error, confidence interval).

• "Models and Sampling Distributions" (formalization of the notion of model, making decisions through probability models, normal and t distributions, central limit theorem).

• "Comparing Two Populations" (application of Chapter 4 concepts to two populations—proportions including the Wilson adjustment and means, medians including bootstrapping).

• "Comparing Several Populations; Exploring Relationships" (chi-square tests, one factor analysis of variance, correlation and simple linear regression).

Chapters follow a similar format as students are led to work through the parts (lettered a, b, c, . . .) of an "Investigation" to learn a topic experientially. There are often written insertions of background information on the study at hand. These motivate the questions being asked and get the student even more involved in the context of the study. Also inserted are definitions that introduce a formula or concept, or discussions that give the student investigator additional statistical interpretation, all to enhance the "Investigation" and allowing it to continue. Following each "Investigation" is a section called "Study Conclusions" which summarizes the findings that the student should have discovered by the end. Often a "Terminology Detour" appears that introduces new terms. Some "Practice Problems" follow, then a "Summary of the Investigation." At the end of the chapter are a "Chapter Summary" and a "Technology Summary" that pull together all of the "Investigations." Each chapter ends with at least 40, and as many as 65, well-designed exercises. An extensive list of "References" for the "Investigations, Practice Problems, and Exercises" is also included in each chapter. There are two appendices: one is an overview of Minitab's basics, subcommands, and macros used in the text; the second describes the superb Java applets, 13 regarding conceptual demonstrations and simulations and 3 on statistical analyses.

One potential drawback of *ISCAM* may be that it is virtually impossible to cover all the material in a three-credit, one-semester course. The authors indicate that it takes two to three weeks per chapter. Chapter 1 seems mandatory as it contains concepts and principles that are inherent in the practice of statistics and required for the rest of the text. If one wanted to concentrate on categorical analysis, then it is suggested to cover all of Chapter 3 with 5.1, 5.2, and 6.1. A course involving quantitative variables would include Chapters 1 and 2 with sections chosen from the remaining four chapters.

The Web site for this text is *www.rossmanchance.com/iscam/* and contains additional information. In addition to the wealth of information that is available there (including answers to 13 "Frequently Asked Questions"), there are extensive instructor resources available by obtaining a password from the publisher. As an aside, the text is paperback. All pages are perforated for easy removal and three-hole punching to put in a binder, if desired.

There is no doubt that some mathematics majors become enamored with statistics as a result of a traditional mathematical statistics course. I was one such student, thrilled by counting techniques, moment-generating functions, and the Jacobian technique for finding the distributions of functions of two random variables. But I now wonder how many more students would be attracted to the profession by an initial experiential exposure to statistics, which immerses them in the investigative process of statistics, as does *ISCAM*. I am excited about the prospect!

Jerry L. Moreno
John Carroll University

A Crash Course in SPSS for Windows: Updated for Versions 10, 11, 12 And 13 (3rd ed.).

Andrew Colman and Briony Pulford. Oxford, UK: Blackwell Publishers, Ltd., 2006, xiii+204 pp., $26.95 (P), ISBN: 1-4051-4531-5.

This is a new edition of a book of the same title by Corston and Colman (2003). Besides updating the material to reflect changes in SPSS Versions 12 and 13, this spiral-bound third edition adds two short chapters about manipulating variables (recoding and computing), handling large data files, and using the syntax editor. A much-welcomed appendix on exporting and importing Excel files completes the new material.

This textbook provides an inexpensive and quick way to become familiar with SPSS. True to its title, the book presents an abbreviated introduction to SPSS without being brusque. The authors expect novices to be able to complete the book in under 10 hours. For those familiar with the Windows operating system and spreadsheet programs (e.g., Excel), the material can completed in a few hours less. (I am an occasional SPSS user and I finished the book in less than four hours.) The pace and level of the text are such that any student should be able to complete the text independently.

The first two chapters cover the most rudimentary aspects of using SPSS: starting the software, moving windows, scrolling, and loading, printing, and saving data. The next 12 chapters describe how to carry out various statistical procedures and tasks including: descriptive statistics; correlation coefficients; chi-square tests of association and independence; independent samples, paired samples, and one-sample t tests; Mann-Whitney and Wilcoxon tests; one-way, multifactor, and repeated measures analysis of variance; multiple regression; log-linear models; and factor analysis. There is also a chapter on creating bar and pie charts, multiple line graphs, and scatterplots.

The book is replete with screen shots of SPSS menus and printed output. The examples are based on real data from research articles or other textbooks, and many of the same datasets are used throughout the book. The authors give sufficient description of the data to motivate the analysis and provide a basis for interpretation of the subsequent results that they provide. Pedagogically, I appreciate this because it reinforces how output must be translated into contextually meaningful and statistically accurate statements.

