H. A. David, *The Method of Paired Comparisons*, Number 12 of Griffin's Statistical Monographs and Courses, Hafner Publishing Company, New York, 1963. $4.75, 124 pp.

Review by E. N. Gilbert

Bell Telephone Laboratories, Murray Hill

In a typical paired comparison experiment a judge or panel of judges examines certain pairs of objects. Each judge picks one object of each pair as his preference. Such experiments occur commonly in consumer preference studies, personnel rating, and psychological investigations; a sports tournament is a kind of paired comparison experiment in which the pairs represent matches and the preferred object is the winner of the match. The purpose of the experiment may be to find the best object of all, to rank the objects in order of merit, to decide whether there is any noticeable difference between the objects, or to test the perception of the judges. The outcome of a paired comparison may be difficult to interpret because the judges contradict one another or even themselves (by preferring i to j, j to k, but k to i). Inconsistent data indicate that the preferences expressed are not strong. The mathematical treatment in this book is based on several models which describe the strengths of judges' preferences in terms of probabilities π_{ij}. A judge is supposed to make comparisons independently with probability π_{ij} of preferring i to j.

As presented in this monograph, the method of paired comparisons offers challenging mathematical problems involving model building, parameter estimation, hypothesis testing, and experiment design and has obvious practical applications. I therefore assumed at first that paired comparisons must be a familiar subject to all statisticians. However, a search for other statistical books on the subject and a few discussions with statistician friends have convinced me that this first impression was entirely wrong. If so, this short, readable survey of the field is a particularly worthwhile project.

The mathematical models of paired comparison experiments differ mainly in certain "transitivity" restrictions on π_{ij}. In the most restricted case one obtains a *linear model* in which each object i is assigned a real number V_i and $\pi_{ij} = H(V_i - V_j)$ where $H(V)$ is a suitable positive monotone increasing function. In a complete experiment, n judges each make all $\binom{t}{2}$ possible paired comparisons between t objects. The *score* a_i is the number of times object i is preferred. Surprisingly, when $n = 1$, the number of *circular triads* (triples i, j, k for which i is preferred to j, j to k, and k to i) can be deduced just from the scores. The distribution of the scores is obtained when $\pi_{ij} = \frac{1}{2}$ for all i, j. These results are used in significance tests. Typical questions are: Is object #1 significantly better than average? Is object #1 significantly better than object #2? Is the highest scoring object significantly better than average? Chapter 4 is concerned with using experimental data to estimate the parameters V_i in a linear model and also