

An Introduction to Modern Bayesian Econometrics

Tony Lancaster

May 26, 2003

Contents

Introduction	xi
1 The Bayesian Algorithm	1
1.1 Econometric Analysis	1
1.2 Statistical Analysis	2
1.3 Bayes' Theorem	3
1.3.1 Parameters and Data	8
1.3.2 The Bayesian Algorithm	9
1.4 The Components of Bayes' Theorem	9
1.4.1 The Likelihood $p(y \theta)$	10
1.4.2 The Prior $p(\theta)$	28
1.4.3 The Posterior $p(\theta y)$	39
1.5 Conclusion and Summary	56
1.6 Exercises and Complements	57
1.7 Appendix to Chapter 1: Some Probability Distributions	63
1.8 Bibliographic Notes	66
2 Prediction and Model Criticism	69
2.1 Methods of Model Checking	70
2.2 Informal Model Checks	71
2.2.1 Residual QQ Plots	72
2.3 Uncheckable Beliefs?	76
2.4 Formal Model Checks	79
2.4.1 Predictive Distributions	79
2.4.2 The Prior Predictive Distribution.	79
2.4.3 Using the Prior Predictive Distribution to Check your Model	81
2.4.4 Improper Prior Predictive Distributions	83
2.4.5 Prediction from Training Samples	84
2.5 Posterior Prediction	87
2.5.1 Posterior Model Checking	89
2.5.2 Sampling the Predictive Distribution	90
2.6 Posterior Odds and Model Choice	97
2.6.1 Two Approximations to Bayes Factors	99

2.7	Enlarging the Model	100
2.8	Summary	105
2.9	Exercises and Further Examples	106
2.10	Bibliographic Notes	110
3	Linear Regression Models	113
3.1	Introduction.	113
3.2	Economists and Regression Models.	113
3.2.1	Mean Independence	115
3.3	Linear Regression Models	115
3.3.1	Independent, Normal, Homoscedastic Errors	117
3.3.2	Vague Prior Beliefs about β and τ	120
3.3.3	The Two Marginals under a Vague Prior	121
3.3.4	Highest Posterior Density Intervals and Regions	128
3.3.5	The Least Squares Line	131
3.3.6	Informative Prior Beliefs	134
3.3.7	Sampling the posterior density of β	135
3.3.8	An Approximate Joint Posterior Distribution	136
3.4	A Multinomial Approach to Linear Regression	143
3.4.1	Comments on the Multinomial Approach	147
3.5	Checking and Extending the Normal Linear Model	150
3.5.1	Checking	150
3.6	Extending the Normal Linear Model:	155
3.6.1	Criticizing the Gasoline Model	155
3.6.2	Generalizing the Error Distribution	161
3.6.3	Model Choice	172
3.7	Conclusion and Summary of the Argument.	174
3.8	Appendix: Analytical Results in the Normal Linear Model	176
3.9	Appendix: Simulating Dirichlet Variates	178
3.10	Appendix: Some Probability Distributions.	179
3.11	Exercises and Complements	182
3.12	Bibliographic Notes	184
4	Bayesian Calculations	187
4.1	Normal Approximations.	189
4.2	Exact Sampling in One Step.	192
4.2.1	Rejection Sampling	192
4.2.2	Inverting the Distribution Function	193
4.3	Markov Chain Monte Carlo	196
4.3.1	MarkovChains and Transition Kernels	196
4.3.2	The State Distribution, $p_t(x)$	198
4.3.3	Stationary Distributions	198
4.3.4	Finding the Stationary Distribution Given a Kernel	199
4.3.5	Finite Discrete Chains	202
4.3.6	More General Chains	204
4.3.7	Convergence	204

