 whose true value is positive there exist constants h' and h'' such that the true value of $h'\theta_1$ is $<\theta_0$, the true value of $h''\theta_1$ is $>\theta_0$, and yet the same one of the statements (i), (ii) and (iii) will be made for all contrasts $h\theta_1$ with positive h. This difficulty can be avoided by laying our requirements on the suitably normalized contrasts; if we find what happens for the normalized contrasts, we will know what happens for all contrasts. The appropriate definition is that a contrast is normalized if the variance of its estimate is $C^2\sigma^2$, where C is a specified constant; for normalized contrasts and their estimates we then write ϑ and $\hat{\vartheta}$ instead of θ and $\hat{\theta}$.

To see how the above difficulty is met, suppose the structure of an experiment is such that the contrast $\mu_1-\mu_2$ is determined with greater precision than the contrast $\mu_4-\mu_5$. This is a property of the design which presumably was desired, and if we guarantee a certain probability P of detecting (that is, making statement (ii) for) a difference $\mu_1-\mu_2$ as great as θ_0, we should be satisfied with the same probability P of detecting a difference $\mu_4-\mu_5$ as great as a certain multiple of θ_0. The multiple assigned by the present method will later be seen to be the ratio of the standard errors of estimate of the two contrasts. Since this property may be expressed by saying that 'the method assures the same probability P of detecting all *normalized* contrasts as great as a certain bound ϑ_0', it indicates that we have a way of normalizing the contrasts which is suitable for this method.

While this motivates our normalizing the contrasts so that their estimates all have the same variance $C^2\sigma^2$, there is still the question of the choice of the constant C. Whether there is a satisfactory universal choice of C is dubious.† In any event, the question is of little practical importance, and we need not settle it in order to analyse the operating characteristic. If for the same experiment two statisticians choose C and C', then equivalent choices of the bounds ϑ_0 and ϑ_0' are related by $\vartheta_0' = (C'/C)\,\vartheta_0$.

A simple example illustrating these considerations will be helpful. Suppose we wish to design an experiment in a 'one-way lay-out' with $k = 6$, that we wish to take twice as many observations on the means μ_1, μ_2, μ_3 as on μ_4, μ_5, μ_6, and that we are primarily interested in the contrasts which are differences $\mu_i-\mu_j$ (for equal numbers of observations per mean we

* In the notation of §2, H may be written $\eta=0$. Thus the F-test accepts H if and only if the confidence sphere $\mathscr{S}$ covers the origin, and this is precisely the case where statement (i) will be made about all the contrasts.

† A possible universal choice of C might be the following: transform from $\hat{\mu}_1, \ldots, \hat{\mu}_k$ to $\hat{\xi}_1, \ldots, \hat{\xi}_k$ by an orthogonal transformation, such that $\hat{\xi}_k$ is the $\hat{\eta}_k$ of §2, and $\hat{\xi}_1, \ldots, \hat{\xi}_{k-1}$ are statistically independent. The set $\hat{\xi}_1, \ldots, \hat{\xi}_{k-1}$ then spans the space of the estimated contrasts. Let $\sigma_i^2 = \mathrm{var}\,(\hat{\xi}_i)$. Define C so that $C\sigma$ is the geometric mean of $\sigma_1, \ldots, \sigma_{k-1}$. C is then a function of the a_{ij} in (6) alone. It may be shown that in the case of independent $\hat{\mu}_i$ with equal variances σ^2/n this gives $C=n^{-\frac{1}{2}}$; more generally if the $\hat{\mu}_i$ are independent with respective variances σ^2/n_i this gives $C^{2(k-1)}=\bar{n}\Big/\prod_1^k n_i$, where $\bar{n}=\sum_1^k n_i/k$. In the example below, this leads to $C^2=(\tfrac{3}{16})^{\frac{1}{5}}/n$, which does not particularly recommend itself from the practical computational point of view.

should then use Tukey's method). These contrasts will then be determined with three kinds of precision, namely, the precision of those like $\mu_1-\mu_2$, like $\mu_4-\mu_5$, and like $\mu_1-\mu_4$, the three standard errors of estimate being in the ratio $2^{\frac{1}{2}} : 4^{\frac{1}{2}} : 3^{\frac{1}{2}}$. Suppose we are satisfied with a probability of $P = 0{\cdot}90$ of detecting a difference $\mu_1-\mu_2$ as great as $1{\cdot}0\sigma$. We will then get the same probability 0·90 for a difference $\mu_4-\mu_5$ as great as $1{\cdot}0\sigma(4^{\frac{1}{2}}/2^{\frac{1}{2}}) = 1{\cdot}4\sigma$, or a difference $\mu_1-\mu_4$ as great as $1{\cdot}0\sigma(3^{\frac{1}{2}}/2^{\frac{1}{2}}) = 1{\cdot}2\sigma$. If for either of the last two differences this sensitivity is not considered sufficient then we should question whether we have chosen the correct design.

