Extending Kemeny and Snell's (1960, pp. 161–167) Markov chain analysis of tennis, Morris determines the importance of winning various points in a tennis game. He defines the importance of a point to be the probability the server wins the game if the point is won minus the probability the server wins the game if the point is lost, and he shows that if the players are of nearly equal ability, the most important point is 30–40. Morris concludes that by trying a little harder on the important points the player will greatly enhance his or her prospects of victory, but he notes, "If a player doesn't try on the unimportant points in his matches, there never will be any important ones" (p. 139).

Two articles were of special interest. One was by Porter, in which he uses expected utility calculations to show that many college football coaches use a dominated extra-point strategy, and another was by Brearley, in which he uses elementary physical principles to prove that Bob Beaman's world-record long jump in the 1968 Olympics was a truly exceptional performance and not the result of the reduced air resistance of Mexico City.

The analysis in the article by Ladany, Humes, and Sphicas was perhaps the most innovative and extensive. The problem considered is that of determining the optimal aiming point for takeoff by a competitor in the long jump. They first develop a complicated expression for the expected length of the jump as a function of the aiming line. Then they approximate this expression through the use of some classical (and obscure) probability estimates and maximize it by using a search technique. A parallel analysis using simulation also is carried out. It is encouraging that the results of these two analyses are in agreement. The assumptions of the model are checked on log-jump data, and the sensitivity of the solutions to the model parameters is analyzed. This article exemplifies sound model building and problem solving, as does the article by Pollock on the fairness of the golf handicap system.

Books such as this, which compile articles from a wide variety of sources, sometimes lack cohesiveness. This is not the case in *Optimal Strategies in Sports* because of the fine editorial work of Ladany and Machol. The articles are well integrated and cross-referenced. Machol's contribution appears significant, and his help is acknowledged in many of the articles in this volume. Another worthwhile feature of this book is the extensive annotated bibliography on mathematical and statistical analysis of sports.

The weak points of this book are few. There are perhaps too many articles on baseball, some of which are redundant and uninspiring. The analyses rely heavily on simulation and descriptive statistics and contribute little in the way of interesting methodology. Also, too few of the applied models in this book actually have been applied. It would have been interesting to find out the results of using some of the suggested strategies.

On the whole, this book could provide valuable motivational material for introductory courses in probability and statistics or management science. Many of the articles illustrate the design and use of elementary probability calculations, yet omit many easy steps that could be completed by the student. The student also could be encouraged to repeat the analyses using new data, to extend some of the results to other sporting events, or even to try out some of the suggested strategies.

MARTIN L. PUTERMAN
University of British Columbia

REFERENCES

Kemeny, J.G., and Snell, J.L. (1960), *Finite Markov Chains*, Princeton, N.J.: D. Van Nostrand Co.

Machol, R.E., Ladany, S.P., and Morrison, D.G. (eds.) (1976), *Management Science in Sports* (Vol. 4), Amsterdam: North-Holland/TIMS Studies in the Management Sciences, 1976.

Inference, Method, and Decision: Towards a Bayesian Philosophy of Science.

Roger D. Rosenkrantz. Boston: D. Reidel Publishing Co., 1977. xv + 262 pp. $26.00 ($11.95 paperback).

Having watched the recent advances of Bayesianity in statistics, it is interesting to see it now moving into another field—philosophy of science—with just the same effects. Non-Bayesian statistics was a collection of separate principles (chi-squared, maximum likelihood, sufficiency, efficiency, bias, minimum variance, shortest confidence interval, fiducial distribution, ancillarity, size and power functions, etc.) each of which had, in itself, an evident intuitive appeal. But these principles were not entirely consistent with one another, and in most cases experience in using them led eventually to recognition of defects or limitations on generality.

Into this situation, Bayesian theory moved as a corrective and unifying force. Needless to say, each of the aforementioned principles contained important elements of truth, or else it could not have appeared reasonable to a reasonable person. But except for the principle of sufficiency, none was free from qualifications. It is now clear that whatever elements of truth were in those principles are retained in Bayesian methods. But Bayesian theory also determines the exact bounds of validity of those principles and shows how they must be modified in cases that overstep them (i.e., presence of nuisance parameters, nonexistence of sufficient or ancillary statistics, prior information, or crucial importance for the inference).

