refsq, 6/15/1994

REFERENCES

A Partially Annotated Bibliography

Abel, N. H. (1826), "Untersuchung der Functionen zweier unabhängig veränderlichen Gröszen x und y, wie $f(x, y)$, welche die Eigenschaft haben, dasz $f[z, f(x, y)]$ eine symmetrische Function von z, x und y ist.", Jour. Reine u. angew. Math. (Crelle's Jour.), **1**, 11–15. First known instance of the associativity functional equation.

Abragam, A. (1961), *Principles of Nuclear Magnetism*, Oxford Science Publications, London. Apparently, he did not believe in phase coherences (as conveyed by off–diagonal elements of a density matrix) in quantum theory; so to predict the evolution of every new quantity, he was obliged to go into a new representation where that quantity is diagonal. But all this information was already present, independently of the representation, in a *single* density matrix with off–diagonal elements.

Aczel, J. (1966), *Lectures on Functional Equations and their Applications*, Academic Press, New York. See also Aczel (1987), *A Short Course on Functional Equations*, D. Reidel, Dordrecht–Holland.

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

Aitken, G. A. (1892), *The Life and Works of John Arbuthnot*, Clarendon Press, Oxford.

Akaike, Hirotugu (1980), "The Interpretation of Improper Prior Distributions as Limits of Data Dependent Proper Prior Distributions", Jour. Roy. Stat. Soc. **B42**, pp 46–52. A failed attempt to deal with the marginalization paradox, which never perceives that the paradox is just as present for proper priors as for improper ones. This has led others into much error and irrelevancy in dealing with time series. Discussed in Chapters 15 and 17.

Andrews, D. R. & Mallows, C. L. (1974), "Scale Mixtures of Normal Distributions", J. Roy Stat. Soc. **B 36**, 99–102.

Anscombe, F. J. (1963), "Sequential Medical Trials", JASA **58**, 365. Declares that the Sequential Analysis of Armitage (1960) is "a hoax".

Arbuthnot, John (1710). "An Argument for Divine Providence", Phil. Trans. Roy. Soc. **27**, pp. 186–190. Reprinted in Kendall & Plackett (1977), Vol 2, pp. 30–34. First known example of rejection of a statistical hypothesis on grounds of the improbability of the data. John Arbuthnot (1667–1735) was physician to Queen Anne and a prolific writer on many topics. Biographical information on Arbuthnot is given in Aitken (1892).

Archimedes (*ca.* 220 B. C.), Works, T. L. Heath, editor, Cambridge University Press, 1897, 1912. Paperback reprint by Dover Publications, Inc., undated (*ca.* 1960).

Aristotle (4th Century B. C.), Organon. Definition of syllogisms.

Aristotle, Physics, Trans. with commentary by H. G. Apostle, Indiana Univ. Press, Bloomington (1969).

Armitage, P. (1960), *Sequential Medical Trials*, Thomas, Springfield, Illinois. Second edition: Blackwell, Oxford, 1975. One of the main origins of the "optional stopping" controversy. See Savage (1962) and Anscombe (1963) for extensive discussion.

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

Bacon, F. (1620), "Novum Organum", in *The Works of Francis Bacon*, vol. 4, J. Spedding, R. L. Ellis & D. D. Heath, editors, 1857–58, London: Longman & Co.

Barber, N. F. & Ursell, F. (1948), "The Generation and Propagation of Ocean Waves and Swell", Phil. Trans. Roy. Soc. London, **A240**, pp. 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, N. Y.

Barnard, G. A. (1983), "Pivotal Inference and the Conditional View of Robustness (Why Have We for So Long Managed with Normality Assumptions?)", in Box, Leonard & Wu (1983). Expresses somewhat the same surprise at the success of the normal distribution as Augustus de Morgan did 145 years earlier. We try to explain this in Chapter 7.

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

Barr, A. & Feigenbaum, E., editors (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", Annals of Probability, Vol. 14, pp. 336–342.

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

Bayes, Rev. Thomas (1763), "An Essay Toward Solving a Problem in the Doctrine of Chances", Phil. Trans. Roy. Soc. pp. 370–418. Photographic reproduction in E. C. Molina (1963). Reprint, with biographical note by G. A. Barnard in Biometrika **45**, 293–313 (1958) and in Pearson & Kendall (1970). The date is slightly ambiguous; this work was read before the Royal Society in 1763, but not actually published until 1764. Further biographical information on Bayes is given by Holland (1962). Stigler (1983) and Zabell (1989) present evidence that Bayes may have obtained his result and communicated it to friends as early as 1748, in response to a challenge from David Hume (1711–1776).

Bell, E. T. (1937) *Men and Mathematics*, Dover Publications, Inc., N. Y. One needs to read this collection of biographical sketches because no substitute for it seems to exist; but let the reader be aware that Eric Temple Bell was also a well–known science fiction writer (under the pseudonym of John Taine) and this talent was not lost here. We can probably trust the accuracy of the names, dates, and documentable historical facts cited. But the other statements tell us very little about the matter under discussion; they tell us a great deal about the fantasies and socio–political views of the writer, and the level of his comprehension of technical facts. For example (p. 167) he *endorses*, on the grounds of "social justice" the beheading of Lavoisier. He makes blatantly false accusations against Laplace, and equally falsely, portrays Boole as a saint who could do no wrong. Displays (p. 256) a ridiculous misconception of the nature of Einstein's work, getting the sequence of facts backward. Tells us (p. 459) that Archimedes never cared for applications of mathematics!

Bellman, R. & Kalaba, R. (1957), "On the Role of Dynamic Programming in Statistical Communication Theory", I.R.E. Trans. PGIT–1, p. 197.

Benford, F. (1938), "The law of anomalous numbers", Proc. Amer. 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.

Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer–Verlag, New York.

Berger, J. O. & Wolpert, R. L. (1988), *The Likelihood Principle*, 2nd edition, Inst. Math. Stat.

Berkson, J. (1977), "My encounter with neo–Bayesianism", Internat. 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 J. D. Barra *et al*, eds, *Recent Developments in Statistics*, North Holland Press, Amsterdam, pp. 345–350.

________ (1979), "Reference posterior distributions for Bayesian inference", J. Roy. Statist. Soc. B **41**, 113–147 (with discussion).

________ (1979), "Expected information as expected utility", Ann. Statist. **7**, 686–690.

________ (1980), "A Bayesian Analysis of Classical hypothesis testing", in Bernardo, J. M. *et al* (1980), pp. 606–647, with discussion.

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

Bernardo, J. M., M. H . deGroot, D. V. Lindley, & A. F. M. Smith, Editors, (1985), *Bayesian Statistics 2*, Elsevier Science Publishers, North–Holland. Proceedings of the Second Valencia International Meeting on Bayesian Statistics, September 6–10, 1983.

Bernoulli, Daniel (1738), *Specimen theoriae novae de mensura sortis*, in Comm. Acad. Scient. Imp. Petropolitanae, Tomus V, pp. 175–192. English translation by Louise Sommer in Econometrica, **22**, 1954.

Bernoulli, Daniel (1777), Mem. St. Petersburg Acad., Acta Acad. Petrop. (1777), pp. 3-33. English translation,"The Most Probable Choice Between Several Discrepant Observations", Biometrika, **45**, pp. 1–18 (1961); reprinted in Kendall & Stewart (1970), pp. 157–167. Questions taking the mean of several observations; discussed in Chapter 7.

Bernoulli, James (1713), *Ars Conjectandi*, Thurnisiorum, Basel. Reprinted in *Die Werke von Jakob Bernoulli*, Vol. 3, Birkhaeuser, Basel (1975), pp. 107–286. First modern limit theorem. English translation of part IV (with the theorem) by Bing Sung, Harvard University Dept. of Statistics Technical Report #2, 1966.

Bertrand, J. L. (1889), *Calcul des probabilités*, Gauthier–Villars, Paris. 2nd. edition, 1907, reprinted (1972) by Chelsea Publishing Co., New York. This work is usually cited only for the "Bertrand Paradox" which appears on pp. 4–5; but it is full of neat, concise mathematics as well as good conceptual insight, both of which are often superior to the presentations in recent works. He understood clearly how much our conclusions from given data must depend on prior information, an understanding that was lost in the later "orthodox" literature. We quote him on this at the end of Chapter 6. However, we disapprove of his criticism of the Herschel–Maxwell derivation that we give in Chap. 7; what he saw as a defect is what we consider its greatest merit, and a forerunner of Einstein's reasoning. Well worth knowing and reviving today.

Billingsley, P. (1979), *Probability and Measure*, Wiley, New York. Contains Borel–Kolmogorov stuff.

Birnbaum, A. (1962), "On the foundation of statistical inference (with discussion)", J. Amer. Stat. Assoc. **57**, 269–326. The first proof of the "likelihood principle" to be accepted by anti–Bayesians.

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

Blackman, R. B. & Tukey, J. W. (1958), *The Measurement of Power Spectra*, Dover Publications, Inc., New York. Periodograms mutilated by *ad hoc* smoothing, which wipes out much of the useful information in them.

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

Bloomfield, P. (1976), *Fourier Analysis of Time Series: an Introduction*, J. Wiley & Sons, New York. Blackman–Tukey methods carried to absurd extremes (some 50 *db* below the noise level!!!)

on alleged astronomical data from Whittaker & Robinson (1924), which he fails to recognize as faked. As a result the bottom of the periodogram is giving zero information about variable stars; it is revealing only the spectrum of the digitizing errors. A potent demonstration of the folly of blind, unthinking application of a statistical procedure.

Boltzmann, Ludwig W. (1871), Wiener Berichte **63**, 397, 679, 712. First appearance of "$p \log p$" type entropy expressions.

Boole, G. (1854), *An Investigation of The Laws of Thought*, Macmillan, London; reprinted by Dover Publications, Inc., New York (1958).

———— (1857), "On the application of the theory of probabilities to the question of the combination of testimonies or judgments", Edinburgh Phil. Trans. v. xxi.

———— (1952–54), *Collected Logical Works*, Vol 1: Studies in Logic and Probability; Vol II: An Investigation of the Laws of Thought; Open Court.

Borel, E. (1909), *Élements de la théorie des probabilités*, Hermann et Fils, Paris. Discusses the Bertrand Paradox at some length and conjectures the correct solution, later found by group invariance arguments.

———— (1924), "Apropos of a Treatise on Probability", review of Keynes (1921). Reprinted in Kyburg & Smokler (1981). Borel, like Bertrand (1889), understood very well how strongly probabilities must depend on our state of prior knowledge. It is a pity that neither undertook to demonstrate the important consequences of this in realistic applications; they might have averted fifty years of false teaching by others.

———— (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.

Boring, E. G. (1955), "The Present Status of Parapsychology", Am. Scientist, **43**, 108–116. Concludes that "the curious phenomenon to be studied is the behavior of parapsychologists."

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

Bortkiewicz, L. V. (1898), *Das Gesetz der Kleinen Zahlen*, Teubner, Leipzig. Contains his famous fitting of the Poisson distribution to the number of German soldiers killed by the kick of a horse in successive years.

Bortkiewicz, L. V. (1913), *Die radioaktive Strahlung als Gegenstand Warscheinlichkeitstheoretischer Untersuchungen*, Berlin

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. & Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, Addison–Wesley, Reading MA. G. E. P. Box is, like L. J. Savage, a curious anomaly in this field; he was an assistant to R. A. Fisher and married his daughter, but became a Bayesian in issues of inference while remaining a Fisherian in matters of significance tests, which he held to be outside the ambit of Bayesian methods. In Jaynes (1985e) we argue that, on the contrary, any rational significance test *requires* the full Bayesian apparatus.

Box, G. E. P. (1982), "An Apology for Ecumenism in Statistics", NRC Technical Report #2408, Mathematics Research Center, University of Wisconsin, Madison. Also in Box, Leonard, & Wu (1983).

Box, G. E. P., Leonard, T. & Wu, C-F, editors (1983), *Scientific Inference, Data Analysis, and Robustness*, Academic Press, Inc., Orlando FL. Proceedings of a conference held in Madison, Wisconsin, November 1981.

Box, Joan F. (1978), *R. A. Fisher: The Life of a Scientist*, Wiley, New York. Joan Fisher Box, being the youngest daughter of R. A. Fisher, gives many personal anecdotes that nobody else could know, interspersed with accounts of the problems he worked on.

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 shall await the Sun's verdict with interest.

Bredin, Jean–Denis (1986), *The Affair: The case of Alfred Dreyfus*, G. Braziller, New York.

