International Journal of Forecasting 28 (2012) 762–763

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Book review

Interpreting Economic and Social Data: A Foundation of Descriptive Statistics, Othmar Winkler. Springer, (2009). 265 pp., Hardcover, ISBN: 978-3-540-68720-7

According to its author, Prof Othmar W. Winkler, this book was written for *"everybody who has to make sense of statistical data and use them"* and as an attempt *"to rehabilitate descriptive statistics, the neglected Cinderella of the profession, nudging statistical theory in the indicated, different direction"*. As someone with interests in business statistics, I wanted to learn more about this direction.

In fact, over 40 years ago Andrew Ehrenberg (1926–2010) nudged Prof Maurice Kendall (1907–1983) in this same direction. His discussion of Kendall (1968) was limited to two short paragraphs describing two small oversights in Kendall's second look at the future of statistics. In Ehrenberg's own words, the second oversight is as follows:

> In paragraph 9 the word "supposedly" has been missed out. In contrasting descriptive statistics and theoretical statistics, Dr Kendall annotates the latter as a branch of scientific methodology. This should read: "Theoretical statistics, supposedly a branch of scientific methodology". It is descriptive statistics and scientific method which have of course to become fully one...

The above quotation helps illustrate the main theme of Winkler's book: there may not be one general theory of statistics that would apply equally well to all fields as an integral part of scientific methodology. Prof Winkler believes that business, economic and social statistics should go beyond sampling and inference, and that this kind of statistics should *"be reoriented towards interpretation of the phenomena of society through all kinds of data, not only from samples"* (p. 9). This is because *"socio-economic statistical data are quite different from the measurements in the sciences"* (p. vii).

Much of what Prof Winkler discusses in his book is surprisingly closely paralleled in Prof Ehrenberg's (early) writings on statistics. For example, the concluding sentence of Ehrenberg (1976) also suggests that statistics should be concerned with *"the analysis of a wider range of data rather than that of an isolated set of readings"*. The introduction of Winkler's book can be summarised well by a sentence from the introduction of Ehrenberg's article: *"Statistics courses are largely irrelevant—not just boring or technically difficult, but irrelevant"*. However, these parallels have not been made explicit in the book.

The book has 12 main chapters. It begins with a description of the developments that have led to the present situation in socio-economic statistics (Chapter 1), clarifying the nature of socio-economic statistical data and the role statistics plays in capturing the socio-economic phenomena (Chapter 2). The notion of *"statistical-counting-units"* is introduced here as a simplified record of *"the relevant 'real-life objects' which portray that social or economic phenomenon: human beings, entities such as corporations, or events, such as births, work accidents or business merges"* (p. vii). The author then discusses the aggregates of statistical-counting-units (Chapter 3) and their ratios (Chapter 4) as things that represent *"the bulk of data in socio-economic statistics"* and *"effective analytical tools"*, respectively.

The topic of the analysis and interpretation of longitudinal data is then discussed in three parts: Part 1—Looking to the past (Chapter 5); Part 2—Looking to the future (Chapter 6); and Part 3—Index numbers (Chapter 7). The next two chapters are devoted to the analysis and interpretation of cross-sectional data, in one dimension (Chapter 8) and in multiple dimensions (Chapter 9). The remaining chapters are concerned with the interfaces between socio-economic statistics and probability (Chapter 10), accounting (Chapter 11) and geography (Chapter 12). Because it does not seem possible to convey the ideas behind each of the above chapters, even briefly, I will only comment on the ideas behind the two chapters that are of particular interest to this journal's readership.

Consistent with the principal direction of the book, in Chapter 5, Prof Winkler challenges an opinion once expressed by Parzen (1962):

> In my opinion, it is of vital importance that the present tendency toward a schism between the statistical literature, the engineering literature and the economics literature be arrested with the aim of developing a true inter-disciplinary field of time series analysis.

0169-2070/$ – see front matter
doi:10.1016/j.ijforecast.2011.12.001

Winkler opines that it is unfortunate that *"statisticians have paid attention only to the numerical aspect of time series, making the search for patterns of variation the main objective. They seem to forget that socio-economic time series are the quantitative history of a situation"*. He suggests that the interpretation of time series *"ought to be above all historiographical"*.

