DOCUMENT RESUME

ED 294 916 TM 011 643


| | |
|---|---|
| AUTHOR | Olejnik, Stephen F.; And Others |
| TITLE | Contrast Analysis for Scale Differences. |
| PUB DATE | Apr 88 |
| NOTE | 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988). |
| PUB TYPE | Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142) -- Reports - Research/Technical (143) |
| EDRS PRICE | MF01/PC02 Plus Postage. |
| DESCRIPTORS | *Comparative Analysis; Sample Size; *Scaling; Statistical Analysis |
| IDENTIFIERS | *Contrast Analysis; Scale Analysis |



ABSTRACT

Research on tests for scale equality have focused exclusively on an overall test statistic and have not examined procedures for identifying specific differences in multiple group designs. The present study compares four contrast analysis procedures for scale differences in the single factor four-group design: (1) Tukey HSD; (2) Kramer-Tukey; (3) Games-Howell; and (4) Dunnett T-3. Two data transformations are considered under several combinations of variance difference, sample sizes, and distributional forms. The data for the study were generated using the Statistical Analysis System computing package. The results indicate that no single transformation or analysis procedure is uniformly superior in controlling the family-wise error rate or in statistical power. The relationship between sample size and variances is a major factor in selecting a contrast analysis procedure. Adopting a multiple analysis strategy is recommended. Four tables show comparisons in these methods. (Author/SLD)



Reproductions supplied by EDRS are the best that can be made from the original document.

ED294916

# Contrast Analysis for Scale Differences

Stephen F. Olejnik
University of Georgia

James Algina
University of Florida

and

Kelly Hern
University of Georgia

Running Head: Contrast Analysis

Paper presented at the meetings of the American Educational Research Association in New Orleans, LA. April, 1988

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEPHEN OLEJNIK

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM011 643

# ABSTRACT

Research on tests for scale equality have focused exclusively on an overall test statistic and have not examined procedures for identifying specific differences in multiple group designs. The present study compares four contrast analysis procedures for scale differences in the single factor four group design. Two data transformations are considered under several conbinations of variance difference, sample sizes and distributional forms. The results indicate that no single transformation or analysis procedure is uniformly superior in controlling the familywise error rate or in statistical power. The relationship between sample size and variances is a major factor in selecting a contrast analysis procedure. Adopting a multiple analysis strategy is recommended.

# Contrast Analysis for Scale Differences

## Introduction

Recent investigations into procedures for comparing population variances (O'Brien, 1978; Conover, Johnson and Johnson, 1981; Olejnik and Algina, 1987) have shown that very few of the proposed tests for scale differences are appropriate when population distributions are non-normal. The only parametric procedures which have been shown to have actual Type I error rates which do no exceed the nominal significance level for a variety of non-normal distrtibutions are those suggested by O'Brien (1978) and Brown and Forsythe (1974). Based on statistical power considerations, O'Brien's procedure is generally recommended when population distributions are normal or platykurcic while the Brown-Forsythe procedure is recommended when the distributions are leptokurtic. Studies investigating tests for scale equality have all focused on the omnibus test of simultaneously examining all possible pairwise and complex contrasts. When tests of scale equality is computed to test whether a data assumption has been met, additional analyses may not be of interest. However, if the test for scale differences is theory based, identifying specifically which populations differ, may be of major importance and possibly of greater interest than the omnibus test. Researchers have not examined procedures that might be used to identify specific differences between populations in multiple group designs.

To test for scale equality both the O'Brien and the Brown-Forsythe procedures require that the original data be transformed but differ in the transformation functions used. The Brown-Forsythe procedure transforms the original data within each comparison group by the absolute difference between each observation in the group and the group median, ($Z_{ij}=|X_{ij}-m_j|$). O'Brien's procedure uses a more complex transformation function involving both deviations from group means and the group variance:

