Bibliography

Akhiezer, N. I. (1965), *The Classical Moment Problem*, Hafner, New York.

Archimedes (*c* 220 BC), in Works, T. L. Heath, ed., Cambridge University Press (1897, 1912).
Paperback reprint by Dover Publications, Inc., New York, undated (*c* 1960).

Aristotle (4th century BC), Organon.
Definition of syllogisms.

Aristotle (4th century BC) *Physics*; translation with commentary by Apostle, H. G., Indiana University Press, Bloomington (1969).

Ash, B. B. (1966), *Information Theory*, John Wiley, New York.

Bacon, F. (1620), 'Novum Organum', in Spedding. J., Ellis, R. L. & Heath, D. D., eds., *The Works of Francis Bacon*, vol. 4, Longman & Co., London (1857–1858).

Barber, N. F. & Ursell, F. (1948), 'The generation and propagation of ocean waves and swell', *Phil. Trans. Roy. Soc. Lond.*, **A240**, 527–560.
Detection of chirped signals in noise.

Barlow, E. R. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart & Winston, New York.

Barndorf-Nielsen, O. (1978), *Information and Exponential Families in Statistical Theory*, J. Wiley & Sons, New York.

Barr, A. & Feigenbaum, E., eds. (1981), *The Handbook of Artificial Intelligence*, 3 vols., Wm. Kaufman, Inc., Los Altos, CA.
Contributions from over 100 authors. Volume 1 surveys search routines, one of the few aspects of AI that could be useful in scientific inference.

Barron, A. R. (1986), 'Entropy and the central limit theorem', Ann. Prob. **14**, 336–342.

Bartholomew, D. J. (1965), 'A comparison of some Bayesian and frequentist inference', *Biometrika*, **52**, 19–35.

Benford, F. (1938), 'The law of anomalous numbers', *Proc. Am. Phil. Soc.* **78**, 551–572.
Benford is probably the one referred to mysteriously by Warren Weaver (1963), p. 270. But, unknown to them, Simon Newcomb (1881) had noticed this phenomenon long before. See Raimi (1976) for many more details and references.

Berkson, J. (1977), 'My encounter with neo-Bayesianism', *Int. Stat. Rev.* **45**, 1–9.

Berkson, J. (1980), 'Minimum chi-squared, not maximum likelihood!', *Ann. Stat.* **8**, 457–487.

Bernardo, J. M. (1977) 'Inferences about the ratio of normal means: a Bayesian approach to the Fieller-Creasy problem', in Barra, J. D., *et al.*, eds., *Recent Developments in Statistics*, North Holland Press, Amsterdam, pp. 345–350.

Bernado, J. M. (1979a), 'Reference posterior distributions for Bayesian inference', *J. Roy. Stat. Soc. B* **41**, 113–147.
With discussion.

Bernado, J. M. (1979b), 'Expected information as expected utility', *Ann. Stat.* **7**, 686–690.

Bernado, J. M., de Groot, M. H., Lindley, D. V. & Smith, A. F. M., eds. (1980), *Bayesian Statistics*, Proceedings of the First Valencia International Meeting on Bayesian Statistics, Valencia, May 28–June 2, 1979, University Press, Valencia, Spain.

Bernado, J. M., de Groot, M. H. & Lindley, D. V., eds. (1985), *Bayesian Statistics 2*, Proceedings of the Second Valencia International Meeting on Bayesian Statistics, September 6–10, 1983, Elsevier Science Publishers, New York.

Billingsley, P. (1979), *Probability and Measure*, Wiley, New York.

Contains more Borel–Kolmogorov stuff that we do not go into.

Bishop, Y., Fienberg, S., & Holland, P. (1975), *Discrete Multivariate Analysis*, MIT Press, Cambridge, MA.

Blanc-Lapierre, A. & Fortet, R. (1953), *Theorie des Fonctions Aleatoires*, Masson et Cie, Paris.

Boole, G. (1916), *Collected Logical Works, Vol. 1: Studies in Logic and Probability; Vol II: An Investigation of the Laws of Thought*; Open Court, Chicago.

Borel, E. (1926), *Traité du Calcul des Probabilités*, Gauthier-Villars, Paris.

The Hausdorff paradox on congruent sets on a sphere is discussed in Tome II, Fasc. 1.

Born, M. (1964), *Natural Philosophy of Cause and Chance*, Dover, New York.

Boscovich, Roger J. (1770), *Voyage Astronomique et Geographique*, N. M. Tillard, Paris.

Adjustment of data by the criterion that the sum of the corrections is zero, the sum of their magnitudes is made a minimum.

Box, G. E. P. (1982), 'An apology for ecumenism in statistics', NRC Technical Report #2408, Mathematics Research Center, University of Wisconsin, Madison.

Box, G. E. P., Leonard, T. & Wu, C-F, eds. (1983), *Scientific Inference, Data Analysis, and Robustness*, Academic Press, Inc., Orlando, FL.

Proceedings of a conference held in Madison, Wisconsin, November 1981.

Bracewell, R. N. (1986), 'Simulating the sunspot cycle', *Nature*, **323**, 516.

Ronald Bracewell is perhaps the first author with the courage to present a definite prediction of future sunspot activity. We await the Sun's verdict with interest.

Brewster, D. (1855), *Memoirs of the Life, Writings, and Discoveries of Sir Isaac Newton*, 2 vols., Thomas Constable, Edinburgh.

Brigham, E. & Morrow, R. E. (1967), 'The fast Fourier transform', *Proc. IEEE Spectrum* 4, 63–70.

Brillouin, L. (1956), *Science and Information Theory*, Academic Press, New York.

Bross, I. D. J. (1963), 'Linguistic analysis of a statistical controversy', *Am. Stat.* **17**, 18.

One of the most violent polemical denunciations of Bayesian methods in print – without the slightest attempt to examine the actual results they give! Should be read by all who want to understand why and by what means the progress of inference was held up for so long. Jaynes (1976) was written originally in 1963 as a reply to Bross, in circumstances explained in Jaynes (1983), p. 149.

Brown, E. E. & Duren, B. (1986), 'Information integration for decision support', *Decision Support Syst.*, **4**, (2), 321–329.

Brown, R. (1828),'A brief account of microscopical observations', *Edinburgh New Phil. J.* **5**, 358–371.

First report of the Brownian motion.

Burg, J. P. (1967), 'Maximum entropy spectral analysis', Proceedings of the 37th Meeting of the Society of Exploration Geophysicists.

Burg, J. P. (1975), 'Maximum entropy spectral analysis', Ph.D. Thesis, Stanford University.

Busnel, R. G. & Fish, J. F., eds. (1980), *Animal Sonar Systems*, NATO ASI Series, Vol. A28, Plenum Publishing Corp., New York.

A very large (1082 pp.) report of a meeting held on the island of Jersey, UK, in 1979.

Cajori, F. (1928), in *Sir Isaac Newton 1727–1927*, Waverley Press, Baltimore, pp. 127–188.

Cajori, F. (1934), *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and his System of the World*, University of California Press, Berkeley.

Carnap, R. (1950), *Logical Foundations of Probability*, Routlege and Kegan Paul Ltd., London.

Chen, Wen-chen, & de Groot, M. H. (1986), 'Optimal search for new types', in Goel, P. & Zellner, A., eds. (1986), *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, Elsevier Science Publishers, Amsterdam, pp. 443–458.

Childers, D., ed. (1978), *Modern Spectrum Analysis*, IEEE Press, New York.

A collection of reprints of early works on maximum entropy spectrum analysis.

Chow, Y., Robbins, H. & Siegmund, D. (1971), *Great Expectations: Theory of Optimal Stopping*, Houghton Mifflin & Co., Boston.

Cobb, L. & Watson, B. (1980), 'Statistical catastrophe theory: an overview', *Math. Modelling*, **1**, 311–317.

We have no quarrel with this work, but wish to add two historical footnotes. (1) Their 'stochastic differential equation' is what physicists have called a 'Fokker–Planck equation' since about 1917. However, we are used to having our statistical work attributed to Kolmogorov by mathematicians. (2) Stability considerations of multiple-valued 'folded' functions of the kind associated today with the name of René Thom are equivalent to convexity properties of a single-valued entropy function, and these were given by J. Willard Gibbs in 1873.

Cohen, T. J. & Lintz, P. R. (1974), 'Long term periodicities in the sunspot cycle', *Nature*, **250**, 398.

