Choosing Between Tukey's hsd and Scheffé's Procedure for
Testing a Set of Non-Orthogonal Linear Contrasts*

D. S. Robson
Cornell University
Ithaca, New York

and

C. Vithayasai
Rice Department
Ministry of Agriculture
Thailand

BU-169-M

October, 1964

ABSTRACT

The hsd may be applied to any linear contrast $\sum_{i=1}^{p} a_i\bar{x}_i$ among p treatment means having equal variances by comparing

$$\sum a_i\bar{x}_i \quad \text{vs} \quad \text{hsd}\left(\frac{1}{2}\sum_{1}^{p}|a_i|\right) .$$

If all of the contrasts to be tested are pairwise comparisons then the hsd procedure is uniformly better than Scheffé's procedure. If only a few of the a_i's are non-zero in each contrast then, again, the hsd procedure may be better, though in this case the choice between the two procedures must be based on a comparison of their calculated critical values.

---

* Paper No. BU-169-M of the Biometrics Unit Mimeo Series, Cornell University

Choosing Between Tukey's hsd and Scheffé's Procedure for Testing a Set of Non-Orthogonal Linear Contrasts*

D. S. Robson
Cornell University
Ithaca, New York

and

C. Vithayasai
Rice Department
Ministry of Agriculture
Thailand

BU-169-M

October, 1964

The Studentized range which, for a set of p independent and identically normally distributed sample means and an error mean square with ν degrees of freedom, takes the form

$$q(p,\nu) = \frac{\bar{x}_{max}-\bar{x}_{min}}{s_{\bar{x}}}$$

has been suggested by Tukey [1] as a basis for making all possible p(p-1)/2 pairwise comparisons among the p means. The probability that none of the differences $|\bar{x}_i-\bar{x}_j|$ will exceed the hsd = $q_\alpha(p,\nu)s_{\bar{x}}$ is the same as the probability that the largest of these differences will not exceed the hsd, and May [2] has tabulated the critical values of $q_\alpha(p,\nu)$ for which

$$P_{H_0}\left(\frac{\bar{x}_{max}-\bar{x}_{min}}{s_{\bar{x}}} > q_\alpha(p,\nu)\right) = \alpha$$

If, in addition to or instead of making all possible pairwise comparisons among the p treatment means, the experimenter wishes to test a set of linear contrasts then a particular form of the hsd procedure still applies. This may be seen by first noting that if a linear contrast is expressible in the form of an average

---

* Paper No. BU-169-M in the Biometric Unit Mimeo Series, Cornell University.

of pairwise treatment differences, as for example

$$\frac{1}{3}\bar{x}_1 + \frac{1}{3}\bar{x}_2 + \frac{1}{3}\bar{x}_3 - \bar{x}_4 = \frac{(\bar{x}_1-\bar{x}_4)+(\bar{x}_2-\bar{x}_4)+(\bar{x}_3-\bar{x}_4)}{3} \quad ,$$

then this contrast may be tested directly against the <u>hsd</u> since, as with any of the individual $p(p-1)/2$ pairwise comparisons,

$$\left| \frac{1}{3}(\bar{x}_1-\bar{x}_4) + \frac{1}{3}(\bar{x}_2-\bar{x}_4) + \frac{1}{3}(\bar{x}_3-\bar{x}_4) \right| \le \frac{1}{3}|\bar{x}_1-\bar{x}_4| + \frac{1}{3}|\bar{x}_2-\bar{x}_4| + \frac{1}{3}|\bar{x}_3-\bar{x}_4|$$

$$\le \bar{x}_{max} - \bar{x}_{min} \quad .$$

Furthermore, since the above inequality holds generally for any weighted mean of pairwise differences, then <u>all</u> linear contrasts which are expressible as averages of pairwise differences will be less than the <u>hsd</u> in absolute value if and only if $\bar{x}_{max} - \bar{x}_{min} <$ <u>hsd</u>.