Let n be the number of observations on μ_4, μ_5, μ_6, and $2n$ the number on μ_1, μ_2, μ_3. While the choice of the normalizing constant C does not matter much and might even be left indefinite, we shall suppose for concreteness in this example that the contrasts are normalized so that their estimates all have the same variance as that of $\mu_1-\mu_2$, namely, σ^2/n. Since for any contrast

$$\operatorname{var}\left(\sum_1^k c_i\hat{\mu}_i\right) = \left(\frac{1}{2}\sum_1^3 c_i^2 + \sum_4^6 c_i^2\right)\sigma^2/n,$$

the coefficients in a normalized contrast $\vartheta = \sum_1^k c_i\mu_i$ must then satisfy the condition

$$\frac{1}{2}\sum_1^3 c_i^2 + \sum_4^6 c_i^2 = 1.$$

Thus the normalized form of the contrast $\mu_4-\mu_5$ is $2^{-\frac{1}{2}}(\mu_4-\mu_5)$, and of $\mu_1-\mu_4$, $(\frac{2}{3})^{\frac{1}{2}}(\mu_1-\mu_4)$. Below we shall find how to choose n so that the probability is 0·90 of detecting a difference of $\mu_1-\mu_2$ equal to a given multiple of σ, say $1{\cdot}0\sigma = \vartheta_0$. The probability will then be the same for the value ϑ_0 of the contrast $2^{-\frac{1}{2}}(\mu_4-\mu_5)$ or the value ϑ_0 of the contrast $(\frac{2}{3})^{\frac{1}{2}}(\mu_1-\mu_4)$. This gives the statements in the last paragraph about the differences $\mu_4-\mu_5$ and $\mu_1-\mu_4$. In such an experiment the contrast consisting of the difference of the average of μ_1, μ_2, μ_3 and the average of μ_4, μ_5, μ_6 might also be of interest. Its normalized form is

$$(\tfrac{2}{9})^{\frac{1}{2}}\left(\sum_1^3 \mu_i - \sum_4^6 \mu_i\right),$$

and we find in a similar way that a difference of the averages $\sum_1^3 \mu_i/3$ and $\sum_4^6 \mu_i/3$ equal to $(\frac{1}{3})(\frac{9}{2})^{\frac{1}{2}}\vartheta_0 = 0{\cdot}7\sigma$ would also be detected with probability 0·90.

In the following calculations we shall symbolize by $t'(f;\,\delta)$ a non-central t variable with f D.F. and non-centrality parameter δ, that is, a variable distributed as the quotient of $z+\delta$ by $f^{-\frac{1}{2}}\chi$, where z is a standard normal deviate and χ is an independent χ variable with f D.F.

Statements (i), (ii) or (iii) are made for a contrast θ according as $\hat{\theta}/\hat{\sigma}_{\hat{\theta}}$ falls in the intervals $(-S, S)$, (S, ∞), or $(-\infty, -S)$. But the variable $\hat{\theta}/\hat{\sigma}_{\hat{\theta}}$ has the non-central t-distribution with ν D.F. and parameter $\delta = \theta/\sigma_{\hat{\theta}}$, where $\sigma_{\hat{\theta}}$ is given by (6). It follows that the probability of each of these statements being made depends only on the true value of $\theta/\sigma_{\hat{\theta}}$ for the particular contrast for which the statement is made, and on the constant S.

If the true value of θ is zero, the probability that we make the desired statement (i) is

$$\Pr\{-S \leqslant t'(\nu;\,0) \leqslant S\},$$

where $t'(\nu;\,0)$ denotes a central t variable with ν D.F. This may also be written

$$\Pr\{F(1,\nu) \leqslant (k-1)\,F_\alpha(k-1,\nu)\},$$

where $F(1, \nu)$ is an F variable with 1 and ν D.F. That this is at least $1-\alpha$ is evident from the confidence statement associated with (8). Indeed, as k increases from 2, this probability increases rapidly from $1-\alpha$ as indicated in Table 7, where 1 minus this probability is tabled for the case $\nu = \infty$.

Table 7. *Probability of not making statement* (i) *about a contrast whose true value is zero* ($\nu = \infty$)

α \ k	2	3	4	5	6	8	10	13	16	20
0·10	0·10	0·032	0·012	0·0053	0·0024	$0{\cdot}0^{3}53$	$0{\cdot}0^{3}13$	$0{\cdot}0^{4}17$	$0{\cdot}0^{5}23$	$0{\cdot}0^{6}18$
0·05	0·050	0·014	0·0052	0·0021	$0{\cdot}0^{3}88$	$0{\cdot}0^{3}18$	$0{\cdot}0^{4}39$	$0{\cdot}0^{5}45$	$0{\cdot}0^{6}57$	$0{\cdot}0^{7}40$
0·01	0·010	0·0024	$0{\cdot}0^{3}76$	$0{\cdot}0^{3}27$	$0{\cdot}0^{3}10$	$0{\cdot}0^{4}17$	$0{\cdot}0^{5}32$	$0{\cdot}0^{6}31$	$0{\cdot}0^{7}32$	$0{\cdot}0^{8}18$

