Studies in Inductive Logic and Probability (Vol. II).

Richard C. Jeffrey (ed.). Berkeley, CA: University of California Press, 1980. 305 pp. $20.00.

This is the second volume of a two-volume set edited by Richard Jeffrey. The main goal of this volume is to present that part of Rudolf Carnap's "A Basic System of Inductive Logic" that did not appear in Volume I. The volume is intended primarily for an audience of philosophers, but as the word "probability" in the title suggests, there is material here for statisticians as well. The problems of inductive logic and probability can be summarized briefly as (a) how one assigns values to one's degree of belief (or credence) in the face of uncertainty and (b) what role learning plays in helping one to reassess these values. To at least some extent, each of the articles in this volume addresses one or both of these questions.

As previously mentioned, Carnap's article (which makes up nearly half of the present volume) is a continuation of an article from Volume I. The article is written in the language of a logician and does not read easily for a statistician. Section 21, however, is somewhat interesting in that Carnap attacks head-on the problem of conditioning on events of zero probability. He leaves some questions open that Douglas Hoover attempts to answer in the last article in the volume, "A Note on Regularity." Hoover claims that Carnap is asking for an application of nonstandard analysis to probability theory of the sort that Parikh and Parnes (1974) and Bernstein and Wattenberg (1969) provide. Stated simply, in nonstandard representations of probability theory all nonempty events have nonzero, but possibly infinitesimal, probability. Since division by infinitesimals is valid, probability can be defined conditional on any nonempty event.

The articles "An Axiomatic Foundation for the Logic of Inductive Generalization," by J. Hintikka and I. Niiniluoto, "A Survey of Inductive Systems," by T. Kuipers, and "The Structure of Probabilities Defined on First-Order Languages," by J. Fenstad are further considerations of Carnap's work and related work by the authors themselves.

Bruno de Finetti's article, "On the Condition of Partial Exchangeability," is accessible to statisticians and is an interesting philosophical piece. In this article, de Finetti offers a middle ground between independence and exchangeability in which events can be deemed partially analogous, but not exchangeable. The influences of partial analogy are stronger than those of independence, but less direct than those of exchangeability. However, a representation theorem of the type de Finetti (1937) proved for exchangeable events also holds in the case of partial exchangeability. Godehard Link, in his article, "Representation Theorems of the de Finetti Type for (Partially) Symmetric Probability Measures," reproves de Finetti's theorems in a more abstract setting. The article by P. Diaconis and D. Freedman entitled "De Finetti's Generalization of Exchangeability" provides numerous examples of partial exchangeability that help to illuminate the discussion of the previous two articles.

The article that is readable by the largest audience is "A Subjectivist's Guide to Objective Chance," by David Lewis. Lewis attempts to distinguish between chance and credence while demonstrating that they both have their roles to play in a system of inductive logic.

Credence is degree of belief and quite subjective, while chance is to be understood as an objective property of the events under consideration. One can hold beliefs about chance just as Bayesians can have distributions over "objective" parameters in an inferential model. Frequency based probability is dismissed with the following remark from page 270, with reference to coin tossing:

> There is no such thing as *the* infinite sequence of outcomes, or *the* limiting frequency of heads that *would* eventuate if some particular coin toss were somehow repeated forever. Rather there are countless sequences, and countless frequencies, that *might* eventuate. . . .

Instructors of elementary probability courses would do well to consider the idea expressed in the above remark before defining probability for their students.

Overall this volume would be of interest to philosophers who are concerned with the foundations of inductive logic, particularly those interested in the work of Carnap. Only a few of the articles would be of interest to statisticians in general.

MARK J. SCHERVISH
Carnegie-Mellon University

REFERENCES

BERNSTEIN, A., and WATTENBERG, F. (1969), "Non-Standard Measure Theory," in *Applications of Model Theory of Algebra, Analysis, and Probability*, ed. W. Luxemburg, New York: Holt, Rinehart and Winston.

DE FINETTI, B. (1937), "La prevision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, 7, 1–68.

PARIKH, R., and PARNES, M. (1974), "Conditional Probabilities and Uniform Sets," in *Victoria Symposium on Nonstandard Analysis*, eds. A. Hurd and P. Loeb, New York: Springer-Verlag, 180–194.

Finite Generalized Markov Programming.

