At first glance, it seems odd to find a text of more than 330 pp. devoted to three SAS® procedures that are focused not on analytics but on creating visualizations. However, a careful reading shows just how many details there are in the SAS graphic procedures that need to be discussed. Sanjay Matange and Dan Heath of SAS provide an in-depth look at the PROC SGPILOT, PROC SGPANEL, and PROC SGSCATTER procedures in this text that could equally serve as a teaching instrument or a quick tutorial for producing high-quality visuals within the SAS system. The text is easy to read and serves as a useful resource for individuals who need to add analytics to their data visuals or those who wish to use options to create the most professional visualization possible in SAS.

The text is laid out into 16 chapters and starts with a general introduction of data-based visualizations and an overview of the text. Chapter 2 gives a description of each procedure involving the basic syntax and utilization of each of the three procedures. Chapter 3 then gives a brief description of the different graph types available, and while much of the focus is on commonly used visuals (scatterplots, histograms, boxplots, bar charts, etc.) there is also early discussion of domain specific visuals (such as vector plots or step plots). Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.

Chapter 4 brings readers into the basics of creating each of these plots with easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Chapters 5, 6, and 7 look at bivariate association style plots first focusing on two quantitative metrics, followed by one easy to recognize and access SAS code. Early on, the authors provide a nice resource in the form of a table that illustrates exactly which graph types can be combined within each of the procedures. This visual serves as an early indication of the complexity and possibility of layered visuals that the SG procedures can produce.
The hypothesis testing/critical region problem is introduced well, with motivation provided by error probability analysis. They gently present the Neyman–Pearson lemma, although I was surprised they did not use the generalized version with the randomized test function to handle the discrete case; in fact, test functions were not mentioned. P-values, the power function, most powerful tests, the likelihood ratio tests, and the difficulty in finding uniformly most powerful tests were well explained. The proof of the distribution of the P-value under $H_0$ was left as an exercise for the reader (but with solution in Appendix B). The chapter contains good examples, and concludes with an adroit exposition on sequential Wald testing.

The authors handle asymptotic analysis in Chapter 5 by introducing three modes of convergence of sequences of random variables, the weak law of large numbers (including Khinchin), and consistency. For motivation, they provide several examples of distributional convergence, including Poisson, exponential, and extreme value distributions, before presenting the iid central limit theorem, and the Berry–Esseen bound for finite variance problems. After handling the plug-in theorem and asymptotic normality of MLE, they provide a good discussion and nice examples for the Wald and Wilks tests/confidence intervals. They address the geometry of confidence regions in the multiparameter case for exact and asymptotic tests, as well as goodness-of-fit tests.

The next two chapters address Bayesian analysis and elements of statistical decision theory. The Bayesian philosophy itself is well presented, as are the necessary tools such as choice of prior, calculation of posterior, point estimation/credible sets, and hypothesis testing. The examples are good. The decision theory chapter covers loss function optimality, risk, admissibility, minimax ability, etc. Many motivating concepts are explained using Bayesian illustrations, Bayes’ estimators, and rules (tests). As in most treatments of decision theory, the choice of loss function remains that of the decision maker.

The final chapter on linear models is a handy collection of results, intuition, and examples of use of the standard linear model. In it, Abramovich and Ritov concisely develop the least-square solutions and, with an appropriate distributional assumption, the MLE for $\beta$ and $\sigma^2$. Fisher information/Cramer–Rao lower bound, confidence intervals for the coefficients and predictions, etc. They consistently provide geometrical insights and excellent examples. The chapter concludes with a brief overview of one-way and two-way analysis of variance.

Part of the implicit “fine print” of this book is that the first month of material is presented in Appendix A, which includes some basic probability theory, random variables, functions of random variables, distribution theory, univariate parametric families, and some linear algebra projection theory applied to statistics. The second appendix contains solutions to selected exercises in the text, from the good collection of exercises at the end of each chapter. Offered without solution are many thought questions scattered throughout the text.