For any contrast θ the probability that we make statement (ii) is

$$\Pr\{t'(\nu;\ \theta/\sigma_{\hat\theta}) > S\}. \tag{18}$$

This is a strictly increasing function of the value of $\theta/\sigma_{\hat\theta}$. Suppose now the contrast is normalized; we then write it as ϑ and have $\sigma_{\hat\vartheta} = C\sigma$. Writing P for the probability (18) when $\vartheta = \vartheta_0$, we have

$$P = \Pr\{t'(\nu;\ \vartheta_0/(C\sigma)) > S\}. \tag{19}$$

From the tables of non-central t by Johnson & Welch (1940) we can find the value of $\delta = \vartheta_0/(C\sigma)$ for which P attains the values 0·01, 0·05, 0·1 (0·1) 0·9, 0·95, 0·99.

The probability of making statement (iii) for any contrast can of course be calculated as the probability of making statement (ii) for its negative.

In the example introduced above of the experiment with 6 means, $C = n^{-\frac{1}{2}}$. If we take $n = 10$ and use significance level $\alpha = 0{\cdot}05$, then $\nu = 84$, $S^2 = 5F_{0{\cdot}05}(5, 84) = (3{\cdot}41)^2$, and we find from (19) and the Johnson & Welch tables that for $P = 0{\cdot}90$, $\delta = 4{\cdot}72$, so $\vartheta_0 = 4{\cdot}72C\sigma = 1{\cdot}49\sigma$. This poor sensitivity is related to the extremely low risk (0·0009 in Table 7 for $\nu = \infty$) of making the other kind of error, namely, calling significantly different from zero the estimate of a contrast whose true value is zero. We supposed in this example that we desired the n that would give a $P = 0{\cdot}90$ for detecting a $\vartheta = 1{\cdot}0\sigma = \vartheta_0$. Since for large ν the parameter δ found from the tables does not vary much with ν we try for a first approximation for n the solution of $\vartheta_0/(C\sigma) = (1{\cdot}0)\, n^{\frac{1}{2}} = \delta$ with the former $\delta = 4{\cdot}72$. This gives $n = 22$. Calculation of ϑ_0 from the tables for $n = 21, 22$, as above for $n = 10$, shows that $n = 22$ is the required value. In the same way we find that if we desire $\vartheta_0 = 2{\cdot}0\sigma$ for $P = 0{\cdot}90$, a first approximation to the required n is $n = (4{\cdot}72/2{\cdot}0)^2 \doteqdot 6$; use of the tables also gives $n = 6$. In all these calculations we took $\alpha = 0{\cdot}05$.

Although the behaviour of the operating characteristic which we have now determined for any single contrast considered alone is of interest, it seems to the writer that in this method designed for judging all the contrasts it is of greater interest to determine the following two properties of the operating characteristic: *First, the probability P_1 of the event $\mathscr{E}_1$ that the desired statement* (i) *would be made for all the contrasts whose true values are zero, and secondly, the probability P_2 of the event $\mathscr{E}_2$ that the desired statement* (ii) *would be made for all normalized contrasts whose values are $\geqslant \vartheta_0$ (in which case the desired statement* (iii) *would also be made for all normalized contrasts whose values are $\leqslant -\vartheta_0$).*

We shall state here the results for the probabilities P_1 and P_2 and defer the proofs to the next section. It is obvious from (8) that if the hypothesis H in (5) is true, $P_1 = 1-\alpha$. If H is false

$$P_1 = \Pr\{F(k-2, \nu) \leqslant (k-1)(k-2)^{-1} F_\alpha(k-1, \nu)\}, \tag{20}$$

where $F(m, \nu)$ denotes an F variable with m and ν D.F., and $F_\alpha(m, \nu)$, its upper α point. It is interesting to note that P_1 depends only on whether H is true or false and not further on any unknown parameters. It is easy to see that if H is false $P_1 > 1-\alpha$. Table 8 gives some values of $1-P_1$; we remark that for $k = 2$ there are no zero contrasts other than $0\cdot\mu_1 + 0\cdot\mu_2$ if H is false, and so $1-P_1 = 0$ if H is false, trivially. We see now that this analogue of the Neyman-Pearson risk of a 'type I' error, namely, the probability $1-P_1$ of failing to make statement (i) for *all* zero contrasts, is the one the method controls in a satisfactory way, rather than the marginal probability for a *single* zero contrast considered above, which assumes the microscopic values indicated by Table 7.

