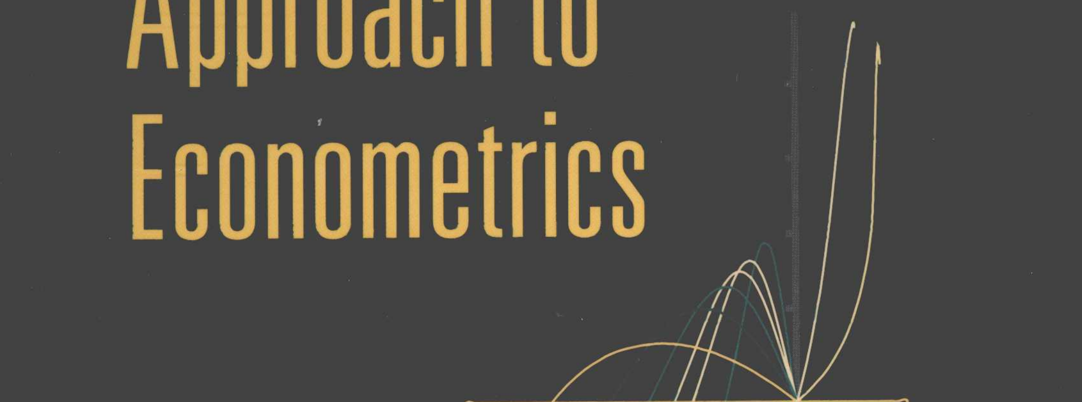
An Information Theoretic Approach to Econometrics



GEORGE G. JUDGE

RON C. MITTELHAMMER

MBRIDGE

An Information Theoretic Approach to Econometrics

GEORGE G. JUDGE
University of California, Berkeley

RON C. MITTELHAMMER
Washington State University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521689731

First published 2012

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication Data

Judge, George G.
An information theoretic approach to econometrics / George G. Judge, Ron C. Mittelhammer.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-86959-1 (hardback) – ISBN 978-0-521-68973-1 (paperback)
1. Econometrics. I. Mittelhammer, Ron. II. Title.
HB139.J795 2011
330.01′5195–dc23 2011018358

ISBN 978-0-521-86959-1 Hardback
ISBN 978-0-521-68973-1 Paperback

Preface

In one sense, the idea for this book started 15 years ago with a six-hour meeting of the two authors between planes at the O'Hare Hilton. At this meeting, maximum entropy and empirical likelihood principles were the major areas of discussion. Since that time the two of us have worked together and have only looked forward. As a result, *Econometric Foundations* appeared in 2000, and a range of related information theoretic articles emerged in the last decade. Pieces of some of these articles appear in this book.

This book was a pleasure to write. We hope the reader will feel our enthusiasm in entering the information theoretic world and leaving behind many conventional econometric methods that we spent a good part of our lives learning.

To write this book, we had the help of many colleagues. Several years of work with Amos Golan provided a base for dealing with pure and noisy inverse problems and the maximum entropy principle. Douglas Miller was involved in this early work and also worked with us on the *Econometric Foundations* book. Marian Grendar, a longtime colleague of one of the authors, worked with us concerning the theoretical underpinning of information theoretic methods and read and commented on many of the chapters in this book. Art Owen was always available to discuss issues relating to the empirical likelihood approach to estimation and inference. Wendy Tam Cho was a creative partner in solving a range of important pure inverse applied problems. To a large group of colleagues, too numerous to mention, we express our warm thanks and appreciation and hope they have been appropriately acknowledged in references throughout the book.

For a book to reach the publication stage, two persons are necessary in addition to the author. One involves a person that takes incoherent words and other symbols and converts them to a working copy. In this context, Danielle Engelhardt, with intelligence, good humor, and word-processing skills, worked with us to get each chapter just right and to turn it into a beautiful copy. A second person is an editor who understands the subject matter and shares your goals. In Scott Parris, we found such a person and a full partner every step of the way. For the two authors, this has been a joint venture; the order of the names has only alphabetical significance.

George G. Judge
Ron C. Mittelhammer

Contents

ONE

Econometric Information Recovery

1.1 Book Objectives and Problem Format

The objectives of this book are to

i) develop a plausible basis for reasoning in situations involving incomplete-partial econometric model information,
ii) develop principles and procedures for learning or recovering information from a sample of indirect noisy data, and
iii) provide the reader with a firm conceptual and empirical understanding of basic information theoretic econometrics models and methods.

What makes the econometric information recovery process interesting is that

- economic-behavioral systems, such as physical and biological systems, are statistical in nature;
- the conceptual econometric model contains parameters and noise components that are unknown and unobserved and, indeed, not subject to direct observation or measurement;
- the recovery of information on the unknown parameters or components requires, for analysis purposes, the use of indirect noisy measurements based on observable data and the solution of an inverse problem that maps the indirect noisy observations, into information on the unknown model and its unobservable components;
- the models may be ill-posed or, in the context of traditional procedures, may be undetermined and the solution not amenable to conventional rules of logic or to being written in closed form.

These problems, taken either individually or in some combination, represent the intellectual challenge of modern econometric analysis and research.

Building on the productive efforts of our precursors in the areas of theoretical economics and inferential statistics, we hope, in this book, to provide an operational understanding of a rich set of information theoretic methods that may be used in theoretical and applied econometrics.

Econometrics is a work in progress. Anyone who doubts this should review a sampling of econometric books starting in the mid-1930s and map the development of econometrics over time. Advances in econometric methodology have been substantial in both content and number, and they continue at a geometric rate.

