

## Multiple Comparisons of Biodiversity

JAMES A. ROGERS and JASON C. HSU

Department of Statistics  
The Ohio State University  
Columbus, OH 43210-1247  
USA

### Summary

A common goal in statistical ecology is to compare several communities and or time points with respect to taxonomic diversity (usually species diversity). For this purpose, the current literature recommends the application of traditional ANOVA techniques to “replicates” of diversity indices. This approach is not even asymptotically correct because diversity index estimates have unequal variances, even when sample sizes are equal and even when the hypothesis of equality of diversity indices is true. It is shown that transformations of the data can *not* be used to remedy this situation. We construct an asymptotically correct method and illustrate its implementation using dinosaur extinction data.

**Key words:** Multiple comparisons; Diversity indices; Transformations.

### 1. Introduction

The study of biodiversity is one of the three fundamental research priorities of the 1991 Sustainable Biosphere Initiative (SBI) of the Ecological Society of America (LUBCHENCO et al., 1991). To assess environmental risks, the SBI asks:

How are community composition and species diversity affected by persistent toxic substances?

A significant decrease in diversity of species may indicate trouble in the ecosystem, due to pollution. Also, ecosystems containing a greater diversity of species may be more likely to thrive than those with less diversity.

Let  $S$  be the total number of species and suppose the diversity of  $k$  communities are to be compared. For each  $i$ ,  $i = 1, \dots, k$ , and  $s$ ,  $s = 1, \dots, S$ , let  $p_s^{(i)} (\geq 0)$  be the proportion of the  $i$ th community constituted by the  $s$ th specie. For each  $i$ , let  $\mathbf{p}_i = (p_1^{(i)}, \dots, p_S^{(i)})$ . Then Simpson's Index  $\varphi_i$  of the  $i$ th community is

$$\varphi(\mathbf{p}_i) = 1 - \sum_{s=1}^S (p_s^{(i)})^2.$$

Simpson's Index is readily seen to be the probability that two randomly selected animals will be of different species. Another commonly used diversity index is

Shannon's Index (or entropy):

$$H'(\mathbf{p}_i) = -\sum_{s=1}^S p_s^{(i)} \log p_s^{(i)}.$$

Multiple Comparisons of Shannon's Index were studied in FRITSCH and HSU (1999). This paper focuses on multiple comparisons with Simpson's Index. In terms of inference, there are some problems which are common to both Shannon's Index and Simpson's Index (for example, they are both functions of several unknown parameters). However, distributional properties of estimators of the two indices are quite different (for example, there is no unbiased estimator of Shannon's Index), and must be addressed separately.

## 2. The Replicate Approach

In order to take advantage of traditional ANOVA techniques, some authors (cf. MAGURRAN, 1988) have suggested obtaining "replicate" diversity index estimates under each experimental condition. This requires a division of each sample into "replicate" subsamples. Whether or not this division into subsamples is arbitrary depends on the sampling scheme and the model: for example if the sample at community  $i$  was obtained using  $r$  different traps, then it is natural to view the contents of each trap as a subsample. However, if each trap is assumed to be a sample from the same evenly distributed population  $\mathbf{p}_i$ , then by sufficiency, one may pool together the contents of all traps for this community. If different traps in the same community are part of a stratified sampling scheme, then the problem is clearly more complicated. We assume each trap is a sample from a well mixed population. Therefore the data may be described as follows:

For each  $i$ ,  $i = 1, \dots, k$ , and each  $s$ ,  $s = 1, \dots, S$ , let  $X_s^{(i)}$  be the sample count for species  $s$  in community  $i$ . For each  $i$ , suppose one takes a sample of size  $m$ , and let  $\mathbf{X}_i = (X_1^{(i)}, \dots, X_S^{(i)})$ . Therefore we have:

$$\mathbf{X}_i \sim \text{Multinomial}(\mathbf{p}_i, m) \text{ independently for } i = 1, \dots, k.$$

Now to obtain the "replicates", one may, for each  $i$ , divide the sample (perhaps arbitrarily) into (say)  $r$  subsamples, each of size  $n$ .

