Book Review in: *Foundations of Physics*, Vol. 14, No. 2. 1984.

E. T. Jaynes: Papers on Probability, Statistics, and Statistical Physics. R. D. Rosenkrantz, ed., D. Reidel [Kluwer], Dordrecht, The Netherlands, 1983, xxiv + 434 pp.

David Hestenes

If I were asked to recommend a single book which every physicist should own and study, this book of collected articles by Edwin T. Jaynes would be that book. I would extend that recommendation to any scientist with sufficient mathematical sophistication. And I would not exclude mathematicians who care about the application of mathematics to the real world.

The serious student of this book will discover a coherent theory of statistical inference of astounding simplicity and power, applicable in every theoretical and experimental domain. The likes of it cannot be found anywhere else. He will be instructed by many examples of masterful qualitative reasoning in support of cogent mathematical analysis. And he will be rewarded with unique insights into important unsolved problems, worthy objects for his own intellectual efforts.

This book chronicles a conceptual revolution beginning with Jaynes' papers on statistical mechanics in 1957. Even in conventional textbooks today, statistical mechanics is a jumble of physical and statistical arguments. Inspired by Shannon's seminal papers on information theory, Jaynes learned how to separate the physics from the statistics. In doing so, he discovered a powerful general principle of statistical inference, now known as the MAXENT principle. This principle tells us how to incorporate any statistical information about state variables of some system into a probability distribution $\{p_k\}$ over the possible states $\{k\}$ of the system. It tells us to choose that distribution which has the largest information entropy $S_I = -\sum_k p_k \log p_k$ consistent with the given statistical information. In other words, maximize the entropy subject to the given statistical constraints.

The marvel of MAXENT is its power and simplicity. It has enabled Jaynes to reformulate statistical mechanics as a theory of statistical prediction from physical laws, the most elaborate and successful application of a general theory of statistical inference that exists. This has purged statistical mechanics of *ad hoc* and irrelevant arguments, such as ergodic arguments to justify the equilibrium theory, and coarse-graining arguments in nonequilibrium theory. It has eliminated serious misconceptions as the belief that Gibbs' canonical distribution describes equilibrium states only. It reveals that there are several different entropy concepts which have been perennially confused in the literature. In particular, the *information entropy* S_I, must be distinguished from the *experimental entropy* S_E which appears in thermodynamics. Jaynes shows that this distinction is just what one needs to prove the second law of thermodynamics. His proof is so clear and simple that even students can understand it.

The record reveals that Jaynes himself has been repeatedly suprised by the power of MAXENT. It took him many years to realize that MAXENT contains all that is needed to generalize equilibrium statistical mechanics to a nonequilibrium theory. The essentials of the nonequilibrium theory are set forth in two of the later articles in this book, but the formulation is so sketchy that the neophyte is not likely to realize that a full-fledged theory has been specified. Unfortunately, many details and applications worked out by Jaynes and his students are still unpublished. But the serious student can learn what he needs from the review article by W. T. Grandy, Jr. [*Phys. Rep.* **62**, No. 3, 175–266 (1980)]. The editor should have included this reference in his supplementary bibliography, which is rather thin on physical applications.

To this day, most specialists in statistical mechanics and thermodynamics have summarily dismissed or overlooked Jaynes' contribution. Hopefully, this book will catch the attention of those who are still looking for a simple and coherent perspective on the field. Jaynes' approach has not yet produced any striking new predictions to prove empirical superiority over its competitors, though some tantalizing possibilities in nonequilibrium theory are waiting to be pursued. In the meantime, abundant evidence for Jaynes' general theory of inference has been rapidly accumulating in other areas.

With his development of MAXENT, Jaynes has been able to abstract from statistical mechanics the powerful statistical methods developed by physicists and integrate them into a general theory which can applied to statistical problems in any field. For many years, only a few applications of MAXENT were known, worked out mainly by Jaynes and his students. But within the last decade the trickle of applications has grown to a torrent—applications by astronomers, chemists, geophysicists, electrical engineers, and economists, among others. The supplementary bibliography to this book points to a few of these. More will be found in the forthcoming volume, *Maximum-Entropy and Bayesian Methods in Inverse Problems*, to be published by Reidel [now called Kluwer].

Naturally, the multitude of new applications has brought new theoretical issues to the fore. Jaynes has been in the midst of these issues, and now, with important contributions from others, powerful methods for solving a large class of inverse problems are emerging. Physicists are just beginning to realize the power of MAXENT in inverse problems such as spectral analysis and inferring crystallographic structure or an interaction potential from scattering data. It seems safe to predict that MAXENT will soon return to physics with a wide range of applications whose conceptual kinships to statistical mechanics will ultimately be widely recognized.

Many scientists and engineers have recently become convinced of the value of MAXENT by its evident superiority over other methods in practical applications. Still, to be impressed by practical success is not to understand, so some of them are beset by nagging doubts and uncertainties. This book has what they need to understand the rationale for MAXENT and related statistical methods.

The simplicity of MAXENT is deceptive. It is simple to state and use, but its justification lies deep in the foundations of probability and information theory. Consequently, it is open to all sorts of criticisms stemming from fundamental misconceptions. Some who see it merely as an optimization principle dismiss it as trivial; after all, they already knew how to optimize with Lagrange multipliers. Others dismiss it as "too good to be true," because they have not understood why it works. Jaynes has taken the trouble to answer all serious criticisms of MAXENT at great length. This has forced him into the deep analysis of the foundations of probability theory found in this book.