4.3.8	Ergodicity	206
4.3.9	Speed	206
4.3.10	Finding Kernels With A Given Stationary Distribution.	207
4.4	Two General Methods of Constructing Kernels:	208
4.4.1	The Gibbs Sampler	209
4.4.2	The Metropolis Method	213
4.4.3	Metropolis-Hastings	215
4.4.4	Practical Convergence	216
4.4.5	Using Samples from the Posterior	217
4.4.6	Calculating the Prior Predictive Density	222
4.4.7	Implementing Markov Chain Monte Carlo	223
4.5	Exercises and Complements	226
4.6	Bibliographic Notes	228
5	Nonlinear Regression Models	231
5.1	Estimation of Production Functions	231
5.1.1	Criticisms of this Model	233
5.2	Binary Choice	235
5.2.1	Probit Likelihoods	236
5.2.2	Criticisms of the Probit Model	238
5.2.3	Other Models for Binary Choice	238
5.3	Ordered Multinomial Choice	240
5.3.1	Data Augmentation	242
5.3.2	Parameters of Interest.	242
5.4	Multinomial Choice	242
5.5	Tobit Models	243
5.5.1	Censored Linear Models	243
5.5.2	Censoring and Truncation:	244
5.5.3	Selection Models	247
5.6	Count Data	249
5.6.1	Unmeasured Heterogeneity in Nonlinear Regression	250
5.6.2	Time Series of Counts	254
5.7	Duration Data	255
5.7.1	Exponential Durations	257
5.7.2	Weibull Durations.	258
5.7.3	Piecewise Constant Hazards	258
5.7.4	Heterogeneous Duration Models	259
5.7.5	Concluding Remarks	259
5.8	Exercises	260
5.9	Appendix to Chapter 5: Some Distributions	261
5.9.1	The Lognormal Family	261
5.9.2	Truncated Normal Distributions	261
5.9.3	The Poisson Family	262
5.9.4	The Heterogeneous Poisson or Negative Binomial Family	263
5.9.5	The Weibull Family	264
5.10	Bibliographic Notes	266

6	Randomized, Controlled and Observational Data	269
6.1	Introduction	269
6.2	Designed Experiments	270
6.2.1	Randomization	270
6.2.2	Controlled Experimentation	271
6.2.3	Randomization and Control in Economics	272
6.2.4	Exogeneity and Endogeneity in Economics	273
6.3	Simpson's Paradox	276
6.4	Conclusions	278
6.5	Appendix: Koopmans' Views on Exogeneity	278
6.6	Bibliographic Notes	279
7	Models for Panel Data	281
7.1	Panel Data	281
7.2	How Do Panels Help?	282
7.3	Linear Models on Panel Data	285
7.3.1	Likelihood	286
7.3.2	A Uniform Prior on the Individual Effects	288
7.3.3	Exact Sampling	292
7.3.4	A Hierarchical Prior	292
7.3.5	Bugs Program	294
7.3.6	Shrinkage	295
7.4	Panel Counts	299
7.4.1	A Uniform Prior on the Individual Effects	300
7.4.2	A Gamma Prior for the Individual Effects.	301
7.4.3	Calculation in the Panel Count Model	302
7.5	Panel Duration Data	302
7.6	Panel Binary Data	304
7.6.1	Parameters of Interest	304
7.6.2	Choices of Prior	305
7.6.3	Orthogonal Reparametrizations	307
7.6.4	Implementation of the model	308
7.7	Concluding Remarks	311
7.8	Exercises	312

313		
7.9	Bibliographic Notes	313
8	Instrumental Variables	315
8.1	Introduction	315
8.2	Randomizers and Instruments	316
8.3	Models and Instrumental Variables	316
8.4	The Structure of a Recursive Equations Model	318
8.4.1	Identification	319
8.5	Inference in a Recursive System	320
8.6	A Numerical Study of Inference with Instrumental Variables	326
8.6.1	Generating data for a simulation study	327
8.6.2	A BUGS model statement	327
8.6.3	Simulation Results	328
8.7	An Application of IV Methods to Wages and Education	332
8.7.1	Is Education Endogenous?	338
8.8	Simultaneous Equations	342
8.8.1	Likelihood Identification	345
8.9	Bibliographic Notes	345
8.10	Instruments via Equilibrium	347
9	Some Time Series Models	349
9.1	First Order Autoregression	349
9.1.1	Likelihoods and Priors	353
9.1.2	BUGS Implementations	355
9.1.3	Prediction	356
9.2	Extensions	358
9.3	Stochastic Volatility	359
9.4	A Second Order Autoregression	360
9.5	Exercises	361
9.6	Bibliographic Notes	361
	Appendix 1: A Conversion Manual	363
.0.1	The Frequentist Approach	363
.0.2	The Bayesian Contrast	364
	Appendix 2: Programming	369
.0.3	S	369
.0.4	WinBUGS	371
.0.5	Formulating the Model and Inputting Data	371
.0.6	Special Likelihoods and the Ones Trick	374
.0.7	Running the Sampler	375
.0.8	Computing References	376

