Book Reviews

Statistics in Biology

Biometry. The Principles and Practice of Statistics in Biological Research. ROBERT R. SOKAL and F. JAMES ROHLF. xxii + 778 pp., illus. $15. Biology Series. **Statistical Tables.** F. JAMES ROHLF and ROBERT R. SOKAL. xiv + 258 pp. Cloth, $7.50; paper, $2.75. Freeman, San Francisco, 1969.

As a first approximation, it is true that statistics is statistics. An introductory course in biometrics is likely to cover much the same material as one in, say, psychometrics or technometrics. There are nevertheless important differences in emphasis and to a smaller degree in content, and these differences are probably greater than most differences among courses within any one of these areas.

In the preceding paragraph I have made an argument in the form of what is, rather misleadingly, called analysis of variance. As Sokal and Rohlf note, this approach is more widely applicable than are precise statistical tests based on it; it "provides an insight into the nature of variation of natural events, into Nature in short, which is possibly of even greater value than the knowledge of the method as such. If one can speak of beauty in a statistical method, analysis of variance possesses it more than any other."

A discursive approach characterizes this book more than almost any other serious text I know in statistics. This is not to say that Sokal and Rohlf neglect the formulas and computational methods that form the meat of their subject. Rather, computational methods are segregated from the running text into "boxes," some of which continue for several pages. The text itself gives the motivation and theoretical background necessary if one is to use these methods intelligently, and explains their general operation.

The mathematics is the minimum consistent with adequate understanding of the context in which the methods can be used. An appendix contains some derivations too long for the text, but even here no calculus is used. The book nevertheless includes one or two original theoretical results, as well as material that makes it unconventional in important respects. There are detailed instructions for operating desk calculators, and 76 pages of computer programs with instructions for their use. These methods are widely available, but good program libraries and efficient procedures are far from ubiquitous, and the demonstration of these modern methods will probably save much anguish for occasional users.

Analysis of variance receives more attention than usual, but unfortunately this relegates the general *t*-test (in my experience the most common statistical test of any kind) to the latter part of the book, where it is recommended only for unusual situations. Randomization tests and the combination of probabilities, also hardly ever mentioned at the introductory level, are important and receive due notice here. A geometrical comparison of the basic multivariate methods would have been illuminating.

I was disappointed to find not even a mention of the general Bayesian approach, which has many simple applications even if it is inordinately complex in its full development, and which provides a way of looking at the world that is more encompassing than the more ad hoc methods of traditional statistics. Frequentist interpretation of probability leads to a narrow interpretation of confidence limits. Sokal and Rohlf take a strictly Neyman-Pearson approach, even to the extreme of applying the accept-reject dichotomy to all cases of statistical testing. If I find a probability value of 0.07, I have less confidence in the hypothesis than if the value is 0.20, but this continuum disappears when fixed significance levels are used even in cases where they are unnecessary.

Although the authors wisely emphasize the robustness of most parametric methods, they do include a nearly adequate treatment of nonparametric methods. The main lack here is of a realization of the sensitivity of estimations and tests on variances to deviation from normality of the underlying distributions. The important coefficient of variation has even poorer treatment than in most texts, which is particularly unfortunate because of the existence of an adequate test, discovered independently a couple of years ago by three biologists. The discussion of the number of significant figures to use would have been more convincing if the principles it sets forth had been followed elsewhere in the book. And a hypothesis suggested by the data cannot, unfortunately, be treated as though it were an initial hypothesis unless there is some way to allow for all other possible results of any kind, not just of the kind noticed, with as low an initial probability. The companion book of tables is largely conventional; their replacement by graphs would in most cases permit more accurate results with the same use of space. All these faults are common ones; it is the more important to understand them.

The authors are practicing biologists, and their approach is more relevant to real biological concerns than is that of most texts. The book seems easily readable, and in this it is nearly unique, although the final test must be by students. The problems seem excellent. The cryptically complicated matter of correlation and regression, often or perhaps even usually misinterpreted by statisticians, receives the best general treatment I have seen. A great deal of biological work is sloppy because many biologists don't understand basic statistical concepts. This book should help to remedy this and in doing so supersede Snedecor, Bliss, and the rest; despite its faults it is easily the best introduction to biometrics, and perhaps to applied statistics generally, that is available.

LEIGH VAN VALEN
Department of Anatomy,
University of Chicago,
Chicago, Illinois

The Operon Reexamined

Nucleus and Cytoplasm. HENRY HARRIS. Clarendon Press (Oxford University Press), New York, 1968. xvi + 142 pp., illus. $7.

The year 1961 was a big one for molecular biology. That year saw the first evidence for the existence of a messenger class of RNA molecules which carried information for protein structure from the nuclear genes to the cytoplasm. In that year, too, the deciphering of the