From *SIAM News*, Volume 37, Number 2, March 2004

The "Only Acceptable Approach" to Probabilistic Reasoning

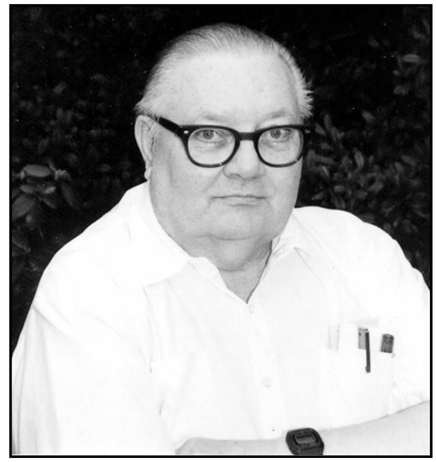
Probability Theory: The Logic of Science. *By E.T. Jaynes, Cambridge University Press, Cambridge, UK, 2003, 758 pages, $60.*

BOOK REVIEW

By Terrence L. Fine

Probabilistic reasoning includes both probability theory and the methodology (statistics) by which we choose specific probability models and apply them. A (not the!) concept of probability requires a stand on its meaning and domain of applicability (chance, uncertainty, indeterminacy), with these two then providing the basis for a suitable mathematical axiomatization.

The different types of meaning that have been held for probability divide as objective and subjective. Objective meanings relate probability to: idealized frequencies of occurrences of events in long runs of repeated, causally unlinked repetitions of a random experiment; propensities for occurrences of events that relate to frequencies through laws of large numbers; and epistemic or knowledge-centered ideas of the degree of inductive support that one proposition provides to another. Subjective meanings, degrees of personal belief about the occurrences of events, have become quite popular and find strong proponents among Bayesian statisticians.



E.T. Jaynes, who died in 1998, devoted forty years to the development of a knowledge-based, epistemic probability that is well known for its espousal of the maximum entropy principle to convert prior information into prior probability distributions.

The usual mathematical choice for probability is that it takes values in the unit interval, satisfies an axiom of additivity for disjoint events or propositions, and may also have certain limiting properties (e.g., σ-additivity or monotone continuity on the standard Kolmogorov account) when infinitely many events are involved. In everyday life we all recognize non-numerical concepts of probability, expressed in such terms as "probably," "likely," "highly likely," "at least as probable as." While almost all agree on probability as a numerical quantity, some, such as myself, believe that we need a wider range of mathematical concepts. There is no consensus as to the meaning of probability. My belief is that all of these accounts of meaning are viable within appropriate domains of application.

The Author

Edwin T. Jaynes, who died in 1998, was a scholar of strong convictions, strongly expressed, sometimes without courtesy for those who disagreed with him. He was also a scholar who devoted himself steadfastly for forty years to the development of a knowledge-based, epistemic, or inductive probability that is well known for its espousal of the maximum entropy principle to convert prior information into prior probability distributions.

His outlook was shaped significantly by his origins as a physicist. Jaynes's position, on his own account, finds its roots in Laplace, in the thought and work of Harold Jeffreys on Bayes methods and invariant priors, in R.T. Cox's derivation of the Bayes formula from a functional equation, and in George Pólya, particularly his *Mathematics and Plausible Reasoning*. Jaynes's death left the completion of the treatise under review to G. Larry Bretthorst, who well executed the sizable task he explains in an editor's foreword.

Jaynes (p. xxii) provides his readers with a clear sketch of his attitudes and goals. First, he addresses mathematical subtleties, particularly those involving infinite sets, when he says that "We sail under the banner of Gauss, Kronecker, and Poincaré rather than Cantor, Hilbert, and Bourbaki." He approves of Gauss's statement (p. 451) that "Infinity is merely a figure of speech, the true meaning being a limit." As Jaynes remarks (p. 43), "In principle, every problem must start with such finite-set probabilities; extension to infinite sets is permitted only when this is the result of a well-defined and well-behaved limiting process from a finite set." He repeats his insistence on this approach at many points, including in his resolution of paradoxes in Chapter 15 and in Appendix B.

