(181)

Rhetoric within the Citadel: Statistics

Donald N. McCloskey
University of Iowa

As a language of numbers growing in prestige, statistics plays a special role. It governs politics and business, fulfilling its name, an 18th-century Latin coinage by German cameralists from Italian statista: "pertaining to the statesman." Scientists in non-experimental sciences view the statistical test as the height of objectivity, a final court of appeal. Behind them the lay public supposes that statistical judgments are timeless and uncontroversial: What causes heart disease? How many murders does an execution deter? What was the population of 5th-century Athens, within three digits of accuracy?

In popular as in scientific rhetoric a statistic is a conversation stopper, half of the pair with mathematics: The Fact to match The Logic. Statisticians know better. They know for instance that outsiders will use statistics badly. The most obvious "rhetoric of statistics" concerns the use and misuse of statistics in arguments about IQ, economic time series, and certain cons of econometrics. The more sophisticated among the practioners -- the "-metricians" of econometrics, psychometrics, sociometrics -- have of course a special interest in how badly statistics can be used in their own fields (Gould 1981; Ames and Reiter 1961; Leamer 1978, 1983), but statisticians themselves have thought a good deal about the matter, under the headings of Statistical Fallacies and How to Lie With Statistics. The more argumentative of their elementary texts (Wallis and Roberts is the classic example; more recently Freedman, Pisani, Purves) concentrate heavily on what can go wrong in a statistical argument by someone who does not understand it. In its professional literature statistics has a well-developed genre of "Idiocies Perpetrated By Non-Statisticians of My Acquaintance," more so than does mathematics or engineering. Statistics is similarly in this respect to the other statecraft, economics.

Though somewhat routine, it is certainly of rhetorical interest to see the way statistics are used by non-statisticians. The main rhetorical point to be made is the routine one that statistics is expertise spoken numerically. Classical rhetoric would analyze the ploy under the heading of Greek ethos, which is to say the character claimed by a speaker, and other arguments from authority. The controversialist waving a number is claiming the authority of Science, as the conspicuous use of Greek claims the authority of scholarship. Study of the rhetoric would examine the contexts in which numbers tell, and the history of their telling.

The other routine rhetorical point about statistics used by non-statisticians concerns the rhetoric of mastery. The statistical amateur waving the Statistical Package for Social Science claims mastery of a technique. Statistical procedures, as distinct from the resulting numbers, are believed by non-statisticians to be techniques for generating truth. Graduate students in economics flock to the course on econometrics because they believe it is the place to learn economic science; the same is true nowadays in the study of politics. Their teachers are embarrassed to disabuse them because they can supply no other formula for science and the young demand formulas. The students are drawn by the illusion that factor analysis or instrumental variable techniques will mechanize scientific persuasion. As William Barrett remarks, and as any statistician could attest from the irrationally hopeful attitude of students, "[t]his worship of technique is in fact more childish than the worship of machines. You have only to find the right method, the definite procedure, and all problems in life must inevitably yield before it" (1979, p. 25).

It startles outsiders to hear that statisticians do not believe most of the results announced in their name. As Louis Guttman is fond of pointing out, for example, a test of statistical significance does not make it unnecessary to replicate experiments. Yet in practice the usual panoply of testing discourages replication. The polite silence of statisticians about the misuse of statistics is taken as indicating that everything is fine, a Type I error. The harshest demands for conformity to standards come from the wrong side, from the less sophisticated of the users of the methods. A little learning is a dangerous thing. The unsophisticated stand ready to attack a paper that uses qualitative variables in a regression or that does not report significance tests. Guttman notes that "believers in unfounded practices" are much more likely to complain that the practices are not followed than is an inventor of the statistic to point out that the practice is silly.

A richer vein of rhetoric, however, lies in the homelands of statistics itself. Disagreement shows the rhetoric. No statistician needs to be told that insiders sometimes disagree, and disagree on grounds they often cannot articulate. This or that statistician is a Bayesian or a decision theorist or an advocate of robustness or whatever. If he is a passionate man he knows passionately that his theory is the correct one; if a sober man he knows it soberly; anyway he knows it. For some reason, however, and despite his writing and talking, he cannot persuade all of his colleagues. He will mutter to himself: They are jealous no doubt of my brilliance, and content to remain fools; or experts in Harry Truman's definition, people afraid to learn anything new because then they wouldn't be experts. The disagreement comes down to "taste" or "philosophy" or "personality" and gets no further.

