International Statistical Review

International Statistical Review (2014), 82, 3, 477–490 doi:10.1111/insr.12085

Book Reviews

Editor: Ananda Sen

Computational Approaches to Studying the Co-evolution of Networks and Behavior in Social Dilemmas
Rense Corten
Wiley, 2014, x + 175 pages, £50.00/€63.20, hardcover
ISBN: 978-1-118-63687-9

Readership: Statistics graduate and advanced undergraduate students, as well as practitioners in social sciences and economics.

If one is familiar with the Prisoner's Dilemma, together with its application to both interpreting and modelling social dilemma games, this book should appeal to him or her. Unlike the classical Prisoner's Dilemma with its focus on the circumstances under which people (players) will cooperate, this book addresses coordination problems (games) for which the players involved can benefit from cooperation but have an incentive to take advantage of each other—leading to less than optimal outcomes at the collective (societal) level. Often, cooperative relations, together with their incentives, do not occur in isolation but are embedded in a social context where the opinions of friends, family and colleagues influence choices and decisions.

The author argues that the value of the coordination game as an explanatory model has been under-valued in comparison with the attention given to the Prisoner's Dilemma. Under what conditions will players behave in such a way that they obtain the socially efficient outcome? Why do people cooperate? Are they motivated by considerations of fairness? Understanding social networks can play an important role in addressing these, and other, questions!

The development of theoretical arguments is appropriately illustrated with practical applications, and the author presents four studies that comprise Chapters 2–5 of the text, which collectively address two fundamental questions:

- How do social network structure and behaviour co-evolve in different social dilemmas?
- Under what circumstances is it more or less likely that networks and behaviour evolve into either optimal or suboptimal interaction structures?

Computer simulation is used in the book as a tool to illustrate the study of complex systems, as it is now almost routinely in, for example, the social sciences. The strategy adopted by the author to explore applications comprises four steps: (i) formulating a model with clearly specified assumptions; (ii) mathematically analysing the model to characterise stable states; (iii) using simulation techniques to explore implications under a wide range of conditions; and (iv) testing some of these predictions.

Chapters 2 and 3 introduce the theoretical models for coordination and cooperation problems, which form the basis for the empirical tests of hypotheses in Chapters 4 and 5. Interestingly, the author devotes Chapter 4 to an empirical study involving laboratory experiments—Appendix A contains the written instructions distributed among participants, and Appendix B describes the computer interface used. Chapter 5 presents a study of alcohol usage by young adolescents assuming that the social dynamics of alcohol use resemble a coordination game (a *friendship network*).

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute. Published by John Wiley & Sons Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Models and simulations with limited information are studied by the author under two *information regimes*: local information (players observing only the behaviour of their neighbours) and global information (players observing the behaviour of all); both as extremes at opposite ends of a general model in which players observe neighbours at a specified distance. The theoretical development and applications are readable and presented without the need for complex mathematical derivations. A basic knowledge of probability theory is necessary to obtain the most from the text and a familiarity with linear functions would be helpful, and the author cites many valuable references where further details of the material presented can be found.

The field of network dynamics is a relatively new research topic; with its roots in sociology and economics, and I am sure that readers of this book that are new to the field will be inspired to address the challenges and promises of further theoretical developments, together with their practical applications. Helpfully, the author concludes in Chapter 6 with possibilities for further research on the theoretical models presented and suggestions for further empirical research.

There is much here of interest to read and study, and once started, it is difficult both to put down and to resist searching the reference lists at the end of each Chapter. I recommend the text to anyone with an interest in social networks and look forward to reading of the further research that will no doubt be undertaken in the coming years.

Carl M. O'Brien: *carl.obrien@cefas.co.uk*
Centre for Environment, Fisheries & Aquaculture Science
Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

Optimal Mixture Experiments
Bikas Kumar Sinha, Nripes Kumar Mandal, Manisha Pal, Premadhis Das
Springer, 2014, xix + 209 pages, €90.09, softcover
ISBN: 978-81-322-1785-5

Readership: Students of the mathematical aspects of design for mixture experiments.

Mixture experiments are used to investigate substances whose makeup consists of ingredients that add to 1, that is, to 100%. It is a very practical topic in general, but also has interesting theoretical aspects. So it admits a wide range of interests.

Over a span of 12 chapters, the authors discuss optimality of mixture experiments under various theoretical constructs. This book is entirely theoretical, without a whiff of data in its pages. In other words, it is not suitable for anyone seeking knowledge of how to actually carry out mixture experiments and analyze the results in a real experimental situation. As a mathematical study of optimality in mixture designs, however, it is superb and can thus be fully recommended for an audience interested mainly in the mathematical aspects and properties of such designs. The writing is excellent throughout.