Bretthorst, G. L. (1987), "Bayesian Spectrum Analysis and Parameter Estimation", Ph.D. Thesis, Washington University, St. Louis, MO. Available from University Microfilms, Ann Arbor, Mich.

Bretthorst, G. L. (1988), *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Notes in Statistics, Vol. 48, Springer–Verlag, Berlin. A revised and expanded version of his thesis.

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. and Duren, B. (1986), "Information integration for decision support", Decision Support Systems, Vol. 4, No. 2, pp. 321–329.

Brown, R. (1828),"A brief account of microscopical observations", Edinburgh New Phil. Jour. **5**, 358–371. First report of the Brownian motion.

Buck, B. & Macaulay, V. A., editors (1991), *Maximum Entropy in Action*, Clarendon Press, Oxford. Eight Lectures given at Oxford University, covering introductory notions and applications in magnetic resonance, spectroscopy, plasma physics, X-ray crystallography, and thermodynamics. The best source to date for an Introduction elementary enough to be useful to beginners; yet proceeding to enough technical detail to be useful to practicing scientists. Be warned that what is called "Maximum Entropy" is in places distorted by *ad hoc* devices such as 'windowing' or 'prefiltering' the data – a practice that we condemn as destructive of some of the information in the data. Probability theory, correctly applied, is quite capable of extracting all the relevant information from the raw, unmutilated data and does best, with the least total computation, when it is allowed to do so freely.

Burg, John Parker (1967), "Maximum Entropy Spectral Analysis", in Proc. 37'th Meet. Soc. Exploration Geophysicists. Reprinted in Childers (1978).

Burg, John Parker (1975), "Maximum Entropy Spectral Analysis", Ph.D. Thesis, Stanford University.

Busnel, R. G. & Fish, J. F., editors (1980), *Animal Sonar Systems*; NATO ASI Series, Vol. A28, Plenum Publishing Corp., New York. A very large (1082 pp.) report of a meeting held at the Isle of Jersey, U. K. in 1979.

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

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.

Carnap, R. (1952), *The Continuum of Inductive Methods*, University of Chicago Press.

Chen, Wen-chen, & deGroot, M. H. (1986), "Optimal Search for New Types", in Goel & Zellner (1986), pp. 443–458.

Chernoff, H. & Moses, L. E. (1959), *Elementary Decision Theory*, J. Wiley & Sons, Inc., New York. When first issued, this work was described as "the only textbook on statistics that is not twenty years behind the times". It is now more than thirty years behind the times, because they could not accept the notion of a probability that is not a frequency, and so did not appreciate the fact that a straight Bayesian approach leads to all the same results with an order of magnitude less formal machinery. Still, it is an interesting and entertaining exposition of Wald's original ideas, far easier to read than Wald (1950).

Childers, D., editor (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", Mathematical Modelling, **1**, pp. 311–317. We have no quarrel with this work, but wish to add an historical footnote. 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. Stability considerations of multiple–valued "folded" functions of the kind associated today with the name of René Thom reduce 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., and Tukey, J. W. (1965), "An Algorithm for the Machine Calculation of Complex Fourier Series," Math, of Computation, **19**, pp. 297–301.

Cooley, J. W., Lewis, P. A. and Welch, P. D. (1967), "Historical Notes on the Fast Fourier Transform," Proc. IEEE **55**, pp. 1675–1677.

Copi, I. M. (1961), *Introduction to Logic*, Macmillan, N. Y. 2nd. edition, 1978.

Cournot, A. A. (1843), *Exposition de la theorie des chances et des probabilités*; L. Hachette, Paris. Reprinted (1984) in *Oeuvres complètes*, J. Vrin, Paris. One of the first of the attacks against Laplace, which were carried on by Ellis, Boole, Venn, Bell, and others to this day.

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 (1965). One prominent statistician, noting this, opined that Cox & Hinkley had "set statistics back 25 years."

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

Cox, R. T. (1946), "Probability, Frequency, and Reasonable Expectation", Am. Jour. Phys. **14**, 1–13. In our view, this article was the most important advance in the conceptual (as opposed to the purely mathematical) formulation of probability theory since Laplace.

——— (1961), *The Algebra of Probable Inference*, Johns Hopkins University Press, Baltimore MD, Baltimore, MD, 1961. An extension of the 1946 article, with additional results and more discussion. Reviewed by E. T. Jaynes, Am. Jour. Phys. **31**, 66 (1963).

——— (1978), "Of Inference and Inquiry", in *The Maximum Entropy Formalism*, R. D. Levine & M Tribus, eds., M.I.T. 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. and Zahner, M. J. (1973), "The maximum–entropy distribution of the future market price of a stock", Operations Research, 21, 1200–1211.

Craig, John (1699), *Theologiae Christianae Principia Mathematica*, Timothy Child, London. Reprinted with commentary by Daniel Titus, Leipzig (1755). See also Stigler (1986) for more comments.

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press. This marks the heyday of supreme confidence in confidence intervals over Bayesian methods, asserted as usual on purely ideological grounds, taking no note of the actual results that the two methods yield. Comments on it are in Jaynes (1976, 1986a).

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

Crick, Francis (1988), *What Mad Pursuit*, Basic Books, Inc., New York. A reminiscence of his life and work, full of important observations and advice about the conduct of science in general, and fascinating technical details about his decisively important work in biology – most of which occurred several years after the famous Crick – Watson discovery of the DNA structure. Almost equally important, this is an antidote to Watson (1968); we have here the other side of the DNA Double Helix story as Crick recorded it in 1974, with a different recollection of events. From our viewpoint, this work is valuable as a case history of important scientific discoveries made without help of probabilistic inference in our mathematical form, but – at least in Crick's mind – obeying its principles strictly, in the qualitative form given by Pólya. We wish that theoretical physicists reasoned as well.

Crow, E. L., Davis, J. A. & Maxfield, M. W. (1960), *Statistics Manual*, Dover Publications, Inc., New York. Has many useful tables and graphs, but expounds straight orthodox methods, never thinking in terms of information content, and therefore never perceiving their weakness in extracting information from the data. We have some fun with it in Jaynes (1976).

Csiszar, I., "Sanov property, generalized I–projection and a conditional limit theorem", Annals of Probability, Vol. 12, pp. 768–793, 1984

Currie, R. G. & Hameed, S. (1986), "Climatically Induced Cyclic Variations in United States Corn Yield and Possible Economic Implications," presented at the Canadian Hydrology Symposium, Regina, Sask.

Czuber, E. (1908), *Wahrscheinlichkeitsrechnung und Ihre Anwendung auf Fehlerausgleichung*, Teubner, Berlin; 2 Vols. Some of Wolf's famous dice data may be found here.

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

Dale, A. I. (1982), "Bayes or Laplace? An Examination of the Origin and Early Applications of Bayes' Theorem", Archiv. f. Hist. of Exact Sciences **27**, pp. 23–47.

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

Daniell, G. J. & Potton, J. A. (1989), "Liquid Structure Factor Determination by Neutron Scattering – Some Dangers of Maximum Entropy", in Skilling (1989), 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.

Dawid, A. P., Stone, M. & Zidek, J. V. (1973), "Marginalization Paradoxes in Bayesian and Structural Inference", J. Roy Stat. Soc. **B35**, pp. 189–233.

Dawkins, R. (1987), *The Blind Watchmaker*, W. W. Norton & Co., New York. A kind of companion piece to Gardner (1981), which concentrates on answering the unceasing attacks on Darwin's theory, by religious fundamentalists who do not understand what Darwin's theory is. Richard Dawkins, Professor of Zoology at Oxford University, goes patiently into much detail to explain why the facts of Nature can be accounted for as the operation of Natural Law, with no need to invoke teleological purpose. Unfortunately, he is not the world's best explainer. While we agree with essentially everything he is *trying* to say, his arguments lack persuasiveness through being obscure, through long–winded irrelevancies; and through his own failure to understand some technical facts and principles of other sciences. For example, there is a strange misapplication of probability theory on p. 159; let it be an exercise for the reader to recalculate the probability of his being struck by lightning in any one minute of his life.

de Finetti, B. (1937), "La prevision: ses lois logiques, ses sources subjectives", Ann. Inst. H. Poincaré, **7**, pp. 1–68. English Translation: "Prevision, its Logical Laws, its Subjective Sources", in Kyburg & Smokler (1981).

_________ (1958), "Foundations of Probability" in *Philosophy in the Mid–century*, La Nuova Italia Editrice, Florence, pp. 140–147.

_________ (1974a), "Bayesianism", Intern. Stat. Rev. **42**, 117–130.

_________ (1974b), *Theory of Probability*, 2 Vols. J. Wiley & Sons, Inc., New York. Adrian Smith's English translation could not hide the wit and humor of this work. Bruno de Finetti was having great fun writing it; but he could scarcely write two sentences without injecting some parenthetic remark about a different topic, that suddenly popped into his mind, and this is followed faithfully in the translation. Full of interesting information that all serious students of the field ought to know; but impossible to summarize, because of its chaotic disorganization. Discussion of any one topic may be scattered over a half–dozen different Chapters without cross–references, so one may as well read the pages at random.

de Groot, M. H. (1970), *Optimal Statistical Decisions*, McGraw–Hill Book Co., New York.

de Groot, M. H. (1986), *Probability and Statistics*, 2nd edition, Addison–Wesley Publishing Co., Reading MA.

de Groot, M. H., Bayarri, M. J., & Kadane, J. B. (1988), "What is the Likelihood Function?" (with discussion). In *Statistical Decision Theory and Related Topics IV*, S. S. Gupta & J. O. Berger, editors, Springer, 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*, Wiley, New York.

de Groot, M. H. & Goel, Prem (1980), "Only Normal Distributions have Linear Posterior Expectations in Linear Regression", J. Am. Stat. Assoc. **75**, 895–200. Still another connection of the kind first found by Gauss (1809) and discussed in Chapter 7.

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

de Moivre, A. (1718), *The Doctrine of Chances: or, A Method of Calculating the Probability of Events in Play*, W. Pearson, London. 2nd edition, Woodfall, London (1738), 3rd edition, Millar, London (1756); reprinted by Chelsea Publishing Co., New York (1967).

de Moivre, A. (1733), *Approximatio ad Summam Terminorum Binomii* $(a+b)^n$ *in Seriem expansi.* Photographic reproduction in Archibald, R. C. (1926), Isis **8**, 671–683.

de Morgan, Augustus (1838), *An Essay on Probabilities*, Longman & Co., London.

——— (1847), *Formal Logic: or the Calculus of Inference Necessary and Probable*, Taylor & Watton, London. An enthusiastic exposition of Laplace's views.

——— (1872), *A Budget of Paradoxes*, 2 Vols. Sophia de Morgan, editor, London. 2nd edition, D. E. Smith, editor (1915); reprinted as one Volume by Dover Publications, Inc. (1954). Augustus de Morgan (1806–1871) was a mathematician and logician, at University College, London from 1828–1866. He collected notes concerning not only logic, but anomalies of logic; the latter are preserved in this delightful account of the activities of circle–squarers, anti–Copernicans, anti–Newtonians, religious fanatics, numerologists, and other demented souls that abounded in 19'th Century England. It gives a vivid picture of the difficulties that serious scholars had to overcome in order to make any forward progress in science. An inexhaustible supply of amusing anecdotes.

de Morgan, Sophia (1882), *Memoir of Augustus de Morgan*, Longman, Green, London. Further biographical and anecdotal material on de Morgan.

Dempster, A. P. (1963), "On a paradox concerning inference about a covariance matrix", Ann. Math. Stat. **34**, 1414–18.

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

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

Dyson, F. J. (1958), Review of Lighthill (1957), Physics Today **11**, 28.

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

——— (1974), "The History of Likelihood", Int. Stat. Rev. **42**, 9–15.

——— (1992), Nature, **352**, pp. 386–387. Commentary on Bayesian methods.

Edwards, H. M. (1987), "An Appreciation of Kronecker", Math, Intelligencer, **9**, pp. 28–35.

——— (1988), "Kronecker's Place in History", in *History and Philosophy of Modern Mathematics*, W. Aspray & P. Kitcher, editors, University of Minnesota Press, Minneapolis.

——— (1989), "Kronecker's Philosophical Views", in Rowe & McCleary (1989); Vol 1, pp. 67–77.

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

——— (1978), "Controversies in the foundations of statistics", Am. Math. Monthly, **85**, 231–246.

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

——— (1979b), "Computers and the theory of statistics: Thinking the unthinkable", SIAM Review, October.