The ideological stance of this treatise is made clear in Jaynes's resolution of the "controversy over 'frequentist' versus 'Bayesian' methods of inference" when he avers that

"There was a strong tendency, on both sides, to argue on the level of philosophy or ideology . . . there is no longer any need to appeal to such arguments. We are now in possession of proven theorems and masses of worked out numerical examples. As a result, the superiority of Bayesian methods is now a thoroughly demonstrated fact in a hundred different areas. One can argue with philosophy; it is not so easy to argue with a computer printout, which says to us: 'Independently of all your philosophy, here are the facts of actual performance.'"

Can mathematical and philosophical concerns be so easily dissolved by examples and computer printouts? Jaynes has little trouble resolving problems that others continue to wrestle with and in doing so shows little respect for their difficulties: On page 22, ". . . and the attempts of physicists to explain quantum theory, are reduced to nonsense by the author falling repeatedly into the mind projection fallacy ['one's own private thoughts and sensations are realities existing externally in Nature']"; on page 23, "A compulsive nitpicker has complained to us . . ."; and on page 674, "On the other hand, to express grave doubts about simple matters that are obvious is the equally standard technique for imputing to one's self deep critical faculties not possessed by others." Jaynes raises many issues about the meaning and proper mathematical formulation of probability; he resolves many of them to his own satisfaction, but not always to the satisfaction of myself and others.

The Book

Probability Theory: The Logic of Science is a detailed treatment of probabilistic reasoning that goes well beyond the elements of probability theory to discuss at length, and through a wide variety of examples, the applications of probability to reasoning about both everyday real-world events and questions in the testing of scientific laws and modeling measurements. Jaynes emphasizes an understanding of the principles of his position, with this understanding built upon detailed discussion of this large number of applications. He takes admirable pains to develop most of the theoretical material as a response to the challenges of particular problems, often drawn from physical science. The mathematical prerequisites are non-measure-theoretic undergraduate calculus.

Chapters 1 and 2 provide the basics of probability. Although Chapter 2 treats only the case of finite sets of propositions, Jaynes asserts (p. 107), "But this is all we ever need in practice." Nonetheless, he makes much use of calculus and continuous distributions, and provides information throughout the remainder of the treatise about particular discrete distributions and continuous distributions. Strongly motivated by his concern with prior information expressed as propositions, he focuses on assigning probability to propositions that are either true or false rather than to sets. As should be expected in a book of more than 750 pages, there are occasional errors; a strange one (pp. 651,652) gives the defining property of a σ-field as closure under complementation, while including closure under countable unions.

The viewpoint of the epistemic Bayesian is to objectively transform all prior information into objective priors. The maximum entropy and invariance principles of Chapters 11 and 12 provide means to effect the transformation of prior information into prior probabilities. This Bayesian viewpoint is systematically applied through detailed discussion of such statistical ideas as tests of significance, hypothesis testing, parameter estimation, decision theory, complete ignorance priors, model selection, and through extensive criticism, generally in the context of examples, of standard statistical methods that Jaynes holds to be unnecessarily ad hoc. This treatise is extraordinarily wide-ranging in its coverage, with contents including lengthy examinations of ordinary experiences, mathematical foundations, particular applications, long-running statistical and philosophical disputes.

Appendix A compares Jaynes's approach to probability with the better-known ones of Kolmogorov and de Finetti and insists upon the comparability of all pairs of conditional propositions. Appendix B, on mathematical formalities and style, will likely draw both sympathetic and antipathetic reactions from readers of *SIAM News* to his positions on measure theory, applied mathematics, mathematics, and their public exposition through an overemphasis on rigorous statement at the expense of clarity and applicability. As he notes on page 674, "On the contrary, much experience teaches us that the more one concentrates on the appearance of mathematical rigor, the less attention one pays to the validity of the premises in the real world, and the more likely one is to reach final conclusions that are absurdly wrong in the real world."

As Bretthorst explains, the References are works cited in the text, while the Bibliography contains other published sources of value. Both lists were drawn from an annotated list created by Jaynes. Often, his summary comments are judgmental, and keyed on whether the book in question agreed with or was a forerunner of his own positions or was otherwise in the wrong. Jaynes (p. 719) dismisses the cogent and detailed arguments developed by Peter Walley in support of a concept of imprecise subjective probability that characterizes knowledge by a set of measures, and not by its individual elements; in Jaynes's misstatement, it is as if imprecise probability amounted to a choice of measures to be used to represent prior knowledge. Nonetheless, I welcome the immediate presence of its author in *Probability Theory: The Logic of Science*. It is not written from the disembodied point of view of much of mathematical and scientific writing.