Summary: An excellent theoretical presentation of mixture knowledge.

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin-Madison
1300 University Avenue, Madison, WI 53706-1532, USA

Genomic Clinical Trials and Predictive Medicine
Richard M. Simon
Cambridge University Press, 2013, xiii + 144 pages, £27.99 (paperback), £70.00 (hardcover)
ISBN: 978-1-1074-0135-8 (paperback); 978-1-107-0880-9 (hardcover)

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

Readership: Statistics graduate students and practitioners with training in statistical modeling and computation.

This text is the first to present a concise explanation of how biomarkers can and should be used in clinical trials to better 'target' the effect of the treatment(s) under investigation to the subjects most likely to benefit. The issue at hand is that often biomarkers are examined in a post-hoc fashion that is susceptible to a high rate of false positive findings. To remedy this, Simon presents clinical trial designs that explicitly incorporate biomarkers in order to maintain the overall Type I error rate. Chapter 1 is a short introduction to clinical trial nomenclature, and Chapter 2 is a discussion of prognostic markers, in contrast to predictive markers, which are the focus of the remainder of the text.

For randomized Phase II trials, Chapter 3 discusses whether biomarkers should be used to randomize patients to treatment or control, or if instead, they should be utilized in interim decision rules based on biomarker-specific response rates when all patients are enrolled regardless of biomarker status.

For Phase III trials, Chapters 4 and 5 focus upon a variety of design topics, including the following: (1) enrichment, that is, the use biomarkers to restrict enrollment; (2) the use of biomarkers in interim analyses of futility and efficacy to determine if the trial should continue; and (3) using the interim analyses to also determine which patients (biomarker positive and/or biomarker negative) should continue to be enrolled.

Chapters 6 and 7 discuss statistical methods for clinical trials that have collected biomarker data and wish to use these biomarkers to model and identify the patient subgroup that is likely to respond to the treatment. Simon presents the use of resampling methods (permutation tests and bootstrapping) to develop a classifier that divides patients into treatment-response groups, or generates a probability of treatment-response for each subject. Methods are also presented for developing biomarker thresholds, as interpretation of biomarkers on a continuum is more difficult than if they are dichotomized as high or low. Related to these analyses, Chapter 8 provides guidelines for retrospective analyses of biomarkers such that an overall nominal rate of false positive findings is maintained, an important issue that still fails to be appreciated by most practitioners.

The author has done an excellent job of synthesizing his series of manuscripts into a single document and it should prove to be an outstanding resource for investigators, both statisticians and non-statisticians, seeking to appropriately incorporate biomarkers when designing a future clinical trial. Although the text contains appendices of statistical computing and modeling concepts and some source code is available through the internet, many of the designs are computationally complex and will test the abilities of non-statisticians trying to understand and implement the designs on their own. My only picky criticism of the text is its use of the word 'genomic' in the title that suggests that the content is only applicable to clinical trials collecting genomic data On the contrary, this work is useful to a much wider audience and applies to any clinical trial collecting patient data that might prove useful for classification and prediction of treatment response.

Thomas M. Braun: *tombraun@umich.edu*
Department of Biostatistics, University of Michigan
1415 Washington Heights, Ann Arbor, MI 48109, USA

Statistics in Action: A Canadian Outlook
Jerald F. Lawless, editor
Chapman and Hall/CRC, 2014, xxiii + 360 pages, £63.99, hardback
ISBN: 978-1-4822-3623-1

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

Readership: Statisticians and those interested in the applications discussed.

From the preface: '... unbeknownst to many, Canada has been for a long time – and remains – a leader in statistical research. To help (others) better appreciate (this) the Statistical Society of Canada has commissioned this book of expository chapters'.

There are 21 chapters in all, written by a total of 45 authors! The chapters cover a wide range of topics: survey methodology, child growth, dependence, sparsity, metropolis algorithm, computer experiments, Bayesian genetic modelling, survival in kidney transportation, health care, financial engineering, e-commerce, salmon and injection drug users, capture-recapture, marine ecology, climate change and problems in a food-hungry world. The writing is excellent throughout.

In summary, Canadians have made many important contributions to science through their use of statistical methods. This excellent book tells you what those contributions are and explains them in detail. It is an excellent presentation of knowledge gained with the help of statistics by Canadian researchers.

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin – Madison
1300 University Avenue, Madison, WI 53706-1532, USA

Clinical Trials with Missing Data: A Guide for Practitioners
Michael O'Kelly and Bohdana Ratitch
John Wiley & Sons, 2014, xxviii + 438 pages, £60.00/€69.70, hardcover
ISBN: 978-1-1184-6070-2

Readership: Statisticians, clinicians and researchers involved in clinical trials.