Efron, B. & G. Gong (1983), "A Leisurely Look at the Bootstrap, the Jackknife, and Cross–Validation", Am. Stat. **37**, pp. 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.

Ellis, R. L. (1842) "On the Foundations of the Theory of Probability", Camb. Phil. Soc. vol. viii. Reprinted in Ellis (1863). Ellis was the British Counterpart of Cournot, in starting the anti–Laplace movement which set scientific inference back a Century.

Ellis, R. L. (1863), *The Mathematical and Other Writings of Robert Leslie Ellis M. A.*, Wm. Walton, editor, Deighton, Bell, Cambridge.

Erickson, G. J. & Smith, C. Ray (1988), editors; *Maximum–Entropy and Bayesian Methods in Science and Engineering*, Vol 1, Foundations; Vol. 2, Applications, Kluwer Academic Publishers, Dordrecht–Holland.

Euler, Leonhard (1749), *Recherches sur la question des inégalitiés du mouvement de Saturne et de Jupiter, sujet proposé pour le prix de l'anneé 1748 pas l'Académie royale des sciences de Paris.* Reprinted in *Leonhardi Euleri, Opera Omnia*, ser. 2, Vol. 25, Turici, Basel, (1960). Euler gave up at the problem of estimating 8 unknown parameters from 75 discrepant observations, but won the prize anyway.

Evans, M. (1969) *Macroeconomic Forecasting*, Harper & Row, N. Y.

Fechner, G. J. (1860), *Elemente der Psychophysik*, two Vols. Volume 1 translated as *Elements of Psychophysics*, E. G. Boring & D. H. Howes, editors, Holt, Rinehart & Winston, New York (1966).

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

Feinberg, S. E. & Hinkley, D. V. (1990), *R. A. Fisher: An Appreciation*, Lecture Notes in Statistics #1, Springer–Verlag, Berlin. This is the second printing of the work, which appeared originally in 1979. A valuable source, if it is regarded as an historical document rather than an account of present statistical principles. Rich in technical details of his most important derivations and gives a large bibliography of his works, including four books and 294 published articles. But in its adulation of Fisher it fails repeatedly to note something that was already well established in 1979: the simpler and unified methods of Jeffreys, which Fisher rejected vehemently, actually accomplished everything that Fisher's methods did, with the same or better results and almost always more easily. In addition, they deal easily with technical difficulties (such as lack of sufficient or ancillary statistics) which Fisher was never able to overcome. Thus this work tends also to perpetuate harmful myths.

Feinstein, A. (1958), *Foundations of Information Theory*, McGraw–Hill, New York. Like the work of Khinchine (1957), a mathematician's view of things, which has almost nothing in common with the physically oriented view of Goldman (1953).

Félix, Lucienne (1960), *The Modern Aspect of Mathematics*, Basic Books, Inc., New York. A Bourbakist view; for the contrary view see Kline (1980).

Feller, W. (1950), *An Introduction to Probability Theory and its Applications, Volume 1*, J. Wiley & Sons., New York. 2nd edition, 1957; 3rd edition, 1968.

Feller, W. (1966), *An Introduction to Probability Theory and its Applications, Volume 2*, J. Wiley & Sons., New York. Second edition, 1971.

Ferguson, T. S. (1982), "An inconsistent Maximum Likelihood estimate", J. Am. Stat. Ass'n **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) 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). For us, it is a straightforward exercise for the reader in our Chapter on Estimation with a Gaussian distribution.

Fine, T. L. (1973), *Theories of Probability*, Academic Press, N. Y.

Fischer, E. P. & Lipson, C. (1988), *Thinking About Science: Max Delbrück and the Origins of Molecular Biology*, Norton, New York. For some time we have seen Max Delbrück referred to as "one of Niels Bohr's greatest and most successful students". It is true that he has played a very important role in the modern development of biology; yet what has emerged was nearly the opposite of his intentions. His original goal – to inculcate the ideas of the Copenhagen interpretation of quantum theory into biology and to learn new principles of physics from biology – has been a failure, all the new developments involving definite, reliable mechanisms that would be understood at once in a machine shop. The role of "quantum effects" in biology seems limited to their role in all of chemistry: to explain the stability of molecules. The uncertainty principle has as yet found no functional role at all in biology, nor have any new physical principles emerged; and we predict with confidence that this will continue to be true.

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

——— (1915), "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population", Biometrika, **10**, pp. 507–521.

——— (1922), "On mathematical foundations of theoretical statistics", Phil. Trans. Roy. Soc. (London), Ser. A, **222**, 309–368. Introduction of the term "sufficient statistic."

——— (1925), *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh. Twelve later editions, to the one by Hafner Publishing Co., New York (1973).

——— (1930a), "Inverse Probabilities", Proc. Camb. Phil. Soc., **26**, 528–535.

——— (1930b), *The Genetical Theory of Natural Selection*, Oxford University Press. Second revised Edition by Dover Publications, Inc., New York (1958). Here Fisher shows that Mendelian genetics is not in conflict with Darwinian evolution theory, as Mendelians supposed in the early 20'th Century; on the contrary, the 'particulate' or 'discrete' nature of Mendelian inheritance clears up some outstanding difficulties with Darwin's theory, resulting from the assumption of blending inheritance which most biologists – including Darwin himself – took for granted in the 1860's. Recall that Mendel's work, with its lore of dominant and recessive genes, *etc.*, was later than Darwin's; but Darwin (1809 – 1882) never knew of it and it was not generally known until about 1900. The reinterpretation of Darwin's theory in these terms, by Fisher and others, is now known as Neo–Darwinism. By the time of Fisher's second (1958) edition the existence of mutations caused by radioactivity was well established, those caused by failures of DNA replication had become highly plausible, and genetic recombination (which had been suggested by August Weismann as early as 1886) was recognized as still another mechanism to provide the individual variations – on which Natural Selection feeds, but whose origin was puzzling to Darwin. Fisher added many new paragraphs, in smaller type, pointing out this newer understanding and its implications; how Darwin would have enjoyed seeing these beautiful solutions to his problems! Fisher's real, permanent contributions to science are in works like this, not in his statistical teachings, which were an advance in the 1920's, but have been a retarding force since the 1940's.

——— (1933), "Probability, Likelihood and Quantity of Information in the Logic of Uncertain Inference", Proc. Roy. Soc. **146**, pp. 1–8. A famous attempt to demolish Jeffreys' work, which we discuss in Chapter 16.

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

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

——— (1950), *Contributions to Mathematical Statistics*, W. A. Shewhart, ed., J. Wiley & Sons, Inc., New York. A collection of his best known early papers.

——— (1956), *Statistical Methods and Scientific Inference*, Oliver & Boyd, London. Second

Revised Edition, Hafner Publishing Co., New York, 1959. Fisher's final book on statistics, in which he tries to sum up his views of the logical nature of uncertain inference. One discerns a considerable shift of position from his earlier works – even admitting, occasionally, that he had been wrong before. He is now more sympathetic toward the role of prior information, saying that recognizable subsets should be taken into account and that prior ignorance is essential for the validity of fiducial estimation. He shows his old power of intuitive insight in his neat explanation of Gödel's theorem, but also some apparent lapses of memory and numerical errors. Every serious student of the subject should read this work slowly and carefully at least twice, because the depth of thinking is often so great that his meaning will not be grasped fully on a single reading. Also, Fisher goes into several specialized topics that we do not discuss in the present work.

________ (1974), *Collected Papers of R. A. Fisher*, J. H. Bennett, editor, University of Adelaide, Australia; Coudrey Offset Press.

________ (1962), "Some Examples of Bayes' method of the Experimental Determination of Probability a priori", J. Roy. Stat. Soc., **B 24**, 118–124.

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.

Fougere, P. F. (1977), J. Geophys. Res. **82**, 1051–1054. Maximum Entropy Spectrum Analysis.

Fraser, D. A. S. (1980), Comments on a paper by B. Hill, in *Bayesian Statistics*, J. M. Bernardo *et al*, editors, University Press, Valencia, Spain, pp. 56–58. Claims to have a counter–example to the likelihood principle. But it is the same as the tetrahedron problem discussed in Chapter 15 above; the correct solution to that problem was not known in 1980.

Galileo Galilei (1638), *Dialogues Concerning Two New Sciences*, Elzevir Press, Holland. English Translation by Henry Crew & Alfonso de Salvio, MacMillan Company, London (1914). Paperback reprint by Dover Publishing Co., undated (*ca.* 1960).

Galton, F. (1863), *Meteorographica*, London; MacMillan. Here this remarkable man invents weather maps and from studying them discovers the "anticyclone" circulation patterns in the northern hemisphere.

________ (1886), "Family Likeness in Stature", Proc. Roy. Soc. London **40**, pp. 42–73.

________ (1889), *Natural Inheritance*, MacMillan, London.

________ (1908), *Memories of My Life*, Methuen, London. More biographical and technical details are in Pearson (1914–1930).

Gardner, M. (1957), *Fads and Fallacies in the Name of Science*, Dover Publications, Inc. A kind of 20'th Century sequel to de Morgan (1872), with attention directed more to fakers in science than to their colleagues in mathematics. Here we meet both the sincere but tragically misguided souls, and the deliberate frauds out to make a dishonest dollar from the gullible.

Gardner, M. (1981), *Science – Good, Bad, and Bogus*, Paperbound edition (1989), Prometheus Books, Buffalo N. Y. A sequel to the previous work, with a sobering message that everyone ought to note. Particular details on several recent trends; the Creationist who utilizes TV to carry attacks on Darwin's theory to millions, while grossly misrepresenting what Darwin's theory is; the ESP advocate who invades scientific meetings to try to invoke Quantum Theory in his support, although he has no comprehension of what Quantum Theory is; the Gee Whiz publicist who turns every tiny advance in knowledge (artificial intelligence, chaos, catastrophe theory, fractals) into a revolutionary crusader cult; the professional Disaster Monger who seeks personal publicity through inventing ever more ridiculous dangers out of every activity of Man; and most frightening of all, the eagerness with which the news media give instant support and free publicity to all this. Today, our airwaves are saturated with bogus science and medieval superstitions belittling and misrepresenting real, responsible science. In the Introduction, Gardner documents the indignant refusal of network executives to correct this, on grounds of its profitability. Then at what point does persistent, deliberate abuse of freedom of speech for profit become a clear and present danger to society? See also Rothman (1989); Huber (1992).

Gauss, K. F. (1809) *Theoria motus corporum celestium*, Perthes, Hamburg. English translation, *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*, Dover Publications, Inc., New York (1963).

Gauss, K. F. (1823) *Theoria combinationis observationum erroribus minimis obnoxiae*; also *Supplementum*, 1826; Dieterich, Göttingen.

Geisser, S. & Cornfield, J. (1963), "Posterior Distribution for Multivariate Normal Parameters," J. Roy. Stat. Soc., **B25**, pp. 368–376. Gives the correct treatment of a problem which was later corrupted into the Marginalization paradox, as explained in Jaynes (1983), pp. 337–339, 374.

Geisser, S. (1980), "The Contributions of Sir Harold Jeffreys to Bayesian Inference", in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner, editor, North–Holland Publishing Co., Amsterdam; pp 13–20..

Gell–Mann, M. (1992) "Nature Conformable to Herself", Bulletin of the Santa Fe Institute, **7**, pp. 7–10. Some comments on the relation between mathematics and physics; like us, he is happy that the plague of Bourbakism is finally disappearing, raising the hope that mathematics and physics may once more become mutually helpful partners instead of adversaries.

Gentleman, W. M. (1968), "Matrix Multiplication and Fast Fourier Transformations," Bell Syst. Tech. Journal, **17** pp. 1099–1103.

Gibbs, J. Willard (1875), "On the Equilibrium of Heterogeneous Substances" reprinted in *The Scientific Papers of J. Willard Gibbs*, Vol. I, Longmans, Green & Co., 1906 and by Dover Publications, Inc., 1961.

Gibbs, J. Willard (1902), *Elementary Principles in Statistical Mechanics*, Yale University Press, New Haven, Connecticut. Reprinted in *The Collected Works of J. Willard Gibbs*, Vol. 2, by Longmans, Green & Co. (1928) and by Dover Publications, Inc., New York (1960).

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

Glymour, C. (1985), "Independence Assumptions and Bayesian Updating", Artificial Intell. **25**, 25–99.

Gndenko, 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.