$$r_{ij}=[(w+n_j-2)n_j(X_{ij}-X_{.j})^2-ws^2(n_j-1)]/[(n_j-1)(n_j-2)],$$

where $s^2$ is the within-group unbiased estimate of variance and w is a weighting factor. Although several values of w could be used, O'Brien (1981) recommends letting w=.5 as a general approach. Both the O'Brien and the Brown-Forstyhe procedures compute an omnibus test for variance equality based on the ANOVA F-ratio using the transformed scores as the dependent measure. To identify specifically which populations differ one of several contrast procedures might be used. The statistical literature contains numerous articles identifying procedures which control Type I error rates at various levels (e.g. Games, 1971; Games & Howell, 1976; Rogan, Keselman, & Breen, 1977; Keselman & Rogan, 1977; Games, Keselman & Rogan, 1981) and which have the greatest statistical power (e.g. Keselman, 1976; Ramsey, 1981). Based on the results of simulation studies contrasting sample means the following recommendations are generally made: 1) when sample sizes are equal and within group variances are equal Tukey's HSD procedure is generally recommended; 2) if sample sizes differ and variances are equal a modification of Tukey's procedure suggested by Kramer (1956) is preferred and 3) if variances are unequal with equal or unequal sample sizes a procedure suggested by Howell and Games (1974) and Games and Howell (1976) based on a Welch-type adjustment for variance inequality has been recommended (Jaccard, Becker and Wood, 1984). Dunnett (1980) has suggested that the Games-Howell procedure might be improved by using the Studentized maximum modulus distribution rather than the Studentized range distribution as the reference distribution. Wilcox (1987, p. 186) recommends Dunnett's T3 approach when the degrees of freedom are less than 50. The extent to which these recommendations generalize to contrasts for scale differences is the purpose of the present study. Table 1 summarizes the computational formulas for the four procedures under investigation.

---

Insert Table 1 about here

---

O'Brien (1981) has cautioned that since the variance of a sample variance is influenced by the square of the population variance (Box, 1954) identifying specific differences in population variances will involve contrasts violating the assumption of variance equality. As a result O'Brien recommends a contrast analysis procedure which is valid when population variances are unequal, such as the procedure suggested by Games and Howell (1976). The purpose of the present study was to investigate the appropriateness of O'Brien's recommendation by comparing several alternative procedures for identifying pairwise differences in population variances under several conditions. There are several reasons however why the results from contrast analysis between means may not generalize to pairwise differences between population variances. First, both the Games-Howell and the Dunnett T3 procedures use a Welch type adjustment for the degrees of freedom. A recent investigation by Wilcox, Charlin and Thompson (1986) has shown that in the case of comparing population means, if sample sizes are unequal and population standard deviations differ in ratios of 4 to 1 that the Welch adjusted ANOVA can lead to a liberal test. O'Brien's transformation leads to much greater variability among observations within groups while the variability of the Brown-Forsythe transformed data is reduced. Thus the procedures using the Welch adjusted degrees of freedom with O'Brien transformed data may lead to a liberal test in some situations. A second concern is the distributional form. Both transformation procedures being studied modify the shape of the data distributions. Only limited work has been conducted investigating the Games-Howell procedure with non-normal distributions. Keselman and Rogan (1978) found the procedure appropriate for comparing means of samples from a chi-square distribution with 3 degrees of freedom. Other distributional forms have not been studied. Finally O'Brien's procedure has not been studied extensively in multiple group designs. The effect of the relationship between sample sizes and population variances has not been investigated for either the omnibus F-ratio or the contrast analysis procedures.

In the present study four contrast analysis procedures considered using both the O'Brien and the Brown-Forsythe transformed data. To evaluate these procedures Type I error rates for the complete and partial null conditions are examined as well as the statistical power when population variances are unequal. For the present study the control of the familywise error rate for pairwise contrasts is assumed to be of primary importance.

Method

Three factors were manipulated in generating the sampling distributions of the test statistics studied: population variances, sample sizes, and distribution form.