This well-written, well-structured and impressive book shows how to effectively deal with the difficulties that arise with the credibility and interpretability of results of clinical trials when there is missing data by providing practical guidance to address these missing data issues. Missing data in clinical trials is becoming increasingly prevalent, and this book will give clinicians and researchers a very broad perspective starting from prevention strategies to formulating clinically plausible assumptions about unobserved data to statistical analysis and interpretation. It serves as an excellent contemporary reference text.

The book is well signposted, and effectively structured. It starts off with introducing the problems associated with missing data, different types of missingness and provides a detailed guide to planning a strategy to prevent or minimize missing data in the design and implementation of a trial. The second part of the book concentrates on specialized statistical methodologies such as mixed models for repeated measures and implementation of multiple imputation in those context. Chapter 7 is dedicated to sensitivity analyses for testing study conclusions for robustness to missing data, especially when data is not missing at random. The examples in the book are taken from real clinical trials, which introduce the reader to realistic case studies and aid to master these concepts well. In addition, the worked out examples in SAS are also of great help, the code samples definitely deepen the understanding of statistical methods.

In general, ideas are well illustrated and interspersed with the theoretical developments. There is a good amount of theoretical details explained with the concepts for the readers interested in the underlying mathematics of the statistical methods, but others more interested in the applications will probably skip the technicalities. The material is presented in a clear and

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

accessible way. There are some inconsistencies noted in the indexing and also on page 270, the word repeated in the SAS code fragment 7.2 is missing the letter 'r'

This text represents a useful addition to the literature and is complementary to existing books on the topic. It is elegant and addresses what it attempts to, without being too overwhelming. I firmly believe that the authors succeed in their basic aim of explaining the nuances associated with missing data in clinical trials. Throughout the book, the authors provide useful pointers as to how to handle missing data though always stressing the point that there does not exist a one-size-fits-all solution. The book is written in a very understandable format and will provide an excellent starting point for graduate students interested in this topic.

Purna Mukhopadhyay: *purna.mukhopadhyay@arborresearch.org*
Arbor Research Collaborative for Health
340 E Huron Street Suite 300, Ann Arbor, Michigan 48108, USA

Stochastic Modeling and Mathematical Statistics
Francisco J. Samaniego
Chapman and Hall/CRC, 2014, 622 pages, $89.95, hardback
ISBN: 978-1-4665-6046-8

Readership: Undergraduate statistics and mathematics students as well as students in the quantitative sciences.

'Stochastic Modeling and Mathematical Statistics' is a new and welcome addition to the corpus of undergraduate statistical textbooks in the market. The singular thing that struck me when I initially perused the book was its lucid and endearing conversational tone, which pervades the entire text. It radiated warmth. Subsequently, on reading the preface, I noticed that another reviewer had made the same observation. This is one feature of the book that simply cannot be missed and, in my mind, is a big positive for an undergraduate text. In fact, with the right amount of balance, the conversational tone can be a positive even for more advanced texts, as Breiman aptly demonstrates in his book 'Probability'. The author also provides candid but friendly advice on how to approach the reading material with detailed discussions and explanations at several points, all of which are very conducive to self-teaching. This, of course, makes the text somewhat verbose, but I can readily envisage a motivated student with little stochastic background, learning quite a lot by working through the book by themselves over the course of a long vacation.

Chapters 1 through 5 are dedicated to a careful development of probability, starting from scratch. Here, the pattern of development is fairly standard. The thing to note is how mathematical discussions, which in the abstract can quickly become dry for the undergraduate, are supplemented by conceptual and philosophical ideas, and often historical references. Core topics for later statistical work are treated in earnest with the chapter on Limit Theorems somewhat shorter than the others. Of course, the notions of limits/asymptotics are not the easiest to get undergraduates to appreciate (and as someone who has taught the statistical inference course for statistics majors at Michigan for many years, I speak from personal experience) but I would, personally, have liked to see more coverage of this material—for example, figures illustrating the convergence of scaled histograms to the limiting probability density, more worked out mathematical examples. I should, however, emphasize that the key ingredients are all there: Chebyshev's inequality, the weak law, Central Limit Theorem (with proof) and especially the inclusion of the delta method, which I really liked. This is used later in the chapter on *Asymptotic*

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

Theory. A nice undergraduate introductory course in probability could easily be designed from the material in Chapters 1 through 5, supplemented by some additional material on stochastic processes.