Gödel, K. (1931), "Über formal unendscheidbare Sätze der Principia Mathematica und verwandter Systeme I", Monatshefte für Mathematik und Physik, **38**, p. 173–198. English translation, "On formally undecidable propositions of Principia Mathematica and related Systems", Basic Books, Inc., New York (1962); Reprinted by Dover Publications, Inc., New York (1992).

Goel, P. & Zellner, A. (1986), editors, *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.

Goldman, S. (1953), *Information Theory*, Prentice–Hall, Inc., New York. We would like to put in a friendly plug for this work, even though it has a weird reputation in the field. The author, in recounting the work of Norbert Wiener and Claude Shannon, explains it for the benefit of beginners much more clearly than Wiener did, and somewhat more clearly than Shannon. Its weirdness is the result of two unfortunate accidents: (1) a misspelled word in the title of Chapter 1 escaped both the author and the publisher, providing material for dozens of cruel jokes circulating in the 1950's; (2) on p. 295 there is a photograph of Gibbs, with the caption: "J. Willard Gibbs (1839–1903), whose ergodic hypothesis is the forerunner of fundamental ideas in information theory." Since Gibbs never mentioned ergodicity, this is a source of more jokes. However, the author is guilty only of trusting the veracity of Wiener (1948).

Good, I. J. (1950), *Probability and the Weighing of Evidence*, C. Griffin & Co., London. A work whose importance is out of all proportion to its small size. Still required reading for every student of Scientific Inference; and can be read in one evening.

———— (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.

———— (1967) "The White Shoe is a Red Herring", BJPS 17, 322, reprinted in Good (1983). Points out the error in the Hempel paradox.

———— (1980), "The contributions of Jeffreys to Bayesian Statistics", in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner, editor, North–Holland Pub. Co., Amsterdam.

———— (1983), *Good Thinking*, University of Minnesota Press. Reprints of 23 articles, scattered over many topics and many years, plus a long bibliography of other works. There are about 2000 short articles like these by Good, found throughout the statistical and philosophical literature starting in 1940. Workers in the field generally granted that every idea in modern statistics can be found expressed by him in one or more of these articles; but their sheer number made it impossible to find or cite them, and most are only one or two pages long, dashed off in an hour and never developed further. So for many years, whatever one did in Bayesian statistics, one just conceded priority to Jack Good by default, without attempting the literature search for the relevant article, which would have required days. Finally, this book provided a bibliography of most of the first 1517 of these articles (presumably in the order of their writing, which is not the order of publication) with a long index, so it is now possible to give proper acknowledgment of his works up to 1983. Be sure to read Chapter 15, where he points out specific, quantitative errors in Karl Popper's work and demonstrates that Bayesian methods, which Popper rejects, actually correct those errors.

Gould, Stephen Jay (1989), *Wonderful Life: The Burgess Shale and the Nature of History*, W. W. Norton & Co., New York. A tiny region in the Canadian Rockies had exactly the right geological history so that soft–bodied animals were preserved almost perfectly. As a result we now know that the variety of life existing in Cambrian time was vastly greater than had been supposed; this has profound implications for our view of evolution. Gould seems fanatical in his insistence that 'evolution' is not synonomous with 'progress'. Of course, anyone familiar with the principles of physics and chemistry will agree at once that a process that proceeded in one direction can also proceed in the opposite one. Nevertheless, it seems to us that at least 99% of observed evolutionary change *has in fact* been in the direction of progress (more competent, adaptable creatures). We also think that Darwinian theory, properly stated in terms conforming to present basic knowledge and present Bayesian principles of reasoning, predicts just this.

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

Graunt, J. (1662), *Natural and Political Observations made upon the Bills of Mortality*, Roycroft, London. Reprinted in *The World of Mathematics*, J. R. Newman, editor, Simon & Schuster, New York (1956); Vol. 3, pp. 1420–1435. First recognition of the useful facts that can be inferred from records of births and deaths; the beginning of sociological inference, as distinguished from the mere collection of statistics. This work is sometimes attributed instead to William Petty; for details see Greenwood (1942).

Greenwood, Major (1942), "Medical Statistics from Graunt to Farr", Biometrika, **32**, pp. 203–225; Part 2 of a three–part work. A lengthy but confusingly disorganized account of John Graunt (1620–1674), William Petty (1623–1687), and Edmund Halley (1656–1742) in the matter of the first mortality tables. Petty (friend of Graunt and one of those restless but undisciplined minds, which dabbles for a short time in practically everything but never really masters anything) attempted to make a survey of Ireland many years before Halley, but did not reason carefully enough to produce a meaningful result. Greenwood ends in utter confusion over whether Petty is or is not the real author of Graunt's book, apparently unaware that Petty's connection is that he edited the fifth (posthumous) edition of Graunt's work; and it was Petty's edition that Halley referred to and saw how to correct. All this had been explained long before, with amusing sarcasm, by Augustus de Morgan (1872, I, 113–115).

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

Griffin, D. R. (1958), *Listening in the Dark*, Yale University Press, New Haven; see also *About Bats*, R. H. Slaughter & D. W. Walton, editors, SMU Press, Dallas, Texas (1970).

Grosser, M. (1979), *The Discovery of Neptune*, Dover Publications, Inc.

Gull, S. F. & Daniell, G. J. (1978), "Image Reconstruction from Incomplete and Noisy Data", Nature, **272**, p. 686.

Gull, S. F. & Daniell, H. J. (1980), "The Maximum Entropy Algorithm Applied to Image Enhancement", Proc. IEEE (E), **5**, p. 170.

Gull, S. F. & Skilling, J. (1984), "The Maximum Entropy Method", in *Indirect Imaging*, J. A. Roberts, editor, Cambridge University Press, U. K.

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

——— (1973), *The Emergence of Probability*, Cambridge Univ. Press.

——— (1984), "Historical Models for Justice", Epistemologia, VII, Special Issue on Probability, Statistics, and Inductive Logic, pp. 191–212.

Haldane, J. B. S. (1932), Proc. Camb. Phil. Soc., **28**, 58. Improper priors advocated and used. Harold Jeffreys (1939) acknowledges this as the source of some of his own ideas on them.

Haldane, J. B. S. (1957), "Karl Pearson, 1857 – 1957", Biometrika **44**, 303–313. Haldane's writings, whatever the 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 Lenin quoted approvingly from Karl Pearson. Haldane was Professor of Genetics at University College, London in the 1930's, 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 politics instead of genetics, hoped to bring about.

Halley, Edmund (1693), "An estimate of the degrees of mortality of mankind —", Phil. Trans. Roy. Soc. **17**, 596–610, 654–656. Reprinted in *The World of Mathematics*, J. R. Newman editor, Simon

& Schuster, New York (1956); Vol. 3, pp. 1436–1447. First mortality table, based on records of births and deaths in Breslau, 1687–1691. See also Greenwood (1942).

Hamilton, A. G. (1988), *Logic for Mathematicians*, Cambridge Univ. Press, Revised 2nd edition.

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 Univ. Press.

Hansel, C. E. M. (1980), *ESP and Parapsychology – a Critical Re–evaluation*, Prometheus Books, Buffalo N. Y.; Chap. 12

Heath, D. & Sudderth, Wm. (1976), "de Finetti's Theorem on Exchangeable Variables", The American Statistician, **30**, 188. An extremely simple derivation.

Hedges, L. V. & Olkin, I., (1985), *Statistical Methods for Meta– Analysis*, Academic Press, Inc., Orlando FL.

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

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

Hempel, C. G. (1967), BJPS **18**, 239–240. Reply to Good (1967).

Herschel, J. (1850), Edinburgh Review **92**, 14. A two–dimensional "Maxwellian velocity distribution" before Maxwell (1860).

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, Francis W. (1926), *Life and Letters of Thomas Jefferson*, New York.

Hobson, A. and Cheung, B. K. (1973), "A comparison of the Shannon and Kullback information measures", Jour. 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. Amer. 16–26 (May 1983).

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

Holland, J. D. (1962), "The Reverend Thomas Bayes F.R.S. (1702–1761)", J. Roy. Stat. Soc. (A), **125**, 451–461.

Howson, C. & Urbach, P. (1989), *Scientific Reasoning: The Bayesian Approach*, Open Court Publishing Co., La Salle, Illinois. A curiously outdated work, which might have served a useful purpose 60 years earlier. Mostly a rehash of all the false starts of philosophers in the past, while offering no new insight into them and ignoring the modern developments by scientists, engineers, and economists which have made them obsolete. What little positive Bayesian material there is, represents essentially a level of understanding that Harold Jeffreys had surpassed 50 years earlier, minus the mathematics needed to apply it.

Howson, C. & Urbach, P. (1991), "Bayesian Reasoning in Science", Nature, **350**, 371–374. An advertisement for the previous work, with the same shortcomings. Since they expound Bayesian principles as they existed 60 years ago, it is appropriate that Anthony Edwards responded (Nature, **352**, 386–387) with the standard counter–arguments given by his teacher, R. A. Fisher, 60 years ago. But to those actively engaged in actually *using* Bayesian methods today, this exchange seems like arguing over two different systems of epicycles.

Hoyt, Wm. G. (1980), *Planets X and Pluto*, University of Arizona Press, Tucson.

Huber, P. J. (1981), *Robust Statistics*, J. Wiley & Sons, Inc., New York.

Huber, P. (1992), *Galileo's Revenge: Junk Science in the Courtroom*, Basic Books, Inc., N. Y. Documents the devastating effects now being produced by charlatans and crackpots posing as scientists. They are paid to give 'expert' testimony that claims all sorts of weird causal relations that do not exist, in support of lawsuits that waste billions of dollars for consumers and businesses. The phenomena of pro–causal and anti–causal bias are discussed in Chapters 5, 16, 17. At present we seem to have no effective way to counteract this; as noted by Gardner (1981), the News Media will always raise a great wind of publicity, giving support and encouragement to the charlatans while denying responsible scientists a hearing to present the real facts. It appears that the issue of what is and what is not valid scientific inference must soon move out of Academia and become a matter of legislation – a prospect even more frightening than the present abuses.

Hume, D. (1739), *A Treatise of Human Nature*, London. As revised by P. H. Nidditch, Clarendon Press, Oxford (1978)

Hume, D. (1777), *An Inquiry Concerning Human Understanding*, Clarendon Press, Oxford

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

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

Jansson, P. A., editor (1984), *Deconvolution, with Applications in Spectroscopy*, Academic Press, Inc., 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. (1957a), "Information Theory and Statistical Mechanics", Phys. Rev. **106**, pp. 620–630; **108**, pp. 171–190. Reprinted in Jaynes (1983)

——— (1957b) *How does the Brain do Plausible Reasoning*? Stanford University Microwave Laboratory Report 421. Reprinted in Erickson & Smith (1988), Vol. 1, pp. 1–23.

——— (1963a), "New Engineering Applications of Information Theory", in *Engineering Uses of Random Function Theory and Probability*, J. L. Bogdanoff and F. Kozin, editors, H. Wiley & Sons, Inc., N. Y. , pp 163–203.

——— (1963b), "Information Theory and Statistical Mechanics", in *Statistical Physics*, K. W. Ford, editor, W. A. Benjamin, Inc., pp. 181–218. Reprinted in Jaynes (1983).

——— (1963c), Comments on an article by Ulric Neisser, Science, **140**, 216.

——— (1963d), Review of *Noise and Fluctuations*, by D. K. C. MacDonald, Am. Jour. Phys **31**, 946. Notes that there is no area of physics in which the phenomenon of noise does not present itself.

——— (1965), "Gibbs vs. Boltzmann Entropies", Am. J. Phys. **33**, 391. Reprinted in Jaynes (1983).

——— (1967), "Foundations of Probability Theory and Statistical Mechanics", in *Delaware Seminar in Foundations of Physics*, M. Bunge, editor, Springer–Verlag, Berlin. Reprinted in Jaynes (1983).

——— (1968), "Prior Probabilities", IEEE Trans. Systems Science and Cybernetics SSC–4, 227–241. Reprinted in V. M. Rao Tummala and R. C. Henshaw, eds, *Concepts and Applications of Modern Decision Models*, Michigan State University Business Studies Series (1976); and in Jaynes (1983).

——— (1973), "Survey of the Present Status of Neoclassical Radiation Theory", in *Proceedings of the 1972 Rochester Conference on Optical Coherence*, L. Mandel & E. Wolf, editors, Pergamon Press, New York.

——— (1973), "The Well–Posed Problem", Found. Phys. **3**, 477–493. Reprinted in Jaynes (1983).