Thus, for each  $i$  we have:

$$\mathbf{X}_{im} \sim \text{Multinomial}(\mathbf{p}_i, n) \text{ iid for } m = 1, \dots, r.$$

Let

$$\hat{\varphi}_{im} = \frac{n}{n-1} \varphi\left(\frac{1}{n} \mathbf{X}_{im}\right) \quad \text{for all } i \text{ and } m,$$

$$\bar{\varphi}_i = \frac{1}{r} \sum_{m=1}^r \hat{\varphi}_{im} \quad \text{for all } i.$$

Then for each  $i$ ,  $E[\hat{\varphi}_{im}] = \varphi(p_i)$  for all  $m$  (the  $\frac{n}{n-1}$  term adjusts the estimator to be exactly unbiased). The  $\hat{\varphi}_{im}$  are smooth functions of sample means, and so follow asymptotically normal distributions as  $n \rightarrow \infty$ , as has been recognized in the literature (cf. MAGURRAN, 1988). It is also true that for fixed  $n$ , the distribution of  $\bar{\varphi}_i$  is asymptotically normal (with exact mean  $\varphi_i$ ) as  $r \rightarrow \infty$ , although this has less relevance to our discussion.

On the basis of these asymptotic results, several authors in the ecology literature have assumed that conventional multiple comparisons techniques are applicable. For example, TAYLOR (1978) used analysis of variance (ANOVA) to compare the discriminant ability of eight diversity measures; LIBER et al. (1994) used repeated measures ANOVA to assert there is a significant difference between Shannon's Index of the total phytoplankton community in control and DIATOX treated limnecoralls; SHEEHAN et al. (1991) used the Waller-Duncan test to assert the biodiversity of dinosaurs did not decline during three time intervals in the Late Cretaceous period; GAUDREAU et al. (1986) used ANOVA to assert there were no significant differences between months in the diversity of the diets of sticklebacks and juvenile brook charr in Quebec; In a section titled "Statistical tests", MAGURRAN (1988, p. 76) states,

Repeated estimates of diversity are usually normally distributed. This was the property that allowed Taylor and others to investigate the discriminatory ability of diversity measures. It also means that analysis of variance can be used to test for significant differences in the diversity of sites.

Roughly speaking, these authors applied or suggested standard ANOVA techniques with the following substitutions to the usual one-way model described, for example, on page 1 of HSU (1996):

$$\hat{\mu}_i \leftarrow \bar{\varphi}_i = \frac{1}{r} \sum_{m=1}^r \hat{\varphi}_{im}, \quad (1)$$

$$\hat{\sigma}^2 \leftarrow \frac{1}{k(r-1)} \sum_{i=1}^k \sum_{m=1}^r (\hat{\varphi}_{im} - \bar{\varphi}_i)^2. \quad (2)$$

## 2.1 The unequal variance problem

The assertion of these authors that repeated estimates of diversity indices are asymptotically normally distributed is correct. However, this alone is not sufficient for standard ANOVA techniques to be valid. Analysis of variance and multiple comparison methods such as Tukey's All-Pairwise Comparisons (MCA) and Dunnett's Multiple Comparison with Control (MCC) require the additional assumption that parameter estimates share a variance  $\sigma^2$  which is common up to known multiplicative constants (cf. HSU, 1996 Chapter 7). If the random errors are normally distributed with known variance ratios, then  $p$ -values for tests and critical values

for simultaneous confidence intervals can be computed. (The LSMEANS option under PROC GLM and PROC MIXED in SAS versions 7 and 8 implements Dunnett's method using the algorithms in Hsu (1992) and Hsu and NELSON (1998) by default, for example.)

However, variances of point estimators of diversity indices are inherently unequal. The variance  $\sigma_{\bar{\varphi}_i}^2$  of the estimate  $\bar{\varphi}_i$  of the diversity of the  $i$ th population depends on the unknown  $p_1^{(i)}, \dots, p_S^{(i)}$  of that population. Specifically, using the expression provided by SIMPSON (1949), one gets

$$\sigma_{\bar{\varphi}_i}^2 = \frac{1}{r} \frac{2}{n(n-1)} \times \left\{ \sum_{s=1}^S (p_s^{(i)})^2 + 2(n-2) \sum_{s=1}^S (p_s^{(i)})^3 + (3-2n) \left( \sum_{s=1}^S (p_s^{(i)})^2 \right)^2 \right\}.$$

(Simpson did not provide a derivation, however it is fairly easy to verify the above expression using moment generating functions). Therefore, ratios of variances of estimates of the  $k$  diversity indices are unknown. Even if the sample sizes are the same, the variances are unknown and unequal. Even if all  $k$  populations have exactly the same diversity index ( $\varphi_1 = \varphi_2 = \dots = \varphi_k$ ), the variances of their estimates can be different because the probability vectors  $(p_1^{(i)}, \dots, p_S^{(i)})$  may be different for different  $i$ .

