Principles of Data Analysis

Harvey S. Leff

Physics Today **56** (12), 64–65 (2003);
https://doi.org/10.1063/1.1650233

molecular fields, that is, the local effective fields due to the magnetic nature of the material. In addition, Spaldin has done an especially good job of writing the chapter on ferromagnetic domains and hysteresis.

The practical realities associated with the field of magnetism are well explained in Spaldin's book. A whole chapter is devoted to magnetic data storage, and many examples throughout the text illustrate applications of the different categories of magnetic materials, even diamagnets. Spaldin also provides a timely discussion of the possible advantages of magneto-optical recording and magnetic semiconductors. On a practical and essential note, she covers SI and cgs systems. She then offers explicit interconversions, an important addition because of the persistent lack of consensus in the literature and in the workplace.

Spaldin imbues her narrative with the history of the subject. She gives names, dates, and original figures and references. For example, the Bitter technique is illustrated with a figure, taken from Francis Bitter's original 1931 publication, showing magnetite deposits on a crystal of nickel. His method helped researchers to see, through microscopy, ferromagnetic domain boundaries. Even the quotations that open chapters are of historical, philosophical, and scientific significance.

The book works as a self-contained undergraduate course on magnetic materials. The exercises are imaginative and pitched at the appropriate level for beginners. The fully worked solutions are mathematically complete and contain a significant amount of helpful prose. Overall, Spaldin brings a nice literary style to the book. She refuses to stick to the language of dusty authoritarian texts and occasionally flirts with varying degrees of informality. At the risk of offending those who may actually think in the language of dusty authoritarian texts, Spaldin captivates readers with her style, which makes learning a pleasurable and effective experience.

Returning to the matter of substance, one criticism is the omission of classic subjects in magnetism, such as permanent magnets, the Landau theory of phase transitions, and the Ising and Heisenberg models. But omitting these subjects is acceptable because until now, I believe, it has been impossible to find a concise and accessible magnetics book for novices. But the long wait is over—*Magnetic Materials* is the missing book.

Neil D. Mathur
University of Cambridge
Cambridge, England

Principles of Data Analysis

Prasenjit Saha
Cappella Archive, Great Malvern, UK, 2003. $12.00 paper (109 pp.). ISBN 1-902918-11-8

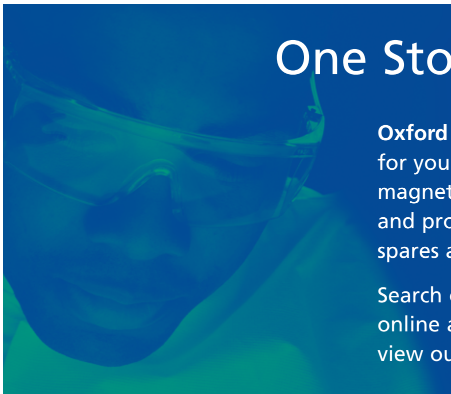
Available free online at http://ankh-morpork.maths.qmw.ac.uk/~saha/book

Probabilities are ubiquitous in physics. Physicists use them routinely in statistical and quantum physics, analyses of experimental data, and elsewhere. Encounters with probabilities typically involve abstract ensembles of identical systems or repeatable measurements, with probabilities interpreted as frequency ratios. This "frequentist" approach is the only way many physicists understand probabilities.

More generally, even if frequency data are unavailable, one may regard a probability as a likelihood based on prior information. The Bayesian approach, which goes back more than 200 years to Thomas Bayes and Pierre-Simon Laplace, enables one to consider probabilities without introducing frequency ratios.

In *Principles of Data Analysis*, theoretical physicist Prasenjit Saha focuses on Bayesian statistics and the maximum-entropy approach, a framework in which one identifies probabilities consistent with prior information such as average values and maximizes the entropy function to obtain best values for the probabilities. To get the Bayesian flavor, consider a data set D, a set of specified parameters ω, and a mathematical model M. If $P(\omega|D,M)$ is the probability of ω given D and M, then Bayes's theorem is $P(\omega|D,M) = P(D|\omega,M)\,P(\omega|M)/P(D|M)$. $P(\omega|M)$ is the *prior* probability of ω, given the model M without any data. $P(D|M)$ is constant for given D and M and $P(\omega|D,M)$ is a *posterior* probability that accounts for information contained in D. Specifying a useful prior $P(\omega|M)$ is a challenging aspect of the Bayesian approach. When Bayesian methods are applied to a given data set and two candidate models, they allow one to evaluate which model is favored by the data.