Each of the analysis "how to" chapters follow the same mold. A background section is first presented that gives a concise description of the statistical procedure(s), who developed it, key assumptions, and caveats. The background does not contain a mathematical description of the method but more of a broad overview that helps orient a student who may have recently learned what seemed like a plethora of methods. This is followed by a description of the dataset to be used in demonstrating the particular procedure. The raw data are then presented along with step-by-step instructions on how to input it for analysis. All of the datasets are very small and can be entered quickly by hand, but they may also be downloaded from the publisher's Web site. Commands for carrying out the analysis are given, along with screen shots, and the output is described in detail. The authors close with a short contextual interpretation of the results.

Overall, the book delivers a quick and easy to follow introduction to SPSS suitable for novices. It covers the very basics of common methods, but the selection of methods leans toward the social sciences—factor analysis and log-linear models are covered, while logistic regression and survival analysis are not. I recommend it as a supplementary text for a high school or undergraduate introductory statistics course.

J. Wade Davis
University of Missouri–Columbia

REFERENCES

Corston, R., and Colman, A. M. (2003), *A Crash Course in SPSS for Windows* (2nd ed.), Malden, MA: Blackwell.

R Graphics.

Paul Murrell. Boca Raton, FL: Chapman & Hall/CRC, 2006, xix+301 pp., $69.95(H), ISBN: 1-58488-486-X.

1. OVERVIEW

R Graphics is an excellent technical introduction to drawing graphics with R. It pulls together information currently scattered throughout various R documents and help pages. The organization and writing is clear and coherent, which is especially welcome when dealing with the intricacies of R graphics. It serves as a very useful reference book.

2. WHAT IS IN THIS BOOK?

The book starts with an introduction to R graphics including a gallery of many different graphics made using R, demonstrating the power of this system for generating static, high-quality plots. These examples range from basic statistical plots, to cleverly annotated plots, to sophisticated 3-D diagrams, and even whimsical art pieces! It is a very inspiring beginning, and one feels well motivated to do battle with R.

The main material in the book is divided into two parts. The first describes the traditional, or base, graphics, while the second part introduces the newer grid graphics system, including lattice graphics. Grid was designed and written almost entirely by the author of the book, so Dr. Murrell speaks with clear authority here.

The base graphics system was the first graphics system developed for R. If you have used R in the past, it is likely to be an old friend (or enemy). Base

graphics has a simple metaphor: ink on paper. Just like drawing with your pen, you cannot undo your mistakes, except to start afresh. More formally, there is no representation of the graphics independent of their presence on the plot so you can only add, not edit or delete, existing output. This makes base graphics simple and easy to understand but fundamentally limited. This limitation is best seen when trying to customize graphics, where you either need to start from scratch or grapple with many arcane settings. *R Graphics* provides an excellent summary of these details.

In terms of functionality, but not yet popularity, base graphics has been superseded by the grid graphics system. The explanation of grid is the strength of this book. The section begins with a description of lattice graphics, a system that illustrates the power of grid. Lattice is an implementation of trellis graphics (Becker, Cleveland, and Shyu 1996), which provides an easy way to produce multiple plots based on different subsets of a dataset. The plots typically share the same scales and allow one to investigate relationships between two variables conditional on one (or more) other variables. Lattice graphics present a higher level of abstraction than base graphics, but configuring lattice can be difficult due to the multitudinous (378 at last count) and repetitive options. This book provides a handy reference to some of these options as well as a brief discussion of annotating lattice plots. The description of creating new plot types is briefer still, largely a reflection of the limitations of lattice.

The framework provided by grid graphics has a number of advantages over the ink on paper design of base graphics:

- Grid objects have an independent representation as R objects not just pixels on a screen (described in sec. 7.3).

- A system of viewports allows for extremely flexible layout (described in sec. 5.5).

- A range of coordinate systems makes it easy to draw what you want where you want it (described in sec. 5.3).

Although it is easy to draw directly to the screen with grid, its real power lies in creating objects that can be drawn at a later time. This allows much greater flexibility as objects can be extensively modified and even deleted. Unfortunately, of the few R packages that use grid, even fewer return grid objects, which makes building on top of them almost as hard as building on top of base graphics. We hope that this will change as more people read this book. With base (and even lattice) graphics, it is hard to write reusable graphical functions, and solving this problem is the promise of the grid system.

An R package, `ggplot`, released after *R Graphics*, further demonstrates this potential. It builds on grid graphics using the principles described by Wilkinson (2005), providing R users with a higher level compositional language and good defaults for generating both basic and sophisticated trellised plots.

3. WHAT IS NOT IN THIS BOOK?

R Graphics is a technical book. It does not attempt to explain the how or why of statistical graphics. It covers the technicality of plot production, not purpose, giving the reader enough rope to be extremely creative or to fail spectacularly. It does, however, point readers to other books where they might learn about good graphics.