Cooley, J. W. & Tukey, J. W. (1965), 'An algorithm for the machine calculation of complex Fourier series', *Math. Comp.*, **19**, 297–301.

Cooley, J. W., Lewis, P. A. & Welch, P. D. (1967), 'Historical notes on the fast Fourier transform', *Proc. IEEE* **55**, 1675–1677.

Cook, A. (1994), *The Observational Foundations of Physics*, Cambridge University Press.

Notes that physical quantities are defined in terms of the experimental arrangement used to measure them. Of course, this is just the platitude that Niels Bohr emphasized in 1927.

Cox, D. R. & Hinkley, D. V. (1974), *Theoretical Statistics*, Chapman & Hall, London; reprints 1979, 1982.

Mostly a repetition of old sampling theory methods, in a bizarre notation that can make the simplest equation unreadable. However, it has many useful historical summaries and side remarks, noting limitations or extensions of the theory, that cannot be found elsewhere. Bayesian methods are introduced only in the penultimate Chapter 10; and then the authors proceed to repeat all the old, erroneous objections to them, showing no comprehension that these were ancient misunderstandings long since corrected by Jeffreys (1939), Savage (1954), and Lindley (1956). One prominent statistician, noting this, opined that Cox and Hinkley had 'set statistics back 25 years'.

Cox, D. R. (1970), *The Analysis of Binary Data*, Methuen, London.

Cox, R. T. (1978), 'Of inference and inquiry', in Levine, R. D. & Tribus, M., eds., *The Maximum Entropy Formalism*, MIT Press, Cambridge, MA, pp. 119–167.

Notes that, corresponding to the logic of propositions, there is a dual logic of questions. This could become very important with further development, as discussed further in Jaynes (1983), pp. 382–388.

Cozzolino, J. M. & Zahner, M. J. (1973), 'The maximum-entropy distribution of the future market price of a stock', *Operations Res.*, **21**, 1200–1211.

Creasy, M. A. (1954), 'Limits for the ratio of means', *J. Roy. Stat. Soc.* **B 16**, 175–185.

Csiszar, I. (1984), 'Sanov property, generalized I-projection and a conditional limit theorem', *Ann. Prob.*, **12**, 768–793.

Czuber, E. (1908), *Wahrscheinlichkeitsrechnung und Ihre Anwendung auf Fehlerausgleichung*, 2 vols., Teubner, Berlin.

Some of Wolf's famous dice data may be found here. In the period roughly 1850–1890, Wolf, a Zurich astronomer, conducted and reported a mass of 'random' experiments; an account of these is given here.

Daganzo, C. (1977), *Multinomial Probit: The Theory and its Application to Demand Forecasting*, Academic Press, New York.

Dale, A. I. (1982), 'Bayes or Laplace? An examination of the origin and early applications of Bayes' theorem', *Arch. Hist. Exact Sci.* **27**, 23–47.

Daniel, C. & Wood, F. S. (1971), *Fitting Equations to Data*, John Wiley, New York.

Daniell, G. J. & Potton, J. A. (1989), 'Liquid structure factor determination by neutron scattering – some dangers of maximum entropy', in Skilling, J., ed. (1989), *Maximum Entropy and Bayesian Methods*, Proceedings of the Eighth Maximum Entropy Workshop, Cambridge, UK, August 1988, Kluwer Academic Publishers, Dordrecht, pp. 151–162.

The 'danger' here is that a beginner's first attempt to use maximum entropy on a complex problem may be unsatisfactory because it is answering a different question than what the user had in mind. So the first effort is really a 'training exercise' which makes one aware of how to formulate the problem properly.

Davenport, W. S. & Root, W. L. (1958), *Random Signals and Noise*, McGraw-Hill, New York.

David, F. N. (1962), *Games, Gods and Gambling*, Griffin, London.

A history of the earliest beginnings of probability theory. Notes that in archaeology, 'the farther back one goes, the more fragmentary is the evidence'. Just the kind of deep insight that we could find nowhere else.

de Finetti, B. (1958), 'Foundations of probability', in *Philosophy in the Mid-century*, La Nuova Italia Editrice, Florence, pp. 140–147.

de Groot, M. H., Bayarri, M. J. & Kadane, J. B. (1988), 'What is the likelihood function?' (with discussion), in Gupta, S. S. & Berger, J. O., eds., *Statistical Decision Theory and Related Topics IV*, Springer-Verlag, New York.

de Groot, M. H. & Cyert, R. M. (1987), *Bayesian Analysis and Uncertainty in Economic Theory*, Chapman & Hall, London.

de Groot, M. H., Fienberg, S. E. & Kadane, J. B. (1986), *Statistics and the Law*, John Wiley, New York.

Deming, W. E. (1943), *Statistical Adjustment of Data*, John Wiley, New York.

Dempster, A. P. (1963), 'On a paradox concerning inference about a covariance matrix', *Ann. Math. Stat.* **34**, 1414–1418.

Dubois, D. & Prade, H. (1988), *Possibility Theory*, Plenum Publ. Co., New York.

Dunnington, G. W. (1955), *Carl Friedrich Gauss, Titan of Science*, Hafner, New York.

Dutta, M. (1966), 'On maximum entropy estimation', *Sankhya,* ser. A, **28**, (4), 319–328.

Dyson, F. J. (1979), *Disturbing the Universe*, Harper & Row, New York.

A collection of personal reminiscences and speculations extending over some 50 years: 90% of it is irrelevant to our present purpose; but one must persist here, because Freeman Dyson played a very important part in the development of theoretical physics in the mid-20th century. His reminiscences about this are uniquely valuable, but are unfortunately scattered in small pieces over several chapters. Unlike some of his less thoughtful colleagues, Dyson saw correctly many fundamental things about probability theory and quantum theory (but in our view missed some others equally fundamental). Reading this work is rather like reading Kepler and trying to extract the tiny nuggets of important truth.

Eddington, Sir A. (1935), *The Nature of the Physical World*, Dent, London.

Another distinguished scientist who thinks as we do about probability.

Edwards, A. W. F. (1972), *Likelihood*, Cambridge University Press.

Anthony Edwards was the last student of R. A. Fisher; although he understands all the technical facts pertaining to Bayesian methods as well as anybody, some mental block prevents him, as it did Fisher, from accepting their obvious consequences. So we must, sadly, part company and proceed with the constructive development of inference without him.

Edwards, A. W. F. (1992), *Nature* **352**, 386–387.

Commentary on Bayesian methods.

Edwards, H. M. (1987), 'An appreciation of Kronecker', *Math. Intelligencer*, **9**, 28–35.

Edwards, H. M. (1988), 'Kronecker's place in history', in Aspray, W. & Kitcher, P., eds., *History and Philosophy of Modern Mathematics*, University of Minnesota Press.

Efron, B. (1975), 'Biased versus unbiased estimation', *Adv. Math.* **16**, 259–277

Efron, B. (1978), 'Controversies in the foundations of statistics', *Am. Math. Monthly* **85**, 231–246.

Efron, B. (1979a), 'Bootstrap methods: another look at the jackknife', *Ann. Stat.* **6**, 1–26.

Efron, B. (1979b), 'Computers and the theory of statistics: thinking the unthinkable', *SIAM Rev.* **21**, 460–480.

Efron, B. & Gong, G. (1983), 'A leisurely look at the bootstrap, the jackknife, and cross-validation', *Am. Stat.* **37**, 36–48.

Orthodox statisticians have continued trying to deal with problems of inference by inventing arbitrary *ad hoc* procedures instead of applying probability theory. Three recent examples are explained and advocated here. Of course, they all violate our desiderata of rationality and consistency; the reader will find it interesting and instructive to demonstrate this and compare their results with those of the Bayesian alternatives.

Evans, M. (1969), *Macroeconomic Forecasting*, Harper & Row, New York.

Fechner, G. J. (1860), *Elemente der Psychophysik*, 2 vols.; Vol. 1 translated as *Elements of Psychophysics*, Boring, E. G. & Howes, D. H., eds., Holt, Rinehart & Winston, New York (1966).

Fechner, G. J. (1882), *Revision der Hauptpuncte der Psychophysik*, Breitkopf u. Härtel, Leipzig.

Feinstein, A. (1958), *Foundations of Information Theory*, McGraw-Hill, New York.

Like the work of Khinchin (1957), a mathematician's view of things, which has almost nothing in common with the physically oriented view of Goldman (1953).

