THE HISTORY OF STATISTICS IN THE CLASSROOM

H. A. David, Department of Statistics, Iowa State University

Abstract

I strongly believe that statistics classes at any level can be enlivened by highlighting colorful contributors to our field. Outlines of some of their research should be supplemented by historical, biographical, and anecdotal material. Laplace (1745-1827) provides a good example. Famous as a theoretical astronomer and mathematician, he is probably best known to statisticians as originator of the central limit theorem. But apparently independently of the publication of Bayes's theorem ten years earlier, he proposed that (in modern language) posterior density is proportional to likelihood. Laplace lived in turbulent times in France. He examined the young Napoleon, but lasted only six weeks in Napoleon's cabinet. Similar introductions will be given of Gauss and R. A. Fisher, the latter with personal comments.

1. Introduction

In an earlier paper (David, 1999) I outlined a short course on the history of statistics given to mainly graduate students at Iowa State. The theme of the present article is that the history of statistics, judiciously used by the instructor, can enliven regular courses, whatever their level. There are many colorful contributors to our field for whom a brief indication of some of their brilliant pioneering work can be accompanied by historical, biographical, or anecdotal material. This will add spice to the instructor's main dish and like all spice should be used sparingly, as appropriate to those partaking.

To illustrate these ideas I have chosen three giants in the development of statistics, Laplace, Gauss, and R. A. Fisher. For each, material at different levels is provided for selection by the instructor. Finally, I list some references. With the recent publication of a number of excellent books, this is a good time to introduce some history of statistics into statistics courses.

2. Laplace (1749-1827)

A look at the dates tells us that Pierre-Simon Laplace was born in turbulent times. He attained fame early, showing that the solar system was stable on the Newtonian model whereas Newton himself had thought it needed an occasional divine nudge. As professor in the Paris Military Academy he examined the 16-year old Napoleon for his commission in 1785, but in 1799 he lasted only six weeks in Napoleon's cabinet, removed as Minister of the Interior for trying to "carry the spirit of the infinitesimal into administration." However, he was made a Senator and became Senate President. Earlier, he was lucky to survive the French Revolution in which several of his friends, including the famous chemist Lavoisier, were guillotined. After the revolution Laplace was a major force in the commission of weights and measures set up to introduce the metric system. His fame was such that when the Bourbon monarchy was restored Louis XVIII made him a Marquis. Today a Paris Metro Station carries his name.

Highly regarded as a theoretical astronomer and mathematician, Laplace is best known to statisticians for the central limit theorem which he arrived at non-rigorously in 1810. This theorem has, of course, played a tremendous role in the development of asymptotic theory. He uses it, for example, to introduce what is essentially the asymptotic relative efficiency of two estimators. A special case of his results is that in a random sample the median is asymptotically more efficient than the mean if

$$f(0) > 1/(2\sigma),$$

where the parent density $f(x)$ is symmetric about zero, with variance σ^2 (Laplace, 1818). Laplace goes on to show that for a normal distribution no linear combination of mean and median improves on the mean alone, a precursor to the sufficiency of the mean established by R. A. Fisher a century later.

But already in 1774 and apparently independently of the posthumous publication of Bayes's theorem ten years earlier, Laplace had proposed, in modern terminology, estimating a parameter θ by taking posterior density $\propto$ likelihood, or

$$f(\theta \mid x_1,\ldots,x_n) \propto f(x_1,\ldots,x_n \mid \theta).$$

This amounts to assuming the prior density $f(\theta)$ to be uniform, an assumption having little effect for large n on the estimate of θ obtained by maximizing the likelihood. Laplace's proposal found wide acceptance under the term "inverse probability". It was only in the second half of the 20th century that the term "Bayesian inference" took over, Fisher (1950, p. 1.2b) apparently having been first to use the adjective "Bayesian".

See also David and Edwards (2001) and Stigler (1986) for translations with commentary of Laplace (1818) and Laplace (1774), respectively. Laplace's (1812) book contained an astonishing wealth of material and remained the most important text in probability and mathematical statistics throughout the 19th century.

3. Gauss (1777-1855)

Carl Friedrich Gauss was a younger contemporary of Laplace. Born in Brunswick, in the German Duchy of the same name, his mathematical talent manifested itself at an early age. There is the oft-told tale of how Gauss spotted an error in his father's addition of a column of figures, at age 3, without any instruction in arithmetic. When he was barely 9 his teacher gave Gauss's class the task of adding 1 + 2 + … + 100, expecting to keep the class busy for some time. But Gauss had the answer almost at once. He paired the numbers as follows

$$\begin{array}{cccc} 1 & 2 & \ldots & 50 \\ \underline{100} & \underline{99} & & \underline{51} \\ 101 & 101 & & 101 \end{array}$$

giving 50 x 101 = 5050 (Bieberbach, 1938).