Next, we note that any linear contrast of the form $\sum_{i=1}^{p} a_i\bar{x}_i$ where $\sum_{i=1}^{p} a_i=0$ is expressible as a linear function of the $p(p-1)/2$ differences,

$$(1) \qquad \sum_{i=1}^{p} a_i\bar{x}_i = \sum_{i<j}^{p} b_{ij}(\bar{x}_i-\bar{x}_j) \quad ,$$

and that for a fixed set of coefficients on the left hand side the set of coefficients on the right hand side can generally be chosen in an infinite variety of ways. For any particular choice of the b_{ij} coefficients we have the relation

$$\left|\sum_{i=1}^{p} a_i\bar{x}_i\right| = \Big(\sum_{i<j}^{p}|b_{ij}|\Big)\left|\frac{\sum\limits_{i<j}^{p} b_{ij}(\bar{x}_i-\bar{x}_j)}{\sum\limits_{i<j}^{p}|b_{ij}|}\right|$$

$$\leq \Big(\sum_{i<j}^{p}|b_{ij}|\Big)\frac{\sum\limits_{i<j}^{p}|b_{ij}|\;|\bar{x}_i-\bar{x}_j|}{\sum\limits_{i<j}^{p}|b_{ij}|}$$

$$\leq \Big(\sum_{i<j}^{p}|b_{ij}|\Big)\Big(\bar{x}_{max}-\bar{x}_{min}\Big)\;.$$

Furthermore, since a variety of choices for the b_{ij} coefficients is generally available for a fixed set of a_i coefficients, we may write

$$\left|\sum_{i=1}^{p} a_i\bar{x}_i\right| \leq \inf\Big(\sum_{i<j}^{p}|b_{ij}|\Big)\Big(\bar{x}_{max}-\bar{x}_{min}\Big)\;.$$

Equality may hold only if $\Sigma a_i\bar{x}_i$ is itself a pairwise comparison.

With the a_i coefficients fixed, the minimum value which can be taken by $\Sigma|b_{ij}|$ is the sum of the positive a_i coefficients. If we identify the positive a_i coefficients as $a_{i_1}, a_{i_2}, \cdots, a_{i_k}$ $(1 \leq k \leq p-1)$ then, noting that the coefficients of $\bar{x}_{i_\nu}$ on the left and right hand sides of (1) must be equal, we see immediately that $\Sigma|b_{ij}|$ cannot be less than $\sum\limits_{\upsilon=1}^{k} a_{i_\upsilon}$, and if the remaining non-positive a_i coefficients are designated as $a_{j_1}, a_{j_2}, \cdots, a_{j_{p-k}}$ so that

$$\sum_{\upsilon=1}^{k} a_{i_\upsilon} = - \sum_{\varkappa=1}^{n-k} a_{j_\varkappa} = \frac{1}{2} \sum_{i=1}^{p} |a_i|$$

then we see that the choice of b_{ij} as

$$b^*_{i_\upsilon j_\varkappa} = \frac{2 \left| a_{i_\upsilon} a_{j_\varkappa} \right|}{\sum_{i=1}^{p} |a_i|} \quad \text{for } \upsilon = 1, \cdots, k \text{ and } \varkappa = 1, \cdots, n-k$$

and

$$b^*_{ij} = 0 \text{ otherwise}$$

would give us this minimum value,

$$\sum_{i<j} b^*_{ij} = \sum_{i<j} |b^*_{ij}| = \inf \sum_{i<j} |b_{ij}| = \sum_{\nu=1}^{k} a_{i_\nu} \quad .$$

We then arrive at the conclusion if the experimenter were to test each of his linear contrasts $\Sigma a_i \bar{x}_i$ by comparing

$$\frac{2 \left| \sum a_i \bar{x}_i \right|}{\sum |a_i|} \quad \text{vs} \quad q_\alpha(p, \nu) s_{\bar{x}}$$

then under the null hypothesis the probability of erroneously finding one or more of the contrasts to be significant is less than or equal to α. This probability is exactly α only if the set of linear contrast contains all possible pairwise comparisons.

