JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **17**: 415–418 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.675

BOOK REVIEWS

Probability Theory and Statistical Inference: Econometric Modelling with Observational Data, ARIS SPANOS, Cambridge University Press, 1999, pp. 815 + xxviii, Price (pb) $49.95, (hb) $105.00

Although the book claims to be a suitable text for a first course in probability and statistics, Spanos includes in his intended audience students who have had at least a semester course in calculus. The principal problem with this claim is that, as will be shown shortly, there is much that is both omitted and that does not belong in such a text. With regard to the former, Type I and Type II errors, two-sample test of means, simple vs. composite null hypotheses, do not appear. However, the Reimann-Stieltjes integral is included but is nowhere defined, and many integration and differentiation concepts that do not appear in a typical first semester of calculus are frequently used. The above notwithstanding, the entry-requirements for accessing this book are extremely low. Linear algebra and theorem/proof methods are not used. Many theorems are stated and explained, but very few are proved formally given that the emphasis throughout is placed more on ideas and concepts. This is not to suggest that undergraduates, or even first-year graduate students, will find the book an easy read.

The book is a useful companion text, alongside, for example, John Freund, Mathematical Statistics (2001). At the University of Cambridge, UK, the two texts provide the primary references for a second year course in probability and statistics given to economists. Given the more formal exposition in Freund, students appreciate the additional insight afforded by Spanos. In addition, Spanos could be profitably assigned in tandem with Mittelhammer (1996): the former more theoretical and the latter more applied.

Chapter One: An Introduction to Empirical Modelling, introduces basic concepts, including the relation between data and a statistical model. The taxonomy of statistical information (D—dependence, M—memory, H—heterogeneity) provides a useful framework with which to understand the differences between the different types of stochastic processes.

Chapter Two: Probability Theory: a Modelling Framework, introduces basic definitions, such as continuous and discrete random variables, sets and set-theoretic operations, random experiments, events, sigma and Borel-fields, probability set functions, and a statistical space. As a prelude to reading, for example, White (1986) or econometric journals, it is much easier to learn these concepts from the present book than from, say, Shiryaev (1986), and other such texts.

Chapter Three: The Notion of a Probability Model, begins by introducing the concept of a simple random variable, which is carefully related to an event space. Well-worked examples show how to

* Correspondence to: B. D. McCullough, Department of Decision Sciences, Academic Building, Room 230, Drexel University, Philadelphia, PA 19104, USA.

derive the field generated by a random variable. An extensive discussion explains the abstract idea of a space $(S, \mathcal{F}, P(\cdot))$ and how it becomes the probability space induced by the random variable X, $(\mathrm{IR}, \mathcal{B}(\mathrm{IR}), P_X(\cdot))$;* the set function $P_X(\cdot)$ leads to the cdf $F_X(x) = P_X(-\infty, x)$ and finally to the density function defined by

$$F_X(x) = \int_{-\infty}^{x} f_x(u)\,\mathrm{d}u\; f_x(x) \geq 0.$$
$$\forall \quad x \in \mathrm{IR}$$

Spanos' extensive use of graphs also begins in this chapter, with the depiction of numerous pdfs and cdfs. Moments, cumulants, skewness and kurtosis along with measures of location and measures of dispersion are treated in detail. An appendix gives mathematical descriptions (i.e. formulae and moments) of 25 or so of the most common discrete and continuous distributions.

Chapter Four: The Notion of a Random Sample, treats joint, marginal, and conditional distributions (discrete and continuous) identifiability, independence and identicalness. Mixture distributions are also mentioned. The section on Families of Distributions does not simply list Student's-*t*, Laplace, etc., but properly places them in the Pearson family. The Exponential, Pareto-Levy, and Johnson Transformation families are also described.