Ferguson, T. S. (1982), 'An inconsistent maximum likelihood estimate', *J. Am. Stat. Assoc.* **77**, 831–834.

Fieller, E. C. (1954), 'Some problems in interval estimation', *J. Roy. Stat. Soc.* **B 16**, 175–185.

This and the contiguous paper by Creasy (1954; this bibliography) became famous as 'the Fieller–Creasy problem' of estimating the ratio μ_1/μ_2 of means of two normal sampling distributions. It generated a vast amount of discussion and controversy because orthodox methods had no principles for dealing with it – and for decades nobody would deign to examine the Bayesian solution. It is a prime example of an estimation problem, easily stated, for which only Bayesian methods provide the technical apparatus required to solve it. It is finally considered from a Bayesian standpoint by José Bernardo (1977).

Fisher, R. A. (1912), 'On an absolute criterion for fitting frequency curves', *Messeng. Math.* **41**, 155–160.

Fisher, R. A. (1915), 'Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population', *Biometrika* **10**, 507–521.

Fisher, R. A. (1930), 'Inverse probabilities', *Proc. Camb. Phil. Soc.* **26**, 528–535.

Fisher, R. A. (1935), *The Design of Experiments*, Oliver & Boyd, Edinburgh; six later editions to 1966.

Fisher, R. A. (1938), *Statistical Tables for Biological, Agricultural and Medical Research* (with F. Yates), Oliver & Boyd, Edinburgh; five later editions to 1963.

Fisher, R. A. & Tippett, L. H. C. (1928), 'Limiting forms of the frequency distribution of the largest or smallest member of a sample', *Proc. Camb. Phil. Soc.* **24**, 180–190.

Fougeré, P. F. (1977), 'A solution to the problem of spontaneous line splitting in maximum entropy power spectrum analysis', *J. Geophys. Res.* **82**, 1051–1054.

Galton, F. (1863), *Meteorographica*, MacMillan, London.

Here this remarkable man invents weather maps and, from studying them, discovers the 'anticyclone' circulation patterns in the northern hemisphere.

Galton, F. (1889), *Natural Inheritance*, MacMillan, London.

Gentleman, W. M. (1968), 'Matrix multiplication and fast Fourier transformations', *Bell Syst. Tech. J.*, **17**, 1099–1103.

Gillispie, C. C., ed. (1981), *Dictionary of Scientific Biography*, 16 vols., C. Scribner's Sons, New York.

The first place to look for information on any scientist.

Glymour, C. (1980), *Theory and Evidence*, Princeton University Press.

Gnedenko, B. V. & Kolmogorov, A. N. (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Cambridge, MA.

On p. 1 we find the curious statement: 'In fact, all epistomologic value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, non-random regularity.' This was thought by some to serve a political purpose in the old USSR; in any event, the most valuable applications of probability theory today are concerned with incomplete information and have nothing to do with those so-called 'random phenomena' which are still undefined in theory and unidentified in Nature.

Goel, P. & Zellner, A. (1986), eds., *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, Elsevier Science Publishers, Amsterdam.

Gokhale, D. and Kullback, S. (1978), *The Information in Contingency Tables*, Marcel Dekker, New York.

Goldberg, S. (1983), *Probability in Social Science*, Birkhaeuser, Basel.

Good, I. J. (1965), *The Estimation of Probabilities*, Research Monographs #30, MIT Press, Cambridge, MA.

Jack Good persisted in believing in the existence of 'physical probabilities' that have some kind of reality independently of human information; hence the (to us) incongruous title.

Grandy, W. T. & Schick, L. H., eds. (1991), *Maximum Entropy and Bayesian Methods*, Proceedings of the Tenth annual Maximum Entropy workshop, Kluwer Academic Publishers, Holland.

Grenander, U. & Szegö, G. (1957), *Toeplitz Forms and their Applications*, University of California Press, Berkeley.

Griffin, D. R. (1958), *Listening in the Dark*, Yale University Press, New Haven.

See also Slaughter, R. H. & Walton, D. W., eds. (1970), *About Bats*, SMU Press, Dallas, Texas.

Gull, S. F. & Daniell, G. J. (1978), 'Image reconstruction from incomplete and noisy data', *Nature* **272**, 686.

Gull, S. F. & Daniell, G. J. (1980), 'The maximum entropy algorithm applied to image enhancement', *Proc. IEEE (E)* **5**, 170.

Gull, S. F. & Skilling, J. (1984), 'The maximum entropy method', in Roberts, J. A., ed., *Indirect Imaging*, Cambridge University Press.

Hacking, I. (1965), *Logic of Statistical Inference*, Cambridge University Press.

Hacking, I. (1973), *The Emergence of Probability*, Cambridge University Press.

Hacking, I. (1984), 'Historical models for justice', *Epistemologia, Special Issue on Probability, Statistics, and Inductive Logic*, **VII**, 191–212.

Haldane, J. B. S. (1957), 'Karl Pearson, 1857–1957', *Biometrika* **44**, 303–313.

Haldane's writings, whatever the ostensible topic, often turned into political indoctrination for socialism. In this case it made some sense, since Karl Pearson was himself a political radical. Haldane suggests that he may have changed the spelling of his name from 'Carl' to 'Karl' in honor of Karl Marx, and from this Centenary oration we learn that V. I. Lenin quoted approvingly from Karl Pearson. Haldane was Professor of Genetics at University College, London in the 1930s, but he resigned and moved to India as a protest at the failure of the authorities to provide the financial support he felt his Department needed. It is easy to imagine that this was precisely what those authorities, exasperated at his preoccupation with left-wing politics instead of genetics, hoped to bring about. An interesting coincidence is that Haldane's sister, Naomi Haldane Mitchison, married a Labour MP and carried on the left-wing cause. James D. Watson was a guest at her home at Christmas 1951, about a year before discovering the DNA helix structure. He was so charmed by the experience that his 1968 book, *The Double Helix*, is inscribed: 'For Naomi Mitchison.'

Hampel, F. R. (1973), 'Robust estimation: a condensed partial survey', *Zeit. Wahrsch, theorie vrw. Beb.* **27**, 87–104

Hankins, T. L. (1970), *Jean d'Alembert: Science and the Enlightenment*, Oxford University Press.

Heath, D. & Sudderth, W. (1976), 'de Finetti's theorem on exchangeable variables', *Am. Stat.* **30**, 188.

An extremely simple derivation.

Helliwell, R. A. (1965), *Whistlers and Related Ionospheric Phenomena*, Stanford University Press, Palo Alto, CA.

Hellman, M. E. (1979), 'The mathematics of public-key cryptography', *Sci. Am.* **241**, 130–139.

Hewitt, E. & Savage, L. J. (1955), 'Symmetric measures on Cartesian products', *Trans. Am. Math. Soc.* **80**, 470–501.

A generalization of de Finetti's representation theorem to arbitrary sets.

Hirst, F. W. (1926), *Life and Letters of Thomas Jefferson*, Macmillan, New York.

Hobson, A. & Cheung, B. K. (1973), 'A comparison of the Shannon and Kullback information measures', *J. Stat. Phys.* **7**, 301–310.

Hodges, J. L. & Lehmann, E. L. (1956), 'The efficiency of some nonparametric competitors of the *t*-test', *Ann. Math. Stat.* **27**, 324–335.

Hofstadter, D. R. (1983), 'Computer tournaments of the Prisoner's dilemma suggest how cooperation evolves', *Sci. Am.* **248**, (5), 16–26.

Holbrook, J. A. R. (1981), 'Stochastic independence and space-filling curves', *Am. Math. Monthly* **88**, 426–432.

Jagers, P. (1975), *Branching Processes with Biological Applications*, John Wiley, London.

James, W. & Stein, C. (1961), 'Estimation with quadratic loss', Proc. 4th Berkeley Symp., Univ. Calif. Press, **1**, 361–380.

Jansson, P. A., ed. (1984), *Deconvolution, with Applications in Spectroscopy*, Academic Press, Orlando, FL.

Articles by nine authors, summarizing the state of the art (mostly linear processing) as it existed just before the introduction of Bayesian and maximum entropy methods.

Jaynes, E. T. (1963), 'Review of *Noise and Fluctuations*, by D. K. C. MacDonald', *Am. J. Phys.* **31**, 946.

Cited in Jaynes (1976) in response to a charge by Oscar Kempthorne that physicists have paid little attention to noise; notes that there is no area of physics in which the phenomenon of noise does not present itself. As a result, physicists were actively studying noise and knew the proper way to deal with it, long before there was any such thing as a statistician.

