FINITARY PROBABILISTIC METHODS IN ECONOPHYSICS

Econophysics applies the methodology of physics to the study of economics. However, whilst physicists have a good understanding of statistical physics, they are often unfamiliar with recent advances in statistics, including Bayesian and predictive methods. Equally, economists with knowledge of probabilities do not have a background in statistical physics and agent-based models. Proposing a unified view for a dynamic probabilistic approach, this book is useful for advanced undergraduate and graduate students as well as researchers in physics, economics and finance.

The book takes a finitary approach to the subject. It discusses the essentials of applied probability, and covers finite Markov chain theory and its applications to real systems. Each chapter ends with a summary, suggestions for further reading and exercises with solutions at the end of the book.

UBALDO GARIBALDI is First Researcher at the IMEM-CNR, Italy, where he researches the foundations of probability, statistics and statistical mechanics, and the application of finite Markov chains to complex systems.

ENRICO SCALAS is Assistant Professor of Physics at the University of Eastern Piedmont, Italy. His research interests are anomalous diffusion and its applications to complex systems, the foundations of statistical mechanics and agent-based simulations in physics, finance and economics.

FINITARY PROBABILISTIC METHODS IN ECONOPHYSICS

UBALDO GARIBALDI
IMEM-CNR, Italy

ENRICO SCALAS
University of Eastern Piedmont, Italy

CAMBRIDGE
UNIVERSITY PRESS

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

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521515597

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-51559-7 Hardback

Cambridge University Press
978-0-521-51559-7 - Finitary Probabilistic Methods in Econophysics
Ubaldo Garibaldi and Enrico Scalas
Frontmatter
More information

Contents

Cambridge University Press
978-0-521-51559-7 - Finitary Probabilistic Methods in Econophysics
Ubaldo Garibaldi and Enrico Scalas
Frontmatter
More information

Foreword

What is this book about?

The theme of this book is the allocation of n objects (or elements) into g categories (or classes), discussed from several viewpoints. This approach can be traced back to the early work of 24-year-old Ludwig Boltzmann in his first attempt to derive Maxwell's distribution of velocity for a perfect gas in probabilistic terms.

Chapter 2 explains how to describe the state of affairs in which for every object listed 'alphabetically' or in a sampling order, its category is given. We can consider the descriptions of Chapter 2 as facts (taking place or not), and events as propositions (true or not) about facts (taking place or not). Not everything in the world is known, and what remains is a set of possibilities. For this reason, in Chapter 3, we show how events can be probabilized and we present the basic probability axioms and their consequences. In Chapter 4, the results of the previous two chapters are rephrased in the powerful language of random variables and stochastic processes.

Even if the problem of allocating n objects into g categories may seem trivial, it turns out that many important problems in statistical physics and some problems in economics and finance can be formulated and solved using the methods described in Chapters 2, 3 and 4. Indeed, the allocation problem is far from trivial. In fact, in the language of the logical approach to probability, traced back to Johnson and, mainly, to Carnap, the individual descriptions and the statistical descriptions are an essential tool to represent possible worlds. A neglected paper written by Brillouin showed that the celebrated Bose–Einstein, Fermi–Dirac and Maxwell–Boltzmann distributions of statistical physics are nothing other than particular cases of the one-parameter multivariate accommodation process, where bosons share the same predictive probability as in Laplace's succession rule. These are particular instances of the generalized multivariate Pólya distribution studied in Chapter 5 as a sampling distribution of the Pólya process. The Pólya process is an n-exchangeable stochastic process, whose characterization leads us to discuss Pearson's fundamental problem of practical statistics, that is, given that a result of an experiment has been observed

r times out of m trials, what is the probability of observing the result in the next (the $(m+1)$th) trial?

Up to Chapter 5, we study the allocation of n objects into g categories either as an accommodation process or as a sampling process. The index set of the finite individual stochastic processes studied in Chapters 4 and 5 may represent either the name of the objects or their order of appearance in a sampling procedure. In other words, objects are classified once for all. If time enters and change becomes possible, the simplest probabilistic description of n objects jumping within g categories as time goes by is given in terms of finite Markov chains, which are an extension in several directions of the famous Ehrenfest urn model. This is the subject of Chapter 6. For the class of irreducible Markov chains, those chains in which each state can be reached from any other state, there exists a unique invariant distribution, that is a probability distribution over the state space which is not affected by the time evolution. If the Markov chain is also aperiodic, meaning that the initial conditions are forgotten forever during time evolution, then the invariant distribution is also the equilibrium distribution. This means that, irrespective of the initial conditions, in the long run, the chain is described by its invariant distribution, which summarizes all the relevant statistical properties. Moreover, the ergodic problem of statistical physics finds a complete solution in the theory of finite Markov chains. For finite, aperiodic and irreducible Markov chains, time averages converge to ensemble averages computed according to the invariant (and equilibrium) distribution. Incidentally, in this framework, many out-of-equilibrium properties can be derived. The Ehrenfest–Brillouin Markov chain of Chapter 7 encompasses many useful models of physics, finance and economics. This is shown not only in Chapter 7, but also in Chapter 8, where a detailed review of some recent results of ours is presented. These models describe a change of category as a sequence of an Ehrenfest destruction (where an element is selected at random and removed from the system) followed by a Brillouin creation (where an element re-accommodates in the system according to a Pólya weight).