Chapter Five: Probabilistic Concepts and Real Data, focusses on the graphical analysis of data, including Q-Q and P-P plots. While much of this will not be new to experienced researchers, it is very well presented and will be of great use to the novice. As an example, this is the first time I have seen aspect ratio mentioned in an econometrics text. The aspect ratio of a graph is the length of vertical axis divided by the length of the horizontal axis. Figures 5.1 and 5.2 show how a judicious choice of aspect ratio can turn a graph of seemingly random data into a dataset with a decidedly obvious pattern. (The implied moral is instructive: do not accept the default aspect ratio given by your software; if your software won't let you change the aspect ratio, you've got a problem!)

Chapter Six: The Notion of a Non-random Sample, begins by revisiting the centrality of the assumption, which reduces a complicated joint distribution function to the product of marginal distributions; sequential conditioning is shown to be the means for successfully handling dependence. This is followed by an extended discussion of joint dependence between two (and three) variables, including correlation and moments, conditional correlation and conditional moments. Dependence for categorical variables is also treated. Several copulae are presented, though they are not referred to by that name.

Chapter Seven: Regression and Related Notions, begins by again revisiting the importance of random sampling, and elaborates on the use of sequential conditioning to treat the various types of dependence explored in the previous chapter. The scedastic function is defined, and pairs of regression/scedastic functions and graphs thereof for several types of distributions are presented. An extended discussion of stochastic conditioning gives a very thorough treatment of $E[Y|\sigma(X)]$. where $\sigma(X)$ is the sigma-field generated by X. The chapter concludes with several pages on the biometric tradition in statistics, especially the contributions of Galton and Pearson.

Chapter Eight: Stochastic Processes, extends to the case of several variables the topics of the previous two chapters. Partial sums lead to heuristic discussions of Markov, random walk, independent increment and martingale processes. Dependence restrictions (including mixing and ergodicity) are followed by homogeneity restrictions (including weak and strict stationarity, as well

* $\mathcal{B}(\mathrm{IR})$, a Borel-field, denotes a sigma-field on the real line.

as exchangeability). Extended treatments are then given to the heuristically-discussed processes, as well as Gaussian, Brownian motion, and point processes.

Chapter Nine: Limit Theorems, begins by dispelling five common misconceptions about limit theorems and traces the roots of limit theorems (Bernoulli, de Moivre, Poisson) before proceeding to consider weak laws (seven varieties), strong laws (six varieties), three laws of the iterated logarithm and central limit theorems (more than ten varieties). Sections on functional central limit theorems and modes of convergence end the chapter. Again the basic taxonomy (D, M, H) makes it very easy to see the differences between the various laws and limit theorems, and the tradeoffs involved in appealing to one rather than another.

Chapter Ten: From Probability Theory to Statistical Inference, covers a number of philosophical issues, including definitions of probability. Since both von Mises (via the "collective") and De Finetti (via "exchangeability") attempted to form bridges between probability and observed data, their respective approaches are presented. Spanos attempts to build his own bridge from probability not to data *per se*, but to statistical inference, arguing that the empirical cumulative distribution function represents this bridge.

Chapter Eleven: An Introduction to Statistical Inference, defines basic ideas such as point and interval estimation. Several pairs of concepts are compared and contrasted: Bayesian and classical approaches; observational and experimental data; and specification and misspecification. Sampling distributions, and functions of one and several random variables are treated in detail. The chapter concludes with a discussion of computer-intensive methods for approximating sampling distributions.

Chapter Twelve: Estimation I: Properties of Estimators, covers the usual topics: definition of an estimator, unbiasedness, Cramer-Rao, Consistent Asymptotic Normality, sampling distributions, sufficiency, completeness, etc. It also has a useful description of the jackknife as a method for reducing bias. It is noted with regret that ancillarity is nowhere mentioned.

Chapter Thirteen: Estimation II: Methods of Estimation, considers first the method of moment matching (which Spanos clearly states is not to be confused with the method of moments) and necessarily follows with a discussion of sample moments and functions thereof. Least squares and the method of moments are treated succinctly, with the bulk of the chapter focussing on maximum likelihood.

