PALGRAVE TEXTS IN ECONOMETRICS

A Primer for Spatial Econometrics

With Applications in R

Giuseppe Arbia

A Primer for Spatial Econometrics

PalgraveTexts in Econometrics

General Editors: **Kerry Patterson and Terence Mills**
Regional Editors: **Northern Europe: Niels Haldrup; Far East: In Choi; USA: William Greene; Southern Europe and Oceania: Tommaso Proietti**

The pace of developments in econometrics has made it increasingly difficult for students and professionals alike to be aware of what is impotent and likely to be lasting in theory and practical applications. This series addresses important key developments within a unified framework, with individual volumes organised thematically. The series is relevant for both students and professionals who need to keep up with the econometric developments, yet is written for a wide audience with a style that is designed to make econometric concepts available to economists who are not econometricians.

Titles include:
Simon P. Burke and John Hunter
MODELLING NON-STATIONARY TIME SERIES

Michael P. Clements
EVALUATING ECONOMETRIC FORECASTS OF ECONOMIC AND FINANCIAL VARIABLES

Lesley Godfrey
BOOTSTRAP TESTS FOR REGRESSION MODELS

Terence C. Mills
MODELLING TRENDS AND CYCLES IN ECONOMIC TIME SERIES

Kerry Patterson
A PRIMER FOR UNIT ROOT TESTING

Kerry Patterson
UNIT ROOTS TESTS IN TIME SERIES VOLUME 1
Key Concepts and Problems

Kerry Patterson
UNIT ROOTS TESTS IN TIME SERIES VOLUME 2
Extensions and Developments

Terence C. Mills
ANALYSING ECONOMIC DATA
A Concise Introduction

Giuseppe Arbia
A PRIMER FOR SPATIAL ECONOMETRICS
With Applications in R

Palgrave Texts in Econometrics
Series Standing Order ISBN 978–1–403–90172–9 (hardback)
978–1–403–90173–6 (paperback)
(*outside North America only*)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in case of difficulty, write to us at the address below with your name and address, the title of the series and the ISBN quoted above.

Customer Services Department, Macmillan Distribution Ltd, Houndmills, Basingstoke, Hampshire RG21 6XS, England

---

A Primer for Spatial Econometrics

With Applications in R

Giuseppe Arbia

palgrave
macmillan

Softcover reprint of the hardcover 1st edition 2014 978-0-230-36038-9

First published 2014 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC,
175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978-1-137-42816-5 ISBN 978-1-137-31794-0 (eBook)

DOI 10.1057/9781137317940

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Typeset by MPS Limited, Chennai, India.

Transferred to Digital Printing in 2014

To P, E, F & E

Contents

List of Figures

List of Examples

Foreword

I have had the distinct pleasure to enjoy my own study of spatial econometrics in collaboration with Giuseppe Arbia during his work on this monograph. There is much to be learned in this rapidly growing field. This primer introduces the workhorse of the field, the spatially autoregressive and spatially autocorrelated linear regression model. A chapter is devoted to important terms of art in the field. We then progress through extensions of the linear model to heteroscedasticity and panel data treatments. Recent developments of spatial econometrics have extended the models to many non-linear cases, including binary and multinomial choice, stochastic frontiers, sample selection and models for count data and ordered choices. This primer provides a gateway to that literature through a presentation of a spatial binary choice model. Readers will appreciate the extensive presentation of real numerical examples in R, which has emerged as the software of choice for model builders in this area.

Spatial econometrics is enjoying widely dispersed growth spurt in many of the social sciences. Readers of this primer will find it to be a very approachable, informative springboard to their own ongoing study and research, as I have.

William Greene
Stern School of Business
New York, August, 2013

Preface and Acknowledgements

Spatial econometrics is a rapidly expanding topic with applications in so many diverse scientific fields that it is almost impossible to fully enumerate all the disciplines. Indeed, in recent years we can find applications in fields such as regional economics, criminology, public finance, industrial organization, political sciences, psychology, agricultural economics, health economics, demography, epidemiology, managerial economics, urban planning, education, land use, social sciences, economic development, innovation diffusion, environmental studies, history, labor, resources and energy economics, transportation, food security, real estate, marketing, and many others.