The question arises, why? The persistence of disagreement in science is usually bracketed as temporary. Studies of science usually pass quickly over the persistence of disagreement, lingering instead on the moment when plate tectonics won the day or when we all became Keynesians. Yet diagreement in some sciences persists for decades, as in geology and economics. So in statistics. The story implicit in statistical textbooks, for instance, bundles the Neyman and Pearson theory together with Fisher's approach as a unified "modern inferential statistics," fitting smoothly together in a history of progress. Fisher's personality probably explains some roughness of

fit. Yet how did he succeed in his opposition to Neyman and Pearson, so costly in alternative hypotheses mislaid? What audiences, what arguments, what figures of speech and logic made the opposition successful? Disagreement is as puzzling and illuminating for the study of science as is scientific progress or simultaneous discovery.

The puzzle is deepened by the way statisticians talk about argument, at any rate in their formal papers. (The arguments are described more critically in the hallways, over a drink at a convention, and in moments of jocular candor in the seminar room.) Statistics comes armed with a rhetoric of mathematical proof. If a new test is "proven," that is supposed to be that. But of course it is not. Reasonable statisticians go on disagreeing on important matters, such as whether or not statistical significance is more mischief than it is worth.

Mathematics settles some arguments but cannot settle all of them, even within mathematics itself. Stanislaw Ulam estimated around 1970 that 200,000 theorems were proven annually in mathematics (1976, p. 288), and he probably did not included mathematical statistics in the total. People do not change their mind about mathematics 200,000 times a year; no one is surveying the mass as a single conversation. Something must be determining which proofs matter. A proof that an estimator has certain asymptotic properties will leave many statisticians cold: "Show me the Monte Carlo studies," they will say, though others will regard such a demand as vulgar. Something more than proof appears to be involved in the adoption of arguments within professional statistics. "Logic" is of course a good thing, but it does not suffice to discriminate. Procedures come to be used or ignored on other grounds than mere adherence to first-order predicate logic, a hurdle of little power (though frequently knocked over). Somehow a choice is made among alternatives, most of which are logical. The loss function of statistical science is sensitive to more than the formal properties of a proof.

It seems plausible that the way of choosing can be thought of as "rhetoric," which is to say the art of argument, good and bad. The word "rhetoric" has the drawback of bringing to mind images of hokum and salesmanship. Hokum and salesmanship is not entirely foreign to science and scholarship, but rhetoric has more room. Some of rhetoric is "mere" rhetoric, which bad people use, but the rest is honest argument.

The quest for certainty casts rhetoric and argument into a disreputable position by comparison with proof. At first a mathematical proof of the Central Limit Theorem does not look much like "argument." Here is an Archimedean point from which to lever up the world. Here is certainty, a refuge from universal doubt. The rhetorical tradition, however, looks upon supposedly apodictic proof as a powerful argument, yet an argument among others nonetheless. Mathematics itself has a rhetorical tradition (Hersh 1987). Standards of proof, for example, change (Steiner 1975, pp. 103-107; David and Hersh 1981, pp. 34, 40, 344, 357, 406). And since the 1930s it has been known that mathematics can never be complete and fully decidable.

In any case, most argument in statistics is notably less decisive than the proof of the Pythagorean Theorem. Or to be exact, in statistics the range of application is what is at stake, and the range of application is always a human and rhetorical question. The Pythagorean Theorem is incorrect on the surface of a sphere, though good enough for the human purpose of land survey on this globe, unless the right triangle is the size of India. The Birkhoff-Khinchin Theorem is incorrect for non-stationary time series, but good enough for the human purpose of estimating the long run price of wheat in Europe 1500-1869, unless the average for detrending is too long to be of interest (Kendall, Stuart, and Ord, 1983, p. 506). This will not disquiet the science of uncertainty. From a statistical point of view the quest for certainty is what looks odd. A rhetoric of statistics would apply the scrutinized uncertainty that is statistics to statistics itself.

