PREFACE

This book could be classified under the heading of "metastatistics." Statistics is the theory of inferences *ideally* drawn from data. Metastatistics is the theory of inferences *actually* drawn from data. Statistical theory takes as given the model, the data, and the purity of the researcher's motives. Metastatistics analyzes how the researcher's motives and opinions influence his choice of model and his choice of data. Metastatistics includes the study of memory and computing failures; it also deals with the social mechanism by which information is transmitted among individuals.

Most of this book deals with a special topic of metastatistics—specification searches. Traditional statistical theory assumes that the statistical model is given. By definition, nonexperimental inference cannot make this assumption, and the usefulness of the traditional theory is rendered doubtful. Moreover, once the model is known, the inferential puzzles that remain are trivial in comparison with the puzzles that arise in the specification of a model. The latter puzzles form the subject of "specimetrics." Specimetrics describes the processes by which a researcher is led to choose one specification of the model rather than another; furthermore, it attempts to identify the inferences that may be properly drawn from a data set when the data-generating mechanism is ambiguous.

My interest in metastatistics stems from my observations of economists at work. The opinion that econometric theory is largely irrelevant is held by an embarrassingly large share of the economics profession. The wide gap between econometric theory and econometric practice might be expected to cause professional tension. In fact, a calm equilibrium permeates our journals and our meetings. We comfortably divide ourselves into a celibate priesthood of statistical theorists, on the one hand, and a legion of inveterate sinner-data analysts, on the other. The priests are empowered to draw up lists of sins and are revered for the special talents they display. Sinners are not expected to avoid sins; they need only confess their errors openly.

In this book I discard the elitism of the statistical priesthood and proceed under the assumption that unavoided sins cannot be sins at all. For several years I have observed colleagues who analyze data. I report herein my understanding of what it is they are doing. The language I use to describe their behavior is the language of the statistical priesthood, and this book may serve partially to bridge the gap between the priests and the sinners. The principal outcome of this effort is an appreciation of why certain sins are unavoidable, even desirable. A new sin does emerge, however. It is a sin not to know why you are sinning. Pointless sin must be avoided.

I began thinking about these problems when I was a graduate student in economics at the University of Michigan, 1966–1970. At that time there was a very active group building an econometric model of the United States. As it happens, the econometric modeling was done in the basement of the building and the econometric theory courses were taught on the top floor (the third). I was perplexed by the fact that the same language was used in both places. Even more amazing was the transmogrification of particular individuals who wantonly sinned in the basement and metamorphosed into the highest of high priests as they ascended to the third floor.

As I struggled against the schizophrenia, I found comforting the Bayesian theory of inference that I was studying in the math department. This truly seemed to be the key toward understanding the difference between the basement and the third floor. Perhaps a complete reconciliation could be achieved.

I am now less optimistic. I am confident that the Bayesian approach helps us understand nonexperimental inference. It helps also to avoid certain errors. But I do not think it can truly solve all the problems. Nor do I foresee developments on the horizon that will make any mathematical theory of inference fully applicable. For better or for worse, real inference will remain a highly complicated, poorly understood phenomenon.

In trying to give a theory of nonexperimental inference, I have accepted an enormous task. The phenomenon of inference is within the purview of statisticians, philosophers, psychologists, and historians. I have found it impossible to master all four fields, though I have tried to understand the central ideas of each. My greatest familiarity is with the theories of statistical inference, but even in that area I am certain that there is much material I have overlooked. My knowledge of philosophy, psychology, and the history of science is spotty, but you will find occasional reference to it. Once I thought I should try to master that literature, but, frankly, I found it inaccessible. Mastering it would have put off the completion of this work for several years. It seemed to be more efficient to say what I have to say, and to be prepared to explore other literature in response to critical comments. Thus I shall welcome the criticism that I am unfamiliar with various materials. Believe me, I have learned much since this project commenced and I do not regard publication as a termination of the process. On the contrary, I feel I had better get it published before I change my mind.

This book is logically divided into three parts. The core of the book consists of Chapters 4 through 9, which analyze six different kinds of specification searches. The first three chapters are introductory, and the last chapter constitutes a footnote describing the inability of anyone actually to behave as described in the first nine chapters.