Given the widespread interest in the subject, this book aims to meet the growing demand by introducing basic spatial econometrics methodologies to a wide variety of researchers. It is specifically designed as a reference for applied researchers who want to receive a broad overview on the topic. In this sense, it is not intended to be a comprehensive textbook on the subject; rather the intention is to keep the details to a minimum and to provide a practical guide which illustrates the potential of spatial econometric modeling, discusses problems and solutions, and enables the reader to correctly interpret empirical results and to start working with the methods.

There are several features that distinguish this book from other existing texts on the subject. First, this book is self-contained and it does not assume any background knowledge apart from a working knowledge of elementary inferential statistics. Chapter 1 contains a summary of basic standard econometric results that are used throughout this text; as such, it can be omitted by the reader who is already knowledgeable on the subject. The treatment of the various topics is rigorous, but with formal proofs that are low level and reduced to a minimum. The book provides the minimum essential basics and intuitions on each topic and refers to other textbooks and to the literature for a more in-depth discussion. Furthermore, the text is integrated with examples, problem sets and practical exercises. To some extent, one may think of this book as an extended chapter of an econometrics textbook. It thus explicitly aims to bridge the gap between the standard textbooks, which still largely ignore the subject, and the more comprehensive and specialized textbooks. Although the book is application-oriented, I have taken care to provide sound methodological developments and to use notations that are generally accepted for the topics being covered. Therefore, the reader will find that this text provides good preparation for the study of more advanced spatial econometrics material.

Secondly, this book only partially overlaps with existing textbooks on the subject. Even if it does not contain a comprehensive treatment of all the topics in spatial econometrics, it nonetheless includes some of the recent advances that are not discussed in other existing textbooks. For example, various estimation alternatives to the traditional Maximum Likelihood paradigm, the treatment of heteroscedastic innovations, spatial discrete choice models and non-stationary models. In addition, even if the bulk of the book deals with synchronic cross-sectional spatial models, section 4.3 contains an introduction to the treatment of spatial panel data, a fast-growing field in spatial econometrics.

Thirdly, people working in this field quickly learn that most of the procedures illustrated in this book encounter severe computational limitations when applied to very large datasets with sample sizes approaching thousands of observations. Computational issues can, indeed, represent a big limitation for many scholars engaged in spatial analysis in non-specialized fields that do not have access to large computer facilities. They can represent a limitation even with the availability of powerful computing machines in all those cases when a real-time decision has to be taken on the basis of the econometric analysis such as, for example, epidemiological and environmental surveillance or computer-assisted surgery, based on medical imaging. To overcome such limitations, this book also includes a chapter entirely devoted to discussing a series of alternative estimation techniques that can help in dramatically reducing the computational time and computer memory requirements.

Finally, the text introduces the reader to the procedures contained in the free statistical software "R". While spatial econometric methodologies are still not integrated in econometric software products (such as, for example, Eviews, Gauss, Gretl, Limdep, Microfit, Minitab, RATS, SAS, SPSS, TSP and many others), there are presently some packages that deal with most of the topics treated in this book (for example, STATA, Matlab) together with more specialized software in the subject (for example, GeoDa). In this book, however, we have decided to illustrate the various methods by making use of the statistical language R for three main reasons. Firstly, the package is freely available, so that the reader of the book can immediately replicate the techniques using available data. Secondly, the R language is intuitive and requires only a small initial investment. Thirdly, since spatial econometrics is a rapidly expanding topic, the language R guarantees almost real-time updating when new procedures are introduced in the literature.

The material presented can be used as a textbook for an introductory course in spatial econometrics which assumes a working knowledge in econometrics at the level of, for example, the introductory chapters of the 7th edition of the textbook by Greene (see W. Greene, *Econometric Analysis* (2011)) or similar. In particular, chapters 1 to 3 could constitute the material for a two- or three-day course (10–12 lecture hours). Chapters 4 (3–4 hours) and 5 could be supplemental material covered on additional days. For chapter 4, the instructor would have the option of covering the entire chapter which could take an entire day or cover a portion of the chapter which would only take three or four hours. The examples, the questions and the exercises contained in the final part of each chapter may be used in combination with a set of computer laboratory sessions that could accompany the lectures.