The investigation was limited to comparisons involving four populations and all 6 possible pairwise comparisons between them. Population variances in the following combinations were studied: 1,1,1,1; 4,1,1,1; 4,4,1,1; and 4,4,4,1. The latter three variance patterns reflect coefficients of variation equal to .74, .60. and .40 respectively. Nine sample size combinations were included: balanced designs of 10, 20, 30 per group;and unbalanced designs of 10,16,24,30; 10,20,40,50; and 6,12,42,60. For the unequal sample size conditions the reverse orderings were also included. Three distributional forms were included: normal (0,0), platykurtic (0,-1) and leptokurtic (0,3.75). Skewed distributions were not included since previous simulations and Box's (1954) analytic work has shown that it is the kurtosis of the sampled distribution which affects the tests for scale equality. The data for the study were generated using the Statistical Analysis System computing package. Scores on the dependent measure (Y) were generated using RANNOR, which is the normal random number generating function in SAS. The generated distributions had a mean of 10 and variance of 1 or 4. To generate non-normal distributions, the normal random variables were transformed using a polynomial power procedure suggested by Fleishman (1978).

For each of the 108 conditions studied (4x9x3 completely crossed factorial design) 1000 replications were generated. For each replication, the original scores were transformed using O'Brien's procedure and the procedure suggested by Brown and Forsythe. The omnibus F-ratio comparing the four populations was computed as well as the test statistics for the 6 possible pairwise contrasts using: the Tukey HSD, the Kramer-Tukey, the Games-Howell and the Dunnett T3 analysis procedures. Critical values for non-tabled Studentized range and the Studentized maximum modulus distributions were estimatel using harmonic interpolation with the inverse of the degrees of freedom as suggested by Harter (1960). For unequal sample sizes the harmonic mean of the sample sizes was used for Tukey's HSD procedure. Although this approach has been criticized in the past it is frequently adopted by researchers and is the approach taken by SAS in the GLM procedure. For all analyses the frequency at which the null hypothesis was rejected was recorded. The rejection rates for the contrast analysis procedures were recorded both ignoring the results for the overall test as well as the conditional results, that is, contrasts were tested only when the overall test was found significant.

## Results

The results reported here are only for the .05 nominal level of significance with the normal distribution. Similar patterns of results were obtained for th other



ERIC

distributional forms studied and for nominal significance levels of .10 and .01. In addition the results for the conditional tests and the analyses ignoring the overall test were very similar so only the results for the nonconditional tests are reported here. Additional results can be obtained from the first author. In evaluating the results, an analysis procedure was judged unacceptable when the observed proportion of Type I errors exceeded two standard errors (.014) from the theoretical Type I error rate. Observered Type I error rates less than .064 was judged acceptable and the usefullness of the procedure was evaluated based on a comparisons of power estimates.

Omnibus F-ratio. Type I error rates for the complete null condition and power estimates for the overall ANOVA F-ratio using O'Brien's approach and the Brown-Forsythe technique are reported in Table 2 for the normal distribution. Both procedures had acceptable observed Type I error rates for all conditions studied. Statistical power advantages depended both on the

Insert Table 2 about here

sample sizes and the pattern of variance differences. When sample sizes differed, O'Brien's procedure provided greater power when sample size and population variances were negatively related. When sample size and population variances were positively related the Brown-Forsthe procedure had greater statistical power. With equal sample sizes O'Brien's procedure had the power advantage only with the variance pattern of 4,1,1,1. With the platykurtic distribution a similar pattern of results were observed with the exception that O'Brien's procedure provided greater power when sample sizes were equal. With the leptokurtic distribution the Brown-Forsythe procedure provided greater statistical power except when the variance pattern was 4,1,1,1 and sample sizes were unequal and negatively related. Under those conditions O'Brien's procedure was slightly more powerful.

Contrast Analysis. Table 3 presents the observed Type I error rates for the four contrast analysis procedures using the O'Brien and Brown-Forsythe transformed data. Under the complete

Insert Table 3 about here

null condition the Kramer-Tukey procedure had acceptable Type I error rates for all conditions studied using either the O'Brien or Brown-Forsythe transformed data. Tukey's HSD, Games-Howell, and Dunnett's T3 procedures all had appropriate error rates when sample sizes were equal, but these procedures had inflated Type I errors when sample sizes differed considerably and O'Brien's transformation was used. Using the Brown-Forsythe transformation, Tukey's procedure had an acceptable Type I error rate for the unequal sample size conditions studied but the Games-Howell and Dunnett's T3 had inflated Type I error rates when sample sizes were unequal.