——— (1976), "Confidence Intervals vs Bayesian Intervals", in W. L. Harper & C. A. Hooker, eds, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. II, Reidel Publishing Co., Dordrecht–Holland, pp. 175–257. Reprinted in Jaynes (1983).

——— (1978) "Where do we Stand on Maximum Entropy?" in *The Maximum Entropy Formalism*, R. D. Levine & M. Tribus, eds, M.I.T. Press, Cambridge Mass., pp. 15–118. Reprinted in Jaynes, (1983).

——— (1980a), "The Minimum Entropy Production Principle", Ann, Rev, Phys. Chem. vol 31, pp. 579–601. Reprinted in Jaynes (1983).

——— (1980b), "Marginalization and Prior Probabilities", in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner, editor, North–Holland Publishing Co., Amsterdam. Reprinted in Jaynes (1983).

——— (1980c), "What is the Question?" in Bernardo, *et al* (1980), pp. 618–629. Discussion of the logic of questions, as pointed out by R. T. Cox (1978), and applied to the relation between parameter estimation and hypothesis testing. Reprinted in Jaynes (1983), pp. 382–388.

——— (1981), "What is the Problem?", Proceedings of the Second SSSP Workshop on Spectrum Analysis, McMaster University, S. Haykin, editor. The following article is an enlarged version.

——— (1982), "On the Rationale of Maximum–Entropy Methods", Proc. IEEE, **70**, 939–952.

——— (1983) *Papers on Probability, Statistics, and Statistical Physics*, R. D. Rosenkrantz, ed., D. Reidel Publishing Co., Dordrecht–Holland. Reprints of 13 papers dated 1957–1980. Second paperbound edition, Kluwer Academic Publishers (1989).

——— (1984a), "Prior Information and Ambiguity in Inverse Problems", in SIAM–AMS Proceedings, Vol. 14, American Mathematical Society, pp. 151–166.

——— (1984b), "The Intuitive Inadequacy of Classical Statistics", with discussion; Epistemologia, VII; Special Issue on Probability, Statistics, and Inductive Logic; Tilger–Genova, pp. 43–73

——— (1985a) "Where Do We Go From Here?", in Smith & Grandy (1985), pp. 21–58.

——— (1985b), "Entropy and Search Theory", in Smith & Grandy (1985), 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.

——— (1985c) "Macroscopic Prediction", in *Complex Systems – Operational Approaches*, H. Haken, editor, Springer–Verlag, Berlin.

——— (1985d), "Generalized Scattering", in Smith & Grandy (1985), pp. 377–398. Some of the remarkable physical predictions contained in the comparison of two MAXENT distributions, before and after adding a new constraint.

——— (1985e), "Highly Informative Priors", with discussion, in Bernardo, *et al*, (1985), pp 329–360. An historical survey, followed by a worked–out example (seasonal adjustment in econometrics) showing how much prior information can affect our final conclusions in a way that cannot even be stated in the language of orthodox statistical theory, because it does not admit the concept of correlations in a posterior distribution function.

——— (1986a), "Bayesian Methods: General Background", in Justice (1986). A general, non–technical introductory tutorial for beginners, intended to explain the terminology and viewpoint, and warn of common pitfalls of misunderstanding and communication difficulties.

——— (1986b), "Monkeys, Kangaroos, and N", in Justice (1986). Preliminary exploration of deeper hypothesis spaces in image reconstruction, with Dirichlet priors.

——— (1986c), "Some Applications and Extensions of the de Finetti Representation Theorem", in Goel & Zellner, (1986); 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.

——— (1986d), "Predictive Statistical Mechanics", in *Frontiers of Nonequilibrium Statistical Physics*, G. T. Moore & M. O. Scully, editors, Plenum Press, N. Y., pp. 33–55.

——— (1987), "Bayesian Spectrum and Chirp Analysis," in *Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems*, C. Ray Smith & G. J. Erickson, editors, D. Reidel, Dordrecht–Holland, pp. 1–37. A reply to Tukey (1984), carried much further by Bretthorst (1988).

——— (1988a), "The Relation of Bayesian and Maximum Entropy Methods", in Erickson & Smith (1988), Vol 1, pp. 25–29.

——— (1988b) "Detection of Extra–Solar System Planets", in Erickson & Smith (1988), Vol 1, pp. 147–160.

——— (1989), "Clearing up Mysteries – The Original Goal", in *Maximum Entropy and Bayesian Methods*, J. Skilling, editor, Kluwer Publishing Co., Holland, pp. 1–27. This contains what we think is the first application of Bayes' theorem to kinetic theory, the first recognition of hidden assumptions in Bell's theorem, and the first quantitative application of the Second Law of Thermodynamics to Biology.

——— (1990a), "Probability in Quantum Theory", in *Complexity, Entropy and the Physics of Information*, W. H. Zurek, editor, Addison–Wesley Pub. Co., Redwood City CA, pp. 381–404. Use of probability theory as logic makes the meaning of quantum theory appear very different, and hints at possible future resolution of its conceptual difficulties.

——— (1990b), "Probability Theory as Logic", in Proceedings of the Ninth Annual Workshop on Maximum Entropy and Bayesian Methods, P. Fougere, editor, Kluwer Publishers, Holland. Shows by a nontrivial example that conditional probabilities need not express any causal influence of the Popper type – a fact highly relevant to the hidden assumptions in the Bell theorem, discussed in Jaynes (1989).

——— (1991a), "Notes on Present Status and Future Prospects", in Grandy & Schick (1991). A general summing–up of the situation as it appeared in the Summer of 1990.

——— (1991b), "Scattering of Light by Free Electrons as a Test of Quantum Theory", in *The Electron*, D. Hestenes & A. Weingartshofer, editors, Kluwer Academic Publishers, Dordrecht, Holland, pp. 1–20. A sequel to Jaynes (1986d), (1989), and (1990b). The ultimate role of probability in quantum theory hangs on resolving the issues raised in these discussions; for this we require new experimental evidence as well as deeper theoretical analysis.

——— (1992a), "Commentary on Two Articles by C. A. Los", in Volume 3 of special issues, "On System–theoretic Methods in Economic Modelling", S. Mittnik, editor, Computers and Mathematics with Applications, **24**, pp. 267–273. This astonishing economist condemns not only our Bayesian analysis, but virtually every useful thing ever done in data analysis, going back to Gauss.

——— (1992b), "The Gibbs Paradox", in Proceedings of the 11'th Annual MAXENT Workshop, *Maximum Entropy and Bayesian Methods, Seattle, 1991*, C. R. Smith & G. Ericksen, editors, Kluwer Academic Publishers, Holland; pp. 1–21. There is no paradox; Gibbs explained it all in his early work on *Heterogeneous Equilibrium*, but this was missed by later readers who examined only his *Statistical Mechanics*. The range of valid applications of classical thermodynamics is far greater for one who understands this. Note a misprint: the text equation preceding Eq. (14) should be: $f(1) = \log(Ck^{3/2})$.

_________ (1993), "A Backward Look to the Future", in *Physics and Probability: Essays in honor of Edwin T. Jaynes*, W. T. Grandy & P. W. Milonni, eds., Cambridge University Press, pp. 261–275. A response to the contributors to this *Festschrift* volume marking the writer's 70'th birthday, with 22 articles by my former students and colleagues.

Jefferson, Thomas (1944), *The Life and Selected Writings of Thomas Jefferson*, A. Koch & W. Peden, editors, The Modern Library, New York. Reprinted 1972 by Random House, Inc.

Jefferys, Wm. H. (1990), "Bayesian Analysis of Random Event Generator Data", Jour. Scientific Exploration, **4**, pp. 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.

Jeffrey, R. C. (1983), *The Logic of Decision*, 2nd edition, Univ. of Chicago Press. Attempts to modify Bayes' theorem in an *ad hoc* way; as discussed in Chapter 5, this necessarily violates one of our desiderata.

Jeffreys, Harold (1931), *Scientific Inference*, Cambridge Univ. Press. Later editions, 1937, 1957, 1973. Be sure to read his introductory section with a Galilean dialogue showing how induction is actually used in science.

_________ (1932), "On the Theory of Errors and Least Squares", Proc. Roy. Soc., **138**, pp. 48–55. A beautiful derivation of the $d\sigma/\sigma$ prior expressing complete ignorance of a scale parameter, fiercely attacked by Fisher (1933) and discussed in Chapters 7, 16.

_________ (1939), *Theory of Probability*, Clarendon Press, Oxford; Later editions, 1948, 1961, 1967, 1988. Appreciated in our Preface.

_________ (1963), Review of Savage (1962); Technometrics **5**, 407–410.

Jeffreys, Lady Bertha Swirles (1992) "Harold Jeffreys from 1891 to 1940", Notes Rec. R. 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 30's. 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 1920's 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.

Jevons, Wm. S. (1874), *The Principles of Science: A Treatise on Logic and Scientific Method*, 2 vols., Macmillan, London. Reprinted by Dover Publications, Inc., N. Y. (1958). Jevons was a student of de Morgan, and also expounds the Laplacean viewpoint; for this they both came under attack from Venn and others in what Zabell (1989) calls "The Great Jevonian Controversy".

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

Johnson, R. W. (1985), "Independence and Bayesian Updating Methods", U. S. Naval Research Laboratory Memorandum Report 5689, dated November 1985.

Johnson, W. E. (1924), *Logic, Part III: The Logical Foundations of Science*, Cambridge University Press, U. K. Reprinted by Dover Publications, Inc., (1964).

Johnson, W. E. (1932), "Probability, the Deduction and Induction Problem", Mind, **44**, 409–413.

Justice, J. H. (1986), editor, *Maximum Entropy and Bayesian Methods in Geophysical Inverse Problems*, Cambridge University Press, Cambridge U. K. Proceedings of the fourth annual "MAXENT Workshop", held in Calgary in August 1984.

Kadane, J. B., Schervish, M. J. & Seidenfeld, T. (1986), "Statistical Implications of Finitely Additive Probability", in *Bayesian Inference and Decision Techniques*, Prem K. Goel & A. Zellner, editors, Elsevier Science Publishers, Amsterdam. The 'KSS' work discussed at length in Chapter 15.

Kahneman, D. & Tversky, A. (1972), "Subjective Probability: a Judgment of Representativeness", Cognitive Psychology, **3**, 430–454. See also "The Framing of Decisions and the Psychology of Choice", Science, **211**, 453–458 (1981).

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

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

________ (1989), "A Theory for the Identification of Linear Relations", in *Proceedings, Colloque LIONS*, H. Brezis & P. G. Ciarlet, editors

________ (1990), *Nine Lectures on Identification*, Springer Lecture Notes on Economics and Mathematical Systems.

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

Kay, S. & Marple, S. L. (1979), Record IEEE ICASSP; pp. 151–154. Spectrum Analysis.

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

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 20'th Century.

Kendall, M. G. (1963), "Ronald Aylmer Fisher, 1890 – 1962", Biometrika **50**, pp 1–15. Reprinted in E. S. Pearson & M. G. Kendall (1970). Like the previous reference, this tells us more about the author than the subject.

Kendall, M. G. & Moran, P. A. P. (1963), *Geometrical Probability*, Griffin, London. Much useful mathematical material, all of which is readily adapted to Bayesian pursuits.

Kendall, M. G. & Plackett, R. L. (1977), *Studies in the History of Statistics and Probability*, 2 vols., Griffin, London.

Kendall, M. G. & Stuart, A. (1961), *The Advanced Theory of Statistics*: Volume 2, Inference and Relationship, Hafner Publishing Co., New York. This represents the beginning of the end for the confidence interval; while they continued to endorse it on grounds of "objectivity", they noted so many resulting absurdities that readers of this work were afraid to use confidence intervals thereafter. In Jaynes (1976) we explained the source of the difficulty and showed that these absurd results are corrected automatically by use of Bayesian methods.

Kendall, M. G. & Stuart, A. (1977), *The Advanced Theory of Statistics: Volume 1, Distribution Theory*, Macmillan, New York.

Kennard, E. H. (1938), *Kinetic Theory of Gases*, McGraw–Hill Book Co., N. Y.

Keynes, J. M. (1921), *A Treatise on Probability*, MacMillan, London. Reprinted by Harper & Row, New York (1962). The first clear explanation of the distinction between logical independence and causal independence. Important today because it served historically as the inspiration for the work of R. T. Cox. For an interesting review of Keynes, see Borel (1924).

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 planned for Volume 2.