The idea of an introductory, easy-to-use, textbook for applied researchers was developed during the many occasions throughout the last decade, in which I had the opportunity to teach introductory courses in spatial econometrics at different universities and institutions in Milan, Barcelona, Fortaleza, Salvador-Bahia and in the summer school, 'Spatial Econometrics Advanced Institute', held yearly in Rome since 2008.

The first plan of the book was written while I was visiting the Economics Department at the Stern School of Business, at New York University during the spring semester of 2011. I am grateful to Bill Greene for inviting me on that occasion and in the following two years in spring when I worked on this project. The final draft was completed during the period I spent as a *professeur invité* at ERMES, Université Pantéon Assas, Paris II for which I am indebted to Alain Pirotte, who was so kind in hosting me and providing me with a suitable environment to finish my work. The rest of the work was done at the Catholic University of the Sacred Heart in Rome, where I moved in late 2011. Section 4.3 is written by Giovanni Millo of Assicurazioni Generali, Trieste, Italy and I am grateful to him for his substantial contribution to the preparation of the book and for his patience in working with me. I am also grateful to Carrie Dolan of the College of William & Mary, Williamsburg, USA for carefully proofreading my draft. I am obviously entirely responsible for all the remaining errors. Carrie was also invaluable in providing and editing some of the maps used in the text and in the examples. Thanks are due also to Miguel Flores of Tecnologico de Monterrey, Mexico for providing the data and the maps on Mexico on which I based some of the examples in chapter 5, and to Diego

Giuliani, Francesca Petrarca and Myriam Tabasso for their rapid, online assistance and advice in respect of some of the R procedures. Giovanni, Carrie, Miguel, Diego, Francesca and Myriam are all former participants of the aforementioned 'Spatial Econometrics Advanced Institute' and I would also wish to express my gratitude to all the students who took part at the school in the last six years (more than 200!) because their active presence in class was certainly the greatest stimulus I received in writing this book.

This book is dedicated to my family. If I look back to the forewords I have written for my previous books, I find expressions of thanks to Paola and to my three children for their presence, patience, help and encouragement. Twenty-five years have now passed since the publication of my first book, and seven since my last one. The children have grown up and many things have changed, but my family is still here with me and to them my thoughts gratefully go on this rainy and gloomy late winter day, when I am here sitting in front of my computer writing the final words of this book.

Rome, Ash Wednesday, 2014

1
The Classical Linear Regression Model

1.1 The basic linear regression model

Let us consider the following linear model

$$ {}_{n}y_{1} = {}_{n}X_{kk}\beta_{1} + {}_{n}\varepsilon_{1} \tag{1.1}$$

where ${}_{n}y_{1} = \begin{bmatrix} y_1 \\ \dots \\ \dots \\ y_n \end{bmatrix}$ is a vector of n observations of the dependent variable y, ${}_{n}X_{k} = \begin{bmatrix} 1 & X_{11} & & X_{1k-1} \\ \dots & & & \dots \\ 1 & & & \\ 1 & X_{n1} & & X_{nk-1} \end{bmatrix}$ a matrix of n observations on $k-1$ non-stochastic exogenous regressors including a constant term, ${}_{k}\beta_{1} = \begin{bmatrix} \beta_1 \\ \\ \\ \beta_k \end{bmatrix}$ a vector of k unknown parameters to be estimated and ${}_{n}\varepsilon_{1} = \begin{bmatrix} \varepsilon_1 \\ \dots \\ \\ \varepsilon_n \end{bmatrix}$ a vector of stochastic disturbances. We will assume throughout the book that the n observations refer to territorial units such as regions or countries.

The classical linear regression model assumes normality, identicity and independence of the stochastic disturbances conditional upon the k regressors. In short

$$\varepsilon \mid X \approx i.i.d.N(0, \sigma_\varepsilon^2 \, {}_nI_n) \tag{1.2}$$

${}_nI_n$ being an n-by-n identity matrix. Equation (1.2) can also be written as:

$$E(\varepsilon \mid X) = 0 \tag{1.3}$$

$$E(\varepsilon\varepsilon^T \mid X) = \sigma_\varepsilon^2 \, {}_nI_n \tag{1.4}$$

Equation (1.3) corresponds to the assumption of *exogeneity*, Equation (1.4) to the assumption of *spherical disturbances* (Greene, 2011).

