PCSTAT: Statistical Analysis of Paired-Comparison Data

C. P. WHALEY

University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Frequently in psychology it is not possible to behaviorally assess a phenomenon along a well established interval or ratio scale, such as response time, proportion correct, and so on. Under such circumstances, the method of paired comparisons or ranking procedures is often relied upon. In the former case, all possible pairs of the "objects" under consideration are presented to each subject one pair at a time. The subject simply responds in a binary fashion indicating which of the two objects of each pair he/she "prefers." The term preference is applied very loosely here, insofar as the judgment could involve any dimension (e.g., size, weight, effectiveness, strength, etc.). The only aspect that is critical is that the subject can use the instruction-specific dimension in a reasonable and meaningful way.

Similarly, with a few minor assumptions (see Laming, 1973, pp. 32-34), ranks assigned to a set of objects according to a similar instruction-specific dimension can be thought to contain the same sort of information present in a paired-comparison incidence matrix. That is, if a subject assigns to object A a higher rank than object B, one can assume that when presented with A and B alone, the same subject should choose A over B. The reason for pointing out this seemingly trivial relationship between ranking and paired-comparison data is that there are some fairly robust parametric measures for analyzing paired-comparison data, whereas psychologists generally rely on the much less powerful nonparametric techniques for analyzing ranking data. The program described here is suitable for analyzing both types of data in that ranks are first transformed into paired-comparison incidence matrix form.

David (1963) presents an extensive review of the data-analytic techniques available for paired-comparison data. Maxwell (1974) covers the most widely applicable tests discussed by David, and points out the suitability of incorporating Cox's (1970, pp. 33-35) variant of the logistic transformation to the paired-comparison probabilities prior to statistical analyses. The row means $[A(1), A(2), \dots, A(K)]$ of the transformed data can then be treated as interval scaled preference values for each of the K objects in the design. Provided that the scale values can be shown to be additive, that is, $[A(1) - A(2)] + [A(2) - A(3)] = [A(1) - A(3)]$, except for minimal random error, an analysis of variance can be carried out to test for significant differences among the scaled object values. Then, multiple comparisons among the pairs of scaled values can be carried out after finding a significant F.

PCSTAT is a FORTRAN program written to carry out these analyses. The computational formulas for the

various statistics are readily available in David (1963) and Maxwell (1974).

PCSTAT Input. Data may be entered in either of two forms: an N by K rank matrix or a K by K incidence matrix, where K equals the number of objects and N equals the number of subjects. The entries in the rank matrix represent the rank of the jth object that was assigned by the ith subject. Entries in an incidence matrix represent the number of judges who preferred the ith object over the jth object.

PCSTAT Output. PCSTAT computes and prints the following: (1) The raw data and the data following transformation via the Cox (1970, pp. 33-35) variant of the logistic transformation. (2) Maxwell's (1974) critical ratio (CR) for each entry in the transformed matrix. Each CR value may be independently evaluated through reference to a table of standard normal Z scores, provided N is sufficiently large. (3) A vector of interval scaled object values which (assuming additivity) best characterize the matrix of preference values. (4) A chi-square test of additivity, including the associated p value. (5) An analysis of variance on the differences among scaled objects, including F and associated p values. (6) Least significant differences among the scaled object values for .01 and .05 alpha levels, as determined from a generalization of the Fisher (1949, pp. 56-58) l.s.d. test. (7) A matrix of estimated data values based on the scaled object values. In the additive case, the estimated values should approximate those of the original (transformed) data matrix. In the case of nonadditivity, it may be possible to establish the source of nonadditivity by systematically comparing the two matrices. (9) A graphic representation of the scaled object indices.

Program Specifications. PCSTAT was written in FORTRAN IV for use on an IBM 370/158. It is currently being adapted to run on a PDP-11/40. The present limitations are set at K = 10 objects and N = 999 subjects. These limits can be extended or reduced to accommodate machine memory.

Availability. A listing of PCSTAT including documentation and test data may be obtained from the author without charge. Address all inquiries and requests to C. P. Whaley, Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

REFERENCES

- COX, D. B. *Analysis of binary data*. London: Methuen, 1970.
- DAVID, H. A. *The method of paired comparisons*. London: Griffin, 1963.
- FISHER, R. A. *The design of experiments*. Edinburgh: Oliver & Boyd, 1949.
- LAMING, D. *Mathematical psychology*. London: Academic Press, 1973.
- MAXWELL, A. E. The logistic transformation in the analysis of paired comparison data. *British Journal of Mathematical and Statistical Psychology*, 1974, 27, 62-71.

(Received for publication March 31, 1977.)