This precocious talent led to an introduction to the Duke of Brunswick who supported the impecunious Gauss from high school to Ph.D. In 1801 Gauss scored two triumphs which made him instantly famous. On the basis of a small number of available observations he correctly predicted the location of the minor planet Ceres after it had been lost out of sight by astronomers. Of more lasting importance, he published a famous book on the theory of numbers. Written in German, the book was translated into Latin for wider readability, at the insistence of the publisher. Gauss, who had wavered between a career in philology or mathematics, did this and later translations himself, with little outside help.

Whereas Laplace, in dealing with astronomical or geodetic data, minimized the sum of the absolute deviations between observed and theoretical values, Gauss minimized the sum of squares. There is little doubt that Gauss used the method of least squares years before Legendre published and named the method in 1805. Gauss had not published the method, regarding it as obvious, but unfortunately could not restrain himself from bringing up his prior use, calling it "our method". The older Legendre was greatly upset, writing that Gauss had already acquired such great fame that he did not need to claim the method of least squares as well (see also Hald, 1998, p. 394). Gauss remained slow to publish.

Legendre had provided a computational method that caught on immediately. We owe to Gauss the theoretical underpinning (Gauss, 1809, 1823). Of course, the Gauss linear model was not presented in matrix form (matrices did not enter the statistical literature until the 1930s). Linear functions were written as, e.g., $cx + c'x' + c''x'' + \text{etc.}$

Gauss's work on the linear model is widely recognized. Much less well known is that Gauss also seems to be the first to have studied estimators of variability (Gauss, 1816). Using mainly large-sample theory he examines, in modern terminology, estimators of the precision $h = 1/\sigma\sqrt{2}$, where σ is the standard deviation of a normal $N(0,\sigma^2)$ sample $X_1,...X_n$. Of estimators based on $S_k = \sum_1^n |X_i|^k$, $k = 1,...,6$, he finds best $\hat{h} = (n/2S_2)^{\frac{1}{2}}$, which is the maximum likelihood estimator. But for ease of calculation he also considers med$|X_i|$, the median absolute deviation (MAD)! For an English translation of Gauss's German paper, with commentary, see David and Edwards (2001).

An extended account of Gauss's statistical work is given by Sheynin (1979).

4. R. A. Fisher (1890-1962)

Ronald Aylmer Fisher arguably contributed more to the theory and practice of statistics than anyone else. An idea of his influence may be gained from the statistical terms he has left us. He was usually the originator of the area described by each term, and if not, was a major developer.

Terms introduced by R. A. Fisher

analysis of variance, ancillary statistics, Bayesian (!), covariance, confounding, consistency, degrees of freedom, factorial design, fiducial, hierarchical, level of significance, likelihood, maximum likelihood, null hypothesis, pivotal quantity, randomization, randomized blocks, sampling distribution, scaling, score, test of significance, statistic, sufficient statistics, variance

With such a wealth of contributions his name comes up often. To illustrate his brilliance we sketch an example not covered in the above list, an early paper on extreme-value theory (Fisher and Tippett, 1928).

Let $X_1, X_2, \ldots$ be i.i.d. variates with cdf $F(x)$. What is the limiting form of the distribution of $X_{(n)} = \max(X_1, \ldots, X_n)$, suitably normalized (if such a limiting form exists)? The limiting form, $\Lambda(x)$ must satisfy the functional equation

$$\Lambda^n(x) = \Lambda(a_n x + b_n),$$

where $a_n > 0$ and b_n are constants.

If $\alpha \neq 1$, then $x = ax + b$ when $x = b/(1-\alpha)$. At this point $\Lambda^n = \Lambda$, i.e., $\Lambda = 0$ or 1.

Consequently the solutions fall into 3 classes:

1. $\alpha = 1$ $\qquad \Lambda^n(x) = \Lambda\ (x + b_n)$
2. $\Lambda = 0$ when $x = 0$ $\qquad \Lambda^n(x) = \Lambda(\alpha_n x)$
3. $\Lambda = 1$ when $x = 0$ $\qquad \Lambda^n(x) = \Lambda(\alpha_n x)$

For 1. $\Lambda(x) = e^{-e^{-x}}$ $\qquad -\infty < x < \infty$.

