Statistical Methods in Medical Research 1996; **5**: 101–104

Book reviews

Hinkelman K, Kempthorne O 1994: *Design and analysis of experiments. Volume I: Introduction to experimental design.* New York: John Wiley and Sons, Inc. 495 pp. US \$59.95; UK £49.50. ISBN 0 471 55178 3.

It is necessary to emphasize the fact at the outset that very much is common between this book and Professor Kempthorne's original pioneering classic text *Design and analysis of experiments*, published in 1952; the authors wanted to rearrange and update the material in that book, as the subject has developed considerably. My first impression, based on a review of Volume I (there is a second volume which will be published in 1996), is that the authors have succeeded quite well in this task.

This volume has 13 chapters devoted to: 'The Process of Science', 'Principles of Experimental Designs', 'Survey of Experimental Designs', 'Linear Model Theory', 'Randomization', 'The Completely Randomized Designs', 'Comparison of Treatments', 'Use of Supplementary Information', 'Randomized Block Designs', 'Latin Square Type Designs', and 'Split-Plot Designs'.

The authors felt it 'absolutely necessary' to add this first chapter on the process of science dealing with the role of experiments, the role of data analysis and the ideas of probability in a population of repetitions. This is a very scholarly and philosophical contribution, rich in content, but how many students are going to read it and how many instructors are going to teach this subject matter, is a debatable question.

The second chapter, on the principles of experimental designs, is very well written, summarizing and providing an overview of many aspects of the subject. The schematic description of four experimental situations is novel and illuminating. The third chapter gives an account of and distinguishes between error-control designs, treatment designs and sampling designs.

Chapter 4 is an expository account of the whole Linear Model Theory, which is the backbone of the analysis of experimental data. It is a complete account, with almost all details such as linear models with constraints on parameters, the conditioned linear model, identifiability and estimability, Aitkin's generalized least squares and the theory of statistical inference associated with the parameters of a linear model. All these details of the general linear model theory are probably not necessary in an introductory volume as only a weak generalized inverse and only some of its properties are required for the analysis of designs. The authors seem to prefer the matrix B, defined by $X'XB = X'$, rather than the matrix $(X'X)^{-}(X'X) = H$, as in Searle's 'Linear Models'. I prefer H as it is simpler to use for verification of estimability by the condition $\lambda' H = \lambda'$ for a parametric function $\lambda'\beta$. The role of $(X'X)^{-}$ in obtaining the variances and covariances of BLUES is not explicitly and clearly brought out. The properties of normal equations are very important in obtaining quickly the BLUES, the standard error and the $S \cdot S$; but this aspect is neglected. Estimation space and error space – two very important concepts are not brought out, nor a formal definition of the sum of squares due to hypothesis given. The role of idempotent matrices and the use of Craig's theorem for independence are also not stressed. The introduction of the concept of BLUPS (best linear unbiased predictors) developed by Harville and others would have been a welcome addition as it has a relationship with 'recovery of interblock information' in incomplete block designs.

Chapter 5 on randomization, as expected from the authors, is a very lucid exposition. The sections in Chapter 6 on the number of replications, the power of the F test and the tables of Δ^* (a measure of noncentrality) are extremely useful and valuable. The section on transformations is a thoroughly updated account of the corresponding material in the older book and the reviewer liked it then and now too. Chapter 7 dealing with multiple comparisons, orthogonal polynomials and contrasts is a welcome addition.

Chapter 8, which deals with supplementary information and analysis of covariance, is probably not of much importance these days, as computer programs that handle linear models can very well account for covariates.

The first eight chapters thus provide a comprehensive foundation and then in Chapter 9 one encounters the first important design – randomized block design. Estimation of missing observations is handled by the use of covariance analysis, rather than the principle of minimizing the error $S \cdot S$. This latter method is not even mentioned. It appears this topic is going to be introduced again in Volume 2.