Conflicting intellectual values within the statistical fraternity are often said to be matters of "philosophy," leading X to be a frequentist and Y a subjectivist. "Philosophy," though, is at once too grand and too narrow a word for what usually amount to differences in vision, training, characteristic images, argumentative habits, and mathematical tastes. The issues are not narrowly philosophical: they do not involve the mind-body problem or the epistemological status of sense impressions. Philosophy is to the point only in a much broader sense, which would include the ways of persuading people honestly. The analogy that best describes the statistician arguing for a new procedure is not a philosopher in an oven doubting the evidence of his senses but a lawyer in a courtroom arguing a case. Rhetoric is the ancient word for such argument, understood to include argument from good mathematics and observation, of course, but one understood also not to exclude good arguments from authority, from ethical appeal, from analogy, from consequence, from precedent.

In the main philosophical tradition these further arguments are set to the side as fallacies. Inference itself is a fallacy in philosophy, a philosophical problem not yet solved. But practical reasoning uses daily the fallacy of valuing evidence from a thief less than evidence from a bishop, or the fallacy of inferring a parameter from a mere sample.

The statistical sciences give signs of being ready for rhetorical scrutiny. True, most meta-statistical thinking has called itself philosophical, concerned with interpretations of probability or decision theory, constructing ideal languages to do the job. Yet the informal talk about why an argument is good or bad will fit awkwardly into an overarching philosophy. Most of the rhetorical matters in science are not overarching.

Among theoretical and applied statisticians the advocacy of Bayesian arguments (in econometrics, Zellner 1971 and Leamer 1978) and, more recently, of robustness (Mosteller and Tukey 1977; Hogg 19xx) indicates a willingness to rethink the way of talking. Statisticans like

William Kruskal, John Tukey, Louis Guttman, Edward Leamer call for data analysis as against inference (describing the data rather than subordinating them to a supposed test of scientific law). They are appealing to a rhetorical standard. They say, "What is the argument here? Why are we doing all this? What is the practical purpose?" It is pragmatic, rhetorical talk, accepting the human purposes people have for the data and using a jack-knife to accomplish the purposes. It downplays the grandeur of testing hypotheses.

John W. Tukey told the story recently of another statistician visiting his class at Princeton and being scandalized that Tukey "talked about practical things to do." The visitor asked, "How could you tell that to students? I would do it myself, but I would shut and lock the office door first." Tukey remarks that such lack of rhetorical candor -- an unwillingness "to think hard about what our real strategies . . . have been" -- is one reason "why we have been unable to teach statistical data analysis" (1986, p. 74).

The recent shift in the history of statistics away from amateur and profession-building history has shown how rich and pragmatic were the arguments that led to the received techniques (Stigler 1978, 1986; MacKenzie 1981). Any statistical technique has a history in the conversations of statisticians, rising slowly out of disputed cases in gambling, mortality, social statistics, and agronomy, and codified over the past century or so from arguments more or less persuasive and more or less verbal. The mathematical expression of probabilities, of curve fitting, and above all of hypothesis testing wonderfully sharpened the points at issue. Yet the sharpening has led also to trimming of the official rhetoric, concealing the rich arguments that give a statistic meaning.

Statistics is rhetorical. That is to say, it depends on a conversation, an agreement among human beings about what they are going to take seriously. It is not immanent in nature. Louis Guttman put it this way: "Statistics . . . belongs to the scaffolding [of scientific theory.] . . . . [It] might be termed artificial . . . . The probabilities . . . are generated by a researcher-made sampling process. Should the researcher change his sampling design, he will change the probability structure of his statistics. There is nothing fixed in nature about the probabilities discussed in statistical inference" (1985, p. 1). By this he means that the researcher agrees to speak as if the sample were selected in a certain way, from an urn of nature configured in a certain (convenient) way: the urn would not for instance choose red balls out of proportion to the reds actually in the urn, except with a calculable probability. But all this is subjunctive (not "subjective": that is not the problem, for the agreement about speaking is already intersubjective). It is about how things would turn out if the sample size were to go to infinity, or if the same sample size were to be replicated many times. The subjunctive is the mood of social agreements.

Many of the fully 53 fallacies that Louis Guttman assembled in a curmudgeonly paper a decade ago turn on rhetorical matters. In the narrow sense of "the impact of vocabulary on thought" (as when we call the Germans "Huns" or the Contras "Freedom Fighters") the list of rhetorical problems is impressive: "statistical significance" becomes confused with scientific significance; the "normal" distribution is abnormal, but is convenient for use and normally gets used; "partial" correlation should be "conditional" correlation, which would emphasize that it does not "hold constant" anything but rather takes the variable conditional on stratification of other variables; averages are not "central tendencies" (in a U-shaped distribution the tendency is peripheral, not central, but the average for a symmetric loss function is the bottom of the U); "path analysis" and "causal modeling," despite their seductive names, do not solve the problem that correlation is not causation; substituting "explanation" or "determination" for "cause" does not solve it either.