P.J. Weeda. Amsterdam: Mathematisch Centrum, 1979. vii + 127 pp. Dfl. 16, 80 (paperback)

This monograph, the author's doctoral thesis from the University of Amsterdam, presents a thorough analysis of a policy iteration method for the Markov renewal decision model (MRDM). The material is quite technical and of interest primarily to researchers in Markov decision processes. While this book contains several interesting results, a nonstandard problem formulation, an extremely formal presentation, and difficult notation will discourage all but the most enthusiastic reader.

The first five chapters are concerned with the computation of average (or gain) optimal policies for finite state and action MRDM's. The basic model is an uncontrolled process and a set of states in which interventions (or controls) can possibly be applied. The effect of an intervention is to instantaneously transfer the system to another (possibly random) state and accrue a reward to the decision maker. Average optimal policies for this model can be found using the policy iteration algorithm (Howard 1960), an iterative, two-phase procedure involving an evaluation step and an improvement step. The latter is carried out by lexicographically maximizing a pair of vector expressions. In this report, a "cutting operation" augments the improvement step. Its purpose is to obtain an improved policy by possibly decreasing the set of states in which interventions are applied. Modifications of methods of Denardo and Fox (1968) are used to demonstrate convergence of the algorithm.

Chapter 4 shows that the cutting operation is equivalent to solving an optimal stopping problem for a Markov chain and presents methods for its implementation. They are explicit solution of the stopping problem by policy iteration and two forms of approximate solution that avoid matrix inversion. Tests on randomly generated problems and production control problems of practical size show that methods based on approximate solution of the stopping problem lead to reductions in execution time of at least 50 percent when compared to the usual policy iteration algorithm.

In the concluding two chapters, a discounted version of the MRDM is studied. Laurent expansions in two parameters, the discount rate, and the length of time of intervention are used to develop methods for finding policies that are bias optimal in the discount rate and sensitive optimal in the intervention time.

MARTIN L. PUTERMAN
University of British Columbia

REFERENCES

DENARDO, E.V., and FOX, B. (1968), "Multichain Markov Renewal Programs," *SIAM Journal on Applied Mathematics*, 16, 468–487.

HOWARD, R.A. (1960), *Dynamic Programming and Markov Process*, Cambridge, Mass.: The MIT Press.

Correlation and Causation.

David A. Kenny. New York: John Wiley, 1979. xiii + 277 pp. $19.95.

The author of this book holds the view, apparently shared by many social scientists and economists, that causal inference from correlational data (in the absence of controlled experiments) is a valid form of statistical and scientific inference. My view has been that the methods and techniques, developed and applied under that premise, for causal inference (e.g., path analysis) are at best a form of statistical fantasy. Kenny's book failed to convince me otherwise.

The book is meant to be an introduction to causal or structural modeling. The reader is assumed to have no prior knowledge of causal analysis, but familiarity with multiple regression and factor analysis. The book is implicitly organized into four major sections, each with three chapters. Chapter 1 gives an introduction to causal analysis and its role in the social sciences; Chapter 2, "Covariance Algebra," presents a set of elementary rules for manipulating covariances; and Chapter 3 introduces the "Principles of Path Analysis." Classical econometric methods for structural parameter estimation, including multiple regression models, feedback models, and two-stage least squares, are discussed in Chapters 4 to 6. The third section, Chapters 7 to 9, discusses several factor analysis models in relation to causal inference. Chapters 10 through 12 deal with the application of correlational methods to experimental and quasi-experimental designs, with heavy emphasis on "cross-lagged panel correlation analysis" as the unifying concept. The final chapter, titled "Loose Ends," reiterates certain technical and conceptual issues concerning causal modeling.

The author acknowledges certain limitations or questionable practices in the book and offers reasons for their presence, for example,

> To some purists I may seem to be rather careless. First, I have often unnecessarily limited generalization by employing standardized variables. I did so because I did not want to confuse the beginners. . . . Second, at some points the distinction between population value and estimate is blurred. I did so in order not to have a text filled with distracting symbols. . . . I felt that if I had to sacrifice elegance for the experts in order to obtain clarity for the beginner, I would choose clarity. (p. ix)

> Although only introductory, [the book] has taken a rather broad sweep. Some of these topics have been covered rather superficially and tangentially. . . . The text has not employed matrix algebra. Thus, many of the methods discussed have not been adequately explained for the general case. (p. 263)

I personally find those "limitations," as well as the explanations for them, perfectly acceptable and not at all disturbing. In fact, the author is laudable for his clarity of expression and writing style throughout the book. The deliberate sacrifice of some mathematical rigor (or clutter, depending on the reader's view) will indeed make the ideas covered in the book more accessible to a sizable part of the intended audience, the "beginner," the "novice," and certain social scientists who would not find any diet abundant with mathematics or even mathematical symbols very palatable.