Saha's coverage includes Bayes's theorem; the binomial and Poisson distributions, with an example showing effects of choosing different priors; Gaussian distributions; the central limit theorem; random walks; the Monte Carlo technique (without explicit use of Bayesian concepts); least squares and distribution function fitting, both within a Bayesian context; information entropy; the maximum-entropy principle; and entropy in thermodynamics. The chapter on entropy and thermodynamics provides a succinct and clear, though relatively abstract, exposition of classical equilibrium thermodynamics and some statistical mechanics. Based on the maximum-entropy technique, the chapter does not require the material on Bayesian statistics that constitutes much of the book.



Saha's writing style, though spirited, is terse. His thrust is on presenting a potpourri of illustrative examples and problems. The problems are graded according to difficulty, and Saha gives hints or answers for them.

Terseness can be a strength that enables a reader to glean the essence of the Bayesian approach relatively quickly. The price is that many details are omitted, and the reader is forced to fill in many steps or seek other sources. To learn the fundamentals, Saha recommends Edwin T. Jaynes's *Probability Theory: The Logic of Science* (Cambridge U. Press, 2003; see also http://bayes.wustl.edu). From about 1958 until his death in 1998, physicist Jaynes wrote articulately about Bayesian statistics. He also did important work using the maximum-entropy framework. Saha also recommends Devinderjit Singh Sivia's *Data Analysis; A Bayesian Tutorial* (Oxford U. Press, 1996). I suggest three additional sources written by physicists: Giulio D'Agostini, *American Journal of Physics*, volume 67, page 1260, 1999; Robert Cousins, *American Journal of Physics*, volume 63, page 398, 1995; and Volker Dose, *Reports on Progress in Physics*, volume 66, page 1421, 2003.

Some of Saha's presentation could be more clear. For example, figure 2.1 displays computer-generated graphs illustrating the results of virtually flipping a biased coin. But the text does not specify either the degree of bias or the simulated data, namely, the number of heads. The mathematical level fluctuates from introductory to reasonably sophisticated and some of the manipulations lack sufficient motivation.

The book misses opportunities to link mathematics and physics. For example, Saha introduces the principle of indifference without mentioning its connection with the principle of *equal a priori* probabilities in statistical mechanics. Many of the 18 examples and 30 problems are not directly related to physics, and citations to the substantial literature on Bayesian methods in physics are lacking.

I commend Saha for making his book available by free download and minimal-cost paper copy. Despite some wrinkles, it provides a fresh, succinct view of data analysis at a level suitable for working physicists, graduate students, and very advanced undergraduates. In combination with the suggested supplements, this book could well serve as a stimulus and springboard for in-depth study and practical application of Bayesian and maximum-entropy techniques.

Harvey S. Leff
California State Polytechnic University Pomona

Physics of Fractal Operators

Bruce J. West, Mauro Bologna, and Paolo Grigolini
Springer-Verlag, New York, 2003. $69.95 (354 pp.). ISBN 0-387-95554-2

Derivatives in ordinary calculus describe local properties of functions such as slope and curvature. Fractional-order derivatives, by contrast, are nonlocal operators best suited to describe systems dominated by nonlocality. That nonlocality could be spatial—for example, long-range interactions or jumps. Or it could be temporal—say, long-time memories or nonstationary behavior. Not so long ago, physicists saw fractional calculus, the branch of calculus considering generalizations of the usual derivatives and integrals to an arbitrary order, as a mathematical toy and object of curiosity. Nowadays we are coming to understand the potential power of fractional-calculus methods. Systems such as viscoelastic polymeric solutions or melts, strongly disordered