

Review: Statistics in Genetics

Author(s): Richard R. Hudson

Review by: Richard R. Hudson

Source: *Science*, New Series, Vol. 250, No. 4980 (Oct. 26, 1990), p. 575

Published by: American Association for the Advancement of Science

Stable URL: <http://www.jstor.org/stable/2878459>

Accessed: 09-01-2016 15:17 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

<http://www.jstor.org>

some cases, it must be admitted, this journey to the boundaries of current understanding is not really such a long trip.

At the outset, Bard deals convincingly with the view, all too widely held these days, that what is most needed for the solution of developmental problems is simply the accumulation of more and more data, preferably molecular data. Is it possible, perhaps, that those whose research concentrates on molecular self-assembly phenomena, where subunits associate spontaneously to form larger structures, thereby come to expect that the facts themselves, once accumulated in sufficient quantity, will likewise assemble together spontaneously to form general principles—all without human intervention or the proposal of alternative hypotheses—and that to wait patiently for this to happen is a form of objectivity? Bard is certainly not a victim of any such fallacy. He proves by example how much more useful it can be to gather together facts already known (often from fields so disparate as not otherwise to be known by the same subsets of biologists), to analyze carefully what these facts mean and how they support or contradict alternative hypotheses, and then to consider which further facts would be most decisive.

The specific subjects cover a wide range: they include the formation of kidney tubules, the eversion of imaginal discs in fly metamorphosis, alternative means by which cells can rearrange themselves into hollow tubes, the spreading of epithelia, the rolling up of the neural tube, cell adhesion molecules, tendon and muscle formation, the different types of collagen, the alignment of fibroblasts, and the topological rules obeyed by arrays of aligned fibroblasts, to name just a few. Indeed, very few topics come to mind that deserved inclusion but were left out. One is histotypic cell sorting, which Bard deliberately skips past by saying that the topic was well enough reviewed in the 1984 edition of Trinkaus's book *Cells into Organs* (Prentice-Hall). Nevertheless, this old and contentious topic (what makes cells rearrange by differentiated cell type?) could surely profit from a little of Bard's style of dispassionate dissection. Incidentally, this book overlaps surprisingly little either with Trinkaus's book or with Lackie's *Cell Movement and Behavior* (Allen and Unwin, 1986) and makes an excellent complement to both.

Bard's usual approach is that of "case studies." For each new phenomenon, he reviews the relevant literature and describes alternative hypotheses. He then weighs the evidence carefully and thoughtfully. He reaches his conclusions with sufficient fairness that, even though I disagreed with many of them, I always felt that my understanding of the alternative points of view

had been improved. He does not hesitate to point out where the evidence does not seem conclusive. In fact, his several lists of what we do not yet know are among the best features of this book. He even has a special appendix on this subject containing three dozen well-chosen questions. Whether or not these will ever rival David Hilbert's famous mathematical list, they do help to unify the book around what needs to be done in the future.

In the search for solutions to these problems many different kinds of techniques are considered. These range from electron microscopy to studies of force exertion. Nor is molecular information omitted where it is relevant; for example Bard considers the effects on mouse development of deletion of the gene for type I collagen.

Among the techniques mentioned is computer simulation. Alas, Bard touches only lightly on this topic, and states his opinion that "if computing is required, it is not for the amateur." If he means by this that biologists themselves cannot reasonably hope to write useful programs, then I would strongly disagree; and for a conclusive disproof, I would point to Bard's own important work on the simulation of zebra and other striping patterns or to the general workability of Turing's reaction-diffusion system. Those familiar with that aspect of Bard's research will be disappointed to find that he has totally omitted it from his book.

The well-chosen bibliography lists no fewer than 544 articles and books, with citations continuing up through 1989. In addition to a satisfactory index of the ordinary kind, an innovative half-page index of "morphogenetic systems"—for example, disc eversion (*Drosophila*) and epiboly (*Xenopus*)—is included.

ALBERT K. HARRIS

Department of Biology,
University of North Carolina,
Chapel Hill, NC 27599-3280

Statistics in Genetics

Genetic Data Analysis. Methods for Discrete Population Genetic Data. BRUCE S. WEIR. Sinauer, Sunderland, MA, 1990. xiv, 377 pp., illus. \$48; paper, \$27.

As the author accurately states in the preface, this is "a guide to methods of traditional analyses of Hardy-Weinberg and linkage disequilibrium and to the methods for characterizing population structure and estimating genetic distance." When dealing with these traditional analyses, it is indeed a clear, carefully written, and thorough guide to the most rigorous and powerful methods for the analysis of genetic variation. More

than that, it is an excellent book for teaching the fundamentals of statistical analysis, even if one is not interested in genetic data *per se*. The author provides a clear, concise review of genetic terminology and experimental methods that makes the book accessible to any biologist or statistician. Also provided are complete descriptions of the models upon which the statistics are based that make it clear what the assumptions and limitations of the methods are. It should be noted that evolutionary models, which might account for the genetic variation observed in today's populations, are not discussed in any detail, and methods of testing evolutionary models or estimating parameters of these models are not described.

Weir explains indicator variables, maximum likelihood methods, moment estimators, and the jackknife and the bootstrap, among other statistical ideas. His forte is calculating variances and testing null hypotheses, and it shows in his clear, careful treatment of the methods. The methods are illustrated with a variety of real data sets, from Mendel's data on peas to the base composition of bacteriophage and mitochondria.

In addition to the traditional analyses, Weir wishes "to review some of the statistical techniques appropriate for restriction fragment length polymorphisms and DNA sequences." These topics occupy about 25 percent of the text. In this rapidly developing area, it is perhaps inevitable that the treatment is not as thorough or definitive. Some recent developments are not treated, and important references are left out. For example, the methods of Engels and Nei for estimating nucleotide variation with restriction site data are not even mentioned. There is no discussion of the analyses of codon usage. Also, in the 30-page review of tree-building methods, covering parsimony, distance matrix, and maximum likelihood methods, there is no mention of phylogenetic invariants (evolutionary parsimony) or neighbor-joining methods. These omissions can be forgiven, given the limited space devoted to modern molecular data and phylogenetic reconstruction. What Weir has chosen to cover is described in his typical careful manner. Besides, most of the methods of the first three-quarters of the book apply equally well to restriction site and nucleotide sequence data.

Weir has produced an excellent book about traditional genetic analyses. I hope he writes a sequel soon, expanding the final two chapters into a full-length book.

RICHARD R. HUDSON
Department of Ecology and
Evolutionary Biology,

University of California, Irvine, CA 92717