Jaynes, E. T. (1973a), 'Survey of the present status of neoclassical radiation theory', in Mandel, L. & Wolf, E., eds., *Proceedings of the 1972 Rochester Conference on Optical Coherence*, Pergamon Press, New York.

Jaynes, E. T. (1973b), 'The well-posed problem', *Found. Phys.* **3**, 477–493.

Reprinted in Jaynes (1983).

Jaynes, E. T. (1980a), 'The minimum entropy production principle', *Ann. Rev. Phys. Chem.* **31**, 579–601.

Reprinted in Jaynes (1983).

Jaynes, E. T. (1980b), 'What is the question?', in Bernardo, J. M., de Groot, M. H., Lindley, D. V. & Smith, A. F. M., eds., *Bayesian Statistics*, University Press, Valencia, Spain, pp. 618–629.

Discussion of the logic of questions, as pointed out by R. T. Cox (1978), and applied to the relationship between parameter estimation and hypothesis testing. Reprinted in Jaynes (1983), pp. 382–388.

Jaynes, E. T. (1981), 'What is the problem?', in Haykin, S., ed., *Proceedings of the Second SSSP Workshop on Spectrum Analysis*, McMaster University.

The following article is an enlarged version.

Jaynes, E. T. (1982), 'On the rationale of maximum-entropy methods', *Proc. IEEE* **70**, 939–952.

Jaynes, E. T. (1984), 'Prior information and ambiguity in inverse problems', in *SIAM-AMS Proceedings*, Vol. 14, American Mathematical Society, pp. 151–166.

Jaynes, E. T. (1985a) 'Where do we go from here?', in Smith, C. & Grandy, W. T., eds. *Maximum-Entropy and Bayesian Methods in Inverse Problems*, D. Reidel Publishing Co., Dordrecht, pp. 21–58.

Jaynes, E. T. (1985b), 'Entropy and search theory', in Smith, C. & Grandy, W. T., eds. *Maximum-Entropy and Bayesian Methods in Inverse Problems*, D. Reidel Publishing Co., Dordrecht, pp. 443–454.

Shows that the failure of previous efforts to find a connection between information theory and search theory were due to use of the wrong entropy expression. In fact, there is a very simple and general connection, as soon as we define entropy on the deepest hypothesis space.

Jaynes, E. T. (1985c) 'Macroscopic prediction', in Haken, H., ed., *Complex Systems – Operational Approaches*, Springer-Verlag, Berlin.

Jaynes, E. T. (1985d), 'Generalized scattering', in Smith, C. & Grandy, W. T., eds., *Maximum-Entropy and Bayesian Methods in Inverse Problems*, D. Reidel Publishing Co., Dordrecht, pp. 377–398.

Some of the remarkable physical predictions contained in the comparison of two maximum entropy distributions, before and after adding a new constraint.

Jaynes, E. T. (1986), 'Some applications and extensions of the de Finetti representation theorem', in Goel, P. & Zellner, A., eds., *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, Elsevier Science Publishers, Amsterdam, pp. 31–42.

The theorem, commonly held to apply only to infinite exchangeable sequences, remains valid for finite ones if one drops the non-negativity condition on the generating function. This makes it applicable to a much wider class of problems.

Jaynes, E. T. (1988a), 'The relation of Bayesian and maximum entropy methods', in Erickson, G. J. & Smith, C. R., eds. (1988), *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Vol. 1, pp. 25–29.

Jaynes, E. T. (1988b), 'Detection of extra-solar system planets', in Erickson, G. J. & Smith, C. R., eds. (1988), *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Vol. 1, pp. 147–160.

Jaynes, E. T. (1991), 'Notes on present status and future prospects', in Grandy, W. T. & Schick, L. H., eds. (1991), *Maximum* Entropy and Bayesian Methods, Proceedings of the Tenth annual Maximum Entropy Workshop, Kluwer Academic Publishers, Holland.

A general summing-up of the situation as it appeared in the summer of 1990.

Jaynes, E. T. (1993), 'A backward look to the future', in Grandy, W. T. & Milonni, P. W., eds., *Physics and Probability: Essays in Honor of Edwin T. Jaynes*, Cambridge University Press, pp. 261–275.

A response to the contributors to this *Festschrift* volume marking the writer's 70th birthday, with 22 articles by my former students and colleagues.

Jefferys, W. H. (1990), 'Bayesian analysis of random event generator data', *J. Sci. Expl.* **4**, 153–169.

Shows that orthodox significance tests can grossly overestimate the significance of ESP data; Bayesian tests yield defensible conclusions because they do not depend on the intentions of the investigator.

Jeffreys, H. (1963), 'Review of Savage (1962)', *Technometrics* **5**, 407–410.

Jeffreys, Lady Bertha Swirles (1992), 'Harold Jeffreys from 1891 to 1940', *Notes Rec. Roy. Soc. Lond.* **46**, 301–308.

A short, and puzzlingly incomplete, account of the early life of Sir Harold Jeffreys, with a photograph of him in his 30s. Detailed account of his interest in botany and early honors (he entered St John's College, Cambridge as an undergraduate, in 1910; and that same year received the Adams memorial prize for an essay on 'Precession and nutation'). But, astonishingly, there is no mention at all of his work in probability theory! In the period 1919–1939, this resulted in many published articles and two books (Jeffreys, 1931, 1939) of very great importance to scientists today. It is, furthermore, of *fundamental* importance and will remain so long after all his other work recedes into history. Bertha Swirles Jeffreys was also a physicist, who studied with Max Born in Göttingen in the late 1920s and later became Mistress of Girton College, Cambridge.

Jerri, A. J. (1977), 'The Shannon sampling theorem – its various extensions and applications', *Proc. IEEE* **65**, 1565–1596.

A massive tutorial collection of useful formulas, with 248 references.

Johnson, R. W. (1979), 'Axiomatic characterization of the directed divergences and their linear combinations', *IEEE Trans.* **IT-7**, 641–650.

Kale, B. K. (1970), 'Inadmissibility of the maximum likelihood estimation in the presence of prior information', *Can. Math. Bull.* **13**, 391–393.

Kalman, R. E. (1982), 'Identification from real data', in Hazewinkel, M. & Rinnooy Kan, A., eds., *Current Developments in the Interface: Economics, Econometrics, Mathematics*, D. Reidel Publishing Co., Dordrecht-Holland, pp. 161–196.

Kalman, R. E. (1990), *Nine Lectures on Identification*, Lecture Notes on Economics and Mathematical Systems, Springer-Verlag.

Kandel, A. (1986), *Fuzzy Mathematical Techniques with Applications*, Addison-Wesley, Reading, MA.

Kay, S. & Marple, S. L., Jr (1979), 'Source of and remedies for spectral line splitting in autoregressive spectrum analysis', Proceedings of the 1979 IEEE International Conference on Acoustics, Speech Signal Processing, October 1978, pp. 469–471.

Kemeny, J. G. & Snell, J. L. (1960), *Finite Markov Chains*, D. van Nostrand Co., Princeton, NJ.

Kendall, M. G. (1956), 'The beginnings of a probability calculus' *Biometrika* **43**, 1–14; reprinted in Pearson & Kendall (1970).

A fascinating psychological study. In the attempt to interpret the slow early development of probability theory as caused by the unfounded prejudices of others, he reveals inadvertently his own unfounded prejudices, which in our view are the major cause of retarded – even backward – progress in the 20th century.

Khinchin, A. I. (1949) *Mathematical Foundations of Statistical Mechanics*, Dover Publications, Inc., New York.

An attempt to base the calculational techniques on the central limit theorem, not general enough for problems of current interest. But the treatment of the Laplace transform relation between structure functions and partition functions is still valuable reading today, and forms the mathematical basis for our own development in Part 2.

Kiefer, J. & Wolfowitz, J. (1956), 'Consistency of the maximum likelihood estimation in the presence of infinitely many incidental parameters', *Ann. Math. Stat.* **27**, 887–906.

Kindermann, R. & Snall, J. L. (1980), *Markov Random Fields*, Contemporary Mathematics Vol. 1, AMS, Providence, RI.

Kuhn, T. S. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press; 2nd edn, 1970.

Kullback, S. (1959), *Information Theory and Statistics*, John Wiley, New York.

A beautiful work, never properly appreciated because it was 20 years ahead of its time.