For the partial null condition all procedures had acceptable Type I error rates when sample sizes were equal for both the O'Brien and Brown-Forsythe transformed data. With unequal sample sizes, the results differed depending on the pattern of variance differences. For the 4,1,1,1 pattern all procedures had acceptable Type I errors for all conditions except the Games-Howell and the Dunnett T3 procedures when sam·le sizes were extremely different and postitively related to the variances. Under these condtions the two contrast analysis procedures were liberal with both O'Brien and Brown-Forsythe transformed data. When the variance pattern was either 4,4,1,1 or 4,4,4,1 a different set of conclusions were evident. Using either O'Brien or B.own-Forsythe transformed data, Tukey's HSD and the Kramer-Tukey procedures were extremely liberal when sample sizes and variances were negatively related and conservative when samole size and variances were postively related. The Games-Howell and Dunnett T3 procedures generally had acceptable Type I error rates for both transformation approaches. The Games-Howell procedure was liberal using the O'Brien transformation and sample sizes differed greatly. Dunnett's procedure was also liberal when sample sizes of 6,12,42,60 were sampled under the 4,4,4,1 variance pattern.

Table 4 reports the average power estimate per contrast. When the observed Type I error rate was unacceptably high the power estimate for that condition is not reported. The proportion of times in which all true differences were identified in a given replication was also recorded but the pattern of results were similar as those reported here. Power estimates for all true differences were much lower than the average per contrast

Insert Table 4 about here

power. When sample sizes were equal, the average power estimate per contrast for all procedures was less than the overall power estimate from the overall omnibus F-ratio. The average power per contrast was greater than the overall F-test for the Games-Howell and Dunnett T3 procedures however when sample sizes differed greatly and were positively related to the variance pattern of 4,4,4,1. With variance pattern 4,1,1,1, the average per contrast power of Tukey's HSD procedure was slightly greater than the overall F-test when sample size and variance were negatively related.

A comparison of power estimates for the contrast analysis procedures indicates that neither transformation procedure nor contrast analysis procedure is uniformly most powerful for the conditions studied. Tukey's HSD procedures using the O'Brien transformed data tended to be more sensitive when the variance pattern was 4,1,1,1 except when sample size and variance was positively related. Under that condition the Games-Howell procedure had greater power. For the other variance patter s studied, the Brown-Forsythe transformation led to greater statistical power when sample sizes were equal or when the relationship between sample size and variances was negative. For positive relationships between sample size and variances, the O'Brien transformation led to a more sensitive test when the Games-Howell approach was used. Using the Brown-Forsythe transformed data, the Games-Howell procedure generally had greater statistical power. Dunnett's T3 procedure often had similar power estimates as the Games-Howell, but never greater sensitivity.

In an effort to increase the statistical power for the Games-Howell and Dunnett T3 procedures, additional analyses were run in which the degrees of freedom for these tests were based on the sample sizes from the populations being compared minus 2 ($n_j+n_k-2$). These analyses had similar Type I error rates as those reported in Table 2 and had only slightly greater statistical power. For example for the 4,4,4,1 variance pattern with sample sizes of 10,20,40,50, the Games-Howell and Dunnett's T3 procedures without the Welch adjusted degrees of freedom when used with Brown-Forsythe transformed data had a power estimates of 49 and 47 respectively while with the Welch adjusted degrees of freedom the power estimates were 45 and 40. Similar differences were obtained using the O'Brien transformed data.

## Discussion

The statistical literature on tests of scale equality is extensive but limited to the omnibus test for population differences. Often in multiple group designs researchers are interested in more than the overall test and seek to identify specifically which populations differ. The present study has shown that when selecting a contrast analysis procedure both the sample sizes and the population variance pattern are important factors to consider. Based on the results presented the following conclusions and recommendations are made:

1. Both the O'Brien and the Brown-Forsythe tests for scale equality have appropriate Type I error rates for single factor multiple group designs with normal or

non-normal population distributions. The statistical power of these procedures depends on the pattern of variance differences and sample sizes. The Brown-Forsythe procedure can he considerably more sensitive than O'Brien's approach when sample size and variance pattern are positively related.