Comparison with Scheffé's procedure

Scheffé [3] has shown that if each one of a set of linear contrasts $\sum_1^p a_i X_i \left(\sum_1^p a_i = 0\right)$ among p independent normal and identically distributed random variables $X_1, \cdots, X_p$ is tested against a corresponding critical value of

$$\sqrt{(p-1)F_\alpha(p-1,\nu)s_\nu^2 \sum_1^p a_i^2} \equiv S_\alpha s_\nu \sqrt{\sum_1^p a_i^2}$$

then the probability that no contrast will exceed its critical value in absolute size is $1-\alpha$. Here $F_\alpha(p-1,\nu)$ denotes the $100(1-\alpha)$ percentage point of an F-distribution on p-1 and ν degrees of freedom, while s_ν^2 is an estimate of the variance of an X_i and is distributed as $\sigma^2\chi_\nu^2/\nu$ independently of $X_1, \cdots, X_p$.

We have now seen that although Tukey's hsd was developed only for testing pairwise comparisons the procedure may be extended to test any linear contrast in the same manner as Scheffé's procedure. Since both methods have an experiment-wise error rate of α they are directly comparable, and the comparison is between the critical values

$$\frac{1}{2} q_\alpha \sum_1^p |a_i| s_{\bar{x}} \quad \text{vs} \quad S_\alpha \sqrt{\sum a_i^2} \; s_{\bar{x}}$$

or more simply

$$q_\alpha \quad \text{vs} \quad 2S_\alpha \frac{\sqrt{\sum a_i^2}}{\sum |a_i|} .$$

For pairwise comparisons $\left(\sum_{1}^{p} a_i^2 = 1^2 + (-1)^2 = 2\right)$ the <u>hsd</u> is always smaller than the corresponding critical value for Scheffé's procedure; that is

$$q_\alpha(p,\nu) < S_\alpha(p-1,\nu)\sqrt{2}$$

for all α, p, ν. This follows from the fact that

$$(\bar{x}_{max} - \bar{x}_{min})^2 \leq \sum_{i<j}^{p} (\bar{x}_i - \bar{x}_j)^2 = 2 \sum_{i=1} (\bar{x}_i - \bar{\bar{x}})^2$$

and hence

$$\alpha = P\left\{ \frac{\bar{x}_{max} - \bar{x}_{min}}{s_{\bar{x}}} > q_\alpha \right\} < P\left\{ \sqrt{\frac{2\sum (\bar{x}_i - \bar{\bar{x}})^2}{s_{\bar{x}}^2}} > q_\alpha \right\} .$$

Thus, as would be expected, Scheffé's procedure is never an admissible substitute for the <u>hsd</u> in testing a set of linear contrasts involving <u>only</u> pairwise comparisons.

For a set of contrasts involving only three means at a time, as in $\bar{x}_{i_1} - (\bar{x}_{i_2} + \bar{x}_{i_3})/2$, Scheffé's procedure may again be inadmissible depending now, however, on the values of α, ν and p. In fact, the same may be said more generally of a set of contrasts involving at most k means at a time, $k < p$, though k must remain small relative to p if all critical values in the hsd procedure are to be smaller than the corresponding critical values in Scheffé's procedure. Among contrasts of the type

$$\sum_{i=1}^{p} a_i\bar{x}_i = \sum_{j=1}^{k} a_{i_j}\bar{x}_{i_j} \quad , \quad |a_{i_j}| \neq 0 \text{ for } j=1,\cdots,k$$

involving exactly k non-zero a_i's the smallest possible value for the ratio $2\sqrt{\Sigma a_i^2}/\Sigma|a_i|$ is $\sqrt{k/\left(\left[\frac{k}{2}\right]\left[\frac{k+1}{2}\right]\right)}$; hence with k even, for example, if $q_\alpha(p,\nu) < 2S_\alpha(p-1,\nu)/\sqrt{k}$ then the hsd procedure should certainly be used in preference to Scheffé's method for testing this set of contrasts. Table 1 displays this relation between the two kinds of critical values for fixed $\alpha = .05$ and $\nu = 60$ with varying values for p and k; the pattern remains essentially the same for other values of α and ν.