Chapter 6 develops the central concepts of classical parametric statistics, treating unbiasedness in substantial detail. There is a brief foray into completeness territory and its implication for constructing Minimum Variance Unbiased Estimators (MVUE). I have mixed feelings about this since completeness is difficult to appreciate intuitively—a point the author does note at the very outset—and its definition is bound to be somewhat fuzzy without some idea of 'almost sure' concepts. In Definition 6.4.2., for example, what remains implicit is the fact that the support set of the complete sufficient statistic T could depend on the parameter θ. Of course, without the notion of completeness, one cannot really complete the discussion on unbiased estimation because there is no other systematic way to construct the best unbiased estimator. I liked the discussion in Section 6.6 where the notion of the MSE is introduced, and in particular, the little discussion before Definition 6.6.1 showing why a globally best estimator cannot be achieved in reasonable statistical models. It is exactly the argument that I present to my undergraduate class!

Chapter 7 deals with asymptotic theory as opposed to the fixed sample size considerations of the previous chapter and introduces, in particular, the two popular generic strategies for constructing estimators in broad classes of models: method of moments and maximum likelihood (ML) estimation. Very thoughtfully, two very useful tools in the context of ML estimation: the Newton Raphson method and EM algorithm are also introduced. My only quibble with this chapter is the lack of a proof-sketch of Theorem 7.3.1: the asymptotic efficiency of the MLE. While the rigorous proof is graduate-level material, a proof-sketch starting from the fact that the MLE solves the score-equation is typically accessible to advanced undergraduates. It needs Taylor expansion, but Taylor expansion has already been used by the author in connection with the delta method. Further analytical derivations appear at some other places in the text. A sketch of the key steps in a supplement/appendix might not be a bad idea for subsequent editions.

Chapter 8 deals with confidence intervals using the method of pivots for both small and large samples, the latter using asymptotic pivots. Chapter 9 on Bayesian estimation is very well-written and thought-provoking, and I was glad to see a more than cursory exposition of the topic. Unsurprisingly, given the philosophical scope of the Bayes paradigm, the author engages in substantial discussions on its pros and cons. Sections 9.3 and 9.4 are particularly interesting as they compare and contrast Bayes and Frequentist estimators: the inclusion of Theorem 9.4.1 is a really good idea. In addition to being an insightful result, it also gives motivated undergraduates the opportunity to see an original result from the author, which is inspiring. Chapters 10 and 11 deal with hypothesis testing and linear models, respectively. In both, the important topics are well-covered. I particularly liked Section 10.7 on fatherly advice on hypothesis testing. I wonder whether the author considered the possibility of including somewhere—maybe in this section—the problem of simultaneous testing of multiple hypotheses. This is typically not dealt with in undergraduate texts (and for good reason, I suppose) but given that the cutting edge work on hypothesis testing now lies in the problem of multiple testing and substantial developments have occurred over the last two decades, I suspect that time has come to introduce undergraduates to very basic principles of the 'many hypotheses' area.

Chapter 12 introduces basic nonparametric methods with emphasis on the rank test. I liked the inclusion of the Kaplan–Meier estimator in connection with the censored data problem. Personally, I would have liked to see the Kolmogorov–Smirnoff test discussed in the text, because it deals with the basic goodness-of-fit problem in nonparametric statistics and given that a section was devoted to this idea in the parametric set-up in Section 10.6.

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

All in all, I am very positive about this book. The usual question will arise as to how it compares with currently existing undergraduate texts—for example, John Rice's book, which is quite widely used. In terms of content and quality, both are great books at this level, and the adoption of any particular book would probably have to depend on the instructor. In my course at University of Michigan, I rely primarily on my own lecture notes and have used Rice as supplementary material. Having gone through this text, I am strongly inclined to add this to the supplementary list as well. I have little doubt that this book will be very successful as a course textbook in the years to come.

Moulinath Banerjee: *moulib@umich.edu*
Department of Statistics, University of Michigan
439 West Hall, 1085 South University Ave, Ann Arbor, MI 48109, USA

References

Breiman, L. (1992). *Probability.* Philadelphia, PA: Society for Industrial and Applied Mathematics.

The Skew-Normal and Related Families
Adelchi Azzalini
Cambridge University Press, 2013, 270 pages, £50.00, hardback
ISBN: 978-1-1070-2927-9

Readership: Students and researchers interested in probability distributions and their statistical aspects.

