Introduction to the Mathematical and Statistical Foundations of Econometrics

Herman J. Bierens

Pennsylvania State University

November 13, 2003

Revised: March 15, 2004

Contents

Chapter 7:

Dependent Laws of Large Numbers and Central Limit Theorems

Appendix:

Chapter 8:

Maximum Likelihood Theory

III.5. Complex integration

Appendix IV:

Tables of Critical Values

References

Preface

This book is intended for use in a rigorous introductory Ph.D. level course in econometrics, or in a field course in econometric theory. It is based on lecture notes that I have developed during the period 1997-2003 for the first semester econometrics course "Introduction to Econometrics" in the core of the Ph.D. program in economics at the Pennsylvania State University. Initially these lecture notes were written as a companion to Gallant's (1997) textbook, but have been developed gradually into an alternative textbook. Therefore, the topics that are covered in this book encompass those in Gallant's book, but in much more depth. Moreover, to make the book also suitable for a field course in econometric theory I have included various advanced topics as well. I used to teach this advanced material in the econometrics field at the Free University of Amsterdam and Southern Methodist University, on the basis of the draft of my previous textbook, Bierens (1994).

Some chapters have their own appendixes, containing the more advanced topics and/or difficult proofs. Moreover, there are three appendixes with material that is supposed to be known, but often is not, or not sufficiently. Appendix I contains a comprehensive review of linear algebra, including all the proofs. This appendix is intended for self-study only, but may serve well in a half-semester or one quarter course in linear algebra. Appendix II reviews a variety of mathematical topics and concepts that are used throughout the main text, and Appendix III reviews the basics of complex analysis which is needed to understand and derive the properties of characteristic functions.

At the beginning of the first class I always tell my students: "Never ask me how. Only ask me why." In other words, don't be satisfied with recipes. Of course, this applies to other economics fields as well, in particular if the mission of the Ph.D. program is to place its graduates at research universities. First, modern economics is highly mathematical. Therefore, to be able to make original contributions to economic theory Ph.D. students need to develop a "mathematical mind." Second, students who are going to work in an applied econometrics field like empirical IO or labor need to be able to read the theoretical econometrics literature to keep up-to-date with the latest econometric techniques. Needless to say, students interested in contributing to econometric theory need to become professional mathematicians and statisticians first. Therefore, in this book I focus on teaching "why," by providing proofs, or at least motivations if proofs are too complicated, of the mathematical and statistical results necessary for understanding modern econometric theory.

Probability theory is a branch of measure theory. Therefore, probability theory is introduced, in Chapter 1, in a measure-theoretical way. The same applies to unconditional and conditional expectations in Chapters 2 and 3, which are introduced as integrals with respect to probability measures. These chapters are also beneficial as preparation for the study of economic theory, in particular modern macroeconomic theory. See for example Stokey, Lucas, and Prescott (1989).

It usually takes me three weeks (on a schedule of two lectures of one hour and fifteen minutes per week) to get through Chapter 1, skipping all the appendixes. Chapters 2 and 3 together, without the appendixes, usually take me about three weeks as well.

Chapter 4 deals with transformations of random variables and vectors, and also lists the most important univariate continuous distributions, together with their expectations, variances, moment generating functions (if they exist), and characteristic functions. I usually explain only the change-of-variables formula for (joint) densities, leaving the rest of Chapter 4 for self-tuition.

The multivariate normal distribution is treated in detail in Chapter 5, far beyond the level found in other econometrics textbooks. Statistical inference, i.e., estimation and hypotheses testing, is also introduced in Chapter 5, in the framework of the classical linear regression model. At this point it is assumed that the students have a thorough understanding of linear algebra. This assumption, however, is often more fiction than fact. To tests this hypothesis, and to force the students to refresh their linear algebra, I usually assign all the exercises in Appendix I as homework before starting with Chapter 5. It takes me about three weeks to get through this chapter.

Asymptotic theory for independent random variables and vectors, in particular the weak and strong laws of large numbers and the central limit theorem, is discussed in Chapter 6, together with various related convergence results. Moreover, the results in this chapter are applied to M-estimators, including nonlinear regression estimators, as an introduction to asymptotic inference. However, I have never been able to get beyond Chapter 6 in one semester, even after skipping all the appendixes and Sections 6.4 and 6.9 which deal with asymptotic inference.

Chapter 7 extends the weak law of large numbers and the central limit theorem to stationary time series processes, starting from the Wold (1938) decomposition. In particular, the martingale difference central limit theorem of McLeish (1974) is reviewed together with preliminary results.

Maximum likelihood theory is treated in Chapter 8. This chapter is different from the standard treatment of maximum likelihood theory in that special attention is paid to the problem of how to setup the likelihood function in the case that the distribution of the data is neither absolutely continuous nor discrete. In this chapter only a few references to the results in Chapter 7 are made, in particular in Section 8.4.4. Therefore, Chapter 7 is not prerequisite for Chapter 8, provided that the asymptotic inference parts of Chapter 6 (Sections 6.4 and 6.9) have been covered.

Finally, the helpful comments of five referees on the draft of this book, and the comments of my colleague Joris Pinkse on Chapter 8, are gratefully acknowledged. My students have pointed out many typos in earlier drafts, and their queries have led to substantial improvements of the exposition. Of course, only I am responsible for any remaining errors.