**Example:** Let  $\mathbf{p}_1 = (1/5, 1/5, 3/5)$  and let  $\mathbf{p}_i = (7/15, 7/15, 1/15)$  for  $i = 2, \dots, 5$ . Note that all five populations have exactly the same value of Simpson's Index ( $\varphi = 0.56$ ). Suppose that for each  $i$ , we have  $r = 5$  traps, each of which traps an essentially infinite number of organisms ( $n = \infty$ ). Since  $n = \infty$ , the distribution of  $\hat{\varphi}_{im}$  will be normal for all  $i$  and  $m$ . However, the variance of  $\bar{\varphi}_1$  limits to approximately 3.86 times the variance of  $\bar{\varphi}_i$  for  $i = 2, \dots, 5$ , as  $n \rightarrow \infty$ . If one samples from normal distributions with these variance ratios, and applies Dunnett's MCC with nominal family-wise error rate 5%, the resulting true family-wise error rate will be approximately 11.2% (estimated using 25000 Monte-Carlo replicates; standard error is about 0.2%). This example demonstrates that the commonly recommended implementation of the "replicate" approach is not even asymptotically correct, even when the hypothesis of equality is true.

## 2.2 Nonexistence of variance stabilizing transformations

A frequently suggested remedy for the problem of unequal variances is to find (or estimate) a variance stabilizing transformation. Such transformations allow the construction of a pivotal (or nearly pivotal) quantity with which one may construct a confidence interval. When estimator variances differ between populations, transformed estimators may have approximately equal variances across populations.

This strategy is unlikely to work for diversity indices because the variance  $\sigma_{\bar{\varphi}_i}^2$  of  $\bar{\varphi}_i$  is a complicated function of  $p_1^{(i)}, \dots, p_s^{(i)}$ , not just of  $\varphi_i$ . As a result, different populations can have the same value of  $\varphi$ , while having different values for  $\sigma_{\bar{\varphi}_i}^2$ . Note that standard transformation theory is based on the assumption that the variance of an estimator is a function of the mean of that estimator (BARTLETT, 1947).

### 3. Proposed Asymptotically Correct Approach

We propose a solution which is asymptotically correct and does not assume that estimators have equal variances.

Suppose a sample of size  $n_i$  is taken from each community  $i = 1, \dots, k$  ( $n_i$  is now understood to be the total sample size for community  $i$ , that is, summed across all traps/replicates for that community), and again let  $X_i$  be the vector of observed cell counts. Then one may use the following unbiased estimator of  $\varphi_i$ :

$$\hat{\varphi}_i = \frac{n_i}{n_i - 1} \varphi \left( \frac{1}{n_i} X_i \right),$$

which has variance,

$$\sigma_i^2 = \frac{2}{n_i(n_i - 1)} \times \left\{ \sum_{s=1}^S (p_s^{(i)})^2 + 2(n_i - 2) \sum_{s=1}^S (p_s^{(i)})^3 + (3 - 2n_i) \left( \sum_{s=1}^S (p_s^{(i)})^2 \right)^2 \right\}.$$

A reasonable estimator of  $\sigma_i^2$  simply replaces each  $p_s^{(i)}$  in the above expression with  $\hat{p}_s^{(i)} = \frac{1}{n_i} X_s^{(i)}$ . The resulting estimate we denote by  $\hat{\sigma}_i^2$ .