The book's sole chapter on forecasting, namely Chapter 6, is introduced with a quotation from Drucker (1993, p. 118):

> We must start out with the premise that forecasting is not a respectable human activity and not worthwhile beyond the shortest periods. Strategic planning is necessary precisely because we cannot forecast.

The chapter's focus is on the predictive value of statistical data, and the key concept is that of data obsolescence. The gist of the chapter is that the central premise of statistical sampling theory—larger samples allow for more reliable conclusions about a population—does not translate directly to time series forecasting, where longer time series do not necessarily mean better forecasts. This is because the data contained in a time series are not sample observations drawn at random from an immutable population, but a collection of populations ordered in time.

Prof Winkler goes on to argue (p. 90) that *"the failure to understand the true, descriptive nature of socio-economic time series data has kept forecasting methods and forecasters from limiting their time series to only the relevant, more recent and often short parts of a time series"*, and that *"although forecasters have been intuitively aware of data obsolescence, nothing has been written or done about it"*. In this respect, it should be pointed out that the idea of discounting the value of more distant observations was introduced to time series forecasting in around the 1950s under the name of "exponential smoothing", and business forecasters have been making good use of it since then.

Prof Winkler mentions exponential smoothing a couple of times, but in somewhat negative contexts. He argues (p. 90) that the weights should be assigned by *"informed expert judgement, not mechanically by a mathematical formula, like in 'exponential smoothing'"*. Clearly, this is only possible for very important time series, conditional on the experts' charges.

The earliest paper in my collection which is most relevant to this chapter's ideas is that of Gilchrist (1967). Its author wrote:

> A choice of forecasting method thus consists of an examination of past data to select possibly a number of reasonable models. These are then examined from the point of view of goodness of fit not only to the whole data but also to smaller regions of the data.

While focusing on *"the exponential weighting system"*, Prof Gilchrist emphasised that *"the choice of a reasonable system of weights to use"* is as important as choosing the parameters in a given system of weights. For example, the simple predictor $\hat{x}_{t+1} = (3x_t + 2x_{t-1} + 1x_{t-2})/6$ features a system of weights which makes use of the last three observations only, and in such a way that the observation's influence on the forecast decreases as the observation ages.

The last of the main chapters is followed by an unnumbered chapter, titled "Afterthoughts", in which the author tells the reader that the ideas presented in his book *"evolved during six decades of teaching and doing applied statistical research"*, and that the book itself has become *"a statistical autobiography"* (p. 229). Prof Winkler also reveals that the original idea of publishing the book under the title "A foundation of statistics" was not welcomed by the editors. They wrote: *"This is not even statistics, forget about publishing it to avoid the embarrassment for the publisher and the author"*. The book has eventually been published, with a version of the original title as its subtitle. It is this subtitle alone which is responsible for the present review having been written.

In my view, most statisticians (including those of the types of both Kendall and Ehrenberg) would find the book somewhat pedestrian and simplistic. What I personally disliked about it was not limited to the fair number of typos and the insufficient linkages to the relevant literature, including the author's own publications in refereed journals. What I personally liked about it included the author's careful choice of quotations to introduce each chapter and the fact that he has a strong belief in his ideas. As John Stuart Mill once said, *"One person with a belief is equal to a force of 99 who have only interests"*.

In conclusion, the book is written to share the author's belief that *"social and economic statistics, though numeric, is essentially quantified history of society, not a branch of mathematics"* (p. 232). Those who are close to this belief (or those who are yet to form their views on the subject) may find the book interesting.

References

Drucker, P. F. (1993). *Management: tasks, responsibilities, practices*. Harper-Row.

Ehrenberg, A. S. C. (1976). We must preach what is practised: a radical review of statistical teaching. *Journal of the Royal Statistical Society, Series D*, *25*(3), 195–208.

Gilchrist, W. G. (1967). Methods of estimation involving discounting. *Journal of the Royal Statistical Society, Series B*, *29*(2), 355–369.

Kendall, M. G. (1968). On the future of statistics–a second look (with discussion). *Journal of the Royal Statistical Society, Series A*, *131*(2), 182–204.

Parzen, E. (1962). Book review: Time series analysis by E.J. Hannan. *Journal of the American Statistical Association*, *57*(298), 505–506.

Andrey Kostenko
Monash University, Australia
E-mail address: akoste01@gmail.com.