The result was a great increase in the power and scope of statistical methods. Equally important, Bayesian methods appear simpler and easier to apply than the mass of "ad hockeries" that they replaced.

Now we see this story being repeated in philosophy of science, also concerned for the most part with problems of plausible inference. Also here, therefore, if one does not recognize the validity or general applicability of Bayes's theorem as the basic principle of inference, then all sorts of gratuitous, ad hoc notions and principles have to be invented to take its place. And so philosophy of science developed its own list of them: simplicity, induction, confirmation, corroboration, support, falsifiability, sample coverage, and so on. Again, each of these contains some truth, but also has limitations; again, Bayesian analysis makes clear at once where those limitations lie and how the original principles need to be modified in cases that overstep them.

It was, perhaps, inadvisable for me to review a book, one of whose stated purposes was the exposition of my own results on determination of priors. Having, so to speak, composed the libretto for this work, I knew in advance how the plot was going to develop and just what characters would be left dead on the stage at the final curtain. But this makes me uniquely incapable of judging what effect the book may have on others. Accordingly, this review deals in the briefest terms with the parts of the book devoted to expounding my work.

Rosenkrantz is one of the very few who have taken the trouble to understand the principles of maximum entropy and transformation groups before breaking into print with commentaries on them, and he has corrected some of the more grotesque misunderstandings of others (e.g., see the discussion of Benford's law on p. 78). In some respects, Rosenkrantz has understood what I was driving at almost better than anyone else, including some of my former students. Departures from the original libretto are generally too trivial to mention (I noted only one of substance: On p. 74, a distribution for the Bertrand problem is stated to be "obvious by rotational symmetry" when in fact it also requires scale invariance).

Perhaps the flavor of the work can be conveyed most clearly and usefully by considering a single example of the confrontation of Bayesianity with one of those "ad hockeries." Chapter 5 examines the notion of simplicity (the famous Ockham's Razor) as a criterion for deciding between hypotheses. Of course, the primary requirement of our hypothesis is that it should fit the facts; but given two hypotheses that do so equally well, Ockham says that we should prefer the simpler, "Entities are not to be multiplied without necessity." As an abstract generality, intuition assents at once. But this only set the stage for centuries of discussion over precisely what is meant by simplicity; so inconclusive that, as the author notes (p. 115), for a time simplicity was given up for dead, relegated to the limbo of the merely aesthetic. Let us, however, examine only the most obvious example: that a hypothesis with fewer parameters seems intuitively simpler. It is interesting to see the mechanism by which Bayes's theorem usually justifies—but in some cases modifies—this intuition.

Given any hypothesis H involving parameters θ and prior information I about them, the predictive probability of observing data D is $p(D|H) = \int p(D|H, \theta)p(d\theta|I)$, the average likelihood over the prior distribution. Denote by H_n a hypothesis for which $\theta \equiv \{\theta_1, \ldots, \theta_n\}$ is n dimensional, ranging over a parameter space S_n. Now introduce a new model H_{n+1} by adding a parameter θ_{n+1} and going to a new parameter space S_{n+1}, in such a way that $\theta_{n+1} = 0$ represents the old model H_n. Then, on the subspace S_n, the likelihood is unchanged. But the prior probability $p(d\theta|I)$ must now be spread over a larger parameter space than before and will therefore, in general, assign a lower probability to a neighborhood of a point in

S_n than did the old model (this might be stated more precisely, but the previous should suffice to make the point). For a reasonably informative experiment, we expect that the likelihood will be rather strongly concentrated in small subregions $R_n \subset S_n$, $R_{n+1} \subset S_{n+1}$ of the respective parameter spaces. Therefore, if the maximum likelihood point occurs at or near $\theta_{n+1} = 0$, R_{n+1} will be assigned less prior probability than R_n, and we shall have $p(D|H_{n+1}) < p(D|H_n)$; the log-likelihood generated by the data will favor H_n over H_{n+1}.

