

Preface

It is the responsibility of scientists never to suppress knowledge, no matter how awkward that knowledge is, no matter how it may bother those in power. We are not smart enough, to decide which pieces of knowledge are permissible and which are not. . .

Carl Sagan*

Introduction¹

On a snowy winter morning I boarded a crowded city bus, unable to use my bicycle for the usual 10 km commute to Georgetown University. In preparation for teaching that morning, I began perusing the textbook on business and economic statistics that I had adopted for this course. At the next stop, a young woman took the seat next to me, the only one remaining in the full bus. Shortly after settling in, she turned to me: ‘Excuse me sir, is this statistics?’ she motioned to the textbook. ‘Yes’, I responded, surprised, ‘Business- and Economic Statistics.’ At this, a look of revulsion overcame her ‘Ugh. . . Statistics was the only subject I could never handle in college. . .’ She trembled at a memory that still haunted and upset her. At this reaction a series of similar, though less dramatic, occurrences came to mind.² Few other academic subject seem to evoke the distaste that the mention of statistics seems to elicit. Does it have to be that way?³ I grappled with possible explanations for a long time This book is my response that evolved gradually over decades of teaching a variety of business, economic and general statistics courses, using the newest textbooks available, and being involved in survey work and statistical consulting. I wondered why these textbooks on business and economic statistics presented the subject matter as a watered-down version of mathematical statistics, which itself evolved from problems of measurement and observation in the natural sciences. These textbooks treat socio-economic data like the measurements in the natural sciences and present the subject as an application of probability, grouped around the Gaussian, Poisson, F, χ^2 and other statistical distributions, sampling and statistical inference. There was no interest or concern about how to interpret the messages about society contained in the wealth of published economic and social data.⁴ They fail to see that this is the *raison d’être* of the entire statistical enterprise which also should be the main purpose of statistics courses for social scientists. These courses fail to present statistics

* Quoted from “Be careful what you pray for. . . You just might get it” Larry Dossey, M.D. Harper, San Francisco, 1997, p. 165.

as the instrument for scanning the economic and social environment and to monitor important aspects of social reality.

It is the aim of this book to re-orient statistics towards making sense of economic and social data. It is an attempt to rehabilitate ‘descriptive statistics’ as a respectable part of statistics, re-orienting it toward the description of society which in fact was its original purpose and still is the ultimate goal of all statistical endeavors. This book is addressed to the literate and numerate public, trying to open their eyes to various basic facts that are commonly overlooked, in short, to lead them to a fuller awareness of simple basics, to encourage asking questions and to look for answers in the fine print that accompanies tabulations of socio-economic data.

It is also the aim of this book to draw attention to the neglected twilight zone, the no-man’s land between the partisan efforts of statisticians who, inspired by applications in the natural sciences, turned to probability, controlled experiments, model building, etc. on one hand, and the applied fields of social sciences on the other. Statistical theorists feel that they are the guardians of a true science, concerned with the purity of its theoretical core with little regard for interpreting economic and social data. On the other side are social scientists and economists concerned about discovering timeless laws of economics, intent on condensing them into mathematical models. They too are less concerned about using statistical data to monitor and also influence events in society. And last, but not least, there are the dedicated statistical foot soldiers who take censuses and surveys, and prepare tabulations. They too have no time for making sense of their data about society.

As you may notice from the ‘Outline,’ this book departs from the usual structure, but instead follows the steps of the statistical process in a rather abstract, theoretical manner, from the very start of conceptualizing the socio-economic phenomenon to be investigated to the final tabulation of the data. Standard topics, like the Gaussian curve, probability theory and symmetrical, well-behaved frequency distributions are treated at the end of the book, if at all. The initial chapters deal at length with topics that are usually missing in textbooks⁵ such as aggregation, statistical aggregates and ratios. They form the backbone for the interpretation of socio-economic data. Then follow three chapters on time series as the most frequent form in which data are published. These chapters are given priority over frequency distributions in one or more dimensions, treated toward the end.

This book, by the way, is not meant as an introduction to statistics, nor as a “how-to-do, hands-on” manual. Its concern is to make sense of socio-economic data. and to shine new light on various misconceptions the reader may have acquired in previous statistics courses. Only a minimum of mathematics will be required. Calculations are relegated to the five optional appendices. Although mathematical statisticians may find this book pedestrian and simplistic, some abstract thinking is involved and the reader is asked to be patient with unfamiliar ideas.

This book is intended for everyone who has to deal with data about society: students and teachers in business, economics and social science courses, economists, social scientists, financial analysts, market researchers, business and economic forecasters, sociologists, managers, demographers, even geographers. It is my hope

that the chapters of this book will open up a new understanding of socio-economic data for their meaningful interpretation, allay bad feelings toward our field, and stimulate further developments in the indicated direction.