Table 8. *Values of* $1-P_1$ *if* H *is false*

α	ν \ k	3	4	5	6	8	10	13	16	20
0·10	5	0·040	0·056	0·065	0·070	0·077	0·082	0·086	0·088	0·091
	7	0·038	0·053	0·061	0·067	0·074	0·079	0·083	0·086	0·089
	10	0·036	0·050	0·058	0·064	0·071	0·076	0·081	0·084	0·087
	20	0·034	0·047	0·055	0·060	0·067	0·072	0·077	0·080	0·083
	40	0·033	0·046	0·053	0·058	0·065	0·069	0·074	0·077	0·080
	∞	0·032	0·044	0·051	0·055	0·062	0·066	0·070	0·073	0·075
0·05	5	0·019	0·027	0·031	0·034	0·038	0·040	0·042	0·044	0·045
	7	0·018	0·025	0·029	0·032	0·036	0·039	0·041	0·042	0·044
	10	0·017	0·024	0·028	0·031	0·035	0·037	0·040	0·041	0·043
	20	0·016	0·022	0·026	0·028	0·032	0·035	0·037	0·039	0·041
	40	0·015	0·021	0·025	0·027	0·031	0·033	0·035	0·037	0·039
	∞	0·014	0·020	0·023	0·026	0·029	0·031	0·033	0·035	0·036
0·01	5	0·0036	0·0051	0·0060	0·0067	0·0074	0·0079	0·0084	0·0087	0·0089
	7	0·0033	0·0047	0·0056	0·0062	0·0070	0·0075	0·0080	0·0083	0·0086
	10	0·0030	0·0044	0·0052	0·0058	0·0066	0·0071	0·0076	0·0080	0·0083
	20	0·0027	0·0039	0·0047	0·0052	0·0060	0·0065	0·0070	0·0074	0·0078
	40	0·0026	0·0037	0·0044	0·0049	0·0056	0·0061	0·0066	0·0070	0·0074
	∞	0·0024	0·0034	0·0041	0·0045	0·0052	0·0056	0·0061	0·0063	0·0067

Let $\vartheta_{\text{max.}}$ denote the largest of the true values of all the normalized contrasts. If $\vartheta_0 > \vartheta_{\text{max.}}$ the probability P_2 is trivially 1. If $\vartheta_0 \leqslant \vartheta_{\text{max.}}$ define the angle γ from

$$\sin\gamma = \vartheta_0/\vartheta_{\text{max.}}. \tag{21}$$

Then it will be shown that in this case P_2 is the probability that

$$(\cot\gamma)\,\chi_1 + (S\nu^{-\frac{1}{2}}\operatorname{cosec}\gamma)\,\chi_2 \leqslant z_1 + (C\sigma)^{-1}\vartheta_0\operatorname{cosec}\gamma, \tag{22}$$

where χ_1, χ_2, z_1 are statistically independent, χ_1 and χ_2 are χ variables with $k-2$ and ν D.F., respectively, and z_1 is a standard normal deviate. The probability P_2 does not seem exactly calculable from any existing tables, but an excellent approximation* may be obtained for

* Johnson & Welch (1940) used a similar method to approximate non-central t.

moderate or large values of ν by replacing $\nu^{-\frac{1}{2}}\chi_2$ by a normal variable z_2 with mean 1 and variance $1/(2\nu)$. A similar approximation would not work so well for χ_1, since its number $k-2$ of D.F. might be quite small. This way we find P_2 approximately equal to the probability that

$$z+\delta \geqslant t_0(k-2)^{-\frac{1}{2}}\chi_1,$$

where

$$z = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}(z_1-z_2S\operatorname{cosec}\gamma+S\operatorname{cosec}\gamma)$$

is a standard normal deviate and independent of χ_1;

$$\delta = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}[(C\sigma)^{-1}\vartheta_0\operatorname{cosec}\gamma-S\operatorname{cosec}\gamma], \tag{23}$$

$$t_0 = [1+(2\nu)^{-1}S^2\operatorname{cosec}^2\gamma]^{-\frac{1}{2}}(k-2)^{\frac{1}{2}}\cot\gamma. \tag{24}$$

To this approximation P_2 can thus be expressed in terms of the non-central t-distribution as

$$\Pr\{t'(k-2;\,\delta)>t_0\}, \tag{25}$$

where δ and t_0 are given by (23) and (24). In the case where σ is known we can put $\nu=\infty$ in (22), so that the left member becomes $(\cot\gamma)\chi_1+S\operatorname{cosec}\gamma$; we may also put $\nu=\infty$ in (23) and (24), and the above approximation for P_2 then becomes exact.