The literature on families of probability distributions has grown enormously in recent years, especially where the major interest is on distributions that overcome some limitations of the Gaussian distribution, both in the univariate and the multivariate case. This book is devoted to the treatment of one of the approaches to constructing such distributions that has become quite popular over the years. Briefly, the construction principle presented in the book 'starts from a symmetric density function and, by suitable modification of this, generates a set of non-symmetric distributions'. Hence, the main essence of the method is to obtain distributions with a possible departure from symmetry, with the skewness parameter representing one of the focal points of interest

The book is divided into eight chapters dealing with univariate and multivariate skew-normal distributions. It presents a range of results that underline both the probabilistic and the statistical aspects of the considered distributions. Each chapter also contains a list of problems that stimulate the reader to work with these new distributions. Moreover, I found quite interesting the historical and bibliographic notes that are present in each chapter, because they provide a helpful guideline for the vast literature related to the topic.

In my opinion, the book is a must-read for anyone interested in entering the world of skew-normal distributions. Further, for the specialist, it can serve as a good reference book.

Fabrizio Durante: *fabrizio.durante@unibz.it*
School of Economics and Management
Free University of Bozen-Bolzano, 39100 Bolzano, Italy

Examples and Problems in Mathematical Statistics
Shelemyahu Zacks
John Wiley & Sons, 2014, xx + 622 pages, £86.95/€104.40 hardcover
ISBN: 978-1-118-60550-9

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

Readership: Graduate students, their teachers, and anybody doing research in statistics.

From the beginning of the preface, 'I have been teaching probability and mathematical statistics to graduate students for close to 50 years. In my career I realized that the most difficult task for students is solving problems. Bright students can generally grasp the theory easier than apply it. In order to overcome this hurdle, I used to write examples of solutions to problems and hand it to my students. I often wrote examples for the students based on my published research. Over the years I have accumulated a large number of such examples and problems. This book is aimed at sharing these examples and problems with the population of students, researchers, and teachers.'

According to the back cover, the book 'provides the necessary skills to solve problems in mathematical statistics' and 'uniquely bridges the gap between theory and application and presents numerous problem-solving examples that illustrate the related notations and proven results'.

Although I believe that I have understood the author's point and that I agree with him, I cannot avoid being tempted to wonder at the meaning of 'problem solving'. If a bright student meets difficulties in solving examination problems, how can he/she show his/her brightness? A trivial answer is that problems are different: theory-oriented problems in a familiar scene are often easier to him/her than application-oriented problems in a foreign scene. The book is abundant with examples and problems, but many of them seem to be of the former type. Certainly, there are also problems of the latter type, but the bridge over the gap described earlier does not seem to be particularly large.

To look at this matter from another point of view, I again cite the back cover. The book features 'over 160 practical and interesting real-world examples from a variety of fields including engineering, mathematics, and statistics to help readers become proficient in theoretical problem solving'. All these examples are definitely interesting, but it seems to me that the words 'practical' and 'real-world' are here used in a very broad sense. Overall I agree that this book is a good guide in the field of 'theoretical problem solving'.

The book has nine chapters. Each of them is divided in four parts as follows:

Part I: Theory
Part II: Examples
Part III: Problems
Part IV: Solutions of selected problems

The examples in Part II of Chapter n have labels Example $n.1$, Example $n.2$ and so on. So the reader who is interested in examples concerning Section $n.m$ cannot immediately find them. (On the other hand, many examples contain material from different sections.) Instead, the problems have numbers $n.m.1$, $n.m.2$ and so on, and are easily found. Indicating harder problems and also such problems that are solved in Part IV would have been convenient to the reader.

The beginning chapters of the book are on basic probability theory (Chapter 1) and distributions (Chapter 2). Information in samples including sufficiency is covered in Chapter 3. Statistical hypothesis testing is taken up in Chapter 4 while Chapters 5 and 6 deal with point and interval estimation, respectively. Large sample inference is discussed in Chapter 7, followed by a chapter on Bayesian inference. The concluding chapter in the book is devoted to advanced topics in theory of estimation.

Part I of each chapter covers almost half of the content. This book is thus not simply a collection of 'examples and problems in mathematical statistics', as its title reads. It can be used

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

as a textbook of a 'learning by doing'-oriented graduate course in probability and statistical inference. It can also be easily used as a supplementary reading and a reference book.

Jorma K. Merikoski: *jorma.merikoski@uta.fi*
School of Information Sciences
University of Tampere, FI-33014, Finland

Incomplete Categorical Data Design: Non-Randomized Response Techniques for Sensitive Questions in Surveys
Guo-Liang Tian and Man-Lai Tang
Chapman and Hall/CRC, 2013, xix + 301 pages, $89.95, hardcover
ISBN: 978-1-4398-5533-1

Readership: Statistics graduate and advanced master or PhD students, as well as applied statisticians.