Statistical argument depends on words. As William Kruskal argued in his "Formulas, Numbers, Words: Statistics in Prose": "Suppose that Sir R.A.Fisher -- a master of public relations -- had not taken over from ordinary English such evocative words as "sufficient," "efficient," and "consistent" and made them into precisely defined terms of statistical theory. Would his work have had the same smashing influence that it did? I think not, or at least not as rapidly"(1978, p.98). That is, the mere name of a statistical idea matters to its persuasiveness. A test for something called "significance," for example, can be expected to attract the attention of workers in the field. Statisticians have recognized since the phrase was coined that "statistical significance" can be misused, taken as substantive significance. Theoretical and applied statisticians have warned about this and other abuses of argument for sixty years (e.g. Boring 1919). In the 1960s a "Significance Test Controversy" developed in statistical psychology and sociology; the same controversy is breaking out now in economics; and statisticians continue to warn, unheeded, about the misuse of the tests (see again Freedman, Pisani, and Purves 1978). The rhetoric of the applied fields, which regard statistical argument as a machine for persuasion, makes the practice difficult to stop. A sociometrician apprehended in a meaningless significance test is liable to plead, "I have to do something." Weak as the argument is, it overrides better reasoning because the better rhetoric is hidden.

The language of the statistical conversation has consequences, from the logic of randomness to the misapplication of charts. The solution is frequently suggested to be dull language. In the Advanced Theory of Statistics Kendall and Stuart recommended that we speak not of "level of significance" but "size of the test. The sociometricians Morrison and Henkel recommended an even less sexy name for the corresponding test, "sample error decision procedure." No one is going to fight and die for a sample error decision procedure of arbitrary size.

The rhetoric of statistics, though, is more than a matter of truth in packaging. The issue is one of language more generally, and therefore one of audiences, conventions, characteric tropes, tones of voice, authorial intentions, and other pieces of language. The linguistic matter includes mathematics, too, for mathematics is a language.

For instance, the choice of null hypothesis in a test is a rhetorical choice, about which hypothesis is the maintained one. The question is, where is the burden of proof? The notion of burden of proof is rhetorical. (It is non-philosophical. Like the the rest of us a philosopher will assert a burden of proof on the other person, but she has no rhetorical rules for assigning it).

Or again, consider the argument from authority, the heart of modern statistical practice. To speak of the outsiders first, the consumers of mathematical statistics assume that the "producers" have worked it all out. "Just tell me how to do it," the consumer says, "not why it works or might not work." The attitude is frustrating to a statistician wanting to teach the students to become serious analysts of data. As Louis Guttman notes, "non-professional mathematicians and non-mathematicians . . . are shocked to learn that popular computer programs in internationally distributed packages are statistically or mathematically incorrect" (c. 1977, p.84). They assumed that anything so august as SPSS must be properly checked out, the same way one assumes that Rolls Royces will have infrequent repairs. Inside and outside professional statistics the authority of the mathematical statisticians is great. The central limit theorem and the like are impressive theoretical achievements. The problems arise from the practical achievement: that the authority of the theoretical achievements have certified a way of doing science for many thousands of people, some of whom would have been at a loss without the statistical techniques. It is comparable to the invention of the experimental method (and was intended to be so). Certain techniques are privileged by the authority of a Pearson or a Fisher. As Stephen Stigler has shown, the choices were not inevitable.

The ultimate in authority, oddly, is proverbial status, without acknowledgment. More sinned against than sinning (Lear, III, 2, 59); what's mine is yours (Measure, V,1, 539); the rest is silence (Hamlet, V, 2, 372). Gerd Gigerenzer has pointed out that "[t]ypically all the ideas (stemming from either Fisher or Neyman and Pearson) were presented in [25 textbooks he examined] . . . anonymously as the corpus of inferential statistics" (p. 20). The problem of a lack of rhetorical self-consciousness is illustrated by the case: "[T]he anonymous presentation of a monolithic `inferential statistics' facilitated the suppression of controversial issues. . . ." (pp. 20-21).