Appendix 3: BUGS Code	379
.09 Heteroscedastic Regression	379
.10 Regression with Autocorrelated Errors	379
.11 CES production Function	380
.12 Probit Model	380
.13 Tobit Model	381
.14 Truncated Normal	381
.15 Ordered Probit	382
.16 Poisson Regression	382
.17 Heterogeneous Poisson Regression	382
.18 Right Censored Exponential Data (using the ones trick)	383
.19 Right Censored Weibull Data	383
.20 A Censored Heterogeneous Weibull Model	384
.21 A Simultaneous Equations Model	384
.22 A Panel Data Linear Model	385
.23 A Second O	
.24 Stochastic Volatility	385

Introduction

This book is an introduction to the Bayesian approach to econometrics. It is written for students and researchers in applied economics. The book has developed out of teaching econometrics at Brown University where the typical member of the class is a graduate student, in his second year or higher. If he is an economics student he has taken in his first year a semester course on probability and random variables followed by a semester dealing with the elements of inference about linear models from a classical point of view. It is desirable that the reader is familiar with the laws of probability, the ideas of scalar and vector random variables and the notions of marginal, joint and conditional probability distributions and the simpler limit theorems. It could, therefore, be studied by upper level undergraduates, particularly in Europe and other countries with European style undergraduate programs. The mathematics used in the book rarely extends beyond introductory calculus and the rudiments of matrix algebra and I have tried to limit even this to situations where mathematical analysis clearly seems to give additional insight into a problem.

Some facility with computer software for doing statistical calculations would be an advantage because the book contains many examples and exercises that ask the reader to simulate data and calculate and plot the probability distributions that are at the heart of Bayesian inference. For simple cases these sums can be done in, for example, Matlab or one of the several variants of the S language. I supply code written in S for many of the examples. More complicated calculations rely on purpose built Bayesian software, specifically a package with the unlikely name of BUGS, and to make full use of this book it is necessary to obtain and learn to use this package.

Whether it useful to have previous knowledge of econometrics is debatable. On the one hand it is helpful to have some understanding of the method of least squares and of regression, and of fundamental econometric notions such endogeneity and structure. On the other hand this book deals exclusively with Bayesian econometrics and this is a radically different approach to our subject than that used in all¹ existing introductory texts. Because Bayesian inference is different from what is customary it is, in my experience, extraordinarily difficult for ordinary mortals to change their way of thinking from the traditional way to the Bayesian way or vice versa. At least it is for me, and I notice that most of

¹1947 as of January 2003

my students face the same problem. This means that someone whose training has been confined to the conventional approach may find this immersion to be a barrier to understanding the Bayesian method.

This book is about the Bayesian approach to inference; it is not a book about comparative methods and it contains little about traditional approaches which are covered in many textbooks. My aim has been to answer two rather simple questions. The first is “What is Bayesian Econometrics?” and the second is “How do I do it?” In the first chapter I explain that Bayesian Econometrics is nothing more than the systematic application of a single theorem, Bayes’ theorem. I also provide a brief answer to the second question, namely that to apply this theorem in an econometric investigation the best method, in general, is to use our new computer power to sample from the probability distributions that the theorem requires us to calculate. This is the meaning of the word “Modern” in the book’s title. In 1989 the methods described here were scarcely known; in 1995 they would have been difficult for a beginner to apply; in 2003 application of these computer intensive methods is little, if any, more difficult than application of the methods traditionally used in applied econometrics. The remainder of the book essentially provides applications of Bayes’ theorem and illustrations of the method of calculation using mostly the simplest models, extensions to more complex structures will in many cases be fairly obvious.

These illustrations are not comprehensive, indeed, for an (imaginary) reader who gets the point of the opening chapters, they are unnecessary! Bayesian analysis of important economic models has been going on since the 1960’s and significant progress has been made with a number of applications. I do not even deal with all those cases in which the method has been applied, but rather confine my examples to cases that I feel comfortable explaining. My hope is that just a few examples will be sufficient to enable the reader to tackle his own problem using what I shall later call The Bayesian Algorithm.

The book could be used as the basis for a one semester course at graduate or advanced undergraduate level. I have used it as such on several occasions with a teaching style that emphasizes calculations; the practicality of Bayesian methods; and demonstrates sampling algorithms including use of markov chain monte carlo procedures in class and requires students to solve problems numerically.

One way to read the book is to get the gist of the Bayesian method from chapters one and two, without necessarily going into the more detailed discussion in these chapters; then to read chapter three to get a broad understanding of markov chain monte carlo methods. The reader could then choose among the remaining chapters, which are illustrations of the use of Bayesian methods in particular areas of application, according to his or her interests.