

themes in modern econometrics

# Statistics and Econometric Models

VOLUME ONE

---

Christian Gourieroux  
and Alain Monfort

Translated by Quang Vuong

F222  
W 1321

# Statistics and Econometric Models

VOLUME 1

*General Concepts, Estimation, Prediction,  
and Algorithms*

---

*Christian Gourieroux*  
*CREST and CEPREMAP, Paris*  
*Alain Monfort*  
*CREST-INSEE, Paris*  
*Translated by Quang Vuong*

江苏工业学院图书馆  
藏书章



**CAMBRIDGE**  
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge  
The Pitt Building, Trumpington Street, Cambridge CB2 1RP  
40 West 20th Street, New York, NY 10011-4211, USA  
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

Originally published in French as  
*Statistique et modèles économétriques Vol 1* by Économica 1989

and © Ed. ÉCONOMICA, 1989

First published in English by Cambridge University Press 1995  
as *Statistics and econometric models Volume 1*

English translation © Cambridge University Press 1995

Printed in Great Britain at the University Press, Cambridge

*A catalogue record of this book is available from the British Library*

*Library of Congress cataloguing in publication data applied for*

ISBN 0 521 40551 3 (V. 1) 0 521 47162 1 (V. 2)  
ISBN 0 521 47744 1 (V. 1 pb) 0 521 47755 X (V. 2 pb)  
ISBN 0 521 47837 5 (two volume paperback set)

ISBN 0 521 40551 3 (hardback)  
ISBN 0 521 47744 1 (paperback)

# Preface

---

The present textbook was motivated by two principles drawn from our teaching experience. First, the application of statistical methods to economic modelling has led to the development of some concepts and tools that are frequently ignored in traditional statistical textbooks. Second, it is possible to introduce these various tools, including the most sophisticated ones, without having recourse to mathematical or probabilistic concepts that are too advanced. Thus the main goal of this book is to introduce these statistical methods by taking into account the specificity of economic modelling and by avoiding a too abstract presentation.

Our first principle has various consequences. Though we discuss problems, concepts, and methods of mathematical statistics in detail, we rely on examples that essentially come from economic situations. Moreover, we analyse the characteristics and contributions of economic methods from a statistical point of view. This leads to some developments in model building but also in concepts and methods. With respect to model building, we address issues such as modelling disequilibrium, agents' expectations, dynamic phenomena, etc. On conceptual grounds, we consider issues such as identification, coherency, exogeneity, causality, simultaneity, latent models, structural forms, reduced forms, residuals, generalized residuals, mixed hypotheses, etc. It is, however, in the field of methods that the economic specificity brings the most important consequences. For instance, this leads to focus on predictions problems, maximum likelihood or pseudo conditional maximum likelihood type estimation methods,  $M$ -estimators, moment type estimators, bayesian and recursive methods, specification tests, nonnested tests, model selection criteria, etc.

Our second principle concerns our presentation. In general, we have tried to avoid proofs that are too technical. Instead, we have attempted to emphasize the intuition behind the results, which is a condition necessary to a real understanding. In particular, we have not introduced the

concept of  $\sigma$ -algebra and we have systematically left aside issues concerning measurability and negligible sets. We have also tried to strengthen the intuitive understanding of the results by multiplying examples. For the reader interested in technical problems, however, a special chapter (Chapter 24, Volume II) collects rigorous proofs of various asymptotic results. This allows us to lower significantly the mathematical level required for the reading of our book. Lastly, we have included two appendices reviewing basic elements of linear algebra and probability theory. These appendices (A and B, found in Volume II) should provide the reader with a self-contained textbook.

Our textbook can be used in various ways according to the topics covered and levels of difficulty desired. Below, we suggest a three-course sequence. For each course we have defined three sections called “Basic Concepts and Tools,” “Estimation and Prediction,” and “Tests and Confidence Regions.”