Furthermore it is assumed that the k regressors are not perfectly dependent on one another (full rank of matrix X). Under this set of hypotheses the Ordinary Least Squares fitting criterion (OLS) leads to the best linear unbiased estimators (*BLUE*) of the vector of parameters β, say $\hat{\beta}_{OLS} = \hat{\beta}$. In fact the OLS criterion requires:

$$S(\beta) = e^T e = \min \tag{1.5}$$

where $e = y - X\hat{\beta}$ are the observed errors and e^T indicates the transpose of e.

From Equation (1.5) we have:

$$\frac{d}{d\beta} S(\beta) = \frac{d}{d\beta} e^T e = \frac{d}{d\beta} (y - X\hat{\beta})^T (y - X\hat{\beta}) = 2\left(X^T X\hat{\beta} - X^T y\right) = 0$$

whence:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y \tag{1.6}$$

As said the OLS estimator is *unbiased*

$$E(\hat{\beta}_{OLS} \mid X) = \beta \tag{1.7}$$

with a variance

$$Var(\hat{\beta}_{OLS} \mid X) = (X^T X)^{-1} \sigma_\varepsilon^2 \tag{1.8}$$

which achieves the minimum among all possible linear estimators (full *efficiency*) and tends to zero when n tends to infinity (*weak consistency*).

From the assumption of normality of the stochastic disturbances, normality of the estimators also follows:

$$\hat{\beta}_{OLS} \,|\, X \approx N[\beta;\ (X^T X)^{-1} \sigma_\varepsilon^2] \tag{1.9}$$

Furthermore, from the assumption of normality of the stochastic disturbances, it also follows that the alternative estimators, based on the Maximum Likelihood criterion (ML), coincide with the OLS solution.

In fact, the single stochastic disturbance is distributed as:

$$f_{\varepsilon_i}(\varepsilon_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\varepsilon_i^2\right]$$

f being a density function, and consequently the likelihood of the observed sample is:

$$\begin{aligned} L(\beta, \sigma_\varepsilon^2) &= \prod_{i=1}^{n} f_{\varepsilon_i}(\varepsilon_i) = \prod_{i=1}^{n} \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_\varepsilon^2}\varepsilon_i^2\right] \\ &= \left(\sigma_\varepsilon^2\right)^{-\frac{n}{2}} (2\pi)^{-\frac{n}{2}} \exp\left[-\frac{\varepsilon^T \varepsilon}{2\sigma_\varepsilon^2}\right] \\ &= \cos t \left(\sigma_\varepsilon^2\right)^{-\frac{n}{2}} \exp\left[-\frac{\varepsilon^T \varepsilon}{2\sigma_\varepsilon^2}\right] \end{aligned} \tag{1.10}$$

from the assumption of independence of the disturbances. From (1.1) we have that

$$\varepsilon = y - X\beta \tag{1.11}$$

hence (1.10) can be written as:

$$L(\beta, \sigma_\varepsilon^2) = \cos t \left(\sigma_\varepsilon^2\right)^{-\frac{n}{2}} \exp\left[-\frac{(y - X\beta)^T (y - X\beta)}{2\sigma_\varepsilon^2}\right] \tag{1.12}$$

and the log-likelihood as:

$$l(\beta, \sigma_\varepsilon^2) = \ln\left[L(\beta, \sigma_\varepsilon^2)\right] = \cos t - \frac{n}{2}\ln\left(\sigma_\varepsilon^2\right) - \frac{(y - X\beta)^T (y - X\beta)}{2\sigma_\varepsilon^2} \tag{1.13}$$