Even the tone of a statistical work carries a message. When it is not calling on the dignities of mathematical truth a statistical argument will often use an air of down-to-earth common sense. An example occurs at the headwaters of statistics, in Pascal. The Wager goes as follows: "Let us weigh up the gain and the loss involved in calling heads that God exists. Let us assess the two cases: if you win you win everything, if you lose you lose nothing. Do not hesitate, then; wager that he does exist" (p. 151). These sentences occur in the midst of a long passage in which Pascal speaks repeatedly of "you" making a choice: "How will you wager?" "Yes, but you must wager. There is no choice, you are already committed. Which will you choose then?" and so forth. The familiar tone, the very use of "you," is rare in the PensÇes. Pascal explains that it is not a "proof" – "it is by being without proof that [Christians] show they are not without sense"; and "Reason cannot make you choose either, reason cannot prove either wrong" (p. 150). It is a practical business matter; or, better, a counsel of reasonable choice, if not proof, the sort of pragmatic decision that must be made in life but that no certitude governs.

A rhetoric of statistics does not debunk it (revealing it to be Mere Words after all). It shows how statisticians converse (Klamer 1984). The language of the humanities proves useful for the task. In proving a theorem about a sampling distribution, for instance, a mathematical statistician will of course choose an audience for the argument, an implied reader who accepts silently certain conventions of the genre. Rhetoric scrutinizes the choice from the outside – noting the genre used, the authorial voice, the figures of mathematical speech that limit the audience.

The gain to the humanities is to demonstrate the range of rhetoric, the oldest tradition in the humanities applied to the youngest of methodologies in the sciences. The gain to statistics would be self-consciousness. (Admittedly the outcome might be bad, crippling in the way trying to ride a bicycle self-consciously could be; but then it would nonetheless be worth asking, why?) The self-consciousness would not be exactly "philosophical," although presumably the philosophers of statistics might be interested. It would be more "literary," that is, looking on the works of statisticians as texts to be dissected.

It is the sort of thing that historians of statistics often find themselves doing already, though there is no reason to confine it to the past; and, to repeat, it is the sort of thing that philosophers of statistics do, although there is no reason to confine it to the issues that concern them. The texts would be analyzed for their characteristic logical arguments, their figures of speech, their appeals to authority, and the other ways they achieve persuasiveness. The court of scientific opinion becomes persuaded in many ways. A rhetoric of statistics would examine the jurisprudence and the trial practice of the court.

Notes Towards a Bibliography on the Rhetoric of Statistics

Ames, Edward and Reiter, Stanley (1961). "Distributions of Correlation Coefficients in Economic Time Series." Journal of the American Statistical Association 56 : 637-656.

Arrow, Kenneth. (1959). "Decision Theory and the Choice of a Level of Significance for the t-Test." In Ingram Olkin and others, eds. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford: Stanford University Press.

Bakan, David. (1966). "The Test of Significance in Psychological Research." Psychological Bulletin 66 : 423-37. Reprinted in Lieberman, Bernhardt, ed. Contemporary Problems in Statistics: A Book of Readings for the Behavioral Sciences. New York: Oxford University Press, 1971.

Barrett, William. (1979). The Illusion of Technique. Garden City, NY: Anchor

Boring, Edwin G. (1919). "Mathematical versus Scientific Significance." Psychological Bulletin 16 : 335-338.

Cohen, J. (19xx). "The Statistical Power of Abnormal-Social Psychological Research: A Review." Journal of Abnormal and Social Psychology 65 : 145-153.

Cooley, T. F. and LeRoy, S. F. (1981). "Identification and Estimation of Money Demand," American Economic Review 71:825-844.

Davis, Philip J. and Hersh, Reuben. (1981). The Mathematical Experience. Boston: Houghton Mifflin.

Denton, Frank T. (1988). "The Significance of Significance: Rhetorical Aspects of Statistical Hypothesis Testing in Economics." In Arjo Klamer, D.N. McCloskey, and Robert Solow, eds. The Consequences of Rhetoric. New York: Cambridge University Press, forthcoming.

Denton, Frank. (1985). "Data Mining as an Industry." Review of Economics and Statistics 67 : 124-127.