Course I

Sections	Chapters
Basic Concepts and Tools	1, 2, 3, and Appendices A, B
Estimation and Prediction	5, 6, 7.1–7.3, 11.1–11.2
Tests and Confidence Regions	14, 15, 16, 20.1–20.5

Course II

Sections	Chapters
Basic Concepts and Tools	4,13.1–13.3
Estimation and Prediction	7.4–7.5, 8.1–8.4, 9, 10, 11.3, 12
Tests and Confidence Regions	17, 18, 20.6–20.8

## Course III

Sections	Chapters
Basic Concepts and Tools	13.4–13.6, 22.1, 24
Estimation and Prediction	8.5, 21.1–21.2, 23.1, 23.3
Tests and Confidence Regions	19, 21.3–21.4, 22.2–22.3, 23.2

The first course corresponds to a first-year graduate course in statistical methods for econometrics. The topics covered are indeed the basic ones in statistics. An econometric aspect, however, is provided through the weighting of various topics, the introduction of specific notions, and the choice of examples.

The second course completes the first course by covering asymptotic results. The first two courses should constitute the basic background for a statistician/econometrician wishing to apply the most recent econometric tools.

The third course is the most difficult one on technical grounds. It is also a collection of more advanced topics. A good understanding of the contents of the third level should allow the reading of specialized literature and the beginning of methodological research in good conditions.

We owe a special debt to Martine Germond and Beatrice Lejeune who typed respectively the French and English versions of this book. They both performed promptly and accurately a difficult task. We are also grateful to Quang Vuong for his painstaking translation of our work. Financial support for the publication and translation of this book was provided by the French Ministère de la Recherche et de la Technologie and by the French Ministère de la Culture.

C. Gourieroux  
A. Monfort

# Contents

---

<i>Preface</i>	<i>page</i>	xv
<b>1 Models</b>		<b>1</b>
1.1 Modelling		1
1.1.1 Principles		1
1.1.2 Stochastic Models		3
1.1.3 Formulation of a Model		3
1.1.4 Descriptive Models and Explanatory Models		4
1.1.5 Sampling Designs		5
1.2 Statistical Models		6
1.2.1 Definitions		6
1.2.2 A Priori Information		10
1.2.3 Specification Errors		13
1.3 Intermediate Forms		17
1.3.1 A Principle for Formulating Explanatory Models		17
1.3.2 Examples		18
1.3.3 Coherency Conditions		23
1.4 Conditional Models		24
1.4.1 Conditioning		24
1.4.2 Exogeneity		25
1.5 Dynamic Models		27
1.5.1 Definitions		27
1.5.2 Exogeneity and Causality		28
1.5.3 Formulation of Dynamic Models		33
1.6 Exercises		36
1.7 References		40
<b>2 Statistical Problems and Decision Theory</b>		<b>43</b>
2.1 Examples of Statistical Problems		43
2.2 Decision Rules		44

## Contents

2.2.1	Space of Decisions	44
2.2.2	Nonrandomized Decision Rules	45
2.2.3	Randomized Decision Rules	48
2.3	Ordering Decision Rules	48
2.3.1	Correct Decisions	48
2.3.2	Loss Functions	49
2.3.3	Ordering Nonrandomized Decision Rules	51
2.3.4	Ordering Randomized Decision Rules	53
2.3.5	Admissible Decision Rules	54
2.4	An Example	54
2.4.1	The Model	55
2.4.2	The Estimation Problem	55
2.4.3	Risks Associated with Nonrandomized Decision Rules	55
2.4.4	Risks Associated with Randomized Decision Rules	56
2.5	Other Orderings of Decision Rules	59
2.5.1	Bayesian Approach	59
2.5.2	Empirical Bayesian Approach	61
2.5.3	Conditional Approach	62
2.6	Exercises	62
2.7	References	64
<b>3</b>	<b>Statistical Information: Classical Approach</b>	<b>67</b>
3.1	Sufficiency	68
3.1.1	Definition	68
3.1.2	Characterization of Sufficiency	70
3.1.3	Set of Sufficient Statistics	72
3.1.4	Some Properties of Sufficient Statistics	75
3.1.5	Conditioning	76
3.2	Ancillarity	79
3.2.1	Ancillary Statistics	79
3.2.2	Ancillarity and Sufficiency	80
3.3	Information Measures	81
3.3.1	Fisher Information	81
3.3.2	Kullback Information	87
3.4	Identification	88
3.4.1	Definitions	89
3.4.2	Some Examples	90
3.4.3	Application to Exponential Models	92
3.4.4	Identification and Information	95
3.5	Exercises	96