The exact probability P_2 and its approximation (25) depend on the parameter $\vartheta_{\text{max.}}$, which enters (22) through γ. As indicated in the example of the six means, we may be willing to specify ϑ_0 beforehand as a multiple of σ, say

$$\vartheta_0 = A\sigma, \tag{26}$$

where A is a given constant, but it is unpleasant to discover then that P_2 still depends on the unknown parameter

$$\psi = \vartheta_{\text{max.}}/\sigma. \tag{27}$$

The dependence of the exact probability P_2 of (22) on ψ may be made clear by writing (22) as

$$(\psi^2A^{-2}-1)^{\frac{1}{2}}\chi_1+(\nu^{-\frac{1}{2}}S\psi/A)\chi_2 \leqslant z_1+\psi C^{-1}, \tag{28}$$

where the variables χ_1, χ_2, z_1 have the distribution stated below (22). For $\psi>A$, P_2 is the probability of (28); for $0\leqslant\psi\leqslant A$, P_2 trivially equals 1. This problem did not arise when we obtained in (19) the corresponding marginal probability for a single contrast alone whose value $=\vartheta_0$ (the non-centrality parameter δ there may be written A/C), as it does now when we consider the overall probability for all contrasts whose value is $\geqslant\vartheta_0$ (we could write $=\vartheta_0$ here also).

We suggest meeting this problem by using a lower bound P_2' for P_2. We shall prove that $P_2=P_2(\psi)$ is monotone in ψ, decreasing from the value 1 for $\psi\leqslant A$ to a limiting value $P_2(\infty)$ as ψ increases from 0 to ∞. This means that if we are willing to assume ψ does not exceed a known ψ_1, we may use as the bound P_2' the value $P_2(\psi_1)$, which may be accurately approximated by (23), (24), (25) with $\gamma=\arcsin(A/\psi_1)$, and if we are unwilling to set a bound ψ_1 for ψ, or if we are satisfied with a somewhat larger but much more easily calculated bound, we may use $P_2'=P_2(\infty)$, whose value will now be stated.

The limiting value $P_2(\infty)$ is the probability that

$$\chi_1+\nu^{-\frac{1}{2}}S\chi_2\leqslant A/C, \tag{29}$$

where χ_1 and χ_2 are distributed as in (22). This may be accurately approximated by the same method that led to (25), with the result that $P_2(\infty)$ is approximately*

$$\Pr\{t'(k-2;\,\delta)>t_0\}, \tag{30}$$

* It is unfortunate in connexion with (30) and (25) that the Johnson & Welch tables (1940) do not go below 4 D.F., since 1, 2, 3 D.F. are needed for $k=3$, 4, 5; it would be very desirable to have this extension of their tables.

where
$$\delta = (2\nu)^{\frac{1}{2}}[A(CS)^{-1} - 1], \tag{31}$$
$$t_0 = (2\nu)^{\frac{1}{2}}(k-2)^{\frac{1}{2}}/S. \tag{32}$$

For very large values of ν we may use the following simple approximation to $P_2(\infty)$, obtained by replacing $\nu^{-\frac{1}{2}}\chi_2$ by 1 in (29):
$$\Pr\{\chi_1 \leqslant AC^{-1} - S\}.$$

To this approximation we may thus rapidly find for given β the A such that $P_2(\infty)$ attains the value $1-\beta$ for $\vartheta_0 = A\sigma$ by taking the square root of the upper β point from the χ^2 tables for $k-2$ D.F., say $\chi_{\beta;k-2}$, and computing
$$A \doteqdot C(S + \chi_{\beta;k-2}). \tag{33}$$

A still rougher approximation might be obtained by replacing S by its value for $\nu = \infty$, namely, $\chi_{\alpha;k-1}$, to give
$$A \doteqdot C(\chi_{\alpha;k-1} + \chi_{\beta;k-2}). \tag{34}$$

The approximation (34) becomes exact when σ is known; this is obtained by replacing $\hat{\sigma}$ by σ and ν by ∞ in the calculations.

To illustrate these results in the example we have carried along, we recall $k = 6$, $C = n^{-\frac{1}{2}}$, $\alpha = 0{\cdot}05$. Suppose we wish a bound P_2' for P_2 and decide to use $P_2' = P_2(\infty)$. If we take $n = 10$, then $\nu = 84$, and entering the Johnson & Welch tables (1940) with $t_0 = 7{\cdot}61$ from (32) we find that a $P_2(\infty) = 0{\cdot}90$ (to the approximation of (30)) is attained for $\delta = 10{\cdot}8$, and this gives $A = 1{\cdot}98$ from (31), that is, for $\vartheta_0 = 1{\cdot}98\sigma$. The approximation (33) much more quickly gives $A = 6{\cdot}20C = 6{\cdot}20n^{-\frac{1}{2}}$ or $A = 1{\cdot}96$. The approximation (34) gives $A = 1{\cdot}93$. To find the n which gives a $P_2(\infty) = 0{\cdot}90$ for $\vartheta_0 = 1{\cdot}0\sigma$, we first use the previous approximation from (33), $A = 6{\cdot}20n^{-\frac{1}{2}}$, which is calculated with an S based on the now wrong $\nu = 84$, but then S is not sensitive to changes in ν for large ν. This gives $n = 38$ for $A = 1{\cdot}0$; the more correct formula (30) also gives $n = 38$. For most applications we would probably have to compromise on a larger A to get a more feasible n. In a similar way for $n = 5$, $\nu = 39$, $\alpha = 0{\cdot}05$, $C = 5^{-\frac{1}{2}}$, we find first $S = 3{\cdot}50$, then $A = 2{\cdot}81$ from the approximation (33); the more accurate (30) gives $A = 2{\cdot}87$.