Alas, the serious limitations of this book lie not in its lack of mathematical rigor, but in its faulty fundamental logic as well as its faulty presentation and interpretation of certain statistics and statistical methodologies.

The author starts the book with a provocative claim:

> Given the old saying that "correlation does not imply causation," one might wonder whether the stated project of this book—correlational inference—is at all possible. Correlational inference is indeed possible through the application of standard multivariate statistical methods to a stated structural model. (p. 1)

To those of us who still adhere to the "old saying," the author owes an explanation of the logical foundation on which the thesis that "correlation implies causation" is based. It is in this crucial area that the author's thinking appears to be seriously muddled.

From what I can gather from Kenny's exposition about the validity of causal inference from correlation, the key justification lies in the "substantive assumptions of the researcher." Thus, in the case of causal inference from regression analysis (path analysis), if the researcher has reason to believe that X may cause Y, he makes the "substantive assumption" that X causes Y by making the structural model *specification* in terms of a *path diagram* (an arrow pointing from X to Y). Now in the regression analysis phase, if X should be highly correlated (or "significantly" correlated) with Y, then the researcher is supposed to have established causation from the regression study via the statistical significance of the "causal parameter," the standardized regression coefficient (which is the sample correlation between X and Y in this simple case). In this manner, the path analyst is supposed to be able to extract causal information from the data that other statisticians such as myself cannot, simply because of the additional causal assumption (presumably originated from substantive theory) placed on the model.

The logical flaw in the above rationale is that so long as X and Y are "significantly correlated," there is no way for the data to shed light on the truth or falsehood of the "causal assumption," thus making it impossible to validate or invalidate such assumptions; nor is it possible to assess the paucity of the causal conclusion on the basis of statistics or logic.

The logical fallacy underlying path analysis and other forms of causal inference from correlation can be illustrated by the following scenario: A researcher believes that malaria may be caused directly by one's exposure to swampy air ("mal air" or bad air). Having specified his causal assumption by a path diagram, he finds a significant correlation between the incidence of malaria (Y) and the swampiness index (X) of numerous locations sampled in the study. Ergo, the researcher concludes that "mal air" is the direct cause of malaria.

The foregoing example is not atypical of the manner in which causal theories are established by those employing the techniques described in this book. Not infrequently, the causal assumption (theory) is suggested by correlational data, which are then used (tautologically) as if the data were sufficient evidence to confirm the causal theory. Using the path analysis methodology, the researcher can never disconfirm a false causal assumption, regardless of the sample size or evidence, so long as the variables alleged to be causally related are found to be correlated. The fact that some causal conclusions attributed to correlational techniques happened to be true is academic, because the correctness of such conclusions does not diminish the logical flaws inherent in the techniques.

The author's credibility reaches its nadir when he states on page 50 that "Regression coefficients can be interpreted as causal coefficients if certain assumptions are met. These assumptions are the same as those of multiple regression." Having learned the necessary conditions for causal inference, the importance of causal assumptions, path diagrams, causal laws, tracing rules, and so on, we are abruptly informed that we are licensed to practice causal inference via multiple regression if only the regression assumptions are met! I therefore arrive at the inescapable conclusion either that Kenny never did take those other considerations seriously or that he is logically inconsistent.

Causal inference aside, the author seems less than knowledgeable about the topic of regression analysis, on which many of the techniques in this book are based. Consider, for example,

> The assumption of independence can usually be assured by the sampling design of the research. Homoscedasticity and normality of errors are usually not discussed in most social science theories (although they are in some biological theories), but as stated earlier, these are robust assumptions which may only be "approximately" met. (pp. 51–52)

I infer from the above quote that the author is callous about the homoscedasticity and normality assumptions and overly optimistic about the robustness of ordinary least squares against departures from those assumptions.

The statistical sophistication of this book is uniformly low. A conspicuous example of a statistical faux pas is the misinterpretation of hypothesis-testing results. On numerous occasions (e.g., pp. 146, 153, 154), the failure to reject a sharp null hypothesis is interpreted as evidence *supporting* or *confirming* the hypothesis. It is perhaps curious, but not surprising, that of the 63 journal articles cited in the bibliography of this book, only two are from statistical journals, and these two are not related to the subject of causality.