Survey methods for posing sensitive questions to the respondents, while at the same time ensuring that their replies are treated confidentially, have always constituted a challenge to survey research. In 1965, Warner suggested that two questions, constituting a negation of each other, should be presented to the respondent with known probability, but without the interviewer's knowledge of which of the two questions were being answered. As the interviewer had no knowledge of which questions were given a yes or a no reply, but only the reply could be noted by the interviewer, confidentiality was thereby formally achieved and due to the known probability for the choice of the two questions, the right estimate could thus be derived by the researcher. There are a number of non randomized methods based on the same idea where the inference on the some parameter of interest around the sensitive question is sought.

It is the aim of the authors of the current book to describe all aspects of inference to a variety of models dealing with issue described above, including the Crosswise Model, the Triangular Model, the Hidden Sensitivity Model and the Parallel Model. For all these models and generalizations thereof, a detailed discourse of hypothesis testing, estimation, large sample inference and findings from finite sample simulation is presented. Examples accompany the discussion models and methodologies, making the book quite self-contained. The book contains, nonetheless, a comprehensive list of references in the area.

No matter to what extent the researcher is working within this framework, so that the respondent is ensured that his reply is confidential, it is always a challenge to convince the respondent of the same. In this respect, it is not only the model that plays a role. The various techniques applied in posing sensitive questions in a confidential manner can be quite difficult for all respondents to understand. Further, there are respondents who willingly provide replies to sensitive subjects, if they are asked in the right way, and see it as a burden or lack of confidence if the questions are not directly asked. The book also describes that it is a challenge to implement the various techniques, but the main objective of the authors has been the question of a stringent description of the models and their inference. A topic that could be of relevance to look at in other contexts is testing with the different methods, which systematically compare the various models. For example, the incidence of sensitive issues such as abortion, examinations that have not been passed or breaches of the road traffic act could be obtained from outside sources. If the answer to this sensitive question is known beforehand, one could then make a comparison as to which model is the closest. Qualitative methods could also contribute to this.

Another theme is how to create confidence by means of models for posing questions in a world where the authorities know more and more about the citizens, which can pose a

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

challenge to the respondent's confidence as to whether the researcher knows more than he or she has declared and if confidentiality is as great as is asserted by the researcher. This can be essential to the respondent's confidence with the process. Training the interviewer in using the models is also of great importance. Responses often depend on the gender and age of the interviewer-responder pair. Propensity of non-response is heavily associated with confidence in the interviewer or the model, and the issue is compounded in the case of sensitive questions, resulting in bias.

No surveys are better than their weakest link. When sensitive questions are posed, all possibilities of reducing the bias must be considered. The book provides an in-depth description of all aspects surrounding inferences and analyses of the different methods to deal with sensitive survey questions. When this is added to the knowledge of conducting interviews with specific subject subgroups, one is well equipped to conduct a survey with the least possible bias.

Peter Linde: *pli@dst.dk*
Statistics Denmark
Sejrogade 11,DK 2100 Copenhagen, Denmark

Functional and Phylogenetic Ecology in R
Nathan G. Swenson
Springer, 2014, xii + 212 pages, £23.99/€42.39, softcover
ISBN: 978-1-4614-9541-3

Readership: Statistics graduate and advanced undergraduate students, as well as practitioners in bioinformatics and ecology.

The plethora of books on the use and applications of the programming language R is becoming increasingly difficult to keep abreast of. This latest addition to the catalogue introduces phylogenetic and functional trait analyses that can be implemented with ease in R—an advantage and a drawback of the analytical code that is fast becoming a common tool in analyses of ecological data. With its roots in the author's PhD research quantifying function in tropical tree communities and its placing in a phylogenetic context the text assumes a basic knowledge of the underlying community theory.

Books on R may be categorised as (i) either theoretical texts assuming little subject knowledge, which present and develop the theoretical issues together with their application in R using real and simulated data sets; or (ii) purely application-oriented ones, which aim to teach the avoidance of incorrect programming and inference(s), but assume a basic theoretical knowledge of the subject. This book falls in the latter category, but the author provides references to published papers and books from which one can quickly grasp the fundamentals if needed—making it ideal both as a companion to a classroom text and for individual personal study. The reference list is extensive, covering 145 published papers and books, but the book suffers from poor indexing, merely covering a compilation of R functions. The utility of the text would be enhanced if the author included a section collating all R functions with a brief description of their purpose as has often been provided in recently published texts devoted to other programming languages and packages such as SAS, GENSTAT, GLIM and SPSS. Perhaps it is something for the author to consider in any future revision of the text.