The scores functions are defined as:

$$\begin{cases} s(\beta) = \dfrac{d\ln\left[L(\beta, \sigma_\varepsilon^2)\right]}{d\beta} = -\dfrac{1}{\hat{\sigma}_\varepsilon^2}\left(X^T y - X^T X\hat{\beta}\right) = 0 \\ s(\sigma_\varepsilon^2) = \dfrac{d\ln\left[L(\beta, \sigma_\varepsilon^2)\right]}{d\sigma_\varepsilon^2} = -\dfrac{n}{2\hat{\sigma}_\varepsilon^2} + \dfrac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{2\hat{\sigma}_\varepsilon^4} = 0 \end{cases} \tag{1.14}$$

and solving the system of $k+1$ equations, we have:

$$\begin{aligned} \hat{\beta}_{ML} &= (X^T X)^{-1} X^T y \\ \hat{\sigma}^2_{\varepsilon,ML} &= \frac{e^T e}{n} \end{aligned} \tag{1.15}$$

Thus, under the hypothesis of normality of residuals, the ML estimator of β coincides with the OLS estimator. The ML estimator of σ^2_ε on the contrary differs from the unbiased estimator $s^2_\varepsilon = \dfrac{e^T e}{n-k}$ and it is biased, but asymptotically unbiased.

To ensure that the solution obtained is a maximum we consider the second derivatives:

$$\begin{cases} \dfrac{d^2 l(\beta,\sigma^2_\varepsilon)}{d\beta^2} = \dfrac{ds(\beta)}{d\beta^2} = \dfrac{1}{\hat{\sigma}^2_\varepsilon} X^T X \\ \dfrac{d^2 l(\beta,\sigma^2_\varepsilon)}{d{\sigma^2_\varepsilon}^2} = \dfrac{ds(\sigma^2_\varepsilon)}{d\sigma^2_\varepsilon} = \dfrac{n}{2\hat{\sigma}^4_\varepsilon} \\ \dfrac{d^2 l(\beta,\sigma^2_\varepsilon)}{d\beta\, d\sigma^2_\varepsilon} = 0 \end{cases} \tag{1.16}$$

which can be arranged in the Fisher's Information Matrix:

$$_{k+1}I_{k+1}(\beta,\sigma^2_\varepsilon) = \begin{bmatrix} \dfrac{1}{\hat{\sigma}^2_\varepsilon} X^T X & 0 \\ 0 & \dfrac{n}{2\hat{\sigma}^4_\varepsilon} \end{bmatrix} \tag{1.17}$$

which is positive definite.

The equivalence between the ML and the OLS estimators ensures that the solution found enjoys all the large sample properties of the ML estimators, that is to say: asymptotic normality, consistency, asymptotic unbiasedness, full efficiency with respect to a larger class of estimators other than the linear estimators, and invariance.

The OLS estimators also coincide with the Method of Moments estimators (MM). In fact consider the following moment condition:

$$\frac{1}{n} X^T e = E\left(X^T \varepsilon\right) = 0 \tag{1.18}$$

Solving Equation (1.18) gives:

$$\frac{1}{n}X^T(y - X\beta) = 0$$
$$\frac{1}{n}X^T y - \frac{1}{n}X^T X\beta = 0$$

and solving for β, we have:

$$\hat{\beta}_{MM} = \left(X^T X\right)^{-1} X^T y = \hat{\beta}_{OLS} = \hat{\beta}_{ML} \tag{1.19}$$

As for hypothesis testing, let us first consider the following system of hypotheses related to the single parameter β_i:

$$\begin{aligned} H_0 &: \beta_i = 0 \\ H_1 &: \beta_i \neq 0 \end{aligned} \tag{1.20}$$

where β_i is a generic element of the matrix β such that $\hat{\beta}_i \approx N[\beta_i, S^{ii}\sigma_\varepsilon^2]$, and S^{ii} is the i-th element in the main diagonal of matrix $X^T X$. A statistical test can be derived taking the difference between the value of β under the null and under the alternative hypotheses scaled by its standard deviation:

$$test = \frac{\hat{\beta}_i}{\sigma_\varepsilon \sqrt{S^{ii}}} \overset{H_0}{\approx} N(0,1) \tag{1.21}$$