Overall, there are some quirks, as there would be if any of us wrote such a book. For example, the authors disclaim use of measure theory, but in their introduction for minimal sufficiency they have no qualms about presenting an equivalence class explanation. This, like measure theory, will likely lose most engineers. Similarly, some of their treatment of the Hat projection matrix uses geometric concepts from linear algebra; this works, but it is a bit more rigorous than the rest of the text. Most of the multivariate versions of the topics are asterisked as more advanced. They do not present almost sure convergence, but they do use convergence in $\mathbb{P}$-mean. They do not present the Rao (Score) test. And, although I do not begrudge another author’s notation, their moment notation uses the “backward” $\mu_2 = E(Y^2)$, versus the aesthetically superior (in our opinion) $\mu_2 = \mu^2 - \mu_1^2$. Their distributional notation is fine with $f_\theta \in \mathcal{F}_\theta$, $\theta \in \Theta$ and $f_\theta(y; \theta)$.

As teachers of theoretical statistics, we can use a new approach, which this text offers. The book does not aspire to be a compendium of applied techniques, nor a presentation of all the important theoretical developments in these last few centuries. It will be a helpful resource for teachers of mathematical statistics who are looking for an outline of teaching material and useable depth. Their material attains a workable syllabus, which can be easily augmented with the teacher’s preferred emphasis. This volume will make a solid contribution to any theoretical statistics instructor’s collection due to its convenient size, its scope of coverage, judicious use of examples, and clarity of exposition.

John A. Dobelman
Rice University

References


Understanding Advanced Statistical Methods.

This book is designed to replace the prevailing conventional two-course sequence for many academic majors that consists of elementary statistics followed by advanced research methods. Its primary audience is upper-division undergraduates and graduate students from any major area of study. The book emphasizes applications and has a wealth of examples from the social and economic sciences, biological and medical sciences, and physical and engineering sciences. A previous course in statistics is not necessary. The authors claim it can also serve as the text for a lower-division course to satisfy a mathematics general education requirement. The prerequisites are algebra, functions and graphs, and familiarity with spreadsheet software such as Microsoft Excel (the use of dedicated statistical software is encouraged but not required). Calculus is used but is not a prerequisite; the book employs a self-contained “just-in-time” approach to explain mathematical topics such as derivatives and integrals when discussing continuous distributions and expected value, and optimization when discussing maximum likelihood.

The book is similar in approach to Rice (2007) but integrates this approach more intensely throughout the text. The authors’ motivation for writing the book is to remove the gap in understanding that they claim often exists between the first elementary statistics course and the advanced research methods course, with the gap in understanding caused primarily by a formulaic approach to sophisticated topics in the latter course. This book contains just as many formulas as other statistics texts, but with intuitive, engaging, insightful, and irreverent explanations (“We have the famous historical figure R. A. Fisher to blame for the $p < 0.05$ rule of thumb, which we call extremely ugly because it is so overused and abused by researchers”) the authors strive mightily to part the curtain that hides the fundamentals of statistical thinking from most students. The authors take very seriously the word Understanding in the book’s title.

The book has 20 chapters that cover the usual topics, and more, in an undergraduate/graduate math stat text; it is suitable for a fast-paced semester course offered to serious students. The “and more” refers to the strong emphasis throughout the book on thoughtful applications in a wide variety of disciplines. To support this emphasis, each chapter has a generous number of exercises that extend the chapter content and illustrate discipline-specific applications; end-chapter vocabulary lists and formula summaries are also included for each chapter. Chapter 1 is the most important chapter in the book; it explains the statistical science paradigm and the authors’ DATA/data approach, Nature $\rightarrow$ Design and measurement $\rightarrow$ DATA, that contrasts with the more conventional population/sample model. Briefly, DATA include all possible values that could be produced by the process (not the population) being studied, while data denote the values observed in a particular study. While often the surface-level difference appears to be only semantic, there is a real difference resulting from the authors’ model produces data approach as opposed to the traditional data produces model.

The American Statistician, May 2014, Vol. 68, No. 2 123

Downloaded by [Hebrew University] at 04:15 21 May 2014