3.6	References	98
<b>4</b>	<b>Bayesian Interpretations of Sufficiency, Ancillarity, and Identification</b>	<b>101</b>
4.1	Sufficiency	102
4.1.1	Definition	102
4.1.2	Minimal Sufficiency	103
4.1.3	Improving Decision Rules	104
4.1.4	Partial Sufficiency and Bayesian Cut	105
4.2	Ancillarity	106
4.3	Identification	108
4.4	Exercises	109
4.5	References	110
<b>5</b>	<b>Elements of Estimation Theory</b>	<b>113</b>
5.1	Consequences of Decision Theory	113
5.1.1	Improving Nonrandomized Decision Rules	113
5.1.2	Quadratic Risk Functions	114
5.2	Estimation Principles	119
5.2.1	Invariance Principle	119
5.2.2	Unbiasedness Principle	121
5.2.3	Asymptotic Criteria	121
5.3	Search for Good Estimators	122
5.3.1	Search Within a Subclass	123
5.3.2	Criterion Optimization	123
5.3.3	Bayesian Methods	123
5.4	Exercises	124
5.5	References	126
<b>6</b>	<b>Unbiased Estimation</b>	<b>127</b>
6.1	Definitions	127
6.1.1	Unbiased Estimators	127
6.1.2	Existence of Unbiased Estimators	129
6.1.3	Unbiased Estimation and Sufficiency	130
6.1.4	Conditional Models	131
6.2	FDCR Inequality	131
6.3	Best Unbiased Estimators	134
6.3.1	Efficient Estimators	134
6.3.2	Examples of Efficient Estimators	139
6.3.3	Lehmann–Scheffé Theorem	140
6.3.4	Uniqueness of the Best Unbiased Estimator	143
6.4	Best Invariant Unbiased Estimators	144

6.4.1	Gauss–Markov Theorem	145
6.4.2	Best Quadratic Unbiased Estimators	146
6.5	Biased and Unbiased Estimators	150
6.5.1	Nonlinear Functions of an Unbiased Estimator	150
6.5.2	Inadmissible Best Unbiased Estimators	151
6.6	Exercises	154
6.7	References	156
<b>7</b>	<b>Maximum Likelihood Estimation</b>	<b>159</b>
7.1	Principle	159
7.1.1	Definition	159
7.1.2	Three Difficulties	160
7.1.3	An Interpretation of ML Estimation	165
7.2	Likelihood Equations	165
7.2.1	General Remarks	165
7.2.2	Unconstrained Maxima	167
7.2.3	Constrained Maxima	168
7.2.4	Concentrated Likelihood Function	170
7.3	Finite Sample Properties	175
7.3.1	Functional Invariance	175
7.3.2	Relationship with Sufficiency	178
7.3.3	Exponential Models	178
7.4	Asymptotic Properties	180
7.4.1	Existence and Strong Consistency	180
7.4.2	Asymptotic Distribution	183
7.4.3	Concentrated Likelihood Function	186
7.4.4	Direct Derivation of Asymptotic Properties	187
7.4.5	On the Regularity Conditions	189
7.5	marginal and conditional ml estimation	194
7.5.1	Principles	194
7.5.2	Examples	196
7.5.3	Asymptotic Properties	198
7.6	Exercises	203
7.7	References	207
<b>8</b>	<b>M-Estimation</b>	<b>209</b>
8.1	Definition and Asymptotic Properties	209
8.1.1	Definition	209
8.1.2	Existence and Consistency of M-Estimators	210
8.1.3	Asymptotic Normality	212
8.1.4	Quasi Generalized M-Estimators	214