Feige, Edgar. (1975). "The Consequences of Journal Editorial Policies and a Suggestion for Revision." Journal of Political Economy 83 (Dec): 1291-96.

Freedman, David, Robert Pisani, and Roger Purves. (1978). Statistics. NY: Norton.

Gigerenzer, Gerd. (no date). "Probabilistic Thinking and the Fight Against Subjectivity." Unpublished paper, Department of Psychology, UniversititNt Konstanz.

Gould, Stephen Jay. (1981). The Mismeasure of Man. New York: Norton.

Griliches, Zvi. (1976). "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis." In Terleckyj, N.E., ed., Household Production and Consumption. Studies in Income and Wealth, 40. New York: National Bureau of Economics Research.

Guttman, Louis. (1981). "What is Not What in Statistics?" The Statistician 26 (about 1977): 81-107; reprinted in I. Borg, ed. Multidimensional Data Representations: When and Why. Ann Arbor: Methesis Press.

Guttman, Louis. (1985). "The Illogic of Statistical Inference for Cumulative Science." Applied Stochastic Models and Data Analysis 1: 3-9.

Hersh, Reuben and Davis, P.J. (1987). "Mathematics and Rhetoric." In J. Nelson, A. Megill, and D. McCloskey, eds. The Rhetoric of the Human Sciences. Madison: University of Wisconsin Press.

Hogg, Robert. (19xx).

Kendall, Maurice annd Stuart, Alan. (1951). The Advanced Theory of Statistics. 2. London: Griffin.

Kendall, Maurice, Stuart, Alan and Ord, K.J. 1983. The Advanced Theory of Statistics. 3: Design and Analysis, and Time-Series. NY: Macmillan.

Klamer, Arjo. (1984). Conversations with Economists. Totowa, N.J.: Rowman and Allanheld.

Kruskal, William. (1968)(updated 1978). "Significance, Tests of." International Encyclopedia of Statistics. New York: Macmillan.

Kruskal, William. (1978). "Formulas, Numbers, Words: Statistics in Prose." The American Scholar 47 : 223- 229. Reprinted in Fiske, D., ed. New Directions for Methodology in Social and Behavioral Sciences. San Francisco: Jossey-Bass, 1981.

Kurtz, A.K. and Edgerton, H.A., eds. (1939). Statistical Dictionary of Terms and Symbols. New York: Wiley.

Leamer, Edward. (1978). Specification Searches: Ad Hoc Inferences with Nonexperimental Data. NY: Wiley.

Leamer, Edward. (1983). "Let's Take the Con Out of Econometrics." American Economic Review 73 : 31-43.

Mackenzie, Donald A. (1981). Statistics in Britain, 1865-1930: The Social Construction of Scientific Knowledge. NY: Columbia University Press.

Mayer, Thomas. (1975). "Selecting Economic Hypotheses by Goodness of Fit." Economic Journal 85: 877-83.

Mood, A. F. and Graybill, F. A. (1963). Introduction to the Theory of
Statistics. NY: McGraw Hill.

Mosteller, Frederick and Tukey, John W. (1977). Data Analysis and Regression. Reading, Massachusetts: Addison-Wesley.

Neyman, Jerzy and Pearson, E.S. (1933). "On the Problem of the Most Efficient Tests of Statistical Hypotheses." Philosophical Transactions of the Royal Society A 231: 289- 337.

Pascal, Blaise. (1662). Pascal Pensees. trans. A.J. Krailsheimer. Harmondsworth, Middlesex: Penguin, 1966.

Steiner, Mark. (1975). Mathematical Knowledge. Ithaca: Cornell University Press.

Stigler, Stephen. (1978). "Francis Ysidro Edgeworth, Statistician." Journal of the Royal Statistical Society A 141: 187-313.

Tukey, John W. (1986) "Sunset Salvo." The American Statistician 40 : 72-76.

Ulam, S. M. (1976). Adventures of a Mathematician. NY; Scribner's.

Wallis, W. Allen and Harry V. Roberts. (1956). Statistics: A New Approach. NY: Macmillan.

Yule, G. U. and Greenwood, M. (1915). "The Statistics of Anti-typhoid and Anti-cholera Inoculation and the Interpretation of Such Statistics in General." Proceedings of the Royal Society of Medicine 8.

Zellner, Arnold. (1971). An Introduction to Bayesian Inference in Econometrics. NY: Wiley.