This, however, is not a pivotal quantity unless we know the value of σ_ε^2. Since $\frac{(n-k)s_\varepsilon^2}{\sigma_\varepsilon^2} \approx \chi^2{}_{(n-k)}$ (with $\chi^2{}_{(n-k)}$ a chi-squared distribution with n–k degrees of freedom) and using the independence between s_ε^2 and $\hat{\beta}$, we can build up the pivotal quantity:

$$test = \frac{\hat{\beta}_i}{s_\varepsilon \sqrt{S^{ii}}} \approx t_{(n-k)} \tag{1.22}$$

which can be used to test the null hypothesis. If we consider, instead, the multiple null hypothesis

$$\begin{aligned} H_0 &: \beta_2 = \beta_3 = \cdots = \beta_k = 0 \\ H_1 &: \beta_i \neq 0 \end{aligned} \tag{1.23}$$

we can test the significance of the model as a whole through the quantity:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \approx F(k-1;n-k) \tag{1.24}$$

which is distributed as an F with $k-1$ and $n-k$ degrees of freedom and where

$$R^2 = \frac{SSR}{SST} \tag{1.25}$$

In Equation (1.25), SSR represents the sum of squares of regression defined as $SSR=1-SSE=1-e^Te$ (SSE representing the sum of squares of errors), $SST=y^Ty - n\bar{y}$ represents the total sum of squares and $\bar{y}$ the sample mean of y. R^2 is the so-called *coefficient of determination* ($0<R^2\leq 1$), a parameter that measures the degree of fit of the observed data to a linear function. The adjusted version of R^2, which takes into account the number of degrees of freedom of the regression, is given by:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k+1}(1-R^2) \tag{1.26}$$

Alternative measures of the degree of fit are the Akaike Information Criterion

$$AIC = \ln\left(\frac{e^Te}{n}\right) + \frac{2k}{n} \tag{1.27}$$

and the Schwartz (or Bayesian) Information Criterion

$$BIC = \ln\left(\frac{e^Te}{n}\right) + \frac{k\ln n}{n} \tag{1.28}$$

A further approach to hypothesis testing in regression (that could be applied to the system of hypotheses (1.20) and that will be employed later in this book) is based on the general testing procedure known as the *likelihood ratio*. The *likelihood ratio* of a parameter vector, say θ, is given by:

$$\lambda = \frac{L(\theta_0)}{L(\hat{\theta}_{ML})} = \frac{L(\theta_0)}{L(\hat{\theta})} \tag{1.29}$$

with the subscript *0* indicating the value of the parameters under the null. It represents the ratio between the likelihood function evaluated at the parameter's value under the null and the likelihood function at its maximum. A monotonic transformation of the test statistic λ does not change the inferential conclusions, so that it is more common to refer to the *likelihood ratio test* as the quantity:

$$LR = -2\ln(\lambda) = -2\left[l(\theta_0) - l(\hat{\theta})\right] \tag{1.30}$$

Expanding $l(\theta_0)$ as a Taylor series about $\hat{\theta}$ we obtain:

$$\begin{aligned} LR &= -2\left[\left(\theta_0 - \hat{\theta}\right)l'\left(\hat{\theta}\right) + \frac{1}{2}\left(\theta_0 - \hat{\theta}\right)^2 l''\left(\tilde{\theta}\right) + rest\right] \\ &= -2\left[\left(\theta_0 - \hat{\theta}\right)s\left(\hat{\theta}\right) + \frac{1}{2}\left(\theta_0 - \hat{\theta}\right)^2 ni\left(\tilde{\theta}\right) + rest\right] \end{aligned} \tag{1.31}$$

where $\tilde{\theta} \in (\hat{\theta}, \theta_0)$, s(.) is the score function and $ni(\tilde{\theta})$ the element of Fisher's Information Matrix. By definition $s(\hat{\theta}) = 0$, so that:

$$LR = n\left(\theta_0 - \hat{\theta}\right)^2 i\left(\theta_0\right) + o_p(1) \tag{1.32}$$

The approximation:

$$LR \approx W = n\left(\theta_0 - \hat{\theta}\right)^2 i\left(\theta_0\right) \tag{1.33}$$

is called the Wald test statistics. A further approximation of LR:

$$LR \approx LM = \frac{l'(\theta_0)^2}{ni\left(\theta_0\right)} \tag{1.34}$$

is called the "*Rao's score test statistics*" in statistics, but is better known in econometrics as the "*Lagrange multiplier test*". The three test statistics *LR*, *W* and *LM* are asymptotically equivalent and asymptotically distributed as a χ^2 with the number of degrees of freedom equal to the number of parameters to be estimated. With respect to the other two tests, the *LM* test has the advantage that it can be computed without previously obtaining the Maximum Likelihood estimation of the unknown parameters and that it does not require the specification of any explicit alternative hypothesis.