Khinchin, A. I. (1957), *Mathematical Foundations of Information Theory*, Dover Publications, Inc. Attempts of a mathematician to 'rigorize' Shannon's work. But we do not think it was in need of this. In any event, when one tries to work directly on infinite sets from the beginning, the resulting theorems just do not refer to anything in the real world. Khinchin was probably careful enough to avoid actual error and thus produced theorems valid in his imaginary world; but we note in Chapter 15 some of the horrors that have been produced by others who tried to do mathematics this way.

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 Math. Vol. 1, AMS, Providence R.I.

Kline, Morris (1980), *Mathematics: The Loss of Certainty*, Oxford University Press. A fairly complete history, recalling hundreds of interesting anecdotes, but expressing views very different from the Bourbakist ones of Félix (1960).

Knight, P. (1987), "Single–Atom Masers and the Quantum Nature of Light", Nature, **326**, 329. A report on recent advances in our ability to control quantum states of individual atoms.

Kolmogorov, A. N. (1933), *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Math. (2), Berlin. English translation, *Foundations of the Theory of Probability*, Chelsea Publishing Co., New York (1950). Described in Appendix A.

Koopman, B. O. (1936), "On Distributions Admitting a Sufficient Statistic", Trans. Amer. Math. Soc. **39**, 399–509. Proof that the NASC for existence of a sufficient statistic is that the sampling distribution have the exponential form, later recognized as identical with what Maximum Entropy generates automatically. Simultaneous with Pitman (1936).

Kronecker, L. (1901), *Vorlesungen über Zahlentheorie*, Teubner, Leipzig; republished by Springer, 1978.

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

Kullback, S. (1959), *Information Theory and Statistics*, Wiley, New York. A beautiful work, never properly appreciated because it was 20 years ahead of its time.

Kullback, S. & Leibler, R. A. (1951), "On information and sufficiency", Ann. Math. Stat. 22, 79–86.

Kurtz, P. (1985), *A Skeptic's Handbook of Parapsychology*, Prometheus Books, Buffalo N. Y. Several Chapters have relevant material; see particularly Chapter 11 by Betty Markwick.

Kyburg, H. E. & Smokler, H. E. (1981) *Studies in Subjective Probability*, 2nd edition. J. Wiley & Sons, Inc., New York.

Lancaster, H. O. (1969), *The Chi–squared Distribution*, J. Wiley & Sons, Inc., New York.

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**, pp 47–77. Interprets the Burg solution in terms of more general problems in several fields. Highly recommended for a deeper understanding of the mathematics.

Lane, D. A. (1980), "Fisher, Jeffreys, and the Nature of Probability", in *R. A. Fisher, an Appreciation*, S. Fienberg *et al*, editors, Springer–Verlag, New York; pp. 148–160.

Laplace, Pierre Simon (1774), "Mémoire sur la probabilité des causes par les évènemens", Mémoires de l'Académie royale del sciences **6**, 621–656. Reprinted in Laplace (1878–1912), vol. 8, pp. 27–65, Gauthier–Villars, Paris. English translation by S. M. Stigler (1986).

——— (1783), Histoire d l'Acadamie, pp. 423–467. An early exposition of the properties of the "Gaussian" distribution. Suggests that it is so important that it should be tabulated.

——— (1810), "Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités", Mémoires de l'Académie des sciences de Paris, 1809, pp. 353–415, 559–565. Reprinted in Laplace (1878–1912), vol. 12, pp. 301–353. A massive compendium of the origin, properties, and uses of the Gaussian distribution.

——— (1812), *Théorie analytique des probabilités*, 2 vols., Courcier Imprimeur, Paris; (3rd edition with supplements, 1820). Reprinted in Laplace (1878–1912), Vol. 7. Reprints of this rare but very important work and the following one are available from: editions Culture et Civilisation, 115 Ave. Gabriel Lebron, 1160 Brussels, Belgium.

——— (1814), *Essai Philosophique sur les Probabilités*, Courcier Imprimeur, Paris. Reprinted in *Oeuvres complètes de Laplace*, Vol. 7, Gauthier–Villars, Paris, 1886. English translation by F. W. Truscott & F. L. Emory, Dover Publications, Inc., New York (1951). Be warned that this "translation" is little more than a literal *transcription*, which distorts Laplace's meaning on many points. It is essential to check the original French version before accepting any interpretive statement in this work.

——— (1878–1912), *Oeuvres complètes*, 14 vols. Gauthier–Villars, Paris.

Lee, Y. W. (1960), *Statistical Theory of Communication*, J. Wiley & Sons, New York. The usable but watered–down pedagogical work that grew out of Wiener (1949). Masses of well explained examples, but none of the mathematical techniques such as the Paley–Wiener factorization or functional integration over Wiener measure, that were used in the original. Greatly extends the folklore about Gibbs that started in Wiener (1948). Reviewed by E. T. Jaynes, Am. Jour. Phys. **29**, 276 (1961).

Legendre, A. M. (1806), "Nouvelles Méthods pour la Détermination des Orbits des Cométes," Paris.

Lehmann, E. L. (1959) *Testing Statistical Hypotheses*, 2nd. edition, 1986, Wiley, New York.

Leibniz, Gottfried Wilhelm *General Investigations Concerning the Analysis of Concepts and Truths*, Translated by W. H. O'Briant; Univ. of Georgia Press, Athens (1968).

Lessard, S., editor (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, C. I. & Langford, C. H. (1932), *Symbolic Logic*, New York ****** **********

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.

Lighthill, M. J. (1957), *Introduction to Fourier Analysis and Generalised Functions*, Cambridge Univ. Press. Required reading for all who have been taught to mistrust delta–functions. See the review by Freeman Dyson (1958). Lighthill and Dyson were classmates in G. H. Hardy's famous course in 'Pure Mathematics' at Cambridge University, at a time when Fourier analysis was mostly preoccupied with convergence theory, as in Titchmarsh (1937). Now with a redefinition of the term 'function' as explained in our Appendix B, all that becomes nearly irrelevant. Dyson states that Lighthill 'lays Hardy's work in ruins, and Hardy would have enjoyed it more than anybody.'

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

———— (1957), "A Statistical Paradox", Biometrika **44**, 187–192. Mentions Soal & Bateman parapsychology experiments.

———— (1958) "Fiducial Distributions and Bayes' Theorem", J. Roy. Stat. Soc. **B20**, 102–107.

———— (1971) *Bayesian Statistics: A Review*, Soc. Ind. App. Math., 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 *Bayesian Inference and Decision Techniques*, A. Zellner & P. Goel, eds, North–Holland Publishers, Amsterdam.

Little, J. F. and Rubin, D. B. (1987), *Statistical Analysis with Missing Data*, J. Wiley & Sons, NY. Missing data can wreak havoc with orthodox methods because this changes the sample space, and thus changes not only the sampling distribution of the estimator, but even its analytical form; one must go back to the beginning for each such case. But however complicated the change in the sampling distribution, the change in the likelihood function is very sample. Bayesian methods accommodate missing data effortlessly; in all cases we simply include in the likelihood function all the data we have; and Bayes' theorem automatically returns the new optimal estimator for that data set.

Luce, R. D. & Raiffa, H. (1958), *Games and Decisions*, Wiley, N. Y.

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

Macdonald, P. D. M. (1987), "Analysis of Length–frequency Distributions", in *Age and Growth of Fish*, R. C. Summerfelt & G. E. Hall, editors, Iowa State University Press, Ames, Iowa, 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, *et al* (1985).

Machol, R. E., Ladany, S. P. & Morrison, D. G. (eds), (1976), *Management Science in Sports*, Vol. 4, TIMS Studies in the Management Sciences, North–Holland, Amsterdam. Curious applications of probability theory.

Mandel, J. (1964), *The Statistical Analysis of Experimental Data*, Interscience, New York. Straight orthodox *adhockeries*, 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. & D. J. Thompson (1982), "Robust–Resistant Spectrum Estimation", Proc. IEEE, **70**, pp. 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 vastly 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 & Wm. Garrett, *the Life of James Clerk Maxwell*, Macmillan, 1881.

Maxwell, J. C. (1860), "Illustration of the Dynamical Theory of Gases. Part I. On the Motion and Collision of Perfectly Elastic Spheres," Phil. Mag. **56**.

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 P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, N. Y.

Mead, L. R. & Papanicolaou, N. (1984), "Maximum Entropy in the Problem of Moments", J. Math. Phys. **25**, pp. 2404–2417.

Middleton, D. (1960), *An Introduction to Statistical Communication Theory*, McGraw–Hill Book Co., New York. A massive work (1140 pages) with an incredible amount of mathematical material. The title is misleading, since the material really applies to statistical inference in general. Unfortunately, most of the work was done a little too early, so the outlook is that of sampling theory and Neyman–Pearson decision rules, now made obsolete by the Wald decision theory and Bayesian advances. Nevertheless, the mathematical problems – such as methods for solving singular integral equations – are independent of one's philosophy of inference, so it has much useful material applicable in our current problems. One should browse through it, and take note of what is available here.

Miller, R. G. (1974), "The jackknife – a review", Biometrika **61**, 1–15.

Mitler, K. S. (1974), *Multivariate Distributions*, Wiley, New York.

Molina, E. C. (1931), "Bayes' Theorem, an Expository Presentation", Bell Sys. 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.

Molina, E. C. (1963), *Two Papers by Bayes with Commentaries*, Hafner Publishing Co., New York. Contains penetrating historical remarks about the relation of Laplace and Boole, noting that those who have quoted Boole in support of their attacks on Laplace, may have misread Boole's intentions.

Moore, G. T. & Scully, M. O. eds (1986), *Frontiers of Nonequilibrium Statistical Physics*, Plenum Press, N. Y. 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.

Mosteller, F. (1965), *Fifty Challenging Problems in Probability with Solutions*, Addison–Wesley, Reading, Massachusetts.

Munk, W. H. & Snodgrass, F. E. (1957), "Measurements of Southern Swell at Guadalupe Island", Deep–Sea Research, **4**, pp 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.

Nachtigall, P. E. & Moore, P. W. B.,editors (1989), *Animal Sonar: Processes and Performance*, NATO ASI Series, Vol. A156, Plenum Publishing Corp., New York. Proceedings of a meeting held in Helsingor, Denmark in 1986. Some remarkable advances; bats are shown not only to detect targets, but to predict their future course. Dolphins do fast (300 microsecond) acoustic processing very much like that of bats, and with very similar neurophysiological structures.

Newcomb, S. (1881), "Note on the frequency of use of the different digits in natural numbers", Am. J. Math. **4**, 39–40.

Newton, Sir Isaac (1687) *Philosophia Naturalis Principia Mathematica*, Translation by Andrew Motte, 1729, revised and reprinted as *Mathematical Principles of Natural Philosophy*, Florian Cajori, editor, Univ. Calif. Press, Berkeley (1946). See also Cajori (1928, 1934).

Neyman, Jerzy & Pearson, E. S. (1933), "On the problem of the most efficient test of statistical hypotheses", Phil. Trans. Roy. Soc., **231**, 289–337.

Neyman, Jerzy & Pearson, E. S. (1967), *Joint Statistical Papers*, Cambridge Univ. Press. Reprints of the several Neyman–Pearson papers of the 1930's, originally scattered over several different journals.

Neyman, Jerzy (1950), *First Course in Probability and Statistics*, Henry Holt & Co, New York.

——— (1952), *Lectures and Conferences on Mathematical Statistics and Probability*, Graduate School, U. S. Dept. of Agriculture. Contains an incredible comparison of Bayesian interval estimation *vs.* confidence intervals. A good homework problem is to locate the error in his reasoning.

——— (1959), "On the two different aspects of representative method: The method of stratified sampling and the method of purposive selection", Estadistica **17**, 587–651.

——— (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.

——— (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. Jour. **3**, p. 324.

Nyquist, H. (1928), "Certain Topics in Telegraph Transmission Theory," Trans. AIEE, pp. 617.

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 an egregious violation of the principles of rational inference, and signifies only that the problem was improperly defined. 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, straight–out rejection may be a reasonable approximation.

Ore, O (1953), *Cardano, the Gambling Scholar*, Princeton Univ. Press.

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

Pearson, E. S. & Kendall, M. G. (1970), *Studies in the History of Statistics and Probability*, Hafner Publishing Co., Darien, Conn.

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).

——— (1905), "The Problem of the Random Walk", Nature, **72**, 294, 342.

——— (1914–1930) *The Life, Letters and Labours of Francis Galton*, 3 Vols., Cambridge University Press. Francis Galton had inherited a modest fortune, and on his death in 1911 he endowed the Chair of Eugenics at University College, London. Karl Pearson was its first occupant; this enabled him to give up the teaching of applied mathematics to engineers and physicists, and concentrate on biology and statistics.