The author provides an introductory primer for ecologists in Chapter 2, which must be essential reading for anyone contemplating the collection of phylogenetic data, its structuring for analytical purposes, plotting and manipulation. Subsequent chapters cover metrics of biodiversity, null modelling and randomisation. Importantly, the last chapter focuses on interfacing the

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

R environment with the C-based program *Phylocom*; thus recognising that not all programming is best achieved in R alone—clearly, a strength of the author's text and a pointer to a future when integration of multiple programming platforms will become common place. The presentation of commands and programming details are suitable for both users of PC and MAC computers.

I can recommend the book to anyone with an interest in community ecology; both student and researcher alike, with a desire to investigate the distribution, abundance, dynamics and diversity of species.

It would, however, be remiss of me not to mention that there are many detractors of phylogenetic community ecology, and I invite the reader to search the literature to assess the arguments for and against, issues and controversies—something that the author does not address. This noted, the author's text will at least provide the confidence that metrics are correctly programmed before their utility and appropriateness is challenged. I am sure the debate will continue but the author has provided a valuable collection of programming tools that will ease further study and investigation, and consequently facilitate constructive challenges to their application and utility.

Carl M. O'Brien: *carl.obrien@cefas.co.uk*
Centre for Environment, Fisheries & Aquaculture Science
Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

Statistical Applications for Environmental Analysis and Risk Assessment
Joseph Ofungwu
Wiley, 2014, 648 pages, £83.50/€100.20 hardcover
ISBN: 978-1-118-63453-0

Readership: Statistics graduate and advanced undergraduate students, as well as practitioners in earth sciences, geology and hydrology.

Do not be deceived by the title of this book—the first $4/5^{th}$ of the text are an introduction to standard statistical theory and techniques that any undergraduate student might expect to be exposed to in an applied statistics course, the last $1/5^{th}$ of the book covers environmental sampling design and risk assessment. This is not to detract from the usefulness of the text but merely to emphasise that there existed a need to provide a statistical text that is accessible to those environmental professionals trusted with the protection of public health and ecological well-being.

The book describes fundamental statistical concepts and their applications without assuming prior knowledge of statistics. The author presents and discusses the role of statistics in environmental contaminant sampling and exposure risk assessment. The latter are presented in the context of the relevant guidelines of an environmental regulatory agency, for example, USEPA (United States Environmental Protection Agency), providing a real-world situation that should typically be encountered by practitioners. Elements of sampling design, focusing on the required data sample sizes for various types of explicitly defined statistical analyses and tests, along with associated sampling plans are well described by the author. The author considers risk assessment for the monitoring of environmental contaminants —carefully accounting for the presence of background concentrations of chemicals at a site in order to correctly assess overall health risks associated with exposure to the site contaminants that may be in excess of risk guideline limits. Graphical methods and quantitative analyses are used separately and jointly, where appropriate, in order to facilitate routine monitoring and risk assessment. In addition to the conventional (point risk estimation) approach that produces a single estimate of

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

risk attributable to exposure to a contaminant, the author presents probabilistic risk assessment too. This involves dealing with likelihoods of specified risks using Monte Carlo simulation techniques that explicitly incorporate uncertainty in input parameters. The approach is illustrated using the example of inhalation exposure to the airborne volatile organics benzene, trichloroethylene (TCE) and perchloroethylene/tetrachloroethylene (PCE) at a prison facility. The book is worth reading if only for these examples and applications.

The author addresses a historic lack of access to computer programming tools with modern-day illustrations of applications using the R software system; including details of the R functions to call and the R scripts that can be routinely run.

I can recommend the book to those indicated in the readership above as the text highlights the important role of statistics in protecting public and environmental health while hopefully avoiding unnecessary environmental clean-up costs.

Carl M. O'Brien: *carl.obrien@cefas.co.uk*
Centre for Environment, Fisheries & Aquaculture Science
Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

Bayesian Essentials with R
Jean-Michel Marin and Christian P. Robert
Springer, 2014, xiv + 296 pages, €52.99, hardcover
ISBN: 978-1-4614-8686-2

Readership: Statistics graduate students as well as practitioners of applied Bayesian modeling and methodologies with intent of implementation in the R statistical package.

This book, now in its second edition, is designed as an introductory graduate-level text book for applied Bayesian modeling using the R package. Some basic knowledge of probability, distributions, likelihood, familiarity with calculus, and linear algebra is necessary for the reader.

This book provides an application driven treatment of Bayesian methods and guides through practical implementation through R. The book has sufficient description of all Bayesian modeling techniques and Monte Carlo computational tools at a reasonable depth. The authors have also created an R package called *bayess* for easy implementation of most of the discussed techniques and have provided a good selection of exercise problems.