2. Under the complete null condition the Kramer-Tukey contrast analysis procedure provides the best protection against familywise Type 1 error rates for all conditions considered. Tukey's HSD controls the familywise Type I error rate when sample sizes are equal but can be liberal when using O'Brien transformed data and sample sizes are unequal. The Games-Howell and Dunnett's T3 procedures control familywise Type I errors for balanced designs but can be liberal when sample sizes are unequal. Data transformed using O'Brien's approach appears more sensitive to sample size differences than the Brown-Forsythe transformation.

3. For the partial null condition, both Tukey's HSD and the Kramer-Tukey procedures using either the O'Brien or the Brown-Forsythe transformed data can be very liberal when sample size and population variance are negatively related. The Games-Howell and Dunnett's T3 procedures control the familywise Type 1 error rate except for extreme differences in sample sizes that are positively related to the population variance pattern. Under those conditions the latter procedures can also be liberal.

4. Power differences for the contrast analysis procedures also depend on the population variance pattern and sample size relationships. Except for the 4,1,1,1 variance pattern the Brown-Forsythe transformation generally provided greater sensitivity to scale differences in pairwise contrasts. With this transformation the Games-Howell approach generally had a slight power advantage. For the 4,1,1,1 variance pattern, O'Brien's transformation generally had greater statistical power with Tukey's HSD procedure providing the most sensitive test. An exception to this generalization occurs in conditions involving unequal sample sizes that are postively related to the variance pattern. For the latter conditions the Games-Howell approach was more sensitive.

5. While the Games-Howell procedure generally provided the most sensitive test for pairwise contrasts, which is consistent with O'Brien's (1981) recommendation, this approach was consistently more powerful when data are transformed using the Brown-Forsythe procedure. Thus, although an overall F-test using O'Brien's approach can he more powerful than the Brown-Forsythe procedure, when the Games-Howell procedure is selected to identify specific differences greater sensitivity to scale differences is achieved by using Brown-Forsythe transformed data.

6. While the Games-Howell procedure with data transformed using the Brown-Forsythe approach can be

generally recommended, the statistical power of this technique can be extremely low when sample sizes are unequal and are negatively related to the population variances. For this set of conditions none of the contrast analysis procedures considered provided adequate statistical power, although the omnibus F-test had adequate power for this set of conditions except for the 4,4,4,1 variance pattern.

7. When tests of scale equality are theory based and pairwise contrasts are of interest, researchers should strive for a balanced design. If substantial differences in sample sizes cannot be avoided then a multiple analysis strategy should be adopted using the results in the present paper as guidelines in interpreting the results.

References

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variance in the one-way classification. Annals of Mathematical Statistics, 25, 290-302.

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. Journal of the American Statistical Association, 69, 364-367.

Conover, W. J., Johnson, M. E. & Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental self bidding data. Technometrics, 23, 351-361.

Dunnett, C. W. (1980). Pairwise multiple comaprisons in the unequal variance case. Journal of the American Statistical Association, 75, 796-800.

Fleishman, A. I. (1979). A method for simulating non-normal distributions. Psychometrika, 43, 521-532.

Games, P. A. (1971). Multiple comparisons of means. American Educational Research Journal, 8, 531-565.

Games P. A. & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal N's and/or variances: A Monte Carlo study. Journal of Educational Statistics, 1, 113-115.

Games, P. A., Keselman, H. J., & Rogan, J. C. (1981). Simultaneous pairwise multiple comparison procedures for means when sample sizes are unequal. Psychological Bulletin, 90, 594-598.

Harter, H. L. (1960). Tables of range and studentized range. Annals of Mathematical Statistics. 31, 1122-1147.

Howell, J. F. & Games, P. A. (1974). The effects of variance heterogeneity on simultaneous multiple-comparison procedures with equal sample size. British Journal of Mathematical Psychology, 27, 72-81.