Paradoxical as it may seem, my overall impression of this book is that it is a well-written book of poor substantive quality. By that I mean the author expresses well what he has to say, but much of what is said is neither science nor statistics. Toward the end of the book, even the author himself seems to be casting doubt about the validity of the methodology presented. Contrast the confident statement (quoted earlier) made at the beginning of the book (p. 1) to the following statements made near the end:

> No doubt that after finishing this text the reader has a sense of disillusionment. Before starting the text, it was perhaps hoped that one would learn magical methods to correlations to test theory. . . . Such is not the case. (p. 260)

> To some extent internal validity can be defined as all the threats to causal inference that are controlled for by a true experiment. Thus by definition a correlational study lacks internal validity. It would then seem that the more appropriate goals of a correlational study are convergent and discriminant validity. (p. 261)

To me, these statements are among the most sensible ones made in this book. Unfortunately, they appear to be contradictory to the main theme of the book and are likely to be overlooked or lightly regarded by the reader because of their passing appearance without elaboration in the "Loose Ends" chapter.

In summary, I would not recommend this book to its intended audience, for it will do much more harm than good to the novice. To the sophisticated statistical audience, the logical fallacies and statistical errors, however subtle some of them may be, will probably be obvious and this book may prove to be a useful and entertaining source for them to discover for themselves the logical deficiencies underlying a class of pseudo-black-magic methods.

As for myself, I am much less perturbed by the poor substantive quality of this book than by the fact that we are witnessing the emergence of a subculture of economists and social scientists, who are no more qualified or equipped to practice statistics than law or medicine, yet who nonetheless do practice it among their circles of nonstatisticians, without much visible sign of protest from the community of statisticians. In closing this review, I feel obliged to register my strongest protest against the type of malpractice fostered and promoted by the title and content of this book.

ROBERT F. LING
Clemson University

Scientific Strategies to Save Your Life: A Statistical Approach to Primary Prevention (Vol. 35 of Statistics: Textbooks and Monographs).

Irwin D.J. Bross. New York: Marcel Dekker, 1981. v + 259 pp. $19.75.

Bross introduces himself in Chapter 1 as a statistician whose job is to "save lives." (His Ph.D. thesis was written under W.G. Cochran in 1949 and shortly thereafter Bross was associated with C. Winsor, M. Merrell, and A. Lilienfeld. Since 1959 Bross has been at Roswell Park Institute in Buffalo, N.Y.) According to Bross the main impediment to saving lives is not the difficulty of identifying hazards ("During the 25 . . . years that I've been doing public health research I have made at least half a dozen major scientific discoveries with a substantial and immediate potential for saving hundreds or thousands of lives"), but rather the resistance to necessary measures of special interests. Worse, "the persons who have done the most to block or delay needed action . . . have generally been scientists, university professors, doctors, administrators of prestigious institutions, health bureaucrats and other professionals," not businessmen. Bross therefore seeks to alert the public and to give it the self confidence to challenge the "professionals [who] have been lying . . . for years." On a more practical level Bross's advice is to trust "concerned citizens" and "environmental groups" rather than establishment scientists. Somewhat inconsistently the publisher's press release recommends the book for courses in biostatistics, epidemiology, bioethics, and other areas. The typescript (expensive for this format) text consists of lengthy expositions of Bross's views on the dangers of low-level radiation, with verbatim reproductions of Bross's testimony before government agencies, interspersed with a philosophy of science, science-policy, and public safety that is illustrated by Bross's own research and by the rejection or neglect of that research by scientists and regulatory bodies.

Chapters 2, 4, 6, and 7 are mostly devoted to Bross's testimony on the hazards of low-level ionizing radiation before the Nuclear Regulatory Commission (NRC). Bross and two associates concluded on the basis of a reanalysis of the 1962 Tri-State Survey of leukemia that low-level radiation, of the magnitude used in diagnostic x-rays, was causing 10 times as many leukemias as accepted estimates would allow. According to Bross the dose response curve, which is roughly a straight line at higher dosages, flattens out at low dosages, and at a level well above zero. (Bross's curve, ignoring units and scale, looks something like $y = 1 + \frac{1}{4}x^2$ for $.1 \leq x \leq 1$, $y = \frac{3}{4} + \frac{1}{2}x$ for $x > 1$.)