Chapter 1 introduces some basic concepts of programming and data manipulation in R including vector and matrix operations, data structures, probability distributions, and visualization. More advanced tricks like writing R functions, working with R objects are also very nicely presented here. A guideline for using the *bayess* R package specially designed to implement the examples presented in this book is included in this chapter.

Chapter 2 contains basic formulas related to modeling with normal distribution. Fundamental definition and motivation of prior, conjugate priors, and posterior distribution along with some Bayesian inferential tools such as the credible intervals and Bayes factors are provided with full details here. A wonderful explanation of the improper prior and associated issues along with some basic Monte Carlo tricks for approximation are included in this chapter.

Chapter 3 is focused on regression and variable selections. Here, the authors discuss linear regression model, classical least squares estimator, and its properties. A Bayesian formulation of linear model with Jeffreys prior and Zellner's G-prior is introduced in this chapter with full rigor. The authors further touch upon few related concepts such as model comparison, prediction, Gibb's sampling, and the BayesReg R function. The variable selection is elicited

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

through the concept of G-prior and shrinkage, and also through the stochastic search over the model space.

Chapter 4 builds upon the concepts of linear model (presented in Chapter 3) in transitioning the discussion to the generalized linear models. In this chapter, a nice presentation of probit and logit models for binary data is provided. For both models, the authors discuss possible choices of informative and non-informative priors. The last section in this chapter deals with modeling count data, contingency tables, and Bayesian inference under flat priors. Strewn throughout the chapter are several data examples along with accompanying R codes that makes the whole thing very user friendly and easy to implement.

Chapter 5 introduces the concepts related to capture-recapture experiments. Inference in a finite population and several sampling models are described here with great clarity and detailed derivations. Among the discussed sampling models are the binomial capture model, two-stage capture-recapture model, and the T-stage capture-recapture model. The third section of this chapter specially deals with cases when the population size is not fixed over the duration of the experiment ('open population'). A detailed derivation is given for the marginal and conditional posterior distributions of the 'open population' model along with the accept-reject algorithm for model fitting purpose. The final section deals with some more advanced techniques such as the Arnason–Schwarz capture-recapture model with descriptions about associated prior formulation and Gibb's sampling based parameter estimation.

Chapter 6 covers the idea of mixture models. In this chapter, the authors introduce the concepts behind finite mixture models, mixture likelihoods, and the associated posterior distributions. As mixture models are particularly difficult to sample from, the authors devoted an entire section laying out the detailed design of the MCMC algorithm to fit a finite mixture model. Fundamental concepts such as label switching difficulty, prior selection and tempering are well explained and also included in this chapter. The final section of chapter 6 is devoted to dealing with mixtures with an unknown number of components. Here, the authors describe a marginal likelihood based testing perspective to accurately estimate the number of clusters. The chapter is very well organized and helps the reader to easily understand the concepts of mixture models.

Chapter 7 gives a basic introduction of time series data, the associated modeling and inference issues under the Bayesian framework. In Sections 2 and 3, popular time series models like the autoregressive and moving average models are extended to Bayesian paradigm. The last two sections deal with ARMA models, hidden Markov models, their extensions and implementations through MCMC algorithm. The authors successfully present the ideas of Bayesian time series analysis in this chapter without having to go into too much detail.

Chapter 8 is motivated by problems in image analysis. In this chapter, the authors give a nice formulation of image analysis from modeling perspective. A very good description of the spatial dependence in a pixelized image is included in Section 2. The authors describe several modeling techniques for the spatial dependence such as usage of lattices, Markov random fields, Ising model, and Potts model. Each model is clearly illustrated with an application. All R codes are included for model implementation, which makes this chapter really handy for anyone who is starting his or her research and application in areas of image analysis. Methods for image segmentation for the purpose of extracting object image from noises is discussed in the concluding section of this chapter.

In summary, 'Bayesian Essentials with R' by Jean-Michel Marin and Christian P. Robert is an excellent book for anyone who is interested in practical applications of Bayesian methodologies. An introductory course in probability or statistics is necessary to understand the materials presented in this book. The level of this book is suitable for a graduate course. Overall all concepts are discussed very well with full mathematical rigor and underlying practical motivation.

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute

Each discussed methodology is well illustrated by appropriate data examples. All R codes to implement the discussed models are also provided in this book for quick application. The price is extremely reasonable, and I will enthusiastically recommend this book to anyone who wants to learn and apply Bayesian methodology for data analysis purpose.

Sounak Chakraborty: *chakrabortys@missouri.edu*
Department of Statistics, University of Missouri-Columbia
209F Middlebush Hall, Columbia, Missouri 65211-6100, USA

© 2014 The Authors. International Statistical Review © 2014 International Statistical Institute