8.2	Nonlinear Regression Models of Order 1 and 2	217
8.2.1	Definitions	217
8.2.2	Special Cases and Examples	218
8.2.3	First-Order and Second-Order Identification	220
8.3	Nonlinear Least Squares	222
8.3.1	The Method	222
8.3.2	Asymptotic Properties	224
8.3.3	Quasi Generalized Nonlinear Least Squares	227
8.3.4	Linearizing a Nonlinear Model	229
8.4	Pseudo Maximum Likelihood Estimation	234
8.4.1	ML Estimation With Specification Errors	234
8.4.2	Consistent PML Estimation of Order 1	238
8.4.3	Quasi Generalized PML Estimation	248
8.4.4	Consistent PML Estimation of Order 2	252
8.5	Estimation of a Conditional Median	255
8.5.1	The Model	255
8.5.2	Least Absolute Deviations Estimation	255
8.5.3	Maximum Score Estimation	264
8.5.4	Consistent Estimators of Parameters Appearing in the Conditional Median	267
8.6	Appendix	270
8.7	Exercises	272
8.8	References	274
<b>9</b>	<b>Methods of Moments and Their Generalizations</b>	<b>277</b>
9.1	Asymptotic Least Squares	279
9.1.1	Parameters of Interest and Auxiliary Parameters	279
9.1.2	The Estimation Principle	279
9.1.3	Best Asymptotic Least Squares Estimators	282
9.2	Examples	284
9.2.1	Relations of the Form $b = h(a)$	284
9.2.2	Estimation Using the Laplace Transform	284
9.2.3	Exponential Families	285
9.2.4	Estimating Equations Linear in the Parameters of Interest	289
9.3	Seemingly Linear Models	291
9.3.1	Consistency of OLS Estimators	291
9.3.2	Autoregressive Models With Autocorrelated Errors	293
9.3.3	Measurement Errors in the Variables	294

9.3.4	Simultaneous Equation Models	297
9.4	Instrumental Variable Estimation	299
9.4.1	Instrumental Variables	299
9.4.2	Instrumental Variable Estimators	301
9.4.3	Asymptotic Distributions of IV Estimators	303
9.4.4	Best Instrumental Variable Estimation	305
9.4.5	Two-Stage Least Squares as Best Asymptotic Least Squares	310
9.5	Generalized Methods of Moments	311
9.5.1	Definition	312
9.5.2	Best GMM Estimation	315
9.5.3	GMM Estimation After Linearization of the Estimating Equations	317
9.5.4	Nonlinear Two-Stage Least Squares	318
9.6	Exercises	320
9.7	References	322
<b>10</b>	<b>Estimation Under Equality Constraints</b>	<b>325</b>
10.1	Constraints	325
10.1.1	Explicit, Implicit, and Mixed Forms	325
10.1.2	Examples	327
10.1.3	Relation Between Explicit and Implicit Forms	329
10.2	Least Squares Under Linear Constraints	331
10.2.1	Explicit Linear Constraints	331
10.2.2	Implicit Linear Constraints	334
10.3	Asymptotic Properties	336
10.3.1	General Framework	337
10.3.2	Taylor Expansions of the First-Order Conditions	342
10.3.3	Asymptotic Distributions	345
10.4	Constrained Two-Step Estimation	350
10.4.1	Methods Based on Unconstrained Estimators	350
10.4.2	Methods Based on the Constraints	352
10.5	Examples	354
10.6	Exercises	356
10.7	References	359
<b>11</b>	<b>Prediction</b>	<b>361</b>
11.1	General Concepts	361
11.1.1	Statement of the Problem	361
11.1.2	Prediction Criteria	362