Balanced incomplete block (BIB) designs and partially balanced incomplete block (PBIB) designs are briefly introduced. The reviewer would have liked this section in Chapter 9 to be preceded by a general development of the intra, as well as intra and inter block analysis, of incomplete block designs. Perhaps this will be done in Volume 2. The elegant mathematics of PBIB designs with its associate classes is by now obsolete for a practitioner. A general account of two-way designs and the valuable information contained in the C-matrix of a design would have been a better addition here. Kempthorne's original result, that BIB is the design that minimizes the average variance of elementary contrasts, is not mentioned.

The chapters on factorial experiments, Latin squares, split-plot designs, response surface designs and mixture designs are extremely well written, covering all major aspects, without burdening the reader with unnecessary details. The space devoted to 'rotatability' is, in my opinion, a little insufficient and a section on bioassay designs to estimate the relative potency would have enhanced the usefulness of the book. In my experience, the distinction between nonestimability, confounding, partial confounding and unconfounding and their relationships with the eigenvalues and vectors of the C-matrix of a design help the students understand the intricacies. The actual transformation from treatment effects to factorial effects and its reciprocal transformation for some particular 2^2 or 2^3 experiment would have been better, rather than only a mention of it. Extension of split-plot designs are given but strip-plot design is not mentioned. The analysis of split-plot designs is developed by using randomization theory, but the authors should have at least briefly given the alternative approach of using a Helmert's orthogonal transformation and the normal theory approach.

On the whole, I am very happy with this revised, enlarged and updated volume of Kempthorne's original text and look forward to the second volume. The criticism in this review is more in the nature of taste and a preference for a different arrangement and choice of topics. They do not lessen the value of this book in any way.

Reviewed by Anant M Kshivsagar, University of Michigan, Ann Arbor, Michigan, USA.

Le CT, Boen JR 1995: *Health and numbers: basic biostatistical methods.* New York, NY: Wiley-Liss, Inc. 247 pp. US $34.95; UK £22.50. ISBN 0 471 01248 3.

An instructor who is preparing to teach a statistics course generally considers a number of issues. What are the backgrounds and interests of the students? What is the size of the class and how often will homework assignments be collected? Will mechanics or concepts be the main emphasis? Are computers going to be used, or will students be required to do a lot of hand calculations? Successful courses can, with effort, be designed regardless of the answers to these questions. But consideration of these questions can help one decide which textbook to use.

Le and Boen's *Health and numbers: basic biostatistical methods* is designed to be the basis of an introductory course in biostatistics. Overall, its modest length, friendly style of writing and relevant examples should make it an easy text to read for students in the health sciences, especially if their mathematical background includes logarithms and simple algebra. But one aspect of the book that is disconcerting is its somewhat pessimistic attitude. Upon finishing the introductory chapter, the reader is left with the impression that the subject of statistics is difficult and counterintuitive. This impression could be a positive motivator, in that it will cause some students to take the subject seriously, and it will also provide solace to students who are having difficulty grasping statistical concepts. This pessimistic approach may also encourage scientific investigators to include a statistician on their research team, rather than attempting to do all of the statistics themselves. But I am concerned that this pessimism could discourage some interested students from ever attempting a course in statistics, and it may accelerate 'maths anxiety' in others.

Another thing to consider is the authors' views on the use of computers. This text assumes that students will be using hand calculations to perform the methods learned, and the authors go further to argue that statistical packages should only be used by statisticians, since the volumes of output produced by such packages can lead to naive misuse and abuse by nonstatisticians. This argument does have some merit, but it is also true that computers can remove the monotony of hand calculations so that students can concentrate on learning the concepts. Of course, not every institution has the facilities and the teaching assistants required to incorporate computers into an introductory statistics course.

The organization of this book is fairly traditional. Chapters 1 and 2 present methods of data description and display for discrete and continuous data. In contrast to what authors usually insert in their chapters about descriptive methods, Le and Boen wisely refrain from assaulting students with the tedium of the construction of stem-and-leaf diagrams and box plots (saving these details for later). Instead, they supplement the basic topics of descriptive statistics, graphs and tables, with useful examples of screening tests, rates and standardiz-