5. DERIVATION OF P_1, P_2, ETC.

The probabilities P_1 and P_2 can be found neatly by continuing the geometric approach of §2. Extending the notation there, denote by $|\eta|$ the length of the vector η, $|\eta|^2 = \sum_1^{k-1} \eta_i^2$. Since the value ϑ of a normalized contrast is the projection of η on d, it is evident that
$$|\eta| = \vartheta_{\max}. \tag{35}$$

Suppose for the moment $\eta \neq 0$ (equivalent to H false). If we imagine marking off the projection of η on d for each unit vector d to get a polar co-ordinate graph of the totality of values of the normalized contrasts, we see the graph consists of a sphere which has the vector η as a diameter. The zero-valued contrasts for which we desire to make statement (i) constitute the $(k-2)$-dimensional set of vectors d in the tangent plane $\mathscr{P}$ to the sphere at the origin. If $\eta = 0$ this picture collapses in an obvious way.

Consider now the totality of normalized contrasts whose values are $\geqslant \vartheta_0$, for which we wish to make the statement (ii). There are none if $|\eta| < \vartheta_0$. If $|\eta| \geqslant \vartheta_0$ they fill a circular cone $\mathscr{C}$ (by 'cone' we mean throughout *one nappe* of a cone) with axis along η, vertex at the origin, and elements making an angle $\frac{1}{2}\pi - \gamma$ with η, where γ is defined by (22). The cone $\mathscr{C}'$ obtained by reflecting $\mathscr{C}$ in the origin is filled by the normalized contrasts whose values are $\leqslant -\vartheta_0$.

In the same y-space the estimated values of the normalized contrasts have a similar graph, the diameter of the sphere being $\hat{\eta}$. From the interpretation of the confidence interval (13) for the value ϑ of any normalized contrast d as the projection of the sphere $\mathscr{S}$ on d, it is easy to see that if $|\hat{\eta}| < SC\hat{\sigma}$ the statement (i) is made for all the contrasts, while if $|\hat{\eta}| \geqslant SC\hat{\sigma}$ the normalized contrasts for which we make the statement (ii) fill a cone $\mathscr{D}$ with axis along $\hat{\eta}$, vertex at the origin, and elements making an angle arccos $(SC\hat{\sigma}/|\hat{\eta}|)$ with $\hat{\eta}$, and the normalized contrasts for which we make the statement (iii) fill the reflexion $\mathscr{D}'$ of $\mathscr{D}$ in the origin.

If $\eta \neq 0$, the event $\mathscr{E}_1$ of probability P_1 happens if and only if the projection of the sphere $\mathscr{S}$ on the plane $\mathscr{P}$ covers the origin. It is convenient now to rotate the axes in the y-space so that the vector η lies along the positive y_{k-1}-axis. Denote the new coordinates of $\hat{\eta}$ by $\hat{\zeta}_1, \ldots, \hat{\zeta}_{k-1}$, and of η by $\zeta_1, \ldots \zeta_{k-1}$, respectively. Then $\hat{\zeta}_1, \ldots, \hat{\zeta}_{k-1}, \hat{\sigma}^2$ will be independent, the $\hat{\zeta}_1$ will be normal with variance $C^2\sigma^2$, the means $\zeta_1, \ldots, \zeta_{k-2}$ will all be zero, while $\zeta_{k-1} = |\eta|$. The plane $\mathscr{P}$ now has the equation $y_{k-1} = 0$. The distance of the centre of $\mathscr{S}$ from the y_{k-1}-axis is $\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}}$, and so the projection of $\mathscr{S}$ on $\mathscr{P}$ will cover the origin if and only if this distance does not exceed the radius of $\mathscr{S}$,

$$\sum_1^{k-2} \hat{\zeta}_i^2 \leqslant S^2C^2\hat{\sigma}^2,$$

or

$$\frac{\sum_1^{k-2} \hat{\zeta}_i^2/(k-2)}{C^2\hat{\sigma}^2} \leqslant (k-1)(k-2)^{-1} F_\alpha(k-1, \nu).$$

Since the random variable on the left side of the latter inequality has the F-distribution with $k-2$ and ν D.F., we have now succeeded in deriving (20).