For a vector of parameters Equation (1.34) becomes:

$$LM = s(\theta_0)^T I(\theta)^{-1} s(\theta_0) \tag{1.35}$$

which is the general expression of the Lagrange Multiplier test that will be used later in this book. In the case of the linear regression, a simple way of testing hypotheses on the parameters is to define:

$${}_1s_{k+1}(\theta_0) = \left[{}_1s\left(\beta_0\right)_k, s\left(\sigma^2_{0,\varepsilon}\right)\right] \tag{1.36}$$

and

$${}_{k+1}I_{k+1}(\theta_0) = {}_{k+1}I_{k+1}\left(\beta_0, \sigma^2_{0,\varepsilon}\right) \tag{1.37}$$

and substituting (1.14) and (1.17) into (1.36) and (1.37) and both into (1.35), we obtain the *LM* test for hypotheses on the regression parameters.

Finally, a crucial hypothesis to be tested on the model is the hypothesis of the normality of the residuals on which all the previous testing strategies are grounded. A popular parametric procedure was introduced by Jarque and Bera (1987) who suggested building up a test of normality by testing jointly that the third and the fourth moments of the empirical distribution of residuals are not significantly different from those of the Gaussian distribution. The formal expression of the test is the following:

$$JB = \frac{n}{6}\left[SK^2 + \frac{(K-3)^2}{4}\right]$$

with $SK = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(e_i - \overline{e})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(e_i - \overline{e})^2\right]^{3/2}}$ the skewness and, respectively,

$K = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(e_i - \overline{e})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(e_i - \overline{e})^2\right]^2}$ the kurtosis of the residuals. Under the null of normality, this quantity can be shown to be distributed as a χ^2 with 2 degrees of freedom.

Example 1.1 Barro and Sala-i-Martin model of regional convergence

The Barro and Sala-i-Martin (1995) model of regional convergence expresses the growth rate of per capita GDP in one region in a certain moment of time (expressed as the logarithm of the ratio) as a linear function of the per capita GDP at the beginning of the period. If the slope in this linear model is negative, then those regions that are poorer at the beginning of the period experience higher growth rates and, conversely, those regions with the higher per capita GDP at the beginning of the period experience lower growth rates. This indicates convergence of the regions towards the same level of per capita GDP. We can express the model as:

$$\log \frac{y_{it}}{y_{i0}} = \alpha + \beta y_{i0} + \varepsilon_i \qquad t = 1, 2, \ldots , T$$

y_{it} = *being the per capita GDP in year t and region i. The parameter* $b = -\frac{\ln(1+\beta)}{T}$ *represents the so-called "speed of convergence". The following table shows the per capita GDP in year 2000 and the growth of the real per capita GDP in the period 2000–08 as it was observed in the 20 Italian regions.*

Region	Per capita GDP	Growth of GDP (2000–2008)	Region	Per capita GDP	Growth of GDP (2000–2008)
1. Piedmont	130	2.7	11. Marche	125	3.1
2. Aosta Valley	150	2.5	12. Latium	130	2.9
3. Lombardy	140	2.7	13. Abruzzo	100	4.0
4. Trentino-Alto Adige.	170	0.5	14. Molise	90	3.5
			15. Campania	110	2.1
5. Veneto	160	1.5	16. Puglia	95	3.0
6. Friuli Venezia Giulia	160	0.5	17. Basilicata	80	4.2
			18. Calabria	100	3.0
7. Liguria	135	2.0	19. Sicily	100	2.0
8. Emilia Romagna	145	1.6	20. Sardinia	110	2.4
9. Tuscany	135	2.2			
10. Umbria	130	3.2			

Source: http://sitis.istat.it/sitis/html/.

The estimation of the model using the OLS method leads to the following results

Parameter		Standard Error	t-test	p-value
Intercept	6.161369	0.731837	8.419	1.18e−07***
Slope	−0.029510	0.005752	−5.130	7.01e−05***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.*