Description of Chapters

Chapter 1 provides a short view of the developments that led to the present situation in socio-economic statistics. The powerful influence and the band wagon effect of the developments in statistics in biology, agriculture came to dominate all fields of statistical application. This chapter points out that socio-economic statistical data are quite different from the measurements in the sciences.

Chapter 2 traces the statistical process, from the conception and formulation of a socio-economic phenomenon, such as unemployment, poverty, productivity or crime; to the identification and recording 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 mergers. The simplified records of these 'real-life-objects' then become the 'statistical-counting-units'.

In Chap. 3 the subsequent grouping of these 'statistical-counting-units' into suitable aggregates is discussed. These new entities, the statistical aggregates, are defined by their three 'dimensions': the subject matter, the time period, and the extent of the geographic area covered. As to the subject-matter "dimension", the qualitative characteristics of the statistical counting units are important for the formation of a hierarchy of sub-aggregates. The magnitude of each of the three 'dimensions' of an aggregate determines how to interpret the gains and losses from aggregating the 'statistical-counting-units'. These statistical aggregates represent the bulk of the data in socio-economic statistics. They are quite distinct from the data in the natural sciences, an important matter that has not received due attention.

In Chap. 4 a variety of ratios is discussed as simple and effective analytical tools. These ratios allow us to perceive and make sense of the underlying economic and social reality conveyed by these aggregates. Despite their pervasiveness and importance, ratios have rarely been discussed.

Chapters 5, 6 and 7 study the development, over time, of economic and social phenomena through time-series of socio-economic data.

Chapter 5 presents a critical view of the customary decomposition of time series into trend, seasonal pattern, business cycle and randomness. Instead of the mathematical decomposition into the standard components, time-series should be understood as quantitative economic and social history that can be interpreted meaningfully through a hierarchy of simple ratios between aggregates. These figures are not to be understood as abstract algebraic numbers.

Chapter 6 explores the fact that statistical data lose their relevance over time and become obsolete and less relevant for anticipating the future of a situation in society. Good forecasting requires acquaintance with the historic development of the underlying economic or social forces. Much depends on the speed with which

the data become obsolete. The level of aggregation also affects obsolescence. All this requires judicious decisions regarding the weight older data should be given in a forecasting model, and the point in the past from which on the data of a time series should be disregarded.

Chapter 7 has two parts. In the first part, Sect. 7.1, Price-Index-Numbers are discussed as an important type of time series. A simpler, ratio-based approach is presented⁶ that is more transparent and easier to interpret than the historic Price-Index-Number formulations currently in use, allowing for understanding and interpreting the actual changes in price levels.

In the second part, Sect. 7.2, Index-Numbers of Production are critically reviewed. Different production concepts are discussed and simpler ways of measuring production and productivity are developed.

Chapter 8 deals with the interpretation of highly asymmetric frequency distributions that predominate in economic and social data. Simple measures are presented to deal appropriately with these highly asymmetrical data, to assess and interpret centrality, asymmetry and dispersion.

Chapter 9 discusses the puzzling case of one particular regression analysis that changed my views on cross-sectional data in general. Without going into the algebra of their calculation, specific problems in Regression and Correlation with aggregate data are discussed

Chapter 10 explores the relationship between statistics and the calculus of probability. Although socio-economic statistics is numeric, using mathematical symbols, algebra, geometry and graphs, it must not be considered as a branch of mathematics.⁷ Socio-economic statistical data have an important conceptual non-numeric component that defies a numbers-only approach. One must keep in mind that its purpose is the perception of very real economic and social happenings in historic time, and in geographic and subject-matter space. Misuses of probability, foremost the mis-interpretation and misuse of “Statistical Significance,” are critically reviewed⁸.

Chapters 11 and 12, explore areas that social, business, and economic statistics has in common with subjects that do not readily come to mind as linked with statistics. While exploring these areas in these two final chapters the nature of socio-economic statistics is further clarified.

Chapter 11 has more in common than is usually acknowledged.⁹ When statistics is not considered as a branch of mathematics, however, it is easier to see that macro economics, really National Accounting – which is essentially economic statistics – keeps track of the economy like financial accounting keeps track of a business corporation. The discussion reveals surprising affinities between socio-economic statistics and financial accounting.

Chapter 12 discusses the importance of geographic-spatial distributions, a matter that has been absent from the theory of statistics, though not from statistical field-work. Although specialized quantitative-statistical research abounds in geography, the geographic-spatial dimension has not been recognized as belonging to statistics¹⁰ and ought to be included in its theory.