To understand MAXENT, one must first learn to interpret probabilities as descriptions of knowledge states rather than states of nature. This has been a great stumbling block to many who have been taught that every probability must be interpreted as a "frequency in some random experiment." Jaynes handles this problem with incisive explications of the probability concept, and by proving a number of theorems specifying precise relations between probabilities and frequencies.

To understand MAXENT, one must also know why it works. That is a long story patiently expounded by Jaynes. It demands an analysis of the properties of information entropy to explain why it is the only function producing an unbiased probability distribution when it is optimized. It calls for entropy concentration theorems to prove that MAXENT picks out states which are most likely to be realized in nature.

Finally, to master MAXENT one must understand how Lagrange multipliers acquire factual interpretations. How is it, for example, that a physical quantity like temperature comes to identified with the Lagrange multiplier introduced by the energy constraint? That subtle point is explained in Jaynes' first paper.

Jaynes sees MAXENT as a natural extension and unification of two separate lines of historical development in statistical reasoning. The first line includes Bernoulli, Laplace, Jeffreys, and Cox;

the second line includes Maxwell, Boltzmann, Gibbs, and Shannon. Jaynes' own place at the apex of that illustrious group is secure, for MAXENT should be attributed to him, though he is absolutely scrupulous about assigning credit to his predecessors. Jaynes displays a fine capacity for scholarly judgment in assessing the contributions of his predecessors. We are the lucky benefactors of his analysis, for he resurrects many arguments and results which have been overlooked or attacked and discarded by others, and he shows how they fit into a simple and coherent theory of statistical inference.

The second line of development is concerned with statistical mechanics, which we have already discussed.

In respect to the first line of development, Jaynes alone has called attention to the profound work of Richard T. Cox, who, from simple conditions for logical consistency, derived the two fundamental "laws of probability":

$$p(A|B) + P(\overline{A}|B) = 1\,,$$

$$p(AB|C) = p(A|BC)p(B|C)\,.$$

As Jaynes has emphasized, this gives the controversial "Bayes Law" an unassailable logical status and a greater generality than commonly recognized. Moreover, he goes on to show in one place or anonther (unfortunately not all in this book) how most (if not all) useful results in probability and statistics are consequences of these two laws together with a few principles, like MAXENT, for assigning numerical values to probabilities. For example, he shows that additional *ad hoc* constructs like "confidence intervals" are not needed to evaluate the reliability of statistical predictions.

Jaynes has extracted one other general principle from the work of his predecessors. Stimulated by Jeffreys, he has shown that probability distributions for some parameters are determined by *group invariance* under changes of variables consistent with the physical interpretations of the parameters. It is worth noting that both the MAXENT and group invariance principles, like the two fundamental laws of probability, can be regarded as conditions of consistency on the available information. One wonders if all statistical reasoning amounts to no more than the imposition of consistency requirements. Do these principles of consistent inference penetrate more deeply into the foundations of logic than truth tables? Note that the principles are perfectly general. They require no concept of random variable, so they reach far beyond the limited domain of conventional probability theory.

The principles of MAXENT and group invariance provide particular solutions to the "Problems of Priors," which is to assign definite numerical values to the probability $p(A|BC)$ of some prediction A from a theory B and other prior information C. The Problem of Priors is the outstanding problem of statistics, because it arises in every application and a general solution is not known. One can doubt that a general solution can be found, but, as Jaynes remarks, it seems likely that there are other principles besides MAXENT and group invariance which provide solutions for signifcant classes of particular cases. In its most general form, the Problem of Priors merges with the logical problems of proof and decidability. Logicians, pay attention!

Of Jaynes eight notable predecessors, all but Shannon were physicists. This reminds us how the development of mathematics is driven by efforts to solve real scientific problems. But Jaynes has not ignored the probability theory and statistics developed by mathematicians. Indeed, he has actively joined in the "great debate" between orthodox and Bayesian statisticians, and this book displays the potent firepower he brings to the Bayesian cause. Jaynes is ready to tell physicists that they can learn a lot from statisticians. But he argues that (nearly) all significant statistical results can be derived from the simple principles we have already mentioned.

The probabilists and statisticians have a lot to learn from Jaynes. His view is more general and unified than theirs, and he can introduce them to a wealth of applications. This view provides no rational basis for the current division between probabilists and statisticians, which hardly amounts to more than the ancient contempt of the pure for the applied.

This book is evidence that the widely proclaimed "Bayesian revolution" in statistics has been superseded by a *Jaynesian revolution* of far greater consequence. Jaynes rightly minimizes the revolutionary character of his thought and emphasizes its continuity with the past, just as Einstein had done with his theory of relativity. He could say, with an earlier revolutionary of some repute, "I have come not to destroy, but fulfill." I suspect that both Einstein and Jaynes would rather regard their contributions as midcourse corrections in the great voyage of science.

The scientific community owes a debt of gratitude to R. D. Rosenkrantz for editing this volume. His enlightened editorial commentary shows that at least one mathematician appreciates Jaynes. Now it is time for the publisher to offer an inexpensive paper back edition which students can afford.

An important contribution to the foundations of probability theory, statistics and statistical physics has been made by E. T. Jaynes. The recent publication of his collected works provides an appropriate opportunity to attempt an assessment of this contribution. Discover the world's research.Â B. R. Frieden uses a single procedure, called extreme physical information, with the aim of deriving â€˜most known physics, from statistical mechanics and thermodynamics to quantum mechanics, the Einstein field equations and quantum gravityâ€™. His method, which is based on Fisher information, is given a detailed exposition in this book, and we attempt to assess the extent to which he succeeds in his task.