THE RISE OF STATISTICAL THINKING, 1820–1900.
by Theodore M. Porter
Princeton, 1986
333 pp. $35

THE HISTORY OF STATISTICS: The Measurement of Uncertainty before 1900
by Stephen M. Stigler
Harvard, 1986
410 pp. $25

The two most essential tools for modern mathematical statistics—the "method of least squares" and the "normal" curve—were developed by the second decade of the 19th century. Yet the great intellectual synthesis that turned statistics into an academic discipline with a pervasive impact on society was not accomplished until the end of the century. How the synthesis came about, and why it took so long, is the chief puzzle that both of these books seek to solve.

Stigler's book is a statistician's history, full of equations, signally concerned with the immensely difficult task of translating old results into terms that modern statisticians can easily understand, in order to make clear just what each pioneer in the field achieved. His focus is on his subjects' mathematical innovations, not their social or political thought or epistemological positions. Thus he largely ignores men like the "statists," as students of human society were often known during the 18th and early 19th centuries, and the physicists, neither of whom had much *direct* effect on the mainstream of mathematical statistics.

Porter's book, by contrast, is that of a historian of science, concerned not only with the methods that the statisticians invented but also with their social and political thought. Assuming, in effect, that his readers already know the importance of every technical advance, Porter includes very few equations. Instead, he quotes extensively the words of many whose contributions appear to have been primarily rhetorical.

By 1805, a Frenchman, Adrien Marie Legendre, had developed the "method of least squares," which allowed astronomers and geodesists to combine many slightly different observations of the same event in order to determine, for instance, the "true" orbit of a comet or planet. By 1811, mathematicians Pierre Simon Laplace and Carl Friedrich Gauss had presented rigorous mathematical treatments of the "normal" or bell-shaped curve and shown how it could be used to assess errors in estimation.

But this understanding of statistics as a tool to reduce discrepancies in measurements or to determine deviations from an average value was limited and constraining. It forced students of society or of nature to explain away variation. It also led them into a view that all events, natural or human, were governed by a single "normal" law. The breakthrough, pioneered by James Clerk Maxwell in physics and Francis Galton in biometry, was to see statistics as a means of uncovering and explaining variation and of determining the influence of many separate causes on outcomes.

Since Isaac Newton's day, natural scientists, particularly astronomers and physicists, had been able to employ both experimental controls and theories to limit their concern to a small number of variables. Numerical methods had been used chiefly to reduce what would now be termed "measurement error." By contrast, theory as used in the social and biological sciences was weaker. Relevant experiments were not always feasible, and those who studied living things were so conscious of the large number of potentially important influences and the roughness of data that they hesitated to employ the new statistical notions.

When men trained in astronomy, such as Belgian polymath Adolphe Quetelet, sought to replace the *ad hoc* techniques of the statists with rigorous notions drawn from probability, they introduced determinism and a narrow focus on mistakes in observation. Quetelet's concept of the "average man" and his all-too-successful efforts to show that data on almost any subject obeyed the "normal law" actually impeded rather than advanced the development of mathematical statistics. Quetelet quite simply failed to distinguish particular causal factors from one another.

Nonetheless, Quetelet was the pivot of 19th-century statistics and its influence on social thought in two respects. First, his view that variations among individuals were much less important than aggregate averages partly inspired James Clerk Maxwell's kinetic theory of gases, which introduced statistics into physics. (The actions of particular men or molecules might be unpredictable, but those of a "typical" entity of either sort could be described fairly accurately.) Second, social, medical, and physical scientists were moved by a religious fervor to vindicate the doctrine of free will against the exaggerated determinism of Quetelet and his chief popularizer, historian Henry Thomas Buckle. The scientists' efforts, Porter claims, helped lead to the eventual popular acceptance of the idea of uncertainty.

However necessary such criticisms of determinism were to the identification of the problem of variability, they did not, by themselves, provide methods for measuring variation. Those methods were developed by the great English student of heredity Francis Galton (1822–1911) and his followers, especially Karl Pearson. They realized that, among other things, "least squares" could be used to describe and account for individual differences in physical and mental characteristics. Together, they laid the foundations of mathematical statistics.

The chief difficulty with each of these complementary books is that neither author sets forth very clearly, much less attempts to test, any overarching notions about why the changes he describes took place. Stigler emphasizes that statistics served as a substitute for theories and experiments in social science. Porter says that expert opinion seemed to shift automatically from an emphasis on data to one on diversity and variation. But neither makes a fully convincing case.

Modern applied statistics is concerned with testing hypotheses against other hypotheses. It is ironic, then, that the authors of these innovative studies concentrate on interesting individual variations, but have not found ways of appraising explanations for the correlations and regressions in thought that took place during the formative years of statistics.

—J. Morgan Kousser '86