Review of Stigler, Stephen M., The seven pillars of statistical wisdom. Harvard Univ. Press (2016)

Michael Evans
University of Toronto

This interesting book opens with the question "What is Statistics?". Probably most people believe that they have a reasonable answer to this question and might be surprised to learn that statisticians themselves are somewhat uncertain as to the answer. This uncertainty is expressed by the diversity of responses one would obtain from a sampling of professional statisticians. Rather than deal with this question specifically, the author states his goal as "to articulate the central intellectual core of statistical reasoning". This program is achieved via the discussion of seven key ideas/concepts which lie at the heart of the discipline.

It would be easy to regard the disagreements within the discipline concerning, for example, the role of subjectivity, the necessity for frequentist assessment criteria of statistical methodology, decision theory versus inference, etc., as indicative of a subject in crisis or at least one lacking firm foundations. These issues are not simple, however, and, at least for this reviewer, philosophy rather than mathematics lies at the heart of any resolution of most of them. This is probably a slow evolutionary process rather than a quick fix. It is fair to say too that resolution will have significant impact because of the ubiquity of statistical reasoning in many fields. The current well-publicized discussions over the use of p-values to assess evidence demonstrates this very clearly.

The virtue of the book, and it is something that anyone who has studied the subject will very likely attest to, is that, in spite of the differences of opinion about a variety of matters, there is still a lot to be learned that is of enormous value both practically and intellectually. Such material will surely be part of the subject in some fashion no matter what form its future development takes. My favorite example, and it is discussed in the book, is the clarity statistics brings to the discussion of what it means for variables to be related and, moreover, when can it be claimed that such a relationship is a cause-effect relationship. Conditional distributions and randomization play key roles in the discussion of relationships among variables and it is hard to see how one would come up with such a clear and incisive treatment of this undoubtedly important topic without thinking statistically.

The author identifies the seven pillars as aggregation, information measurement, likelihood, intercomparison, regression, design of experiments and residual. Each is treated in a separate chapter and the writing is very clear and engaging. One of the features of the book is that the historical development of each pillar is discussed and, in fact, this is typically the main focus. From each topic treated one gets a sense of just how hard won some of these concepts are as they appeared to fly in the face of "common sense" or at least were contrary to traditional practice.

Even the simple idea of taking an average of observations was not obvious as it was thought it would be better to try and identify the "best" observation

from the group. The story of the simple average as an estimate in the chapter on aggregation turns out to be much more interesting than one might imagine before reading about it. For example, some of the earliest development of this concept is associated with measuring the discrepancy between true and magnetic north which was of immense importance to navigation.

The chapter on information measurement is for the reviewer really about the measurement of the accuracy of an estimate. Certainly an estimate without some assessment of its accuracy is not terribly useful and so this is a necessary part of such a statistical analysis. The use of the word information in this context seems a bit nebulous, however, as it is not clear what is meant by the very word and furthermore it isn't clear that a measure of accuracy is the expression of the information in a data set. Surely it is part of it but the estimate itself could be viewed as the primary expression of the information in the data. After all that is the goal of the statistical reasoning in this context, namely, to express what the information (or better still, the evidence) in the data says about the true value of some quantity of interest. In any case, this is a minor quibble and the chapter makes a number of excellent points about the assessment of accuracy that took a considerable amount of development over many years. The result that the variation in a sum does not increase proportionately with the number of independent terms added is a particular case in point.

The chapter on likelihood focuses on the other grand theme, after estimation, that a theory of statistics is to address, namely, hypothesis assessment or testing. It introduces a number of concepts that play a central role in modern statistics such as the significance test and p-value, Bayesian inference and the likelihood function. The chapter explains very clearly how and why these ideas were developed. There is still a reasonable degree of controversy that surrounds the use of p-values or Bayesian methodology but perhaps one can take hope from the historical developments discussed and imagine a future where these issues become settled after a deeper understanding is obtained of what the problem is.

The remainder of the book is full of interesting commentary on the various pillars discussed. In teaching this material it is often difficult to convey the real significance of these developments, although the existence of this book will surely help in this regard. In the concluding chapter the author discusses the current emphasis on the application of algorithms to "big data" to solve problems and makes the point that sound statistical reasoning is hard and it will take time and careful thought to settle on appropriate methodology. The historical record proves this, as the book demonstrates, even for what we may now consider to be very simple problems. This is an encouraging message for the field.

Learning to reason statistically helps to make one a clearer and more logical thinker about important issues in the world. Part of the achievement of this book is that it makes some of this available to the general reader without the necessity of having to delve into more technical aspects of the subject. Overall, this book is recommended to anyone with even a modest curiosity about the quantitative aspects of science.