Landau, H. J. (1983), 'The inverse problem for the vocal tract and the moment problem', *SIAM J. Math. Anal.* **14**, 1019–1035.

Modeling speech production by a reflection coefficient technique closely related to the Burg maximum entropy spectrum analysis.

Landau, H. J. (1987), 'Maximum entropy and the moment problem', *Bull. Am. Math. Soc.* **16**, 47–77.

Interprets the Burg solution in terms of more general problems in several fields. Highly recommended for a deeper understanding of the mathematics.

Legendre, A. M. (1806), 'Nouvelles méthods pour la détermination des orbits des cométes', Didot, Paris.

Leibniz, G. W. (1968), *General Investigations Concerning the Analysis of Concepts and Truths*, trans. W. H. O'Briant, University of Georgia Press.

Lessard, S., ed. (1989), *Mathematical and Statistical Developments in Evolutionary Theory*, NATO ASI Series Vol. C299, Kluwer Academic Publishers, Holland.

Proceedings of a meeting held in Montreal, Canada in 1987.

Lewis, G. N. (1930) 'The symmetry of time in physics', *Science* **71**, 569.

An early recognition of the connection between entropy and information, showing an understanding far superior to what many others were publishing 50 years later.

Lindley, D. V. (1956), 'On a measure of the information provided by an experiment', *Ann. Math.* **27**, 986–1005.

Lindley, D. V. (1958) 'Fiducial distributions and Bayes' theorem', *J. Roy. Stat. Soc.* **B20**, 102–107.

Lindley, D. V. (1971) *Bayesian Statistics: A Review*, Society for Industrial and Applied Mathemathics, Philadelphia.

Linnik, Yu. V. (1961), *Die Methode der kleinsten Quadrate in Moderner Darstellung*, Deutscher Verl. der Wiss., Berlin.

Litterman, R. B. (1985), 'Vector autoregression for macroeconomic forecasting', in Zellner, A. & Goel, P., eds., *Bayesian Inference and Decision Techniques*, North-Holland Publishers, Amsterdam.

Lukacs, E. (1960), *Characteristic Functions*, Griffin, London.

Macdonald, P. D. M. (1987), 'Analysis of length-frequency distributions', in Summerfelt, R. C. & Hall, G. E., eds., *Age and Growth of Fish*, Iowa State University Press, pp. 371–384.

A computer program for deconvolving mixtures of normal and other distributions. The program, 'MIX 3.0' is available from: Ichthus Data Systems, 59 Arkell St, Hamilton, Ontario, Canada L8S 1N6. In Chapter 7 we note that the problem is not very well-posed; Icthus acknowledges that it is 'inherently difficult' and may not work satisfactorily on the user's data. See also Titterington, Smith & Makov (1985).

Mandel, J. (1964), *The Statistical Analysis of Experimental Data*, Interscience, New York.

Straight orthodox *ad hockeries*, one of which is analyzed in Jaynes (1976).

Mandelbrot, B. (1977), *Fractals, Chance and Dimension*, W. H. Freeman & Co., San Francisco.

Marple, S. L. (1987), *Digital Spectral Analysis with Applications*, Prentice-Hall, New Jersey.

Martin, R. D. & Thompson, D. J. (1982), 'Robust-resistant spectrum estimation', *Proc. IEEE* **70**, 1097–1115.

Evidently written under the watchful eye of their mentor John Tukey, this continues his practice of inventing a succession of *ad hoc* devices based on intuition rather than probability theory. It does not even acknowledge the existence of maximum entropy or Bayesian methods. To their credit, the authors do give computer analyses of several data sets by their methods – with results that do not look very encouraging to us. It would be interesting to acquire their raw data and analyze them by methods like those of Bretthorst (1988) that do make use of probability theory; we think that the results would be quite different.

Masani, S. M. (1977), 'A paradox in admissibility', *Ann. Stat.* **5**, 544–546.
Maxwell, J. C. (1850), Letter to Lewis Campbell; reproduced in L. Campbell & W. Garrett, *The Life of James Clerk Maxwell*, Macmillan, 1881.
McColl, H. (1897) 'The calculus of equivalent statements', *Proc. Lond. Math. Soc.* **28**, 556.
Criticism of Boole's version of probability theory.
McFadden, D. (1973), 'Conditional logit analysis of qualitative choice behavior', in Zarembka, P., ed., *Frontiers in Econometrics*, Academic Press, New York.
Mead, L. R. & Papanicolaou, N. (1984), 'Maximum entropy in the problem of moments', *J. Math. Phys.* **25**, 2404–2417.
Miller, R. G. (1974), 'The jackknife – a review', *Biometrika* **61**, 1–15.
Mitler, K. S. (1974), *Multivariate Distributions*, John Wiley, New York.
Molina, E. C. (1931), 'Bayes' theorem, an expository presentation', *Bell Syst. Tech. Publ.* Monograph B-557.
Stands, with Keynes (1921), Jeffreys (1939) and Woodward (1953), as proof that there have always been lonely voices crying in the wilderness for a sensible approach to inference.
Moore, G. T. & Scully, M. O., eds. (1986), *Frontiers of Nonequilibrium Statistical Physics*, Plenum Press, New York.
Here several speakers affirmed their belief, on the basis of the Bell inequality experiments, that 'atoms are not real' while maintaining the belief that probabilities *are* objectively real! We consider this a flagrant example of the mind projection fallacy, carried to absurdity.
Munk, W. H. & Snodgrass, F. E. (1957), 'Measurements of Southern Swell at Guadalupe Island', *Deep-Sea Res.* **4**, 272–286.
This is the work which Tukey (1984) held up as the greatest example of his kind of spectral analysis, which could never have been accomplished by other methods; to which in turn Jaynes (1987) replied with chirp analysis.
Newton, Sir Isaac (1687) *Philosophia Naturalis Principia Mathematica*, trans. Andrew Motte, 1729; revised and reprinted as *Mathematical Principles of Natural Philosophy*, Florian Cajori, ed., University of California Press (1946).
See also Cajori (1928, 1934; both this bibliography).
Neyman, J. & Pearson, E. S. (1933), 'On the problem of the most efficient test of statistical hypotheses', *Phil. Trans. Roy. Soc.* **231**, 289–337.
Neyman, J. & Pearson, E. S. (1967), *Joint Statistical Papers*, Cambridge University Press.
Reprints of the several Neyman–Pearson papers of the 1930s, originally scattered over several different journals.
Neyman, J. (1959), 'On the two different aspects of representative method: the method of stratified sampling and the method of purposive selection', *Estadistica* **17**, 587–651.
Neyman, J. (1962) 'Two breakthroughs in the theory of statistical decision making', *Int. Stat. Rev.* **30**, 11–27.
It is an excellent homework problem to locate and correct the errors in this.
Neyman, J. (1981), 'Egon S. Pearson (August 11, 1895–June 12, 1980)', *Ann. Stat.* **9**, 1–2.
Novák, V. (1988), *Fuzzy Sets and their Applications*, A. Hilger, Bristol.
Nyquist, H. (1924), 'Certain factors affecting telegraph speed', *Bell Syst. Tech. J.* **3**, 324.
Nyquist, H. (1928), 'Certain topics in telegraph transmission theory', *Trans. AIEE*, **47**, 617–644.
O'Hagan, A. (1977), 'On outlier rejection phenomena in Bayes inference', *J. Roy. Stat. Soc.* **B 41**, 358–367.
Our position is that Bayesian inference has no pathological, exceptional cases and in particular no outliers. To reject any observation as an 'outlier' is a violation of the principles of rational inference, and signifies only that the problem was improperly formulated. That is, if you are able to decide that *any* observation is an outlier from the model that you specified, then that model does not properly capture your prior information about the mechanisms that are generating the data. In principle, the remedy is not to reject any observation, but to define a more realistic model (as we note in our discussion of Robustness). However, we concede that if the strictly correct procedure assigns a very low weight to the suspicious datum, its straight-out surgical removal from the data set may be a reasonable approximation, very easy to do.
Ore, O. (1953), *Cardano, the Gambling Scholar*, Princeton University Press.

Ore, O. (1960), 'Pascal and the invention of probability theory', *Am. Math. Monthly* **67**, 409–419.

Pearson, K. (1892), *The Grammar of Science*, Walter Scott, London.