We will now use the fact that our estimators  $\hat{\varphi}_1, \hat{\varphi}_2, \dots, \hat{\varphi}_k$  are independent of each other, and asymptotically follow normal distributions with means  $\varphi_1, \varphi_2, \dots, \varphi_k$  and variances  $\sigma_1, \sigma_2, \dots, \sigma_k$ , respectively. Independence of the estimators implies that estimates of contrasts will follow the one-way structure of HAYTER (1989). One may therefore apply an analogue of the one-sided Dunnett's method gives the following simultaneous lower confidence bounds for the difference between the diversity  $\varphi_i$  of populations  $i = 1, \dots, k$  and the diversity  $\varphi_k$  of a control population:

$$\varphi_i - \varphi_k > \hat{\varphi}_i - \hat{\varphi}_k - d \sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_k^2} \quad i = 1, \dots, k - 1,$$

or alternatively the upper bounds

$$\varphi_i - \varphi_k < \hat{\varphi}_i - \hat{\varphi}_k + d \sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_k^2} \quad \text{for } i = 1, \dots, k - 1, \quad (3)$$

where  $d = d_{\alpha}$  is the solution to the equation

$$\int_{-\infty}^{+\infty} \prod_{i=1}^{k-1} [\Phi((\lambda_i z + d)/(1 - \lambda_i^2)^{1/2})] d\Phi(z) = 1 - \alpha \quad (4)$$

with

$$\lambda_i = \left( \frac{\hat{\sigma}_k^2}{\hat{\sigma}_i^2 + \hat{\sigma}_k^2} \right)^{1/2}, \quad i = 1, \dots, k-1. \quad (5)$$

Note that if the estimated variances are assumed to be the true variances, then  $\lambda_i \lambda_j = \text{Corr}(\hat{\varphi}_i - \hat{\varphi}_k, \hat{\varphi}_j - \hat{\varphi}_k)$ , as is typically the case when point estimators are independent of each other. An analogue of the two sided method of DUNNETT (1955) provides the following simultaneous confidence intervals for the difference between each new treatment mean  $\varphi_i$  and the control mean  $\varphi_k$ :

$$\varphi_i - \varphi_k \in \hat{\varphi}_i - \hat{\varphi}_k \pm |d| \sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_k^2} \quad \text{for } i = 1, \dots, k-1,$$

where  $|d|$  is the solution to the equation

$$\int_{-\infty}^{+\infty} \prod_{i=1}^{k-1} [\Phi((\lambda_i z + |d|)/(1 - \lambda_i^2)^{1/2}) - \Phi((\lambda_i z - |d|)/(1 - \lambda_i^2)^{1/2})] d\Phi(z) = 1 - \alpha \quad (6)$$

with  $\lambda_i$   $i = 1, \dots, k-1$  defined as above. The PROBMC function in SAS with inputs  $= (\lambda_1, \dots, \lambda_{k-1})$ , infinite degrees of freedom, and  $\alpha$  can be used to compute  $d$  and  $|d|$ . Alternatively, the function `multicomp.default`, implemented in S-Plus versions 4.0 and higher, can be used to determine this critical value directly.

An analogue of the Tukey-Kramer approximate simultaneous confidence intervals for all-pairwise differences (TUKEY, 1953; KRAMER, 1956) is

$$\varphi_i - \varphi_j \in \hat{\varphi}_i - \hat{\varphi}_j \pm |q^*| \sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} \quad \text{for all } i \neq j, \quad (7)$$

where  $|q^*|$  the solution to the equation

$$k \int_{-\infty}^{+\infty} [\Phi(z) - \Phi(z - \sqrt{2} |q^*|)]^{k-1} d\Phi(z) = 1 - \alpha. \quad (8)$$

By the result of HAYTER (1984), this analogue of the Tukey-Kramer approximation is conservative asymptotically. The critical value  $|q^*|$  may be obtained using PROBMC in SAS (in this case the only inputs are number of populations  $k$  and error rate  $\alpha$ ), `multicomp.default` in S-Plus 4.0 and higher or from printed tables (e.g. in HSU, 1996). Note that the conservative Tukey-Kramer approximation can be improved by taking all of the  $\lambda_i$  into account: see BRETZ, GENZ, and HOT-HORN (2001) in this issue.

#### 4. Application to Dinosaur Extinction Data

There is considerable debate regarding the cause(s) of the extinction of the dinosaurs at the end of the Cretaceous period. Among the competing theories is the possibility that extinction was due to a cataclysmic event such as an asteroid impact. It is now widely accepted that there was a large asteroid impact roughly contemporaneous with the end of the Cretaceous, but it is not clear whether this was the primary cause of the demise of the dinosaurs. In an attempt to support the sudden extinction hypothesis, SHEEHAN et al. (1991) argued that, among dinosaur communities represented at the Hell Creek formation in North Dakota and Montana, diversity remained relatively constant throughout the Cretaceous period. The authors reason that if the disappearance of the dinosaurs was gradual, one should observe a decline in diversity prior to extinction.