$b_n = \log n$

Fisher set down his applied ideas in his famous *Statistical Methods for Research Workers,* first published in 1925. It reached 14 editions and was translated into 6 foreign languages. Fisher was a master of the English language, but his concise style did not make for easy reading. Snedecor's *Statistical Methods*, first published in 1937, helped further to popularize Fisher's work.

The symbol F was introduced by Snedecor in honor of Fisher, who however preferred his earlier

$$Z(=\frac{1}{2}\log F).$$

In fact, Table V of the *Statistical Tables* by Fisher and Yates gives percentage points of z and e^{2z} !

This is a very mild example of Fisher's quirkiness. He could, in fact, be very harsh to opponents of his ideas and his battles with Karl Pearson and Jerzy Neyman are legendary. As I can testify, other statisticians were not safe from his invective either, but he was encouraging to the young. An excellent biography, warts and all, has been written by one of his daughters, Joan Fisher Box (Box, 1978). The subtitle, The Life of a Scientist, reminds us that in spite of his many fundamental contributions to statistics, Fisher was also a leading figure in quantitative genetics. This remarkable man even did his own experimental work on plants and small animals.

References Cited

Bieberbach, L. (1938). *Carl Friedrich Gauss: Ein deutsches Gelehrtenleben*. Keil, Berlin.

Box, J. F. (1978). *R. A. Fisher: The Life of a Scientist*. Wiley, New York.

David, H. A. (1999). A 1-credit course on the history of statistics. *Proceedings of the Section on Statistical Education*, pp. 245-248. American Statistical Association, Alexandria, VA.

David, H. A. and Edwards, A. W. F. (2001). *Annotated Readings in the History of Statistics*. Springer, New York.

Davis, C. H. (1857). *Theory of the Motion of Heavenly Bodies*. Translation of Gauss (1809). Little Brown, Boston.

Fisher, R. A. (1950). *Contributions to Mathematical Statistics*. Wiley, New York.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings Cambridge Philosophical Society*, **24**, 180-190. Reproduced in Fisher (1950, Paper 15).

Gauss, C. F. (1809). *Theoria Motus Corporum Coelestium*. In *Werke* **7**, 1-280. Translated by Davis (1857).

Gauss, C. F. (1816). Bestimmung der Genauigkeit der Beobachtungen. In *Werke* **4**, 109-117.

Gauss, C. F. *Werke*, 12 vols., 1863-1933. Königliche Gesellschaft der Wissenschaften zu Göttingen.

Hald, A. (1998). *A History of Mathematical Statistics from 1750 to 1930*. Wiley, New York.

Laplace, P. S. (1810). Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur applications aux probabilités. *Mém. Acad. Roy. Sci.*, Paris, pp. 353-415.

Laplace, P. S. (1812, 1818). *Théorie analytique des probabilités* and *Deuxième supplément*. Courier, Paris.

Legendre, A. M. (1805). Nouvelles méthodes pour la détermination des orbites des comètes. Courcier, Paris.

Sheynin, O. (1979). C. F. Gauss and the theory of errors. *Archive for History of Exact Sciences* **20**, 21-72.

Stigler, S. M. (1986). Laplace's 1774 memoir on inverse probability. *Statistical Science*

1, 359-378.

Other General References on the History of Statistics

Hald, A. (1990). *A History of Probability and Statistics and Their Applicatons before 1750*. Wiley: New York.

Heyde, C. C. and Seneta, E. (eds.) (2001). *Statisticians of the Centuries*. Springer: New York.

Johnson, N. L. and Kotz, S. (eds.) (1997). *Leading Personalities in Statistical Sciences*. Wiley: New York.

Kotz, S. and Johnson, N. L. (eds.) (1991). *Breakthroughs in Statistics*, Vols. I, II. (1997) Vol. III. Springer: New York.

Pearson, E. S. and Kendall, M. G. (eds.) (1970); Kendall, M. G. and Plackett, R. L. (eds.) (1977). *Studies in the History of Statistics*, Vols. I, II. Griffin: London.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Harvard University Press: Cambridge.

Stigler, S. M. (1999). *Statistics on the Table: The History of Statistical Concepts and Methods.* Harvard University Press: Cambridge.

See also the MacTutor History of Mathematics archive, University of St. Andrews, Scotland, created by J. J. O'Connor and E. F. Robertson, at http://www.history.mcs.st-and.ac.uk/history. This gives portraits and much general information on Laplace, Gauss, Fisher and some other important contributors to statistics.