Reprinted 1900, 1911 by A. & C. Black, London, and in 1937 by Everyman Press. An exposition of the principles of scientific reasoning; notably chiefly because Harold Jeffreys was much influenced by it and thought highly of it. This did not prevent him from pointing out that Karl Pearson was far from applying his own principles in his later scientific efforts. For biographical material on Karl Pearson (1857–1936) see Haldane (1957; this bibliography).

Pearson, K. (1905), 'The problem of the random walk', *Nature* **72**, 294, 342.

Pearson, K. (1921–33), *The History of Statistics in the 17'th and 18'th Centuries*, Pearson, E.S., ed., Lectures given at University College, London, Griffin, London (1978).

Penfield, W. (1958), *Proc. Natl Acad. Sci. (USA)* **44**, 59.

Accounts of observations made during brain surgery, in which electrical stimulation of a specific spots on the brain caused the conscious patient to recall various long-forgotten experiences. This undoubtedly true phenomenon is closely related to the theory of the A_p distribution in Chapter 18. But now others have moved into this field, with charges that psychiatrists are causing their patients – particularly young children – to recall things that never happened, with catastrophic legal consequences. The problem of recognizing valid and invalid recollections seems headed for a period of controversy.

Pierce, J. R. (1980) *Symbols, Signals, and Noise: An Introduction to Information Theory*, Dover Publications, Inc., New York.

An easy introduction for absolute beginners, but does not get to the currently important applications.

Poisson, S. D. (1837), *Recherches sur la Probabilité des Jugements*, Bachelier imprimeur-Libraire, Paris.

First appearance of the Poisson distribution.

Pólya, G. (1921), 'Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassennetz', *Math. Ann.* **84**, 149–160.

It is sometimes stated that this was the first appearance of the term 'random walk'. However, we may point to Pearson (1905) and Rayleigh (1919); both in this bibliography.

Pólya, G. (1923), 'Herleitung des Gauss'schen Fehlergesetzes aus einer Funktionalgleichung', *Math. Zeit.* **18**, 96–108.

Pontryagin, L. S. (1946), *Topological Groups*, Princeton University Press, Princeton, NJ.

Popper, K. (1958), *The Logic of Scientific Discovery*, Hutchinson & Co., London.

Denies the possibility of induction, on the grounds that the prior probability of every scientific theory is zero. Karl Popper is famous mostly through making a career out of the doctrine that theories may not be proved true, only false; hence the merit of a theory lies in its falsifiability. There is an evident grain of truth here, expressed by the syllogisms of Chapter 1; and Albert Einstein also noted this in his famous remark: '*No amount of experiments can ever prove me right; a single experiment may at any time prove me wrong.*' Nevertheless, the doctrine is true only of theories which assert the existence of unobservable causes or mechanisms; any theory which asserts observable facts is a counter-example to it.

Popper, K. (1963), *Conjectures and Refutations*, Routledge & Kegan Paul, London.

Popov, V. N. (1987), *Functional Integrals and Collective Excitations*, Cambridge University Press.

Sketches applications to superfluidity, superconductivity, plasma dynamics, superradiation, and phase transitions. A useful start on understanding of these phenomena, but still lacking any coherent theoretical basis – which we think is supplied only by the principle of maximum entropy as a method of reasoning.

Prenzel, H. V. (1975), *Dynamic Trendline Charting: How to Spot the Big Stock Moves and Avoid False Signals*, Prentice-Hall, Englewood Cliffs, NJ.

Contains not a trace of probability theory or any other mathematics: merely plot the monthly ranges of stock prices, draw a few straight lines on the graph, and their intersections tell you what to do and when to do it. At least this system does enable one to see the four-year presidential election cycle, very clearly.

Press, S. J. (1989), *Bayesian Statistics: Principles, Models and Applications*, J. Wiley & Sons, Inc., New York.

Contains a list of many Bayesian computer programs now available.

Preston, C. J. (1974), *Gibbs States on Countable Sets*, Cambridge University Press.

Here we have the damnable practice of using the word *state* to denote a probability *distribution*. One cannot conceive of a more destructively false and misleading terminology.

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, 2 vols., Academic Press, Inc., Orlando, FL; combined paperback edition with corrections (1983).

Puri, M. L., ed. (1975), *Stochastic Processes and Related Topics*, Academic Press, New York.

Quaster, H., ed. (1953), *Information Theory in Biology*, University Illinois Press, Urbana.

Ramsey, F. P. (1931), *The Foundations of Mathematics and Other Logical Essays*, Routledge and Kegan Paul, London.

Frank Ramsey was First Wrangler in Mathematics at Cambridge University in 1925, then became a Fellow of Kings College, where among other activities he collaborated with John Maynard Keynes on economic theory. He would undoubtedly have become the most influential Bayesian of the 20th century, but for the fact that he died in 1930 at the age of 26. In these essays one can see the beginnings of something very much like our exposition of probability theory.

Rayleigh, Lord (1919), 'On the problem of random vibrations, and of random flights in one, two or three dimensions', *Edin. & Dublin Phil. Mag. & J. Sci.* **37**, series 6, 321–347.

Reichardt, H. (1960), *C. F. Gauss–Leben und Werk*, Haude & Spener, Berlin.

Reid, C. (1970), *Hilbert*, Springer-Verlag, New York.

Reid, C. (1959), 'On a new axiomatic theory of probability', *Acta. Math. Acad. Sci. Hung.* **6**, 285–335.

This work has several things in common with ours, but expounded very differently.

Rihaczek, A. W. (1981), 'The maximum entropy of radar resolution', *IEEE Trans. Aerospace Electron. Syst.* **AES-17**, 144.

Another attack on maximum entropy, still denying the possibility of so-called 'super resolution', although it had been demonstrated conclusively in both theory and practice by John Parker Burg many years before, and was by 1981 in routine use by many scientists and engineers, as illustrated by the reprint collection of Childers (1978); see this bibliography.

Rissanen, J. (1983), 'A universal prior for the integers and estimation by minimum description length', *Ann. Stat.* **11**, 416–431.

One of the few fresh new ideas in recent decades. We think it has a bright future, but are not yet prepared to predict just what it will be.

Robbins, H. (1950), 'Asymptotically subminimax solutions of compound statistical decision problems', *Proceedings of the 2nd Berkeley Symposium of Mathematics Statistics and Probability*, University of California Press, pp. 131–148.

An anticipation of Stein (1956); see this bibliography.

Robbins, H. (1956), 'An empirical Bayes' approach to statistics', *Proceedings of the 3rd Berkeley Symposium on Mathematics, Statistics and Probability I*, University of California Press, pp. 157–164.

Robinson, A. (1966), *Non-standard Analysis*, North-Holland, Amsterdam.

How to do every calculation wrong.

Robinson, E. A. (1982), 'A historical perspective of spectrum estimation', *Proc. IEEE* **70**, 855–906.

Robinson, G. K. (1975), 'Some counterexamples to the theory of confidence intervals', *Biometrika* **62**, 155–162.

Rowlinson, J. S. (1970), 'Probability, information and entropy' *Nature* **225**, 1196–1198.

An attack on the principle of maximum entropy showing a common misconception of the nature of inference. Answered in Jaynes (1978).

Sampson, A. R. & Smith, R. L. (1984), 'An information theory model for the evaluation of circumstantial evidence,' *IEEE Trans. Systems, Man, and Cybernetics* **15**, 916.

Sampson, A. R. & Smith, R. L. (1982), 'Assessing risks through the determination of rare event probabilities', *Op. Res.* **30**, 839–866.

Sanov, I. N. (1961), 'On the probability of large deviations of random variables', IMS and AMS Translations of Probability and Statistics, from *Mat. Sbornik* **42**, 1144.

Scheffé, H. (1959), *The Analysis of Variance*, John Wiley, New York.

Schendel, U. (1989) *Sparse Matrices*, J. Wiley & Sons, New York.

Schlaifer, R. (1959), *Probability and Statistics for Business Decisions: An Introduction to Managerial Economics Under Uncertainty*, McGraw-Hill Book Company, New York.

An early recognition of the need for Bayesian methods in the real-world problems of decision; in striking contrast to the simultaneous Chernoff and Moses work (1959) on decision theory.

Schneider, T. D. (1991), 'Theory of molecular machines', *J. Theor. Biol.* **148**, 83–137.

In two parts, concerned with channel capacity and energy dissipation.

Schnell, E. E. (1960), 'Samuel Pepys, Isaac Newton and probability', *Am. Stat.* **14**, 27–30.