The data is organized by dividing the formation into three equally spaced stratigraphic levels, each of which represents a period of approximately 730,000 years. Fossils were cross-tabulated according to the family to which the dinosaur belonged and the stratigraphic level. This data was collected at several sites, but for brevity we present the combined data (see Table 1).

We denote the true value of Simpson's Index at the Lower, Middle, and Upper level by  $\varphi_{\text{lower}}$ ,  $\varphi_{\text{middle}}$  and  $\varphi_{\text{upper}}$  respectively. There are several possible ways to formulate this problem. However, since few would entertain the possibility that dinosaur diversity remained exactly constant for more than two million years, the question is not "Did dinosaur diversity change?", but rather "Can we infer in which direction dinosaur diversity changed, and by how much?". To answer the latter question, one valid approach is to present simultaneous confidence intervals for all of the pairwise contrasts :  $\varphi_{\text{upper}} - \varphi_{\text{middle}}$ ,  $\varphi_{\text{middle}} - \varphi_{\text{lower}}$  and  $\varphi_{\text{upper}} - \varphi_{\text{lower}}$ .

Using the expressions for  $\hat{\varphi}_i$  and  $\hat{\sigma}_i$  from the previous section and applying (7), we obtain simultaneous confidence intervals for all pairwise contrasts. The  $\alpha = 0.05$  quantile of the studentized range distribution for  $k = 3$  is  $|q^*| = 2.344$ . The resulting confidence intervals are provided in Table 2. While the point estimate for  $\varphi_{\text{lower}} - \varphi_{\text{upper}}$  does suggest a decline in diversity, the confidence inter-

Table 1

Dinosaur counts by family and stratigraphic level. Families represented are *Ceratopsidae*, *Hadrosauridae*, *Hypsilophodontidae*, *Pachycephalosauridae*, *Tyrannosauridae*, *Ornithomimidae*, *Saurornithoididae*, *Dromaeosauridae*. Data was provided by Sheehan in personal communication

Interval	Counts
Upper	(50, 29, 3, 0, 3, 4, 1, 0)
Middle	(53, 51, 2, 0, 3, 8, 6, 0)
Lower	(19, 7, 1, 0, 2, 0, 3, 0)

Table 2  
Simultaneous 95% confidence intervals for all pairwise contrasts of dinosaur diversities

Contrast	Estimate	Std. Error	Lower Bound	Upper Bound
$\varphi_{\text{lower}} - \varphi_{\text{middle}}$	-0.035	0.0842	-0.2320	0.162
$\varphi_{\text{lower}} - \varphi_{\text{upper}}$	0.015	0.0842	-0.1820	0.212
$\varphi_{\text{middle}} - \varphi_{\text{upper}}$	0.050	0.0354	-0.0328	0.133

vals indicate that the data is not sufficient to confidently infer the direction of the changes in diversity. The data is sufficient to infer that changes in diversity never exceeded 0.232 (all of the confidence bounds had absolute value less than or equal to 0.232). Therefore, if a difference of (say) 0.25 or less is considered to be practically equivalent to zero difference, then we may infer that all three communities were practically equivalent with respect to Simpson’s Index.

Another valid way of formulating the problem is to consider the lowest (oldest) stratigraphic level as a control. Furthermore, one might be willing to give up the possibility of making claims that diversity increased over time: it may be that only decreases are of interest. In this case, simultaneous upper bounds for  $\varphi_{\text{upper}} - \varphi_{\text{lower}}$  and  $\varphi_{\text{middle}} - \varphi_{\text{lower}}$  are appropriate. Using (5), we calculate  $\lambda_{\text{middle}} = 0.8734$  and  $\lambda_{\text{upper}} = 0.8734$  (it is only a coincidence that these two values of  $\lambda$  are equal). Using these values for  $\lambda$ , we are able to calculate  $d = 1.7904$ . Now applying (3), we get the confidence intervals provided in Table 3.

Although we have obtained tighter bounds (as a result of our giving up the ability to make certain claims), they all fall to the right of zero, so we are still unable to confidently infer that any decrease in diversity occurred.