Acknowledgments I gratefully acknowledges help received from colleagues, Michael Czinkota, International Marketing; Jackie Hoell, Business Statistics; Harvey Iglarsh, Management Science; Denis F. Johnston Sociology, Ursula Kueck, Economic Statistics, Alan Mayer-Sommer, Accounting; Stanley Nollen, Managerial Economics; Kimberly Sellers, Mathematical Statistics; Annette Shelby, Business Communication; Robert Thomas, Marketing; Wilfried Ver Eecke, Philosophy, Daniel Westbrook, Econometrics, Elizabeth Winkler, English, Deepak Sahay, my graduate student assistant. Each has contributed in some way with critical observations, knowledge and encouragement to shape the ideas in this manuscript. And last but not least, Dr. Niels Peter Thomas of Springer Verlag without whose initiative this book would not have been written.

Notes

1. This book is the result of my six decades of teaching all kinds of statistics on graduate and undergraduate levels, as well as the result of my involvement in many interesting statistical studies and surveys. I gradually arrived at the conviction that much of current statistical theory does not apply to economic and social data, a situation that can be imagined as two only partially overlapping circles or ellipses, one representing the general theory of statistics and the other, socio-economic statistics. The small area of intersection where these two geometric figures overlap represents the methods of the general theory of statistics that cover the needs of socio-economic statistics.
2. This is the consequence of the fact that statistics has been identified with mathematics in general and with the calculus of probability in particular. The reaction to mathematics, described in the following address to the German mathematical society, is a familiar story in which the word ‘mathematics’ can readily be substituted with ‘statistics’: “It’s the same old refrain: . . . I hate math . . . Pure torture from the start of school. It’s a total mystery how I ever managed to graduate. . . A nightmare for me. . . Mathematical formulas are. . . pure poison. They just turn me off . . . Complaints such as these are heard all the time . . . educated people express them routinely with a blend of. . . defiance and pride. They assume their listeners sympathy. . .” p. 9 “Drawbridge Up”, translated from “Zugbruecke Ausser Betrieb” Hans Magnus Enzensberger, translated from the German by Tom Astin. Published by A.K. Petus, LTD, Natuck, Ma, 1999.
3. The recent trend toward reducing the required course offerings of statistics in business schools and economic departments in European and also in American universities, seems to be rooted in similar experiences. It appears as a reaction to a sense of frustration with the kind of statistics taught. Though intellectually challenging, it seems to be of insufficient relevance for the social sciences, at least at the undergraduate level, to justify the effort needed to master it. The amount of time dedicated to this kind of statistics is to be allotted to other courses, believed to be more relevant.

Business and economic statistics courses in academic curricula are dominated by mathematical statistics that evolved from the physical-natural sciences. These courses further assume that the problems and data in economics and related fields are like those in the natural sciences, and that all students who take a statistics course want to become professional statisticians. This situation has become entrenched because those, believed to be qualified to teach statistics in economic and business curricula, are expected to have a background in mathematical statistics, taught in mathematics departments. These same academicians also act as reviewers of statistical journals, favoring mathematically oriented manuscripts for publication. That situation has continued for some time, but efforts are beginning to be made to change that.

The recent discussion in statistical journals, e.g. in the 2003 issues of the “Allgemeines Statistisches Archiv”, revealed a concerned questioning by academic statisticians, why statistics, as a core course in economic curricula, is losing ground, and why the required hours of teaching statistics are being curtailed. In short, why the interest in statistics is waning. Peter von der Lippe, Sibylle Schmerbach “Mehr Wirtschaftsstatistik in der Statistikausbildung

fuer Volks- und Betriebswirte" Allgemeines Statistisches Archiv 87, pp. 335–344, 2003 – Hans Peter Litz, Curriculare und fachsystematische Aspekte einer univrsitaeren Wirtschafts- und Sozialstatistik, Allgemeines Statistisches Archiv, 88, pp. 347–361, 2004; – Werner Gruenewald, Hans-Joachim Mittag, Michael Mueller, Reiner Staeglin, Peter Lorscheid und Roland Gnoss, "Diskussionsbeitraege zu "Mehr Wirtschaftsstatistik in der Statistikausbildung fuer Volks- und Betriebswirte", AStA 88, pp. 100–117, 2004- Peter von der Lippe, Sibylle Schmerbach "Antwort zur Discussion um 'Mehr Wirtschaftstatistik in der Statistikausbildung fuer Volks- und Betriebswirte'" AStA 88, pp. 362–367. Evidently colleagues in economics departments and business school have lost patience with this kind of statistics as it has been taught. Yet, the most they feel able to do is to support the reduction of the amount of statistics in the curriculum. In analogy to Lester C. Thurow's highly critical assessment of the (then) current state of economics, the situation of statistics could be named appropriately "Dangerous Currents: the State of Business-, Economic- and Social Statistics." Lester C. Thurow, *Dangerous Currents: The State of Economics*, Random House, New York, 1984 (esp. Chap. 4 "Econometrics: the icebreaker caught in the ice").