Thus, if the old model already is flexible enough to accommodate the data, then, as a general rule, Bayes's theorem will, like Ockham, tell us to prefer the intuitively simpler hypothesis. Generally, the inequality will go the other way only if the maximum likelihood point is far from $\theta_{n+1} = 0$ (i.e., a significance test would indicate a need for the new parameter), because then the likelihood will be so much smaller on R_n than on R_{n+1} that it will more than compensate the lower prior probability of the latter; as noted, Ockham would not disagree.

But, having seen this mechanism, it is easy to invent cases (e.g., if introduction of the new parameter is accompanied by a drastic redistribution of marginal prior probability on the subspace of the old model) in which Bayes's theorem will contradict Ockham because it is taking into account further circumstances undreamt of in Ockham's philosophy.

This example is relevant for a much-discussed current topic in statistics. Bayesian methods enjoy some well-established optimality properties, but those advantages can be lost in practice if we try to apply them beyond the ambit of what follows by rigorous mathematics from the basic rules of probability theory. Because those rules refer to normalized distributions, they cannot justify the use of an improper prior except when it is approached as the limit of sequence $\{p_i\}$ of proper priors, and when the corresponding sequence $\{P_i\}$ of posterior distributions proves to have a mathematically well-behaved limit. All sorts of artificially created paradoxes can ensue if one tries to use an improper prior directly without first investigating this limiting operation.

Thus, in our elucidation of Ockham's rule, everything depended on the lowered prior density caused by introducing the new parameter θ_{n+1}. Our conclusions depend crucially on the prior information as it affects the ratio of prior probabilities of the high-likelihood regions R_n, R_{n+1}; we may pass to the limit of improper priors only if this is done in a way that preserves that ratio. If we tried to insert uniform improper priors directly, the ratio would appear indeterminate (0/0) and meaningless. Likewise, the marginalization paradox is discussed from this point of view in *Studies in Bayesian Econometrics* (Zellner 1979).

It would be interesting to see how Ockham's principle could be justified in orthodox statistical theory, which does not admit a prior probability for a hypothesis. Even in Bayesian theory, the question is subtle enough to have caused trouble.

Undoubtedly, the book under review represents only the first step of Bayesianity into modern philosophy of science, as did that of L.J. Savage (1954) for 20th-century statistics. That is, in both cases, the works have this status only because others chose to ignore the work of Sir Harold Jeffreys, who had said so much of it clearly and correctly, in perhaps less generality but greater depth, many years before. Likewise, in both cases, the inhibitions of previous training prevented the authors from taking more than the first cautious step in the direction pointed out by Jeffreys.

Savage was unable to follow Jeffreys to the conclusion that prior probabilities should, in principle, be determined by logical analysis of the prior information; and so he made no contribution to the problem on which many Bayesians, including this reviewer, have devoted major efforts: how to carry out that logical analysis explicitly. Likewise, Rosenkrantz does not seem to advance to the logical conclusion that Bayesian analysis not only clarifies the aforementioned "ad hockeries" but also supersedes them and removes any need to discuss them at all. Presumably, this development will appear in a later work by Rosenkrantz or one of his readers. At this point, philosophy of science may start to be of use to the practicing scientist, but at the cost of losing its identity and becoming merely a branch of statistical inference.

EDWIN T. JAYNES
Washington University

REFERENCE

Zellner, Arnold (ed.) (1979), *Studies in Bayesian Econometrics*, Boston: D. Reidel Publishing Co.

PUBLICATIONS RECEIVED

Annales de l'Institut National de la Statistique et des Etudes Economiques. *The Econometrics of Panel Data* (Colloque International du CNRS, INSEE, Paris, 22–24 August 1977). Paris: INSEE, 1978. 747 pp. $18.00.