The book does not describe interactive or dynamic graphics, such as linked brushing. Many readers may have experienced the power of a tight coupling between statistical analysis and graphics through XLispStat (Tierney 1991) or DataDesk (Velleman 1988). These systems enable the user to make plots of model diagnostics that are dynamically linked to plots of the data, updating themselves as a model changes. This is very useful for exploring data, but it remains difficult to realize within R. Base graphics and grid graphics are not designed with interaction in mind so, appropriately, *R Graphics* focuses exclusively on static and presentation quality graphics, and the author only points readers toward recent developments in interactive graphics in R.

The book's Web site provides an electronic version of the code used in the book, but little else. It would be useful to add material on features of grid that have appeared subsequent to the release of the book. For example, a recent release includes facilities for converting postscript files to grid objects.

The description of grid graphics is the most important part of this book, explaining a new, more powerful graphics engine. However, this system is not given sufficient emphasis and does not provide enough general examples that could be easily adapted to a reader's data. Our fear is that most readers will not get past the traditional graphics part of the book. While *R Graphics* provides a good summary of this system, it may slow the needful death of a primitive and out-dated system.

4. RECOMMENDATION

This is an excellent book. Everyone who uses R to draw graphics (all R users, we hope!) should have it open on their desk or at least on their shelves! We especially encourage readers to get beyond base graphics and carefully study the grid graphics engine.

Hadley Wickham
Iowa State University
Dianne Cook
Iowa State University

REFERENCES

Becker, R. A., Cleveland, W. S., and Shyu, M. J. (1996) "The Visual Design and Control of Trellis Display," *Journal of Computational and Graphical Statistics*, 5, 123–155.

Tierney, L. (1991), *LispStat: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics*, New York: Wiley.

Velleman, P. F., and Velleman, A. Y. (1988), "DataDesk Professional," Northbrook, IL: Odesta Corporation, *http://www.datadesk.com*.

Wilkinson, L. (2005), *The Grammar of Graphics*, New York: Springer.

Statistics: An Introduction using R.

Michael J. Crawley. Chichester, UK: Wiley, 2005, xiii+327 pp. $130.00 (H), $45.00 (P), ISBN: 0-470-02297-3 (H), ISBN: 0-470-02298-1 (P).

The primary goal of this book is to introduce the essentials of statistical analysis with virtually no theory to the "unwilling" student. Other goals are teaching R, encouraging the readers to practice statistical analysis where models are exposed to rigorous criticism and simplification, and de-emphasizing "recipe teaching" such as "p value ≤ 0.05."

The publisher's summary claims that this "is the first text to offer such a concise introduction to a broad array of statistical methods, at a level that is elementary enough to appeal to a broad range of disciplines." I do not agree that this is the first such textbook. Neither do I find it to be the best one. The two major contenders are Dalgaard (2002) and Verzani (2005) (whereas, e.g., Maindonald and Braun (2003) requires a slightly higher level of mathematical statistical sophistication). However, I like the book's style of introducing and discussing statistical reasoning and thinking in a way that seems typical of "British Statistics." I particularly enjoyed parts of Chapters 1 "Fundamentals" and 7 "Statistical Modelling."

There are 16 chapters ordered as "Fundamentals," "Dataframes," "Central Tendency," "Variance," "Single Samples," "Two Samples," "Statistical Modelling," "Regression," "Analysis of Variance," "Analysis of Covariance," "Multiple Regression," "Contrasts," "Count Data," "Proportion Data," "Death and Failure Data," and "Binary Response Variable," plus an appendix on "Fundamentals of the R Language." This represents fairly standard content and is quite similar to Dalgaard (2002). The text is loosely structured as there are a varying number of *un*numbered sections per chapter and no figure or table captions, which can be slightly confusing at times. A strong point is an extended index of 19 pages.

As nowadays is common, there is an accompanying Web site, *http://www.imperial.ac.uk/bio/research/crawley/statistics/*, which contains the datasets, all the scripts from the book (even those to create extra graphics), and exercises (as a PDF), which the book does not include. Unfortunately, the Web site material is currently organized as collection of text files, instead of as an R package, which would be much more convenient and has become customary for most similar books.

There are several small issues which I do not like. First, there are mistakes (not typos) which could have been avoided with more care. In the regression chapter, the box "Proof that $SSY = SSR + SSE$" contains a blunder claiming that the reason that $\sum(y - a - bx)(bx - b\bar{x}) = 0$ is due to $\sum(y - a - bx) = 0$ and $\sum(bx - b\bar{x}) = 0$. (The author uses no summation indices nor does he differentiate between true (unknown) and estimated parameters, e.g., a and $\hat{a}$). The argument is incorrect, since one can easily have $\sum_i A_i = \sum_i B_i = 0$, but $\sum_i A_i B_i \neq 0$.

In the multiple regression chapter, a whole series of model simplifications are based on p values of estimated coefficients. Only at the end of an example do we learn that a residual analysis reveals that a different model should be fit, namely with response "log(ozone)" instead of "ozone," and that an outlier should be removed. This means that the earlier p values were not trustworthy, and model selection could have led to different conclusions.