Information theoretic methods, which have a base in statistical mechanics in physics, have developed in econometrics over the last two decades. In this book we provide a conceptual and empirical understanding of information theoretic methods in some of the major areas of econometrics. Because in econometrics, and in other subject-matter areas, we must work with indirect noisy observations and ill-posed econometric models, traditional econometric methods may not be applicable in answering many of the quantitative questions we wish to ask.

To be a bit more specific, as noted previously, in econometric analyses the unknown and uncontrolled components of the econometric model cannot generally be observed directly. Thus, the analyst must use indirect noisy observations based on observable data to recover information relative to these unknown and unobservable components. This situation is associated with a concept in systems and information theory called the *inverse problem*, which is the problem of recovering information about unknown and uncontrolled components of a model from indirect noisy observations. The adjective *indirect* refers to the fact that although the observed data are considered to be directly influenced by the values of model components, the observations are not themselves the direct values of these components but only indirectly reflect the influence of the components. Thus, the relationship characterizing the effect of unobservable components on the observed data must be somehow inverted to recover information about the unobservable model components from the data-observations. Because econometric relations generally contain a systematic and a noise component, the problem of recovering information about unknowns and unobservables $(\boldsymbol{\theta}, \boldsymbol{\varepsilon})$ from sample observations $(\mathbf{y}, \mathbf{x})$ within the context of an econometric model $\mathbf{Y} = \eta(\mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ is referred to as an *inverse problem with noise* or as a *stochastic inverse problem*. A solution to this inverse problem is of the general form $(\mathbf{y}, \mathbf{x}) \Rightarrow (\boldsymbol{\theta}, \boldsymbol{\varepsilon})$. Because most econometric analyses are of this form, it would seem natural that solution methods should be used that are consistent with the underlying information recovery problem.

1.2 Organization of the Book

To establish notation and connect with reader knowledge, the book starts with the specification and analysis of the simplest parametric and semiparametric probability models. The book is organized in five parts.

Part I is concerned with estimation and inference procedures for parametric and semiparametric linear probability models. In Chapter 2, a notational basis for the econometric models ahead is specified. Estimation and inference is considered for both parametric and semiparametric models and the idea of extremum estimation is introduced. In Chapter 3, estimation and inference methods are introduced for obtaining information on parameters that are functionally related to moments of the data sampling distribution. The focus is on just and overdetermined cases and the chapter starts with an examination of both the method of moments and the generalized method of moments. Finally, the concept of estimating equations, which subsumes the moment equation approaches, is introduced. With an eye to the chapters ahead, we note how extremum estimation relates to the aforementioned methods.

In Part II, we leave the traditional econometric world and focus on econometric models in the form of ill-posed stochastic inverse problems, where information recovery is based on indirect noisy observations and asymptotic considerations come into play. In Chapter 4, a nonparametric stochastic linear inverse problem (econometric model) is defined and a solution (estimation method) is proposed. In Chapter 5, we consider the problem of inference as it is related to the evolving stochastic empirical likelihood inverse problem. In Chapter 6, we introduce the Kullback-Leibler information criterion and the Shannon-Jaynes maximum entropy principle, and provide an estimation and inference basis for the evolving maximum empirical exponential likelihood estimator.

In Part III, we introduce the Cressie-Read family goodness of fit-power divergence measures and provide, in Chapter 7, a framework for estimation and inference. In Chapter 8, sampling experiments are used to illustrate the finite sampling properties of this family of estimators. Recognizing there may be uncertainty regarding the specification of the estimating equations, a choice rule under quadratic loss is proposed.

In Part IV, we consider an information theoretic approach to the binary-discrete choice stochastic inverse problem. In Chapter 9, a minimum power divergence (MPD) family of distribution functions for the binary response discrete choice model is proposed, an estimation framework is specified, and the corresponding statistical properties are identified. In Chapter 10, an estimation and inference basis for the binary MPD family is demonstrated and sampling experiments are used to illustrate finite sample results.

In Part V, we recognize that a basic limitation of traditional likelihood-divergence approaches is the impossibility of describing or identifying estimators-distributions of an arbitrary form. In Chapter 11, we address this estimation and inference problem by suggesting a loss-function-based way of choosing an optimum member from the Cressie-Read family.

Finally, in Chapter 12, we look back over the preceding chapters with a critical eye and make predictions about the econometric information theoretic road ahead.

1.3 Selected References

Cavalier, L., 2008. "Nonparametric Statistic Inverse Problems," *Inverse Problems*, **24**:1–19.

Mittelhammer, R., M. Judge, and D. Miller, 2000. *Econometric Foundations*, Cambridge University Press, New York.

PART I

TRADITIONAL PARAMETRIC AND SEMIPARAMETRIC ECONOMETRIC MODELS: ESTIMATION AND INFERENCE

In Part I, we use a familiar data sampling process to focus on parametric and semiparametric econometric models. This grouping nicely reflects the information, real or imagined, that the analyst uses in terms of the economic and data sampling process that is being modeled. In contrast to fully defined parametric models, semiparametric models cannot be fully defined in terms of the values of a finite number of parameters. In particular, there is no assertion made that a particular parametric family of probability distribution is known that fully defines the probability distribution underlying the data sampling process. Building on this base, in Part I we move in the direction of extremum formulations for analyzing models of this type. For a more complete discussion of the material and relevant proofs in Chapters 2 and 3, see Mittelhammer, Judge, and Miller (2000) and Mittelhammer (1996).