Bartholomew, David J., and Forbes, Andrew F. *Statistical Techniques for Manpower Planning*. New York: John Wiley & Sons, 1979. xiii + 288 pp. $37.50.

Brooks, Camilla A., and Bailar, Barbara A. *An Error Profile: Employment As Measured by the Current Population Survey* (U.S. Dept. of Commerce Statistical Working Paper No. 3). Washington, D.C.: U.S. Dept. of Commerce, 1978. xix + 84 pp. Free.

Duncan, Joseph W. (ed.). *Statistical Services in Ten Years' Time*. Oxford: Pergamon Press, 1978. viii + 187 pp. $30.00.

Engen, S. *Stochastic Abundance Models*. New York: Halsted Press, 1978. viii + 126 pp. $13.95.

Finkelstein, Michael O. *Quantitative Methods in Law*. New York: Free Press, 1978. ix + 318 pp. $17.95.

Finn, Jeremy, and Mattsson, Ingrid. *Multivariate Analysis in Educational Research: Applications of the Multivariance Program*. Chicago: National Educational Resources, 1978. viii + 138 pp. $12.50 (paperback).

Fishman, George S. *Principles of Discrete Event Simulation*. New York: John Wiley & Sons, 1978. xviii + 514 pp. $24.50.

Frankel, Theodore. *Gravitational Curvature: An Introduction to Einstein's Theory*. San Francisco: W.H. Freeman and Co., 1979. xviii + 172 pp. $18.50 ($8.95 paperback).

Freund, John A. *Modern Elementary Statistics* (5th ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1979. xv + 510 pp. $15.95.

Garey, Michael R., and Johnson, David S. *Computers and Intractability: A Guide to the Theory of* NP-*Completeness*. San Francisco: W.H. Freeman and Co., 1979. x + 338 pp. $18.50 ($10.00 paperback).

Gill, John L. *Design and Analysis of Experiments in the Animal and Medical Sciences* (Vols. 1, 2, and 3). Ames, Iowa: Iowa State University Press, 1979. Vol. 1: xiii + 409 pp. $14.00. Vol. 2: x + 300 pp. $11.00. Vol. 3: viii + 173 pp. $7.00 (paperback).

Haberman, Shelby J. *Analysis of Qualitative Data, Volume 1: Introductory Topics*. New York: Academic Press, 1978. xiii + 368 pp. $19.50.

Hax, Arnoldo C. (ed.). *Studies in Operations Management*. Amsterdam: North-Holland Publishing Co.; New York: Elsevier/North-Holland Publishing Co., 1978. viii + 580 pp. $71.00.

Jacobs, Harold R. *Elementary Algebra*. San Francisco: W.H. Freeman and Co., 1979. x + 872 pp. $12.00.

Kruskal, William H., and Tanur, Judith M. (eds.). *International Encyclopedia of Statistics* (Vols. 1 and 2). New York: Free Press, 1978. xxi + 1349 pp. $100.00 (set).

Leaverton, Paul E. *A Review of Biostatistics, A Program for Self-Instruction* (2nd ed.). Boston: Little, Brown and Co., 1978. xv + 92 pp. $6.50 (paperback).

Levine, Raphael, and Tribus, Myron (eds.). *The Maximum Entropy Formalism*. Cambridge, Mass.: MIT Press, 1978. xi + 498 pp. $22.50.

Lorrain, Paul, and Corson, Dale R. *Electromagnetism: Principles and Applications*. San Francisco: W.H. Freeman and Co., 1979. xiv + 507 pp. $12.50 (paperback).

Mansfield, Ralph. *Introduction to Technical Mathematics*. San Francisco: Harper & Row, 1978. xi + 512 pp. $15.95.

Miller, M. Clinton (ed.). *Mainland's Elementary Medical Statistics*. Ann Arbor, Mich.: University Microfilms International, 1978. xiii + 381 pp. $25.50.

Namboodiri, N. Krishnan (ed.). *Survey Sampling and Measurement*. New York: Academic Press, 1978. xxi + 364 pp. $23.00.

National Center for Health Statistics. *Acute Conditions: Incidence*