Jaccard J., Becker, M. A. & Wood, G. (1984). Pairwise multiple comparison procedures: A review. Psychological Bulletin, 96, 589-596.

Keselman, H. J. (1976). A power investigation of the Tukey multiple comparison statistic. Educational and Psychological Measurement, 36, 97-104.

Keselman, H. J. & Rogan, J. C. (1977). The Tukey multiple comparison test: 1953-1976. Psychological Bulletin, 84, 1050-1056.

Keselman, H. J. & Rogan, J. C. (1978). A comparison of the modified-Tukey and Scheffe methods of multiple comparisons for pairwise contrasts. Journal of the American Statistical Association, 73, 47-52.

Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbe s with replications. Biometrics, 12, 307-310.

O'Brien, R. G. (1978). Robust techniques for testing heterogeneity of variance. Psychometrika, 43, 327-342.

O'Brien, R. G. (1981). A simple test for variance effects in experimental designs. Psychological Bulleltin, 89, 570-574.

Olejnik, S. F. & Algina, J. (1987). Type I error rates and power estimates of selected parametric and nonparametric tests of scale. Journal of Educational Statistics, 12, 45-61.

Ramsey, P. H. (1981). Power of univariate pairwise multiple comparison procedures. Psychological Bulletin, 90, 352-366.

Rogan, R. C., Keselman, H. J., & Breen, L. J. (1977). Assumption violations and rates of Type I error for the Tukey multiple comparison test: A review and empirical investigation via a coefficient of variance variation. Journal of Experimental Education, 46, 20-25.

Wilcox, R., Charlin, V. L., & Thompson, K. L. (1986). New monte carlo results on the robustness of the ANOVA F, W and $F^*$ statistics. Commun. Statist.-Simula., 15, 933-943.

Wilcox, R. (1987). New Statistical Procedures for the Social Sciences. Lawrence Erlbaum Associates, NJ.

Table 1

Contrast analysis procedures considered in identifying scale differences.

| Procedure | Computed Test Statistic | Reference Distribution |
|---|---|---|
| Tukey HSD | $\hat{\Psi} / (MSW/\tilde{n}^{a})^{1/2}$ | $q_{\alpha, J, N-J}$ |
| Kramer-Tukey | $\hat{\Psi} / (MSW/n_j + MSW/n_k)^{1/2}$ | $q_{\alpha, J, N-J}$ |
| Games-Howell | $\hat{\Psi} / (s^2/n_j + s^2/n_k)^{1/2}$ | $q_{\alpha, J, v^{b}}$ |
| Dunnett T3 | $\hat{\Psi} / (s^2/n_j + s^2/n_k)^{1/2}$ | $A^{c}_{\alpha, C^{d}, v}$ |

[a] Harmonic mean of sample sizes.
[b] Welch degrees of freedom:

$$\frac{(s^2/n_j + s^2/n_k)^2}{\dfrac{(s^2/n_j)^2}{n_j - 1} + \dfrac{(s^2/n_k)^2}{n_k - 1}}$$

[c] Studentized maximum modulus distribution.

[d] Number of contrasts examined.



ERIC

Table 2

Type I Errors and Power Estimates for the Omnibus F-Ratio Using the O'Brien (OB) and Brown Forsythe (BF) Transformed Data from Normal Distributions.

| | Variances $\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1,1,1,1 | | 4,1,1,1 | | 4,4,1,1 | | 4,4,4,1 | |
| Sample Sizes $n_1, n_2, n_3, n_4$ | OB | BF | OB | BF | OB | BF | OB | BF |
| 10,10,10,10 | 034 | 028 | 44 | 39 | 35 | 38 | 13 | 19 |
| 20,20,20,20 | 041 | 032 | 86 | 80 | 87 | 87 | 49 | 63 |
| 30,30,30,30 | 047 | 035 | 98 | 95 | 98 | 98 | 81 | 89 |
| 10,16,24,30 | 039 | 026 | 72 | 56 | 91 | 83 | 80 | 81 |
| 30,24,16,10 | | | 86 | 89 | 54 | 75 | 13 | 28 |
| 10,20,40,50 | 047 | 040 | 78 | 61 | 97 | 92 | 98 | 97 |
| 50,40,20,10 | | | 98 | 98 | 72 | 90 | 11 | 30 |
| 6,12,42,60 | 053 | 033 | 65 | 41 | 94 | 80 | 99 | 97 |
| 60,42,12, 6 | | | 97 | 99 | 21 | 65 | 05 | 13 |