11.1.3	Simulations	365
11.2	Examples	367
11.2.1	Predictions in the Linear Model	367
11.2.2	Properties of Predictors in the Linear Model	369
11.2.3	Prediction Intervals	373
11.2.4	Prediction of a Nonlinear Function of $Y_{n+1}$	374
11.2.5	Prediction of a Latent Variable	375
11.2.6	Prediction in a Dynamic Model	377
11.3	Residuals	379
11.3.1	Residuals in an Exponential Model	379
11.3.2	Generalized Residuals	381
11.3.3	Simulated Residuals	385
11.4	Appendix	386
11.5	Exercises	388
11.6	References	390
<b>12</b>	<b>Bayesian Estimation</b>	<b>393</b>
12.1	The Bayesian Approach	393
12.1.1	Review	393
12.1.2	Prior Distributions	394
12.1.3	An Example of Bayesian Estimation	397
12.1.4	Diffuse Priors	399
12.2	Conjugate Priors	400
12.2.1	Models with Sufficient Statistics of Dimension Independent of the Number of Observations	400
12.2.2	Conjugate Priors	403
12.2.3	Posterior Distributions From Conjugate Priors	404
12.2.4	Examples	405
12.3	Asymptotic Results	413
12.3.1	Asymptotic Study of Posterior Distributions	413
12.3.2	Consequences	415
12.4	Diffuse Priors	416
12.4.1	Representations of a Noninformative Prior	416
12.4.2	Application to the Linear Model	422
12.4.3	Marginalizing	425
12.5	Best Linear Unbiased Bayesian Estimation	426
12.5.1	Existence and Form of the Optimal Estimator	426
12.5.2	Shrinkage Estimators	428

12.5.3	Bayesian Linear Estimators and Admissible Linear Estimators in the Classical Sense	429
12.6	Approximate Determination of Posterior Distributions	430
12.6.1	Approximating the Prior Distribution	431
12.6.2	Direct Approximations to Posterior Distributions	433
12.6.3	Numerical Integration	435
12.7	Appendix	437
12.8	Exercises	439
12.9	References	441
<b>13</b>	<b>Numerical Procedures</b>	<b>443</b>
13.1	Numerical Optimization	443
13.1.1	Gradient Methods	444
13.1.2	Gradient Methods and ML Estimation	452
13.1.3	Direct Search Methods	458
13.2	Fixed Point Methods	460
13.2.1	Principles	460
13.2.2	An Example: The Fair Algorithm	461
13.3	EM Algorithm	462
13.3.1	Definition	463
13.3.2	Application to Exponential Latent Models	464
13.3.3	Properties	466
13.4	Kalman Filter	467
13.4.1	State Space Models	468
13.4.2	Kalman Covariance Filter	470
13.4.3	Information Filter	474
13.4.4	Kalman Filter and Computation of Likelihood Functions of State Space Models	477
13.5	Prediction and Smoothing in State Space Models	480
13.5.1	Prediction	480
13.5.2	Smoothing	481
13.6	Recursive Least Squares and Recursive Residuals	484
13.7	Exercises	487
13.8	References	490
	Appendix tables	492
	Index	498

# *Models*

## 1.1 Modelling

### 1.1.1 Principles

The problem of modelling arises as soon as one wants to describe and analyze a real phenomenon. Because reality is often complex, the human mind is unable to comprehend it in its entirety. Thus it is necessary to construct a simplification of reality (or *model*) allowing one to study it partially. This simplification cannot take into account all the characteristics of reality, but only those that seem related to the object of study and that are of sufficient importance. A model suited to a particular purpose often becomes inadequate when the object of study changes (even if the study concerns the same phenomenon) or when there is a need for greater accuracy.