While Bross's polemics are disturbing, one can sympathize with his predicament, given his passionate belief in the validity and life-saving potential of his work and the strongly negative reaction of other scientists. The depths of Bross's bitterness are most evident in Chapter 11, where he relates that in order to get his Tri-State reanalysis published (Bross 1979), he had to accept the simultaneous publication of a critique (Boice and Land 1979) solicited by the editor from Drs. John Boice and Charles Land of the National Cancer Institute. More damaging than the critique, according to Bross, was an editor's note stating that "Dr. Bross stands virtually alone in defense of his data and the interpretation he places on them." Bross also complains that the editor did not allow him space for an adequate reply to Boice and Land. However, nothing prevented Bross from replying in detail in the volume under review. Instead Bross asks the reader to dismiss Land on the basis of Bross's "tactical" rule: "Once is a critique. Twice is a hatchet job. Three times is a professional hatchet man." While Land has indeed published several negative evaluations of research of which Bross approves, it is equally true that Bross himself has repeatedly attacked, for example, the report of the National Academy of Sciences Committee on the Biological Effects of Ionizing Radiation.

Boice and Land explain the flattening out of Bross's dose response curve at a level above zero as a consequence of Bross's failure to properly allow for the leukemias that arise from causes other than radiation. In Chapter 2, without mentioning this criticism, Bross states that his analysis "has already taken into account the base line level of leukemia," implying that this correction is simple and noncontroversial. However, Bross's figure apparently depends on the model he created to analyze the Tri-State Survey. His estimates of the parameters of this model are severely criticized by Boice and Land for, among other reasons, treating some estimates as known parameters to get confidence intervals for other parameters. According to Boice and Land "conventional case-control analyses do demonstrate an association between diagnostic x-rays and leukemia, . . . [but] the excessive x-rays . . . were possibly administered for pre-leukemic states or early stages of leukemia," an interpretation supported by data from Stewart (1973), and by a later paper (Linos et al. 1980), in which no association between leukemia and diagnostic x-rays was found. In any event, the validity of Bross's dose-response curve cannot be sensibly judged by unprejudiced laymen using only their common sense, contrary to the thrust of Bross's book.

Chapter 3 reprints a published letter of Bross's arguing for his "primacy principle": once a health scientist has produced "prima facie evidence of a hazard . . . the burden of proof shifts to the proponents of a technology to show it is safe." Bross gives a reasonable account of some of the objections to his proposal. Chapter 5 is a discussion of Bross's "Galilean rule: A theory must fit the facts," using as illustrations the Tri-State re-analysis and Bross' co-occurence theory (that leukemia should be associated with certain childhood diseases such as asthma, urticaria, eczema, pneumonia, etc., all of which, according to Bross, can result from radiation-induced DNA damage).

In Chapter 7 Bross discusses rebuttals to his NRC testimony. Bross had used as his y-axis the "proportion affected," while his critics wanted dosage to be plotted against actual leukemias. ("Proportion affected" is Bross's estimate of those who have suffered DNA damage which may eventually be expressed as leukemia and other diseases.) Bross concedes that this theory "necessarily involves some speculations that are not [yet] established by factual evidence." But Bross says that his critics were raising a false issue, since if one divides "proportion affected" by nine, one gets the observed number of leukemias. The two curves have the same shape, and the same general conclusions therefore follow, according to Bross. However one can hardly blame Bross's critics for objecting to damage estimates an order of magnitude larger than what can be observed. The critics did not seem to notice that (as Boice and Land point out) Bross's dose-response curve actually implies that incremental doses of low-level radiation (e.g., x-rays) have *less* impact per rad than larger exposures, and Bross's conclusion that diagnostic x-rays should be curtailed does not follow from his own dose response curve, since this curve is flat in the region of typical x-ray dosages. These NRC hearings were conducted using the "adversary science" format, advocated by Bross.

Chapter 8 is called "Judging a Scientific Argument," and uses Bross's NRC presentation as a model. Chapter 9, "Some Practical Strategies in Biostatistics," is supposed to help the reader to "distinguish between the handful of real experts in any area and the swarm of phony experts. . . ." Bross, who betrays no lack of confidence in his own research, advises the reader that "the more authoritatively the opinion is asserted, the less credence you should put in it." Chapter 10 contains his criticism of academic statisticians who worry too much about optimal estimation (". . . a way to bolster the egos of middle aged and deteriorating males"). In Chapter 11 ("Politics and Science") Bross returns to a frequent theme—the "repression" of and "disincentives" for positive findings concerning the effects of low-level radiation; for example, loss of funding and blocking of publication.

Chapter 12 ("How to Tell Washington the Truth") appeared in *The American Statistician* (Bross 1980) under a slightly different title. Bross