——— (1920), "Notes on the History of Correlation", Biometrika, **13**, 25–45. Reprinted in Pearson & Kendall (1970).

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

Penfield, Wilder (1958), Proc. Nat. Acad Sciences (USA), **44**, p. 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. As a mental phenomenon this is closely related to the theory of the A_p distribution. Now others with different motives have moved into this field, with suggestions that psychiatrists are causing their patients to recall things that never happened.

Penrose, O. (1979), "Foundations of Statistical Mechanics", Rep. Prog. Phys. **42**, 1937–2006. Published in "Reports of Progress", although it reports no progress.

Pitman, E. J. G. (1936), "Sufficient Statistics and Intrinsic Accuracy," Proc. Camb. Phil. Soc. **32**, 567–579. Proof, almost simultaneous with Koopman (1936), of the NASC for sufficiency, now known as the Pitman–Koopman theorem.

Poincaré, H. (1899), "L'Oeuvre Mathématique de Weierstraß", Acta Math. **22**, 1–18. Contains an authoritative account of the relation between the works of Kronecker and Weierstraß, pointing out that the difference was more in taste than in substance; to be contrasted with that of E. T. Bell (1937), who tries to make them mortal enemies.

——— (1904), *Science et Hypothesis*, English translation, Dover Publications, Inc., (1952). Poincaré had the gift of being able to say more in a sentence than most writers can in a page. Full of quotable remarks, as true and important today as when they were written.

——— (1909), *Science et Méthode*, English translation, Dover Publications, Inc., (1952). Like Kline (1980), a ringing indictment of the contemporary work in mathematics and logic, for which the Bourbakists have never forgiven him. However, in knowledge and judgment Poincaré was far ahead of his critics, because he was better connected to the real world.

——— (1912), *Calcul des probabilités*, 2nd. edition, Gauthier–Villars, Paris. Contains the first example of the assignment of a probability distribution by the principle of group invariance.

Poisson, S. D. (1837), *Recherches sur la Probabilité des Jugements*. First appearance of the Poisson distribution.

Pólya, G. (1920), "Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem," Math. Zeit., **8**, 171–181; reprinted in Pólya (1984), Vol. IV. First appearance of the term "Central Limit Theorem" in print. He does not actually prove the theorem (which he attributes to Laplace), but points out a theorem on uniform convergence of a sequence of monotonic functions which can be used to shorten various proofs of it.

——— (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 Rayleigh (1919) and Pearson (1905).

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

——— (1945), *How to Solve It*, Princeton University Press. Second paperbound edition by Doubleday Anchor Books (1957).

——— (1954), *Mathematics and Plausible Reasoning*, 2 Vols., Princeton University Press.

——— (1984), *Collected Papers*, 4 Vols. Gian–Carlo Rota, editor, MIT Press, Cambridge MA. Volume IV contains papers on probability theory and combinatorics, several short articles on plausible reasoning, and a bibliography of 248 papers by him. George Pólya always claimed that his main interest was in the mental processes for solving particular problems rather than in generalizations. Nevertheless, some of his results launched new branches of mathematics through their generalizations by others. The present work was influenced by Pólya in more ways than noted in our Preface: most of our exposition is aimed, not at expounding generalities for their own sake, but in learning how to solve specific problems – albeit by general methods.

——— (1987), *The Pólya Picture Album: Encounters of a Mathematician*, G. L. Alexanderson, editor, Birkhäuser, Boston. Over his lifetime, George Pólya collected a large picture album with photographs of famous mathematicians he had known, which he took delight in showing to visitors. After his death, the collection was published in this charming book, which contains about 130 photographs, plus a biography of Pólya by the editor.

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.

Popper, Karl (1974), "Replies to my Critics", in *the Philosophy of Karl Popper*, P. A. Schilpp, ed., Open Court Publishers, La Salle. Presumably an authoritative statement of Popper's position, since it is some years later than his best known works, and seeks to address points of criticism directly.

Popper, K. & Miller, D. W. (1983), "A proof of the impossibility of inductive probability", Nature, **302**, 687–88.

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

Pratt, J. W. (1961) Review of *Testing Statistical Hypotheses* (Lehmann, 1959), J. Am. Stat. Assoc. **56**, pp. 163–166. A devastating criticism of orthodox hypothesis testing theory.

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

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

Quetelet, L. A. (1835), *Essai de Physique sociale.*

Quetelet, L. A. (1869), *L'homme moyen*, Physique Sociale., Vol. 2, Bruxxeles.

Raiffa, H. A. & Schlaifer, R. S. (1961), *Applied Statistical Decision Theory*, Graduate School of Business Administration, Harvard University.

Raimi, R. A. (1976), "The first digit problem", Am. Math. Monthly **83**, 521–538. Review article on "Benford's law" with many references.

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 Twentieth 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.

Rao, M. M. (1993), *Conditional Measures and Applications*, Marcel Dekker, Inc. New York. Noted in Appendix A as indicating how foreign the notion of conditional probability is in the Kolmogorov system.

Rayleigh, Lord (1919), "On the Problem of Random Vibrations, and of Random Flights in One, Two or Three Dimensions", Edinb. & Dublin Phil. Mag. and Jour. of Science, Series 6, **37**, 321–47.

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

Reid, Constance (1970), *Hilbert*, Springer – Verlag, New York.

Reid, Constance (1982), *Neyman – From Life*, Springer–Verlag, N. Y.

Rempe, G., Walther, H. & Klein, N. (1987), Phys. Rev. Lett. **58**, 353. Successful operation of single–atom masers.

Renyi, A. (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 & Electronic Systems **AES–17**, p. 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).

Rissanen, J. (1983), "A Universal Prior for the Integers and Estimation by Minimum Description Length", Annals of Statistics, **11**, pp 416–431. One of the few fresh new ideas in recent decades.

Robbins, H (1950), "Asymptotically subminimax solution of the compound statistical decision problem", Proc. 2nd Berkeley Symp. of Math. Statist. and Prob., Univ. Calif. Press, 131–148. An Anticipation of Stein.

Robbins, H. (1956), "An empirical Bayes' approach to statistics", Proc. 3rd Berkeley Symp. on Math. Stat. and Prob. I, 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**, pp. 855 – 906.

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

Rosenkrantz, R. D. (1977), *Inference, Method, and Decision: Towards a Bayesian Philosophy of Science*, D. Reidel Publishing Co., Boston. Reviewed by E. T. Jaynes in Jour. Am. Stat. Ass'n, Sept. 1979, pp. 740–741.

Rothman, Tony (1989), *Science à la Mode*, Princeton University Press. Accounts of what happens when scientists lose their objectivity and jump on bandwagons. We would stress that they not only make themselves ridiculous, they do a disservice to science by promoting sensational but nonproductive ideas. For example, we think that it will be realized eventually that the 'Chaos' bandwagon has put a stop to the orderly development of a half–dozen different fields without enabling any new predictive ability.

Rowe, D. E. & McCleary, J., editors (1989), *The History of Modern Mathematics*, 2 Vols, Academic Press, Inc., Boston.

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).

Royall, R. M. & Cumberland, Wm. G. (1981), "The Finite–Population Linear Regression Estimator and Estimators of its Variance – An Empirical Study", Jour. Am. Stat. Assoc. **76**, 924–930. A demonstration of the folly of randomization, particularly cogent because the authors are not Bayesian and did not set out with that purpose.

Ruelle, D. (1991), *Chance and Chaos*, Princeton University Press. How not to use probability theory in science; virtually every useful aspect of it is missed.

Sampson, A. R. and 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. and Smith, R. L. (1982), "Assessing risks through the determination of rare event probabilities", Operations Research, 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).

Savage, I. R. (1961), "Probability inequalities of the Tchebyscheff type", J. Res. Nat. Bureau Stand. **65B**, pp 211–222. A useful collection of results, which ought to be made more accessible.

Savage, L. J. (1954), *Foundations of Statistics*, J. Wiley & Sons. Second Revised edition, 1972, by Dover Publications, Inc., New York. This work was attacked savagely by van Dantzig (1957).

——— (1962), *The Foundations of Statistical Inference: A Discussion*, G. A. Barnard & D. R. Cox, editors, Methuen, London. Reviewed by H. Jeffreys (1963).

——— (1981), *The Writings of Leonard Jimmie Savage – A Memorial Selection*, Published by the American Association of Statistics and the Institute of Mathematical Statistics. Jimmie Savage died suddenly and unexpectedly in 1971, and his colleagues performed an important service by putting together this collection of his writings that were scattered in many obscure places and hard to locate. Some personal reminiscences about him are in Jaynes (1984b) and Jaynes (1985e).

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

Schendel, U. (1989) *Sparse Matrices*, J. Wiley & Sons, N. Y.

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 & Moses work on decision theory.

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**.

——— (1947), "The Foundation of the Theory of Probability", Proc. Roy. Irish Acad. (A), 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.

——— (1948), *Statistical Thermodynamics*, Cambridge Univ. Press.

Schuster, A. (1897), "On Lunar and Solar Periodicities of Earthquakes", Proc. Roy. Soc. **61**, pp. 455–465. This could almost be called the origin of orthodox significance tests. He undertakes to refute some claims of periodicities in earthquakes, by considering only the sampling distribution of the periodogram *under the hypothesis that no periodicity exists*! He never considers: what is the probability of getting the observed data if a periodicity of a certain frequency *does* exist? Orthodoxy has been following this nonsensical procedure ever since. We show here that evidence for periodicity is contained in the *shape* of the periodogram, not its sampling distribution. But to show this requires the elimination of nuisance parameters in a way that orthodox ideology does not permit.

Schwartz, Laurent (1950) *Théorie des Distributions*, 2 vols., Hermann et Cie, Paris.

Seal H. (1967), "The Historical Development of the Gauss Linear Model", Biometrika **54**, 1–24. Reprinted in Pearson & Kendall (1970).

Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, N. J. An attempt to develop a theory of two–valued probability.

Shafer, G. (1982), "Lindley's Paradox", J. Am. Stat. Assn., **77**, 325–334. Apparently Shafer was unaware that this was all in Jeffreys (1939; p. 194) some twenty years before Lindley.

Shamir, A (1982), "A polynomial time algorithm for breaking Merkle–Hellman cryptosystems", Research Announcements.

Shannon, C. E. (1948), "A Mathematical Theory of Communication", Bell Systems Technical Jour., **27**, 379, 623. Reprinted in C. E. Shannon & Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

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

Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Products*, van Nostrand, 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, N. Y.

Shore, J. E. and Johnson, R. W. (1980), "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross–entropy", IEEE Trans. Information Theory IT–26, 26–37. Many different choices of axioms all lead to the same actual algorithm for solution of problems.

The authors present a different basis from the one first proposed (Jaynes, 1957). But we stress that maximum entropy and minimum cross–entropy are not different principles; a change of variables converts one into the other.

Siegmann, D., (1985) *Sequential Analysis*, Springer. No mention of Bayes' theorem or optional stopping!!

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

Simon, H. A. & Rescher, N. (1966), "Cause and Counterfactual", Phil. Sci. **33**, 323–340.

Simon, H. A. (1977), *Models of Discovery*, D. Reidel Pub. Co., Dordrecht Holland.

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

Sivia, D. S. & Carlile, C. J. (1992), "Molecular Spectroscopy and Bayesian Spectral Analysis—How Many Lines Are There?", J. Chem. Phys. **96**, pp. 170–178. Successful resolution of noisy data into as many as nine gaussian components by a Bayesian computer program.

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

Smart, W. M. (1947), *John Couch Adams and the Discovery of Neptune*, Roy. Astronomical Society, Occasional Notes No. 11.

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

Smith, C. Ray & Erickson, G. J. (1987), editors, *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.

Soal, S. G. & Bateman, F. (1954), *Modern Experiments in Telepathy*, Yale University Press, New Haven.

Sonett, C. P. (1982), "Sunspot Time Series: Spectrum From Square Law Modulation of the Hale Cycle," Geophys. Res. Lett., **9** pp. 1313–1316.

Spinoza, B., *Ethics*, part 2, Prop. XLIV: "De natura Rationis no est res, ut contingentes; sed, ut necessarias, contemplari."

Spitzer, F. (1964), *Principles of Random Walk*, van Nostrand, N. Y. 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.

——— (1956), "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution", Proc. 3rd. Berkeley Symp. Univ. Calif. Press, **1**, 197–206. First announcement of the "Stein shrinking" phenomenon.