From this we learn that both Pascal and Newton had the experience of giving a correct solution and not being believed; the problem is not unique to modern Bayesians.

Schrödinger, E. (1945), 'Probability problems in nuclear chemistry', *Proc. Roy. Irish Acad.* **51**.

Schrödinger, E. (1947), 'The foundation of the theory of probability', *Proc. Roy. Irish Acad. (A)*, **51**, pp. 51–66, 141–146.

Valuable today because it enables us to add one more illustrious name to the list of those who think as we do. Here Schrödinger declares the 'frequentist' view of probability inadequate for the needs of science and seeks to justify the view of probability as applying to individual cases rather than 'ensembles' of cases, by efforts somewhat in the spirit of our Chapters 1 and 2. He gives some ingenious arguments but, unknown to him, these ideas had already advanced far beyond the level of his work. He was unaware of Cox's theorems and, like most scientists of that time with continental training, he had apparently never heard of Thomas Bayes or Harold Jeffreys. He gives no useful applications and obtains no theoretical results beyond what had been published by Jeffreys eight years earlier. Nevertheless, his thinking was aimed in the right direction on this and other controversial issues.

Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.

An attempt to develop a theory of two-valued probability, by a fanatical anti-Bayesian.

Shafer, G. (1982), 'Lindley's paradox', *J. Am. Stat. Assoc.* **77**, 325–334.

Apparently, Shafer was unaware that this was all in Jeffreys (1939, p. 194) some 20 years before Lindley. But Shafer's other work had made it clear already that he had never read and understood Jeffreys.

Shamir, A (1982), 'A polynomial time algorithm for breaking the basic Merkle–Hellman cryptosystem', in Chaum, D., Rivest, R. L. & Sherman, A. T., eds., *Advances in Cryptology: Proceedings of Crypto 82, 23–25 August 1982*, Plenum Press, New York, pp. 279–288.

Shaw, D. (1976), *Fourier Transform NMR Spectroscopy*, Elsevier, New York.

Sheynin, O. B. (1978), 'S. D. Poisson's work in probability', *Archiv. f. Hist. Exact Sci.* **18**, 245–300.

Sheynin, O. B. (1979), 'C.F. Gauss and the theory of errors', *Archiv. f. Hist. Exact Sci.* **19**, 21–72.

Shiryayev, A. N. (1978), *Optimal Stopping Rules*, Springer-Verlag, New York.

Siegmann, D. (1985) *Sequential Analysis*, Springer-Verlag, Berlin.

No mention of Bayes' theorem or optional stopping!

Simmons, G. J. (1979), 'Cryptography, the mathematics of secure communication', *The Math. Intelligencer* **1**, 233–246.

Sinai, J. G. (1982), *Rigorous Results in the Theory of Phase Transitions*, Akadémiai Kiado, Budapest.

Skilling, J., ed. (1989), *Maximum Entropy and Bayesian Methods*, Proceedings of the Eighth Maximum Entropy Workshop, Cambridge, UK, August 1988, Kluwer Academic Publishers, Dordrecht, Holland.

Smith, C. R. & Grandy, W. T., eds. (1985), *Maximum-Entropy and Bayesian Methods in Inverse Problems*, D. Reidel Publishing Co., Dordrecht, Holland.

Smith, C. R. & Erickson, G. J., eds. (1987), *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems*, D. Reidel Publishing Co., Dordrecht, Holland.

Smith, D. E. (1959), *A Source Book in Mathematics*, McGraw-Hill Book Co., New York.

Contains the Fermat–Pascal correspondence.

Smith, W. B. (1905), 'Meaning of the Epithet Nazorean', *The Monist* **15**, 25–95.

Concludes that prior to the Council of Nicæa, 'Nazareth' was not the name of a geographical place; it had some other meaning.

Sonett, C. P. (1982), 'Sunspot time series: spectrum from square law modulation of the Hale cycle', *Geophys. Res. Lett.* **9**, 1313–1316.

Spinoza, B. (1663), 'Renati des Cartes Principiorum philosophiae pars I, & II, more geometrico demonstratae,' *Ethics*, part 2, Prop. XLIV: 'De natura Rationis no est res, ut contingentes; sed, ut necessarias, contemplari.'

Translated, this proposition reads: 'It is not in the nature of reason to regard things as contingent; instead, they should be regarded as necessary.'

Spitzer, F. (1964), *Principles of Random Walk*, van Nostrand, New York.

Background history and present status.

Stein, C. (1945), 'A two sample test for a linear hypothesis whose power is independent of the variance', *Ann. Math. Stat.* **16**, 243–258.

Stein, C. (1956), 'Inadmissibility of the usual estimator for the mean of a multivariate normal distribution', *Proceedings of the 3rd Berkeley Symposium*, vol. 1, pp. 197–206, University of California Press.

First announcement of the 'Stein shrinking' phenomenon.

Stein, C. (1959), 'An example of wide discrepancy between fiducial and confidence intervals', *Ann. Math. Stat.* **30**, 877–880.

Stein, C. (1964), 'Inadmissibility of the usual estimate for the variance of a normal distribution with unknown mean', *Ann. Inst. Stat. Math.* **16**, 155–160.

Stein's inadmissibility discoveries, while shocking to statisticians with conventional training, are not in the least surprising or disconcerting to Bayesians. They only illustrate what was already clear to us: that the criterion of admissibility, which ignores all prior information, is potentially dangerous in real problems. Here that criterion can reject as 'inadmissible' what is in fact the optimal estimator, as noted briefly in Chapter 13.

Stigler, S. M. (1974a), 'Cauchy and the Witch of Agnesi', *Biometrika* **61**, 375–380.

Stigler, S. M. (1974b), 'Gergonne's 1815 paper on the design and analysis of polynomial regression experiments', *Historia Math.* **1**, 431–477.

Stigler, S. M. (1982a), 'Poisson on the Poisson distribution', *Stat. & Prob. Lett.* **1**, 33–35.

Stigler, S. M. (1982b), 'Thomas Bayes's Bayesian inference', *J. Roy. Stat. Soc.* **A145**, 250–258.

Stone, M. & Springer-Verlag, B. G. F. (1965), 'A paradox involving quasi prior distributions', *Biometrika* **52**, 623–627.

Stromberg, K. (1979), 'The Banach-Tarski paradox' *Am. Math. Monthly* **86**, 151–160.

Congruent sets stuff.

Student (1908), 'The probable error of a mean', *Biometrika* **6**, 1–24.

Takeuchi, K., Yanai, H. & Mukherjee, B. N. (1982), *The Foundations of Multivariate Analysis*, J. Wiley & Sons, New York.

Taylor, R. L., Daffer, P. Z. & Patterson, R. F. (1985), *Limit Theorems for Sums of Exchangeable Random Variables*, Rowman & Allanheld Publishers.

Presents the known versions of limit theorems for discrete-time exchangeable sequences in Euclidean and Banach spaces.

Thomas, M. U. (1979), 'A generalized maximum entropy principle', *Operations Res.* **27**, 1188–1195.

Tikhonov, A. N. & Arsenin, V. Y. (1977), *Solutions of Ill-posed Problems*, Halsted Press, New York.

A collection of *ad hoc* mathematical recipes, in which the authors try persistently to invert operators which have no inverses. Never perceives that these are problems of *inference*, not *inversion*.

Todhunter, I. (1865), *A History of the Mathematical Theory of Probability*, Macmillan, London; reprinted 1949, 1965, by Chelsea Press, New York.

Todhunter, I. (1873), *A History of the Mathematical Theories of Attraction and the Figure of the Earth*, 2 vols., Macmillan, London; reprinted 1962 by Dover Press, New York.

Toraldo di Francia, G. (1955), 'Resolving power and information', *J. Opt. Soc. Am.* **45**, 497–501.

One of the first recognitions of the nature of a generalized inverse problem; he tries to use information theory but fails to see that (Bayesian) probability theory is the appropriate tool for his problems.

Train, K. (1986), *Qualitative Choice Analysis Theory, Econometrics, and an Application to Automobile Demand*, MIT Press, Cambridge, MA.

Truesdell, C. (1987), *Great Scientists of Old as Heretics in 'The Scientific Method'*, University Press of Virginia.

The historical record shows that some of the greatest advances in mathematical physics were made with little or no basis in experiment, in seeming defiance of the 'scientific method' as usually proclaimed. This just shows the overwhelming importance of creative hypothesis formulation as primary to inference from given hypotheses. Unfortunately, while today we have a well-developed and highly successful theory of inference, we have no formal theory at all on optimal hypothesis formulation, and very few successful recent examples of it.