Chapters 2 and 3 make the book relatively self-contained. It is not necessary that the reader be familiar with Bayesian inference, since a complete (though opinionated) introduction is given in these chapters. Familiarity with classical inference, particularly the normal linear regression model, is assumed, but a review of the standard propositions is provided. Mathematical sophistication is not required to read this book. Matrix algebra is used extensively, but graphs are used whenever possible to clarify the logic. The problems in Appendix 4 may also be helpful.

I can hope, but it seems unlikely, that this book will be used as a text at many universities. I myself have used it as a text for special topics courses both at Harvard and at UCLA. Students in these courses had already been through the usual sequence in graduate econometrics. Many had some first-hand knowledge of econometric practice and were happy to learn that someone else thought econometric theory to be rather remote.

Although it may not be used as a principal text, this book would usefully supplement the fare offered in the standard texts. Some of the material from Chapter 5 on interpretive searches has already filtered into the econometrics courses at the better universities. Chapter 7 on proxy searches could supplement the traditional, rather meager treatment of errors-in-variables problems. Chapter 4 contains two easily read sections

on choosing the significance level of a test. Chapter 8, on data selection, has several sections that ought to be valuable reading—especially the unified treatment of multivariate regression, error-components models, and Kalman filtering. The introduction to Chapter 9 on data-instigated models makes informally the important point that there is a fundamental difference between statistical inference and "Sherlock Holmes" inference. What in this book, I ask myself rhetorically, would not make profitable reading?

Which chapters are my favorites? Clearly, the greatest conceptual contribution is the recognition of the difference between statistical inference and "Sherlock Holmes" inference described in Chapter 9. One may doubt whether the discipline suggested there is practical, however. Is the Bode's law example serious? As for practical advice, turn to Chapter 5 on interpretive searches. It includes what I believe to be a correct statement of the problem of collinearity. I also think the distinction between interpretive searches and simplification searches is important. Chapter 10, on judgmental errors (largely a review and interpretation of psychological literature), was the most amusing to write. Chapter 7 was personally useful to me since I learned in writing it that the metaphysical distinction between endogenous and exogenous variables has little to do with the inferences it is proper to draw from a given body of data.

The impossible task of properly listing my debts is now at hand. As far as I can remember, the first time I did any independent thinking was in writing my undergraduate thesis at Princeton under the direction of John Hartigan. It may be interesting to note that I first read Bayes' *Essay* in his course. It made no particular impression on me at the time. At the University of Michigan the statisticians William Ericson and Bruce Hill taught me Bayesian inference at a perfect time in my life. Were it not for that contact, this book would never have been written. Nor would it have been written had I been unable to observe econometricians at work. At the University of Michigan, Daniel Suits, Saul Hymans, and Harold Shapiro were the (unknowing?) rats in my laboratory. Later, at Harvard, Martin Feldstein, Dale Jorgenson, and Zvi Griliches passed under my microscope. My friend Richard Freeman was a constant source of ideas and encouragement. Also at Harvard, I had the good fortune to spend several years with Gary Chamberlain, who served as a student, a collaborator, and a friend. More recently, Herman Leonard has been a stimulating coworker.

Many of my ideas have been influenced by attendees at the semi-annual Seminars in Bayesian Inference in Econometrics. Arnold Zellner especially should be mentioned. He and Jacques Drezé were the first to carry the Bayesian fasces into the econometrics arena. Another attendee at these seminars (and a Harvard colleague), John Pratt, has had a significant influence on my thoughts. Though they may wish to deny it, I have discovered kindred souls in the form of James Dickey and Thomas Rothenberg.

There is a long list of statisticians whose written contributions have influenced me. That debt is acknowledged in the text with references to their work. Parts of Chapters 2 and 4 were originally prepared jointly with Howard Raiffa for an introductory book that remains unpublished. Earlier versions of this book were read in part and commented on by colleagues at UCLA: Mike Darby, John Riley, Bob Clower, and Jack Carr. Walter Vandaele provided useful detailed comments on several chapters. Lynn Shisido and Tom Means checked the entire manuscript and uncovered numerous errors.

I acknowledge also my debt to the National Science Foundation, Social Science Division, headed by Dr. James Blackman, for grants GS-31929 and SOC 76-08863, without which this book would not have been prepared.

Lastly I would like to bring to the attention of interested parties the existence of *SEARCH*, a computer package designed to implement some of the ideas in this book.

EDWARD E. LEAMER

Los Angeles, California
February 1978

CONTENTS