If $|\eta| \geqslant \vartheta_0$, the event $\mathscr{E}_2$ happens if and only if the fixed cone $\mathscr{C}$ lies entirely inside the random cone $\mathscr{D}$. It may be verified that this is equivalent to the confidence sphere $\mathscr{S}$ lying entirely within a cone $\mathscr{F}$ with axis along η, vertex at the origin, and elements making with η an angle γ defined by (21). The sphere $\mathscr{S}$ will touch the cone $\mathscr{F}$ on the inside if the distance of its centre from the y_{k-1}-axis plus $\sec\gamma$ times the radius of the sphere equals $\zeta_{k-1}\tan\gamma$. Thus the event $\mathscr{E}_2$ happens if and only if

$$\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}} + CS\hat{\sigma}\sec\gamma \leqslant \hat{\zeta}_{k-1}\tan\gamma.$$

If we now divide this inequality through by $C\sigma\tan\gamma$ we get the desired result for P_2 stated in connexion with (22).

The monotone behaviour of $P_2(\psi)$ stated in §4 was suggested by the above geometric picture of the sphere $\mathscr{S}$ and the cone $\mathscr{F}$, and may be established by geometric arguments. It will be simpler to give here an analytic proof utilizing the inequality (22). Denote by $f(\gamma)$ the conditional probability, given χ_2, that this inequality is satisfied. Clearly it will suffice to show that for all $\chi_2 > 0$, $f(\gamma)$ is a monotone decreasing function of ψ for $\psi > A$, or since $\operatorname{cosec}\gamma = \psi/A$, that the derivative $f'(\gamma)$ is positive.

Let us write the case of equality in (22) as

$$\chi_1 = z_1 \tan\gamma + B\sec\gamma, \tag{36}$$

where

$$B = AC^{-1} - \nu^{-\frac{1}{2}}S\chi_2.$$

Because of the statistical independence of χ_1, χ_2, z_1, the conditional probability $f(\gamma)$ may be found by treating χ_2 as a constant and working with the joint distribution of χ_1 and z_1 in the z_1, χ_1-plane: $f(\gamma)$ is the amount of probability in this plane above the z_1-axis and below the line (36). We shall drop the subscripts from χ_1 and z_1 in the rest of this calculation. The z-intercept of the line (36) is $-B \operatorname{cosec} \gamma$, and the probability $f(\gamma)$ may thus be expressed as the integral

$$f(\gamma) = \int_{-B \operatorname{cosec} \gamma}^{\infty} p_1(z)\, g(z, \gamma)\, dz,$$

where

$$g(z, \gamma) = \int_0^{z \tan \gamma + B \sec \gamma} p_2(\chi)\, d\chi,$$

and $p_1(z)$ and $p_2(\chi)$ are the densities of a standard normal deviate and of a χ variable with $k-2$ D.F., respectively, and do not depend on γ. Now

$$f'(\gamma) = -p_1(-B \operatorname{cosec} \gamma)\, g(-B \operatorname{cosec} \gamma, \gamma)\, \partial(-B \operatorname{cosec} \gamma)/\partial\gamma + \int_{-B \operatorname{cosec} \gamma}^{\infty} p_1(z)\, [\partial g(z, \gamma)/\partial\gamma]\, dz,$$

and

$$\begin{aligned} \partial g(z, \gamma)/\partial\gamma &= p_2(z \tan \gamma + B \sec \gamma)\, \partial(z \tan \gamma + B \sec \gamma)/\partial\gamma \\ &= (z \sec^2 \gamma + B \sec \gamma \tan \gamma)\, p_2(z \tan \gamma + B \sec \gamma). \end{aligned}$$

But $g(-B \operatorname{cosec} \gamma, \gamma) = 0$, and hence

$$f'(\gamma) = \int_{-B \operatorname{cosec} \gamma}^{\infty} (z \sec^2 \gamma + B \sec \gamma \tan \gamma)\, p_1(z)\, p_2(z \tan \gamma + B \sec \gamma)\, dz. \tag{37}$$

If $B \leqslant 0$ the integrand is $\geqslant 0$ on the range of integration and then clearly $f'(\gamma) > 0$. If $B > 0$ we transform the integral (37) by the substitution $w = z \sec \gamma + B \tan \gamma$ to get

$$f'(\gamma) = \int_{-B \cot \gamma}^{\infty} p_1(w \cos \gamma - B \sin \gamma)\, p_2(w \sin \gamma + B \cos \gamma)\, dw.$$

By utilizing the explicit forms of the densities p_1 and p_2 this may be written

$$f'(\gamma) = \int_{-B \cot \gamma}^{\infty} v(w)\, dw, \tag{38}$$

where

$$v(w) = D[u(w)]^{k-3}\, w\, e^{-\frac{1}{2}w^2},$$

D is a positive constant, and

$$u(w) = w \sin \gamma + B \cos \gamma \tag{39}$$

is positive for $z > -B \cot \gamma$. If we break the range of integration in (38) up into the intervals $(-B \cot \gamma, 0)$, $(0, B \cot \gamma)$, $(B \cot \gamma, \infty)$, and drop the last, where $v(w) > 0$, we get

$$f'(\gamma) > \int_0^{B \cot \gamma} [v(-w) + v(w)]\, dw. \tag{40}$$

For $0 < w < B \cot \gamma$, $u(-w)$ is seen from (39) to be $< u(w)$, hence $|v(-w)| \leqslant v(w)$ since $k \geqslant 3$, and so the integrand in (40) is $\geqslant 0$ and therefore $f'(\gamma) > 0$.