Table 3

Type I Error Rates for the Complete and Partial Null Conditions with Normally Distributed Data.

| Sample Sizes $n_1, n_2, n_3, n_4$ | Variances $\sigma_1^2\ \sigma_2^2\ \sigma_3^2\ \sigma_4^2$ | O'BRIEN | | | | BROWN-FORSYTHE | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T[a] | KT | GH | T3 | T | KT | GH | T3 |
| 10,10,10,10 | 1,1,1,1 | 032 | 032 | 015 | 009 | 027 | 027 | 032 | 029 |
| 20,20,20,20 | | 040 | 040 | 021 | 019 | 032 | 032 | 036 | 027 |
| 30,30,30,30 | | 039 | 039 | 024 | 019 | 038 | 038 | 035 | 025 |
| 10,16,24,30 | | 046 | 037 | 054 | 048 | 046 | 027 | 050 | 045 |
| 10,20,40,50 | | 069 | 041 | 112 | 095 | 053 | 041 | 070 | 061 |
| 6,12,42,60 | | 086 | 057 | 181 | 175 | 062 | 031 | 125 | 116 |
| 10,10,10,10 | 4,1,1,1 | 000 | 000 | 005 | 004 | 000 | 000 | 011 | 009 |
| 20,20,20,20 | | 001 | 001 | 012 | 006 | 003 | 003 | 016 | 012 |
| 30,30,30,30 | | 000 | 000 | 014 | 008 | 000 | 000 | 018 | 016 |
| 10,16,24,30 | | 000 | 002 | 019 | 017 | 001 | 004 | 025 | 021 |
| 30,24,16,10 | | 001 | 001 | 023 | 014 | 002 | 000 | 019 | 011 |
| 10,20,40,50 | | 001 | 005 | 034 | 034 | 002 | 011 | 025 | 023 |
| 50,40,20,10 | | 002 | 000 | 056 | 051 | 003 | 001 | 049 | 051 |
| 6,12,42,60 | | 002 | 011 | 059 | 057 | 002 | 013 | 038 | 035 |
| 60,42,12, 6 | | 000 | 000 | 123 | 112 | 003 | 001 | 075 | 072 |
| 10,10,10,10 | 4,4,1,1 | 043 | 043 | 003 | 001 | 027 | 027 | 008 | 006 |
| 20,20,20,20 | | 060 | 060 | 011 | 007 | 044 | 044 | 013 | 010 |
| 30,30,30,30 | | 060 | 060 | 017 | 014 | 043 | 043 | 015 | 014 |
| 10,16,24,30 | | 205 | 134 | 016 | 015 | 118 | 067 | 020 | 019 |
| 30,24,16,10 | | 005 | 032 | 010 | 008 | 007 | 020 | 023 | 016 |
| 10,20,40,50 | | 264 | 148 | 018 | 013 | 155 | 078 | 025 | 019 |
| 50,40,20,10 | | 000 | 024 | 020 | 015 | 000 | 018 | 023 | 018 |
| 6,12,42,60 | | 397 | 238 | 013 | 009 | 208 | 101 | 026 | 019 |
| 60,42,12, 6 | | 000 | 007 | 010 | 007 | 000 | 006 | 019 | 014 |
| 10,10,10,10, | 4,4,4,1 | 044 | 044 | 007 | 005 | 036 | 036 | 017 | 012 |
| 20,20,20,20 | | 058 | 058 | 012 | 012 | 042 | 042 | 017 | 014 |
| 30,30,30,30 | | 056 | 056 | 023 | 021 | 044 | 044 | 027 | 023 |
| 10,16,24,30 | | 120 | 094 | 029 | 023 | 085 | 060 | 039 | 031 |
| 30,24,16,10 | | 018 | 042 | 023 | 021 | 013 | 033 | 025 | 018 |
| 10,20,40,50 | | 167 | 088 | 065 | 057 | 123 | 071 | 047 | 037 |
| 50,40,20,10 | | 009 | 030 | 032 | 026 | 002 | 030 | 036 | 029 |
| 6,12,42,60 | | 200 | 127 | 119 | 111 | 147 | 088 | 082 | 067 |
| 60,42,12, 6 | | 007 | 025 | 072 | 061 | 005 | 023 | 049 | 037 |