The remaining two chapters constitute a polemic aimed at critics of the LSE methodology. One tenet of the LSE methodology.[1] He advises the researcher to test multiple hypotheses on the same dataset, under the rubrics of "misspecification testing" and the "general-to-specific approach to specification search". LSE misspecification testing has been called datamining by some, and Spanos offers his rebuttal to this charge. Although the argument is quite interesting, space constraints do not permit recapitulation. However, I am compelled to mention the confluence between these two chapters and the article by Hoover and Perez (1999), who revisit Lovell's (1983) classic datamining article. The authors find that, "[E]mpirical size and power of specifications produced from general-to-specific searches, with one caveat, conform well to the theoretical size and power one would expect if one knew—and knew that one knew—the true specification *a priori*. Test statistics based on such searched specifications therefore bear the conventional interpretation one would ascribe to one shot tests." At the very least, the datamining methodological disputes are not waning. I leave adjudication to the philosophers of science. See the discussion following

[1] See Mizon, 1995 for a description of the LSE approach.

the Hoover and Perez paper and a symposium on the methodology of datamining in *Journal of Economic Methodology* **7**(2), 2000).

Now some quibbles must be raised. There are several typographical errors in the text; the reader should consult the errata list (`www.econ.vt.edu/spanos/book1999`). In addition, the index is vexingly incomplete. For example, Galton is mentioned but not Gosset. Neither "experiment" nor "experimental" is listed. Under "mixed (continuous/discrete) random variables" are two entries, the second of which leads to "mixing conditions". Spanos' subscription to the radical feminist prescription to abolish the use of masculine pronominal forms is marred by the omission of the letter "s" before the word "he" on page 166. There is some inconsistent terminology. On page 45 Spanos eschews the use of "sample space" as a misnomer, preferring "outcomes set". Yet later he labels $\mathcal{X} := \{0, 1\}^n$ as the sample space for a Bernoulli model, and explicitly defines the sample space as "the product of the support of the random variables". In the discussion of ergodic processes, he calls an "assemble" what I know of as an "ensemble".

In summary, Spanos has written a text that clearly explains many of the more difficult statistical concepts underlying the practice of econometrics, and it will be of great use to graduate students. Economists who do empirical work will also find much that is useful. Econometricians concerned with the methodological implications of searching for a model will find the last two chapters either comforting or infuriating.

ACKNOWLEDGEMENTS

Thanks to Michael McAleer and Melvyn Weeks for numerous comments on previous versions.

REFERENCES

Davidson R, MacKinnon J. 1993. Estimation and Inference in Econometrics. Oxford University Press: New York.
Engle R, Hendry DF, Richard J-F. 1983. Exogeneity. *Econometrica* **51**: 277–304.
Gallant AR. 1997. *An Introduction to Econometric Theory*. Princeton University Press: Princeton, NJ.
Greene W. 1999. *Econometric Analysis*, 4th edn. Prentice-Hall: Upper Saddle River, NJ.
Hoover KD, Perez SJ. 1999. Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search. *Econometrics Journal* 167–191, with discussion and reply.
Lovell M. 1983. Data Mining. *Review of Economics and Statistics* **65**: 1–12.
Mittelhammer R. 1996. Mathematical Statistics for Business and Economics. Springer: New York.
Mizon G. 1995. Progressive Modelling of Macroeconomic Time Series: The LSE Methodology. In *Macroeconometrics: Developments, Tensions, and Prospects*, Hoover KD (ed.). Kluwer Publications: Dordrecht; pp. 107–170.
Royden HL. 1968. *Real Analysis*, 2nd edn. Macmillan: New York.
Spanos A. 1986. Statistical Foundations of Econometric Modeling. Cambridge University Press: New York.
Shiryaev A. 1984. Probability. Springer-Verlag: New York.
White H. 1984. Asymptotic Theory for Econometricians. Academic Press: Orlando, FL.

B. D. McCULLOUGH
Department of Decision Sciences
Academic Building, Room 230
Drexel University
Philadelphia
PA 19104, USA