Jaynes's Concept of Probability

Jaynes expresses (p. xxii) his view of "probability theory as extended logic." More specifically, "the mathematical rules of probability theory . . . are also the unique consistent rules for conducting inference (i.e., plausible reasoning) of any kind, and we shall apply them in full generality to that end." In this he has forebears in J.M. Keynes, B.O. Koopman, and R. Carnap. His viewpoint is intended to span the applications of both Bayesian and frequentist methods, and he has an unusually broad view of the applicability of his theory of probability. In brief summary, the key elements of Jaynes's position are:

- Probability is conditional upon information.
- This information is expressible in the form of a true-or-false proposition.
- The collection of such propositions forms a Boolean algebra.
- Probability is fundamentally conditional and represents the logical plausibility of a pair of propositions thought of as one proposition conditional upon (given the truth of) another.
- Probability is necessarily numerically valued.
- A unique function for updating, given by the familiar Bayes rule, is the only function that satisfies informal desiderata of correspondence with common sense and maintenance of consistency (an argument first developed by R.T. Cox).
- The prior distribution required by the Bayes rule can be determined objectively through recourse to such as the maximum entropy principle.
- The resulting concept of probability is broadly applicable both to technical or scientific applications and to issues (even "weird" ones) arising

in daily life.

■ Probabilistic reasoning based upon this process is systematic, avoids the "ad hockery" of much of statistical practice, and works very well in many real-world applications.

■ The above is the only acceptable approach to probabilistic reasoning.

Cox's argument (see Chapter 2) for obtaining a quantitative representation for qualitative or comparative axioms about conditioning, a goal essential to Jaynes's program, has been re-examined carefully by Halpern (J.Y. Halpern, "Cox's Theorem Revisited," *Journal of Artificial Intelligence Research*, Vol. 11, 1999, pp. 429–435). When we restrict ourselves to finitely many propositions, as advocated by Jaynes, some technicalities must be addressed if we are to reach the conclusion desired by Jaynes; in my view and Halpern's, the assumptions required are not as compelling as glossing over them makes them appear.

Jaynes notes (p. 372), "The problem of translating prior information uniquely into a prior probability assignment represents the as yet unfinished half of probability theory, though the principle of maximum entropy in the preceding chapter [11] provides one important tool." He goes on to state (on the next page), "The prior probabilities represent our prior *information*, and are to be determined, not by introspection, but by *logical analysis* of that information."

Recommendations

While Jaynes offers his book as both a text and a reference, I cannot recommend it for a first course in probability. He anticipates that his reader has had sufficient prior exposure to probability concepts (including Markov processes) to permit their abrupt introduction; as he notes on page 66, "We can hardly suppose that the reader is not already familiar with the idea of expectation. . . ." Important subjects like probability descriptions through probability mass functions, density functions, cumulative distribution functions, characteristic functions, generating functions, and concepts like expectation and independence are presented in passing and far too briefly. Others, such as the binomial distribution, are developed at length from an unexpected direction.

The myriad issues this treatise raises, and the provocative and sometimes dismissive answers given to many of them, make it beneficial to a reader who has a traditional grounding and an open mind. Such a reader will have the background needed to sort the wheat from the chaff and will find Jaynes's work thought-provoking. This is also a profitably browsable book (e.g., Chapter 7 on various origins of the Gaussian) for a person with a background in probability and its applications.

Jaynes (p. 74) classifies himself with the agnostics: "We agnostics often envy the True Believer, who thus acquires so easily that sense of security which is forever denied to us." Yet he sees things too clearly, even in the murkiest recesses, to be other than a "True Believer." At the present state of evolution of probabilistic reasoning, we need to accept a collection of approaches, objective and subjective, numerical and comparative or qualitative, each having only a roughly recognized domain of usefulness. Jaynes is convincing that his efforts, well described in *Probability Theory: The Logic of Science*, yield an approach that belongs in this collection.

Terrence L. Fine is a professor in the School of Electrical and Computer Engineering and director of the Center for Applied Mathematics at Cornell University.