Table 1. Comparison between the hsd and the smallest possible critical value for Scheffé's procedure when any one contrast involves at most k of the p treatment means.

		$S_{.05}(p-1,60)\sqrt{k/\left(\left[\frac{k}{2}\right]\left[\frac{k+1}{2}\right]\right)}$			
p	$q_{.05}(p,60)$	k=5	k=4	k=3	k=2
20	5.24	5.29	5.79	7.09	8.19
15	5.00	4.66	5.10	6.25	7.21
10	4.65	3.91	4.28	5.25	6.06
5	3.98	2.90	3.18	3.89	4.49

In any practical situation where there does appear to be a question as to which of the two testing procedures is preferable the choice can best be made by comparing q_α with the calculated values of $2S_\alpha \sqrt{\Sigma a_i^2}/\Sigma|a_i|$ for all of the contemplated contrasts. This approach is demonstrated below with a problem arising in a taxonomic investigation.

Illustration

A botanist has collected plants of a single species from the six different localities depicted schematically in Figure 1. The four easternmost locations fall near the eastern seaboard of the North American continent, and of the two inland locations the point e is at the higher elevation. If morphological differences persist when cuttings from these plants are grown together under common environmental conditions, the botanist will tentatively conclude that these evolutionary differences reflect the differing selective forces imposed by climatic and soil conditions at the several locations. Because of the particular configuration of these six geographic points he is therefore interested in the particular set of contrasts given in Table 2.

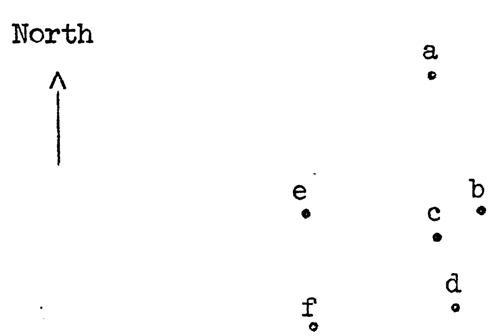

Figure 1. Relative locations of the points of collection of plant specimens

With $p = 6$ and, as in Table 1, $\alpha = .05$ and $\nu = 60$ the tabulated value for $q_{.05}$ is 4.16 and the corresponding value of $S_{.05}$ is $\sqrt{5(2.37)} = 3.44$. The estimated standard error for a contrast of the type $\bar{x}_a - \frac{1}{2}\bar{x}_b - \frac{1}{2}\bar{x}_c$ is $s_{\bar{x}} \sqrt{1^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = 1.225\, s_{\bar{x}}$ and using Scheffé's method the critical value for testing this type of contrast is therefore $3.44(1.225)\, s_{\bar{x}} = 4.21\, s_{\bar{x}}$, compared to $4.16\, s_{\bar{x}}$ for Tukey's method. This comparison is made in Table 2 for each of the nine contemplated contrasts. The one contrast for which Scheffé's method gives the smaller critical value is among the least interesting of the nine since point _e_, though at roughly the same latitude as _b_ and _c_, is also at a substantially higher elevation. We would therefore conclude that in this situation the choice must be the _hsd_.

Table 2. Comparison between $q_{.05}$ and $2S_{.05}\sqrt{\Sigma a_i^2}/\Sigma|a_i|$ for the set of nine contrasts generated from Figure 1. ($\nu = 60$)

Contrast	$q_{.05}$	$2S_{.05}\sqrt{\Sigma a_i^2}/\Sigma\|a_i\|$
a vs (b+c), d vs (b+c), e vs (b+c), a vs (f+d)	4.16	4.21
b vs c, e vs f, f vs d, a vs d	4.16	4.86
(f+d) vs (e+b+c)	4.16	3.14

REFERENCES

1. Tukey, J. W., "The problem of multiple comparisons", Ditto report, Princeton University, Princeton, N. J., 1953.

2. May, J. M., "Extended and corrected tables of the upper percentage points of the 'Studentized' range", Biometrika, 39:192-193, 1952.

3. Scheffé, H., "A method of judging all contrasts in the analysis of variance", Biometrika, 40:87-104, 1953.