**Example 1.1:** To predict the result of an election when there are two candidates  $A$  and  $B$ , one draws with equal probability and replacement a sample of  $n$  voters and observes the number  $Y$  of individuals in this sample intending to vote for  $A$ . The variable  $Y$  can be thought of as a random variable distributed as a binomial  $B(n, p_A)$ , where  $p_A$  is the unknown proportion of individuals voting for  $A$  on the day of the election.

In this example, the model is given by the family

$$\{B(n, p_A), p_A \in [0, 1]\}.$$

Even if this model seems quite natural here, it is only an approximation of reality for it makes several simplifications:

- a) For instance, the sample is drawn with a computer using a program to generate random numbers. Since these programs are based on a deterministic algorithm, one only has an approximation of sampling from a uniform distribution.
- b) The proportion of individuals voting for  $A$  may change between the date of the survey and the date of the election.
- c) Some sampled individuals may not divulge their true voting intentions, etc.

**Example 1.2:** To evaluate the effect on consumption of a change in income, one may propose a model to describe the relationship between these two variables. For instance, one may suppose that they are related through an equality of the form

$$\log C = a \log R + b, \quad a, b \in \mathbb{R}.$$

The parameter  $a$ , called the *consumption elasticity with respect to income*, is equal to the logarithmic derivative  $d \log C / d \log R$ . If income changes by 1 percent, consumption changes approximately by  $a$  percent. The parameter  $a$  provides a natural measure of the effect on consumption of a change in income and the model seems appropriate to the study of this effect. The model is clearly an approximation of reality. Indeed time series data  $(C_t, R_t)$ ,  $t = 1, \dots, T$  on consumption and income will not in general be related by an exact equality such as the one proposed.

**Example 1.3:** A model frequently used to analyze the date of an event is based on the theory of Poisson processes. See Section B.5 in Appendix B, Volume II. The model is more or less suited to the study of unemployment spells (here the event is to find and accept a job). It can be improved and made more realistic by:

- a) no longer assuming independence of the past from the present since the probability of an unemployed person finding a job may depend on the length of the unemployment spell,
- b) introducing various factors affecting an individual's chances of finding a job. These include general factors such as current economic conditions, and individual factors such as the amount of unemployment benefits received by the individual, whether their spouse is employed, etc.



### 1.1.2 Stochastic Models

Some of the previous examples have a stochastic character (see Examples 1.1 and 1.3). On the other hand, the model of consumption behavior (1.2) is purely deterministic, which makes it incompatible with the data. One way to solve this difficulty consists in making the model stochastic. Hence the approximate deterministic relation

$$\log C_t = a \log R_t + b, \quad t = 1, \dots, T,$$

is replaced by

$$\log C_t = a \log R_t + b + u_t, \quad t = 1, \dots, T,$$

where  $u_t, t = 1, \dots, T$ , are random variables with zero means called *disturbances* or *error terms*.

A disturbance measures the discrepancy between the observation  $\log C$  and the approximate mean  $a \log R + b$  proposed by the model. The disturbance may be due to:

- (i) the fact that the relation between  $\log R$  and  $\log C$  is not linear,
- (ii) the fact that the coefficients  $a$  and  $b$  vary over time,
- (iii) the omission (voluntary or not) of secondary variables,
- (iv) measurement errors on the variables  $C$  and  $R$ ...

The disturbance is therefore interpreted as a summary of various kinds of ignorance. In fact, the interpretation of the disturbance is important only when hypotheses are made about its distribution. The form of this distribution depends on the kind of ignorance that the disturbance is supposed to represent.

Apart from its interpretation, the introduction of a disturbance has another purpose. As seen later, in a stochastic model it will be possible to construct measures of the error associated with the use of the model. Indeed an evaluation of the accuracy of the results becomes quite essential.

### 1.1.3 Formulation of a Model

In general, a model has a deterministic component and a random component. How can one obtain an appropriate specification for each of the components? Specifications can rely on observations of the variables or