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

——— (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 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.

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

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

———— (1980), "Stigler's Law of Eponymy", Trans. New York Acad. Sci., **39**, pp. 147–159.

———— (1982a), "Poisson on the Poisson distribution", Stat. & Prob. Lett., **1**, 33–35.

———— (1982b), "Thomas Bayes's Bayesian Inference", J. Roy. Stat. Soc. **A145**, pp. 250–258.

———— (1983), "Who discovered Bayes's Theorem?" Am Stat. **37**, 290–296.

———— (1986), "John Craig and the Probability of History", JASA **81**, 879–887.

———— (1986), Translation of Laplace's 1774 Memoir on "Probability of Causes", Statistical Science **1**, 359.

———— (1986), *The History of Statistics*, Harvard Univ. Press. A massive work of careful scholarship, required reading for all students of the subject. Gives full discussions of many topics that we touch on only briefly.

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

Stone, M. (1970) "Strong inconsistency from uniform priors", J. Am. Stat. Ass'n **71**, 114–116.

Stove, D. (1982), *Popper and After: Four Modern Irrationalists*, Pergamon Press, New York.

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.

Székely, Gabor J. (1986), *Paradoxes in Probability Theory and Mathematical Statistics*, D. Reidel Publishing Co., Dordrecht–Holland. Gives on p. 64 an erroneous solution for a biased coin; no comprehension of the physics. Compare with our Chapter 10.

Takacs, L. (1958), "On a probability problem in the theory of counters", Ann. Math. Stat. **29**, pp 1257–63. There are many earlier papers by Takacs on this topic.

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

Tax, S., Editor (1960), *Evolution After Darwin*, 3 Vols., University of Chicago Press. Volume 1: *The Evolution of Life*; Volume 2: *The Evolution of Man*; Volume 3: *Issues in Evolution*. A collection of articles and panel discussions by many workers in the field, summarizing the state of knowledge and current research directions 100 years after the original publication of Darwin.

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 Research, 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*.

Titchmarsh, E. C. (1937), *Introduction to the Theory of Fourier Integrals*, Clarendon Press, Oxford U. K. The 'state of the art' in Fourier analysis just before the appearance of Lighthill (1957), which made all the lengthy convergence theory nearly irrelevant. However, only a part of this classic work is thereby made obsolete; the material on Hilbert transforms, Hermite and Bessel functions, and Wiener–Hopf integral equations, remains essential for applied mathematics.

Titchmarsh, E. C. (1939), *The Theory of Functions*, 2nd edition, Oxford University Press. In Chapter XI the reader may see – possibly for the first time – some actual examples of nondifferentiable functions. We discuss this briefly in Appendix B.

Titterington, D. M., Smith, A. F. M. & Makov, U. E. (1985) *Statistical Analysis of Finite Mixture Distributions*, Wiley, N. Y.

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

Todhunter, Isaac (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.

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

Tribus, M. (1969), *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.

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", Contrib. to Prob. and Statist., (I. Olkin, editor), Stanford Univ. Press, 448–485.

________ (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.

________ (1977), *Exploratory Data Analysis*, Addison–Wesley, Reading MA. Introduces the word 'resistant' as a data–oriented version of 'robust'.

________ (1978), "Granger on Seasonality", in *Seasonal Analysis of Time Series*, A. Zellner, editor, U. S. Dept. of Commerce, Washington. An amusing view of the nature of Bayesian inference as a sneaky way of concealing one's methodological sins behind a formal apparatus.

________ (1984), "Styles of Spectrum Analysis"; Scripps Institution of Oceanography Reference Series 84–85, March 1984; 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; further comments on this are in Chapter 22.

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) contains parts of it.

Twain, Mark (1900), *The Complete Short Stories and Famous Essays*, P. F. Collier & Son, New York; p. 187. Coming across a French translation of his story: "The Jumping Frog of Calaveras County", he rendered it back literally into English; the result was hilariously funny, telling a grossly distorted story with the meaning altered at many places.

Ulam, S. (1957), Am. Jour. Phys. **25**, 477.

Valavanis, S. (1959), *Econometrics*, McGraw–Hill, N. Y. Modern students will find this useful as a documented record of what Econometrics was like under orthodox statistical teaching. The demand for unbiased estimators at all costs can lead him to throw away practically all the information in the data; he just does not think in terms of information content.

Valery–Radot, René (1923), *The Life of Pasteur*, Doubleday, Page & Co., Garden City, N. Y.

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 Dantzig, D. (1957) "Statistical Priesthood (Savage on personal probabilities)", Statistica Neerlandica **2**, 1–16. Younger readers who find it difficult to understand today how Bayesians could have had to fight for their viewpoint, should read this attack on the work of Jimmie Savage. But one should realize that van Dantzig was hardly alone here; his views were the ones most commonly expressed by statisticians in the 1950's and 1960's.

van den Bos, A. (1971), IEEE Trans. Information Theory, IT–17, pp. 493–494. Reprinted in Childers (1978). Expresses several misgivings about Maximum Entropy spectrum analysis, answered in Jaynes (1982).

Varian, H. (1978), *Microeconomic Analysis*, Norton & Co., N. Y.

Vasicek, O., "A conditional law of large numbers", Ann. Prob. 8, 142–147, 1980.

Venn, John (1866), *The Logic of Chance*, MacMillan & Co., London. Later editions, 1876, 1888. Picks up where Cournot and Ellis left off in the anti–Laplace cause. Some details are given in Jaynes (1986b).

von Mises, R. (1951), *Probability, Statistics and Truth*, Second German edition, Springer–Verlag. English Translation: G. Allen & Unwin, Ltd., London (1957).

von Neumann, J. & Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton Univ. Press.

Vorzimmer, P. J. (1970), *Charles Darwin: The Years of Controversy*, Temple Univ. Press, Philadelphia. An account of the evolution of Darwin's *Origin of Species* from its first publication in 1859 to the sixth and last revision in 1872, as Darwin sought to answer his critics. Some see in this the beginning of the process by which the "naturalist" was replaced by the trained biologist. But others, more perceptive, note that Darwin's first edition makes a better case for his theory than does the last. The contemporary criticisms that he was answering seem today to be mostly incompetent and not deserving of any reply. Even today, as Stephen J. Gould has noted, attacks on Darwin's theory only document the author's misunderstanding of what the theory is.

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.

——— (1942), *On the Principles of Statistical Inference*, Notre Dame Univ. Press.

———— (1943), "Sequential analysis of statistical data: Theory", Restricted report dated September 1943.

———— (1947), *Sequential Analysis* Wiley, New York. Reviewed by G. A. Barnard, J. Am. Stat. Ass'n **42**, 668 (1947).

———— (1950), *Statistical Decision Functions*, Wiley, New York. Wald's final work, in which he now recognized the fundamental role of Bayesian methods and called his optimal methods "Bayes strategies".

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

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. Need any more be said?

Watson, James 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 Francis Crick (1988).

Wason, P. C. & Johnson–Laird, P. N. (1972), *Psychology of Reasoning*, Batsford, London.

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

Weaver, Warren (1963), *Lady Luck, the Theory of Probability*, Doubleday Anchor Books, Inc., Garden City, N. Y.

Weber, B. H., Depew, D. J. & Smith, J. D. editors (1988), *Entropy, Information, and Evolution*, MIT Press, Cambridge MA. A collection of 16 papers given at a symposium held in 1985. An appalling display of the state into which evolution theory has degenerated, due to attempts to explain it in terms of the second law of thermodynamics – by biologists and philosophers in total disagreement and confusion over what the second law is. Discussed briefly in Chapter 7.

A. Wehrl, A. (1978), "General Properties of Entropy", Revs, Mod. Phys. **50**, 220–260.

Whittaker, E. T. & Robinson, G. (1924), *The Calculus of Observations*, Blackie & Son, London. Notable because the fake 'variable star' data on p. 349 were used by Bloomfield (1976), who proceeded to make their analysis, with absurd conclusions, the centerpiece of his textbook on spectrum analysis.

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

———— (1957), "Curve and Periodogram Smoothing", J. Roy. Stat. Soc. **B, 19**, 38–47.

——— (1958), "On the smoothing of probability density functions", JRSS B, **20**, 334–343.

Whyte, A. J. (1980), *The Planet Pluto*, Pergamon Press, N. Y.

Wiener, N. (1948), *Cybernetics*, J. Wiley & Sons, Inc., New York. On p. 109, Norbert Wiener reveals himself as a closet Bayesian, although we know of no work of his that actually uses Bayesian methods. But his conceptual understanding of the real world was in any event too naïve to have succeeded. On p. 46 he gets the effect of tidal forces in the earth–moon system backwards (speeding up the earth, slowing down the moon). The statements about the work of Gibbs on pp. 61–62 are pure inventions; far from introducing or assuming ergodicity, Gibbs did not mention it at all. Today it is clear, from the discovery of strange attractors, chaos, *etc.*, that almost no real system is ergodic, and in perceiving its irrelevancy to statistical mechanics, Gibbs was here a Century ahead of the understanding of others. Unfortunately, Wiener's statements about Gibbs were quoted faithfully by other authors such as S. Goldman (1953) and Y. W. Lee (1960), who were in turn quoted by others, thus creating a large and still growing folklore. Wiener did not bother to proof–read this work, and many equations are only vague hints as to the appearance of the correct equation.

Wiener, N. (1949), *Extrapolation, Interpolation, and Smoothing of Stationary time Series*, J. Wiley & Sons, Inc., New York. Another masterpiece of careless and obscure writing, partially deciphered by N. Levinson in the Appendix, and more fully in the books of S. Goldman and Y. W. Lee.

Wigner, E. (1959), *Group Theory*, Academic Press, Inc., New York

Wigner, E. (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), Phil. Mag. Series 4, Vol. vii. Criticism of Boole's version of probability theory.

Williams, P. M. (1980), "Bayesian Conditionalisation and the Principle of Minimum Information", Brit. Jour. 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.

Woodward, P. M. (1953), *Probability and Information Theory, with Applications to Radar*, McGraw–Hill, N. Y. An interesting historical document, which shows prophetic insight into what was about to happen, but unfortunately just misses the small technical details needed to make it work.

Wrinch, D. M. & Jeffreys, H. (1919), Phil. Mag. **38**, 715–734. This was Harold Jeffreys' first publication on probability theory, concerned with modifications of the Rule of Succession. He must have liked either the result or the association, because for the rest of his life he made reference back to this paper on every possible occasion. Dorothy Wrinch was a mathematician born in Argentina, who studied at Cambridge University and later taught at Smith College in the United States and, in the words of Jeffreys, "became a biologist". Her photograph may be seen in Pólya (1987), p. 85. Two later papers by Wrinch and Jeffreys on the same topic are in the Phil. Mag. **42**, 369–390 (1921); **45**, 368–374 (1923).

Zabell, S. L. (1982), "W. E. Johnson's Sufficientness Postulate", Annals of Statistics **10**, pp. 1091–1099. Discussed in Jaynes (1986b).

———— (1988), "Buffon, Price, and Laplace: Scientific Attribution in the 18'th Century", Archive for History of Exact Sciences, **39**, 173–181.

———— (1989), "The Rule of Succession", Erkenntnis, **31**, 283–321. A survey of the long and tangled history of the subject, with a wealth of unexpected detail and an astonishing number of references, highly recommended. His attempt to assess the past criticisms and present status of induction represents a notable advance over Popper but still fails, in our view, to recognize how induction is used in actual scientific practice. Discussed in Chapter 9.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, J. Wiley & Sons, Inc., New York. Second edition (1987); R. E. Krieger Pub. Co., Malabar, Florida. In spite of the word "Econometrics" in the title, this work concerns universal principles and will be highly valuable to all scientists and engineers. It may be regarded as a sequel to Jeffreys (1961), carrying on multivariate problems beyond the stage reached by him. But the notation and style are the same, concentrating on the useful analytical material instead of mathematical irrelevancies. Contains a higher level of understanding of priors for linear regression than could be found in any textbook for more than 20 years thereafter.

Zellner, A. (1988), "Optimal Information Processing and Bayes' Theorem", American Statistician, **42**, 278–284, with discussion. Points to the possibility of a general variational principle that includes both Maximum Entropy and Bayesian algorithms as solutions. Discussed in Chapter 11.

Zubarev, D. N. (1974) *Nonequilibrium Statistical Thermodynamics*, Plenum Publishing Corp., New York. An amazing work; develops virtually all the MAXENT partition functional algorithm as an *ad hoc* device; but then rejects the MAXENT 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.