4. In the following chapters the term 'data' will be used in the plural when referring to tabulated statistical numbers, (. . .data are. . .) and in the singular when referring to general, not necessarily quantitative, information (. . .data is. . .).
5. The following quotation is applicable to socio-economic statistics, expressing what such a required change in mental outlook entails:

"This experiential horizon is socially and historically, that is, conjecturally conditioned; it shares in the historical character of the whole life of man. With Thomas Kuhn we can register a dual form of progress in our human cogitation: (a) a homogeneous, mostly continuous development within one and the same intellectual model. . . this is an evolutive progress; (b) progress through a fundamental change in the (conjunctural) horizon of experience or the 'intellective model,' whereby the meaning already attained has to be 'translated' anew; this process entails something of a 'revolution' (. . . prior to the 'revolution' there is always a prerevolutionary situation in which for some time past the model had actually ceased to work.) Every sharp change in an intellective and empirical horizon still has its own history! Besides long periods of quiet, homogeneous progress, every so often history exhibits more fundamental jerks: a transition from one historical or conjunctural intellective horizon to another. When a new model has been found (p. 580) it takes time to be accepted by everybody as new evidence. . . for a while old and new culture models will co-exist; the respective champions of the two models often come into conflict; there is even polarization at times: two groups of people, though contemporaries, live in mutually 'alien' worlds, they cease to understand each other. For the result. . . goes deep and reaches wide. As Wittgenstein says: . . . 'What were ducks before the revolution are rabbits afterwards' What in the old model of physics was a solid. . . chair appears in the new atom-model as a kind of empty space with atoms and molecules whirling. . . about inside it. An 'outsider' hearing about this for the first time, will either shake his head in disbelief – or angry protest- over such a new-fangled aberration, since the chair's solidity seems perfectly obvious. . . (p. 581) Schillebeeckx, Edward, *Jesus, an Experiment in Christology*, A Crossroad Book. The Seabury Press, New York, 1979 (A translation of "Jezus").

6. In American textbooks of business- and economic statistics the measurement of price levels usually is presented as a historic peculiarity, separate from the rest of the material that is based on the theory of probability. The straightforward description of the reality of prices is eclipsed by the pseudo-problems of Index number theory. Because socio-economic statistics, in general, has shown little interest in the simple description of reality, no effort was made in price statistics e.g. to clarify, what the things or objects to be reported, the 'statistical counting units', ought to be. See: Winkler, O.W. "Measuring the Price Reality or Measuring a Price Illusion?" *Proceedings of the 11th Tagung fuer Preismessung*, Lutherstadt-Wittenberg, Germany, July 2006.
7. The concept of a 'sample space,' for example, must remain limited to those instances in which data were actually selected as samples. There is a tendency to consider all data as samples,

even those that include the entire population. Similarly, the concept of a ‘super-population’ though useful for sampling theory, is not to be considered for this approach. Another example is the mathematical formulation of a regression surface. It is of interest only within the domain of the actually reported data. The facility with which analysts ‘transform’ and distort their data, is symptomatic for this attitude, e.g. Mosteller, F., Tukey, J. *Data Analysis and Regression*, Addison Wesley Publ. Co, Reading, Mass. 1977. Such manipulations can distort the reality underlying the data, and are to be avoided.

8. The following quotation is but one of the innumerable examples: “The statistical methodology for analyzing data. . . is called statistical inference. . . The logical foundation of statistical inference is the mathematical theory of probability” (p. 75) Neter, John et al. (Boston 1978) Or another example: “The data may be obtained from published sources, through survey research or by designed experiment. However obtained, the data are the observed outcomes or responses of some phenomenon of interest or underlying *random variable*.” (p. 3). (italics added for emphasis). Berenson, et al. (New Jersey, 1983).
9. This chapter follows closely Winkler, Othmar W. “Statistics in Accounting other than Sampling” *Proceedings of the Business and Economic Statistics Section of ASA*, Washington, D.C. 1976, pp. 654–659 as well as a modified version of this article entitled “Secret Allies?” *Management Accounting*, National Association of Accountants, Montvale, NJ, June 1985, pp. 48–53.
10. Such topics are described in publications with titles such as “*Statistical Concepts in Geography*” John Silk, London, Allen & Unwin 1979, esp. pp. 22–26 and 206–249 A list of other pertinent literature in Winkler, O.W. “The Interface of Geography and Economic Statistics” *Proceedings of the Business and Economic Statistics Section, ASA*, Washington, DC, 1981.



<http://www.springer.com/978-3-540-68720-7>

Interpreting Economic and Social Data

A Foundation of Descriptive Statistics

Winkler, O.W.

2009, XVI, 265 p., Hardcover

ISBN: 978-3-540-68720-7