

BOOK REVIEWS

Editor: Petra Macaskill

1. *Joseph L. Fleiss, Bruce Levin and Myunghee Cho Paik*, *Statistical Methods for Rates and Proportions* (3rd edn).
2. *Harvey Motulsky and Arthur Christopoulos*, *Fitting Models to Biological Data Using Linear and Non-Linear Regression: A Practical Guide to Curve Fitting*.

1. STATISTICAL METHODS FOR RATES AND PROPORTIONS (3rd edn). Joseph L. Fleiss, Bruce Levin and Myunghee Cho Paik, Wiley, New Jersey, 2003. No. of pages: xxvii + 760. Price: \$99.95 (hardcover). ISBN: 0-471-52629-0

For those brought up on the first edition (published in 1972) of this book, the third is massively welcome. However, so attached was I to the first that the second almost passed me by, but anyway was available in the library. An interesting recollection of that time was the difficulty in persuading people that analysis of binary data was firstly possible and secondly useful. At that time, inverting a matrix was a formidable undertaking and to most clinicians, for example, the results were perplexing but illuminating. As the fundamental text justifying the statistics for analysing dichotomies it became essential reading for anyone wishing to understand, as the title exclaims, analysis of rates and proportions. The second edition, published 8 years after the first—in 1980, added a great deal of material on confidence limits and exact test among several other vital contemporary aspects.

Now 20 odd years on, this version, which apparently went to press at the same time as Joe Fleiss died, arrives with a strong sense of his original vision. But so much has changed since 1980 that his co-authors have had a formidable task to bring this comprehensive treatment up to date. What they have tried to do is to update without upsetting the original design and purpose, and consequently this is inevitably a different book.

Here one can now find the addition of empirical Bayes methods to assist in complex analyses,

while for example conditional likelihood and exact methods have been considerably advanced. Mostly, we have a fundamental and detailed treatment of almost all aspects for analysing binary data as well as polytomous outcomes. The theory is developed from first principles in a readily understandable manner. In fact the introductions often seem too basic, in that obvious relationships are repeated in different forms, presumably for clarity. However, the level achieved in the full descriptions often goes well above the high school mathematical level, which Fleiss had originally aimed for, but this is inevitable. It is nonetheless eminently absorbable for those with a real interest. By that it is clear that this book is aimed at professional methodologists—mainly statisticians or researchers with a fundamental interest in the justification for common methods used in statistical packages. This book is, for them, both core and comprehensive.

The context in which the theory is set is also a wide-ranging review of the principles and practice behind for example randomization and alternatives, maximum likelihood, statistical inference, study design and many other aspects of research. The examples are generous and mostly biomedical and the literature reviews are likewise comprehensive and useful. The early editions included useful sample size calculations and these are much enhanced here. The numerical problems included in each chapter now have more comprehensive answers and hints—which must be especially useful for students.

My copy had a repeated page, and a missing one, which I hope is ironed out by the publishers, but this book will certainly remain as the definitive text of context, method and application for the efficient

analysis of rates and proportions. Joe Fleiss created a formidable (and much available) tradition in relevant statistical exposition for researchers—which is continued into a wholly different world of statistical and epidemiological analysis—just as cogently.

KLIM MCPHERSON
*Nuffield Department of Obstetrics and
Gynaecology Research Institute
Churchill Hospital Oxford, U.K.*

(DOI: 10.1002/sim.2261)

2. FITTING MODELS TO BIOLOGICAL DATA USING LINEAR AND NON-LINEAR REGRESSION: A PRACTICAL GUIDE TO CURVE FITTING. Harvey Motulsky and Arthur Christopoulos, Oxford University Press, Oxford, 2004. No. of pages: 352. Price: £19.99, \$29.29 (paperback); £40.00, \$65.00 (hardcover). ISBN: 0-19-517180-2 (paperback), 0-19-517179-9 (hardcover)

The book was developed to support and educate users of GraphPad Prism—most of the practical details are illustrated using this program but in a format that is easily transferable to almost any other non-linear regression package. Some assertions, e.g. most non-linear regression programs do not support differential equation defined models, seem to be special pleading for this notable lack in GraphPad Prism. However, in general, the authors avoid the assumption that their program is the only one on the block.

The models are mainly applied to pharmacological topics—ligand binding, concentration effect relationships and some simple pharmacokinetics. These are the basic building blocks of quantitative pharmacology and the authors explore them in depth with detailed examples.

Overall, I was impressed by the clear presentation of what are complex and even arcane topics for a biologist; however, the objectives of modelling biological data are hardly mentioned. The focus of the book is on parameter estimation and choosing among models. There is little or no discussion of the application of models and their parameters for prediction. Procedures such as the posterior predictive check are not mentioned. The primary performance criterion is the precision of the parameter estimate (pp. 97, 171).

In consequence, there is extensive discussion of methods for obtaining confidence intervals of parameter estimates. The simplistic assumption of symmetry is dismissed by examples showing that this is unlikely for real data and common models. They illustrate the use of likelihood profiling and parametric bootstraps to construct more credible

intervals. The resulting intervals are all quite different from each other and there is no discussion of which one might be considered the most reliable (I would prefer the parametric bootstrap method).

Methods for distinguishing among models are also described in some detail. Hypothesis testing based on the *F*-statistic and probabilistic assertions based on the Akaike information criterion are offered but numerical approaches based on the randomization test are not mentioned. It seems an opportunity was missed to show how modern statistical inference can range from methods based on untested distribution-based theory to more empirical but more credible computational approaches.

The application examples reflect the authors' strengths in ligand binding techniques and analysis. Practical tips are offered to improve experimental methods as well as interpret the results of parameter estimation. Although simple cases of predicting unbound concentrations are described, I was surprised that a general method (Feldman, 1972) was not mentioned and is apparently not used by GraphPad Prism.

The section on pharmacokinetics is not well done—the symbol *D* is used for dose with a one compartment model but a few pages later *D* becomes distribution clearance for a two compartment model. There are also serious conceptual errors—a system of differential equations confuses central and peripheral volumes (p. 79). They also mistakenly ascribe the covariance due to estimation error between initial concentration and rate of elimination to a fanciful mechanistic association (p. 102).

I was disappointed that the authors did not indicate awareness of the wider application of non-linear regression to mixed effects modelling. The conceptual framework of starting with an individual is a natural starting place for the beginner in quantitative biology, but the serious biological scientist will quickly want to distinguish fixed and random effects. The authors offer *ad hoc* fixes