RANKINGS FROM PAIRED COMPARISONS[1]

By W. A. Thompson, Jr. and Russell Remage, Jr.

The Florida State University and The University of Delaware

1. Introduction. In many experimental contexts, preference relationships among objects can be obtained where numerical measurement is difficult or impossible. This situation occurs, for example, in regard to individual preferences for pieces of music. Frequently because of memory, fatigue or distance limitations the items will be presented to a subject only in pairs and he will be asked to state his preference. The authoritative statistical paper on this method of paired comparisons is that of Kendall and Smith [7]; they propose a criterion, based on the number of circular preference triads, for judging whether a given set of preferences can reasonably be considered as reflecting a single linear variable. Alternatively, Slater [6] proposes that the items should be ranked so as to minimize the number of violations of observed preference, and that this minimum should then replace Kendall and Smith's circular triad criterion. Thus, more generally than Slater's particular criterion, one reason for wishing to estimate rank order is to obtain a standard whose deviations from the observed set of paired comparisons can then be examined to determine whether or not ranking was justified at all. Other uses of rank order arise in psychological testing and market survey work. We may, for instance, wish to choose the three most preferred blends of coffee or place people into homogeneous groups according to their preferences.

Doehlert's master's thesis [5] was responsible for calling our attention to the subject of this paper. We study two theoretical aspects of the problem of obtaining rankings from paired comparisons. Section 2 treats the problem from the point of view of graph theory. Section 3 then introduces a mathematical model based on the concept of weak stochastic transitivity and uses the graph theoretic results of Section 2 to obtain a maximum likelihood ranking for this model. We show that the maximum likelihood weak stochastic ranking yields Slater's criterion when every pair of objects is compared exactly once.

2. Paired comparison graphs. Let $X = \{x_1, x_2, \cdots, x_m\}$ be a set of $m > 2$ distinct objects. A set of paired comparisons of X is a relation $\mathbf{R}$ in X which is anti-symmetric and anti-reflexive; that is, a subset of $X \times X$ such that $(x_i, x_i) \notin \mathbf{R}$ and if $(x_i, x_j) \in \mathbf{R}$ then $(x_j, x_i) \notin \mathbf{R}$. For brevity such a relation will be called a *comparison of* X. For definiteness the reader may interpret $(x_i, x_j) \in \mathbf{R}$ to mean "in the comparison $\mathbf{R}$, x_i is preferred to x_j"; however, the results of Section 2 in no way depend on this interpretation. A path $\mathbf{K}$ in $\mathbf{R}$ from

Received 1 July 1963; revised 7 November 1963.

[1] Much of this work was supported by the University of Delaware Center for Research on Social Behavior.