Tukey, J. W. (1960), 'A survey of sampling from contaminated distributions', in Olkin, I., ed., *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, Stanford University Press, California, pp. 448–485.

Tukey, J. W. (1962), 'The future of data analysis', *Ann. Math. Stat.* **33**, 1–67.

A potent object lesson for all who try to foretell the future as the realization of their own prejudices.

Tukey, J. W. (1978), 'Granger on seasonality', in Zellner, A., ed., *Seasonal Analysis of Time Series*, US Dept of Commerce, Washington.

An amusing view of the nature of Bayesian inference as a sneaky way of committing indecent methodological sins 'while modestly concealed behind a formal apparatus'.

Tukey, J. W. (1984), 'Styles of spectrum analysis', *Scripps Inst. Oceanography Ref.* Series 84–85, March, pp. 100–103.

A polemical attack on all theoretical principles, including autoregressive models, maximum entropy, and Bayesian methods. The 'protagonist of maximum entropy' who appears on p. 103 is none other than E. T. Jaynes.

Tukey, J. W., Bloomfield, P. Brillinger, D. & Cleveland, W. S. (1980), *The Practice of Spectrum Analysis*, University Associates, Princeton, New Jersey.

Notes for a course given in December 1980.

Tukey, J. W. & Brillinger, D. (1982), 'Spectrum estimation and system identification relying on a Fourier transform', unpublished.

This rare work was written as an invited paper for the IEEE Special Issue of September 1982 on Spectrum Analysis, but its length (112 pages in an incomplete version) prevented its appearing there. We hope that it will find publication elsewhere, because it is an important historical document. Tukey (1984; this bibliography) contains parts of it.

Valery-Radot, R. (1923), *The Life of Pasteur*, Doubleday, Page & Co., Garden City, New York.

Van Campenhout, J. M. & Cover, T. M. (1981), 'Maximum entropy and conditional probability', *IEEE Trans. Info. Theor.* **IT-27**, 483–489.

A rediscovery and generalization of what physicists have, since 1928, called 'the Darwin–Fowler method of statistical mechanics'.

van den Bos, A. (1971), 'Alternative interpretation of maximum entropy spectral analysis', *IEEE Trans. Info. Theor.* **IT-17**, 493–494.

Reprinted in Childers (1978; this bibliography). Expresses several misgivings about maximum entropy spectrum analysis, answered in Jaynes (1982; this bibliography).

Varian, H. (1978), *Microeconomic Analysis*, Norton & Co., New York.

Vasicek, O. (1980), 'A conditional law of large numbers', *Ann. Prob.* **8**, 142–147.

Wald, A. (1941), *Notes on the Theory of Statistical Estimation and of Testing Hypotheses*, Mimeographed, Columbia University.

At this time, Wald was assuring his students that Bayesian methods were entirely erroneous and incapable of dealing with the problems of inference. Nine years later, his own research had led him to the opposite opinion.

Wald, A. (1942), *On the Principles of Statistical Inference*, Notre Dame University Press.

Wald, A. (1943), 'Sequential analysis of statistical data: theory', Restricted report dated September 1943.

Waldmeier, M. (1961), *The Sunspot Activity in the Years 1610-1960*, Schulthes, Zürich.

Probably the most analyzed of all data sets.

Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, London.

Worried about improper priors, he introduces the notion of a 'near-ignorance class' (NIC) of priors. Since then, attempts to define precisely the NIC of usable priors have occupied many authors. We propose to cut all this short by noting that any prior which leads to a proper posterior distribution is usable and potentially useful. Obviously, whether a given improper prior does or does not accomplish this is determined not by any property of the prior alone, but by the joint behavior of the prior and the likelihood function; that is, by the prior, the model, and the data. Need any more be said?

Watson, J. D. (1968), *The Double Helix*, Signet Books, New York.

The famous account of the events leading to discovery of the DNA structure. It became a best seller because it inspired hysterically favorable reviews by persons without any knowledge of science, who were delighted by the suggestion that scientists in their ivory towers have motives just as disreputable as theirs. This was not the view of scientists on the scene with technical knowledge of the facts, one of whom said privately to the present writer: 'The person who emerges looking worst of all is Watson himself.' But that is ancient history; for us today, the interesting question is: would the discovery have been accelerated appreciably if the principles of Bayesian inference, as applied to X-ray diffraction data, had been developed and reduced to computer programs in 1950? We suspect that Rosalind Franklin's first '*A*-structure' photograph, which looks hopelessly confusing to the eye at first glance, if analyzed by a computer program (like those of Bretthorst (1988) but adapted to this problem), would have pointed at once to a double helix as overwhelmingly the most probable structure (at least, the open spaces which say '*helix*' were present and could be recognized by the eye after the fact). The problem is, in broad aspects, very much like that of radar target identification. For another version of the DNA story, with some different recollections of the course of events, see Crick (1988; this bibliography).

Wax, N., ed. (1954), *Noise & Stochastic Processes*, Dover Publications, Inc., New York.

Wehrl, A. (1978), 'General properties of entropy', *Rev. Mod. Phys.* **50**, 220–260.

Whittle, P. (1954), Comments on periodograms, Appendix to H. Wold (1954; this bibliography), pp. 200–227.

Whittle, P. (1957), 'Curve and periodogram smoothing', *J. Roy. Stat. Soc.* **B 19**, 38–47.

Whittle, P. (1958), 'On the smoothing of probability density functions', *J. Roy. Stat. Soc.* **B 20**, 334–343.

Wigner, E. P. (1967), *Symmetries and Reflections*, Indiana University Press, Bloomington.

From the standpoint of probability theory, the most interesting essay reprinted here is #15, 'The probability of the existence of a self-reproducing unit'. Writing the quantum-mechanical transformation from an initial state with (one living creature + environment) to a final state with (two identical ones + compatible environment), he concludes that the number of equations to be satisfied is greater than the number of unknowns, so the probability of replication is zero. Since the fact is that replication exists, the argument if correct would show only that quantum theory is invalid.

Wilbraham, H. (1854), 'On the theory of chances developed in Professor Boole's "Laws of Thought"', *Phil. Mag. Series 4*, **7**, (48), 465–476.

Criticism of Boole's version of probability theory.

Williams, P. M. (1980), 'Bayesian conditionalisation and the principle of minimum information', *Brit. J. Phil. Sci.* **31**, 131–144.

Wilson, A. G. (1970), *Entropy and Urban Modeling*, Pion Limited, London.

Wold, H. (1954), *Stationary Time Series*, Almquist and Wiksell, Stockholm.

Yockey, H. P. (1992), *Information Theory in Molecular Biology*, Cambridge University Press.

Zabell, S. L. (1982), 'W. E. Johnson's sufficientness postulate', *Ann. Stat.* **10**, 1091–1099.

Discussed in Jaynes (1986b).

Zabell, S. L. (1988), 'Buffon, Price, and Laplace: scientific attribution in the 18'th century', *Arch. Hist. Exact Sci.* **39**, 173–181.

Zellner, A. (1984), *Basic Issues in Econometrics*, University of Chicago Press.

A collection of 17 reprints of recent articles discussing and illustrating important principles of scientific inference. Like the previous reference, this is of value to a far wider audience than one would expect from the title. The problems and examples are stated in the context of economics, but the principles themselves are of universal validty and importance. In our view they are if anything even more important for physics, biology, medicine, and environmental policy than for economics. Be sure to read Chapter 1.4, entitled 'Causality and econometrics'. The problem of deciding whether a causal influence exists is vital for physics, and one might have expected physicists to have the best analyses of it. Yet Zellner here gives a far more sophisticated treatment than anything in the literature of physics or any other 'hard' science. He makes the same points that we stress here with cogent examples showing why prior information is absolutely essential in any judgment of this.

Zubarev, D. N. (1974) *Nonequilibrium Statistical Thermodynamics*, Plenum Publishing Corp., New York.

An amazing work; develops virtually all the maximum entropy partition functional algorithm as an *ad hoc* device; but then rejects the maximum entropy principle which gives the rationale for it and explains why it works! As a result, he is willing to use the formalism only for a tiny fraction of the problems which it is capable of solving, and thus loses practically all the real value of the method. A striking demonstration of how useful applications can be paralyzed – even when all the requisite mathematics is at hand – by orthodox conceptualizing about probability.