Table 3  
Simultaneous 95% confidence upper bounds for comparing dinosaur diversity at later time points with dinosaur diversity of oldest time point

Contrast	Estimate	Std. Error	Upper Bound
$\varphi_{\text{middle}} - \varphi_{\text{lower}}$	-0.035	0.0842	0.186
$\varphi_{\text{upper}} - \varphi_{\text{lower}}$	0.015	0.0842	0.136

5. Conclusion

We have shown that a commonly used method for multiple comparisons of diversity indices is not even asymptotically correct, and cannot be remedied using transformations of the data. The approach we have presented is asymptotically correct and can be easily implemented using standard statistical software. Asymptotic correctness is of course only a minimum standard: appropriate small sample adjustments need to be investigated, and this is a future direction for our research.



## References

- BARTLETT, M. S., 1947: The Use of Transformations. *Biometrics* **3**, 39–53.
- BRETZ, F., GENZ, A., and HOTHORN, L. A., 2001: On the Numerical Availability of Multiple Comparison Procedures. *Biometrical Journal* **43**.
- DUNNETT, C. W., 1955: A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association* **50**, 1096–1121.
- FRITSCH, K. and HSU, J. C., 1999: Multiple Comparison of Entropies with Application to Dinosaur Biodiversity. *Biometrics* **55**, 114–119.
- GAUDREAU, A., MILLER, T., MONTGOMERY, W. L., and FITZGERALD, G. J., 1986: Interspecific Interactions and Diet of Sympatric Juvenile Brook Charr *Salvelinus Fontinalis*, and Adult Ninespine Sticklebacks *Pungitius Pungitius*. *J. Fish. Biology* **28**, 133–40.
- HAYTER, A. J., 1984: A Proof of the Conjecture that the Tukey-Kramer Multiple Comparisons Procedure is Conservative. *Annals of Statistics* **12**, 61–75.
- HAYTER, A. J., 1989: Pairwise Comparisons of Generally Correlated Means. *Journal of the American Statistical Association* **84**, 208–213.
- HSU, J. C., 1992: The Factor Analytic Approach to Simultaneous Inference in the General Linear Model. *Journal of Computational and Graphical Statistics* **3**, 151–168.
- HSU, J. C., 1996: *Multiple Comparisons: Theory and Methods*. Chapman and Hall, London.
- HSU, J. C. and NELSON, B., 1998: Multiple Comparisons in the General Linear Model. *Journal of Computational and Graphical Statistics* **7**, 23–41.
- KRAMER, C. Y., 1956: Extension of Multiple Range Tests to Group Means with Unequal Number of Replications. *Biometrics* **12**, 309–310.
- LIBER, K. et al., 1994: Impact of 2,3,4,6-tetrachlorophenol (DIATOX) on plankton communities in limnocorrals. In Graney, R. L., Kennedy, J. H., and Rodgers, J. H., editors, *Aquatic Mesocosm Studies in Ecological Risk Assessment*, SETAC Special Publications Series, chapter 17, pages 257–294. CRC Press, Boca Raton.
- LUBCHENCO, J. et al., 1991: The Sustainable Biosphere Initiative: An Ecological Agenda. *Ecology* **72**, 371–412.
- MAGURRAN, A. E., 1988: *Ecological Diversity and its Measurement*. Princeton University Press, Princeton.
- SHEEHAN, P. M. et al., 1991: Sudden Extinction of the Dinosaurs: Latest Cretaceous, Upper Great Plains, U.S.A. *Science* **254**, 835–839.
- SIMPSON, E. H., 1949: Measurement of diversity. *Nature* **163**, 688.
- TAYLOR, L. R., 1978: Bates, Williams, Hutchinson – a Variety of Diversities. In Mound, L. A. and Warloff, N., editors, *Diversity of Insect Faunas: 9th Symposium of the Royal Entomological Society*, pages 1–18, Oxford. Blackwell.
- TUKEY, J. W., 1953: The problem of multiple comparisons. Dittoed manuscript of 396 pages, Department of Statistics, Princeton University.

JAMES A. ROGERS  
 Department of Statistics  
 The Ohio State University  
 1958 Neil Avenue  
 Columbus, OH 43210-1247  
 (614)292-5375  
 jrogers@stat.ohio-state.edu

Received, September 2000  
 Revised, April 2001  
 Accepted, May 2001