To calculate the limiting value $P_2(\infty)$ we revert to the geometric picture. To make $\psi \to \infty$ we may let $|\eta| \to \infty$ with fixed σ. For $|\eta| > \vartheta_0$, we recall that $P_2(\psi)$ is the probability that the sphere $\mathscr{S}$ lies inside the cone $\mathscr{F}$. The cone $\mathscr{F}$ may be constructed by circumscribing a cone with vertex at the origin about a sphere $\mathscr{T}$ with centre at $|\eta|$ on the y_{k-1}-axis and radius ϑ_0. As $|\eta| \to \infty$ this picture becomes indeterminate. However, we clearly get the same result if we hold the sphere $\mathscr{T}$ and the joint distribution of $\hat{\zeta}_1, \ldots, \hat{\zeta}_{k-1}, \hat{\sigma}^2$ fixed and let the vertex of the circumscribed cone $\mathscr{F}$ go to $-\infty$ on the y_{k-1}-axis. The limiting figure for the cone $\mathscr{F}$ is now the cylinder circumscribed about $\mathscr{T}$ with elements parallel to the y_{k-1}-axis. The probability $P_2(\infty)$ that the sphere $\mathscr{S}$ fall in this cylinder is the probability that

$$\left(\sum_1^{k-2} \hat{\zeta}_i^2\right)^{\frac{1}{2}} + CS\hat{\sigma} \leqslant \vartheta_0.$$

Division of this inequality by $C\sigma$ yields the desired (29) and concludes the derivations.

The writer is indebted to Mr Seiji Sugihara and Mr Judah Rosenblatt for making the numerical calculations. Tables of Merrington & Thompson (1943) were used in the calculation of Tables 1, 2, 4 and 8, tables of Pearson & Hartley (1942) for Tables 1 and 6, tables of May (1952) for Table 2, tables of Merrington (1942) for Table 4, tables of Hartley & Pearson (1950) for Table 5, tables of Thompson (1941) and the W.P.A. tables (1942) for Table 7, and tables of Karl Pearson (1934) for Table 8.

REFERENCES

EISENHART, C. (1947). The assumptions underlying the analysis of variance. *Biometrics*, **3**, 1–21.
FISHER, R. A. (1935). *The Design of Experiments*, §24. Edinburgh: Oliver and Boyd.
FISHER, R. A. & YATES, F. (1943). *Statistical Tables*, Table 23. Edinburgh: Oliver and Boyd.
HARTLEY, H. O. & PEARSON, E. S. (1950). Tables of the χ^2-integral and of the cumulative Poisson distribution. *Biometrika*, **37**, 313–25.
JOHNSON, N. L. & WELCH, B. L. (1940). Applications of the noncentral t-distribution. *Biometrika*, **31**, 362–89.
MAY, JOYCE M. (1952). Extended and corrected tables of the upper percentage points of the 'Studentized' range. *Biometrika*, **39**, 192–3.
MERRINGTON, M. (1942). Table of percentage points of the t-distribution. *Biometrika*, **32**, 300.
MERRINGTON, M. & THOMPSON, C. M. (1943). Tables of the percentage points of the inverted beta (F) distribution. *Biometrika*, **33**, 73–88.
MOOD, A. M. (1950). *Introduction to the Theory of Statistics*, §§14·9–14·11. New York: McGraw-Hill.
NANDI, H. K. (1951). On the analysis of variance test. *Bull. Calcutta Statist. Ass.* **3**, 103–14.
NEWMAN, D. (1939). The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika*, **31**, 20–30.
NEYMAN, J. & PEARSON, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans.* A, **231**, 289–337.
PEARSON, E. S. & HARTLEY, H. O. (1942). The probability integral of the range in samples of n observations from a normal population. *Biometrika*, **32**, 301–10.
PEARSON, K. (1934). *Tables of the Incomplete Beta-Function.* London: Biometrika Office.
SCHEFFÉ, H. (1952). An analysis of variance for paired comparisons. *J. Amer. Statist. Ass.* **47**, 381–400.
THOMPSON, C. M. (1941). Table of percentage points of the χ^2-distribution. *Biometrika*, **32**, 187–91.
TUKEY, J. W. (1951). Quick and dirty methods in statistics. Part II. Simple analyses for standard designs. *Proc. Fifth Annual Convention, Amer. Soc. for Quality Control*, pp. 189–97.
W.P.A. TABLES (1942). *Tables of Probability Functions*, vol. **2**. Washington, D.C.: Nat. Bur. Standards.