[a]T=Tukey HSD, KT=Kramer-Tukey, GH=Games Howell, T3=Dunnett



ERIC Full Text Provided by ERIC

Table 4

Average Power per Contrast for Four Analysis Procedures Using O'Brien and Brown-Forsythe Transformed Data from Normal Distributions.

| Sample Sizes $n_1, n_2, n_3, n_4$ | Variances $\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$ | O'Brien | | | | Brown-Forsythe | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T[a] | KT | GH | T3 | T | KT | GH | T3 |
| 10,10,10,10 | 4,1,1,1 | 29 | 29 | 03 | 02 | 25 | 25 | 08 | 07 |
| 20,20,20,20 | | 72 | 72 | 28 | 24 | 63 | 63 | 39 | 35 |
| 30,30,30,30 | | 92 | 92 | 63 | 57 | 86 | 86 | 70 | 67 |
| 10,16,24,30 | | 71 | 66 | 02 | 01 | 53 | 46 | 07 | 06 |
| 30,24,16,10 | | 37 | 52 | 57 | 53 | 46 | 58 | 59 | 55 |
| 10,20,40,50 | | 79 | 74 | 02 | 01 | 62 | 53 | 06 | 05 |
| 50,40,20,10 | | 45 | 65 | 86 | 84 | 61 | 75 | 80 | 79 |
| 6,12,42,60 | | 67 | 60 | 00 | 03 | 44 | 33 | 00 | 02 |
| 60,42,12, 6 | | 05 | 42 | | | 29 | 59 | | |
| 10,10,10,10 | 4,4,1,1 | 11 | 11 | 03 | 02 | 13 | 13 | 07 | 06 |
| 20,20,20,20 | | 37 | 37 | 29 | 24 | 45 | 45 | 4' | 37 |
| 30,30,30,30 | | 60 | 60 | 61 | 57 | 70 | 70 | 68 | 65 |
| 10,16,24,30 | | | | 08 | 06 | | | 15 | 09 |
| 30,24,16,10 | | 10 | 12 | 47 | 42 | 25 | 26 | 50 | 47 |
| 10,20,40,50 | | | | 15 | 13 | | | 23 | 20 |
| 50,40,20,10 | | 12 | 17 | 77 | 75 | 36 | 40 | 70 | 68 |
| 6,12,42,60 | | | | 03 | 02 | | | 07 | 06 |
| 60,42,12, 6 | | 00 | 02 | 72 | 70 | 12 | 16 | 61 | 64 |
| 10,10,10,10 | 4,4,4,1 | 05 | 05 | 04 | 02 | 07 | 07 | .09 | 07 |
| 20,20,20,20 | | 19 | 19 | 29 | 24 | 30 | 30 | 39 | 36 |
| 30,30,30,30 | | 36 | 36 | 62 | 57 | 55 | 55 | 69 | 66 |
| 10,16,24,30 | | | | 19 | 17 | | 38 | 30 | 27 |
| 30,24,16,10 | | 07 | 04 | 38 | 33 | 20 | 12 | 39 | 38 |
| 10,20,40,50 | | | | 39 | 36 | | | 45 | 40 |
| 50,40,20,10 | | 08 | 02 | 59 | 56 | 30 | 15 | 55 | 52 |
| 6,12,42,60 | | | | | | | | | |
| 60,42,12, 6 | | 03 | 01 | | 43 | 14 | 04 | 37 | 34 |

[a]T=Tukey HSD, KT=Kramer-Tukey, GH=Games-Howell, T3=Dunnett