The Pólya distribution is the invariant and equilibrium distribution for the Ehrenfest–Brillouin Markov chain. In Chapter 9, we study what happens to the Pólya distribution and to the Ehrenfest–Brillouin process when the number of categories is not known in advance, or, what is equivalent, it is infinite. This naturally leads to a discussion of models of herding (an element joins an existing cluster) and innovation (an element decides to go alone and create a new cluster). If the probability of innovation still depends on the number of elements present in the system, the herding–innovation mechanism leads to the Ewens distribution for the size of clusters. On the contrary, in Chapter 10, the more difficult case is considered in which the innovation probability is a constant, independent of the number of elements in the system. The latter case leads to the Yule distribution of sizes displaying a power-law behaviour.

One can always move from the probabilistic description of individuals to the probabilistic description of occupation vectors to the probabilistic description of the frequency of occupation vectors. This is particularly simple if the descriptions have suitable symmetries. In Chapter 5, we show how to do that, for individual descriptions whose sampling distribution is the Pólya distribution. Two methods can be used, and both of them can be traced back to Boltzmann's work: the exact marginalization method and the approximate method of Lagrange multipliers leading to the most probable occupation vector. These methods are then used in the subsequent chapters as they give quantities that can be compared either with the results of Monte–Carlo simulations or with empirical data.

Throughout the book, we have worked with a *finitary* approach and we have tried to avoid the use of the infinite additivity axiom of probability theory. Of course, this is not always possible and the reader is warned that not all our derivations are fully rigorous even if we hope they are essentially correct. In principle, this book could be read with profit by anyone who masters elementary calculus. Integrals and derivatives are seldom used, but limits do appear very often and always at the end of calculations on finite systems. This procedure follows Boltzmann's conviction that the continuum is just a useful device to describe a world which is always finite and discrete. This conviction is shared by E.T. Jaynes, according to whom, if one wishes to use continuous quantities, one must provide the explicit limiting procedure from the finite assumptions.

A final *caveat* is necessary. This is not a textbook on Monte Carlo simulations. However, we do often use them and they are also included in solved exercises. The reader already acquainted with Monte Carlo simulations should not find any major difficulty in using the simulation programs written in *R* and listed in the book, and even in improving them or spotting bugs and mistakes! For absolute beginners, our advice is to try to run the code we have given in the book, while, at the same time, trying to understand it. Moreover, this book is mainly concerned with theoretical developments and it is not a textbook on statistics or econometrics, even if, occasionally, the reader will find estimates of relevant quantities based on empirical data. Much work and another volume would be necessary in order to empirically corroborate or reject some of the models discussed below.

Cambridge legacy

We are particularly proud to publish this book with Cambridge University Press, not only because we have been working with one of the best scientific publishers, but also because Cambridge University has an important role in the flow of ideas that led to the concepts dealt with in this book.

The interested reader will find more details in Chapter 1, and in the various notes and remarks around Chapter 5 and Section 5.6 containing a proof, in the spirit of Costantini's original one, of what, perhaps too pompously, we called the fundamental representation theorem for the Pólya process. As far as we know, both W.E. Johnson and J.M. Keynes were fellows at King's College in Cambridge and both of them worked on problems strictly related to the subject of this book.

Johnson was among Keynes' teachers and introduced the so-called postulate of *sufficientness*, according to which the answer to the fundamental problem of practical statistics, that is, given r successes over m trials, the probability of observing a success in the next experiment depends only on the number of previous successes r and the size of the sample m.

How to use this book?

This book contains a collection of unpublished and published material including lecture notes and scholarly papers. We have tried to arrange it in a textbook form. The book is intended for advanced undergraduate students or for graduate students in quantitative natural and social sciences; in particular, we had in mind physics students as well as economics students. Finally, applied mathematicians working in the fields of probability and statistics could perhaps find something useful and could easily spot mathematical mistakes. Incidentally, we deeply appreciate feedback from readers. Ubaldo Garibaldi's email address is `garibaldi@fisica.unige.it` and Enrico Scalas can be reached at `enrico.scalas@mfn.unipmn.it`.

Enrico Scalas trialled the material in Chapter 1, in Section 3.6, and parts of Chapter 4 for an introductory short course (20 hours) on probability for economists held for the first year of the International Doctoral Programme in Economics at Sant'Anna School of Advanced Studies, Pisa, Italy in fall 2008. The feedback was not negative. Section 3.6 contains the standard material used by Enrico for the probability part of a short (24 hours) introduction to probability and statistics for first-year material science students of the University of Eastern Piedmont in Novara, Italy. Finally, most of Chapters 4, 5 and 8 were used for a three-month course on stochastic processes in physics and economics taught by Enrico to advanced undergraduate physics students of the University of Eastern Piedmont in Alessandria, Italy. This means that the material in this book can be adapted to several different courses.

However, a typical one-semester econophysics course for advanced undergraduate students of physics who have already been exposed to elementary probability theory could include Chapters 2, 4, 5, 6, 7 and selected topics from the last chapters.

Cambridge University Press
978-0-521-51559-7 - Finitary Probabilistic Methods in Econophysics
Ubaldo Garibaldi and Enrico Scalas
Frontmatter
More information

Acknowledgements

Let us begin with a disclaimer. The people we are mentioning below are neither responsible for any error in this book, nor do they necessarily endorse the points of view of the authors.

Domenico Costantini, with whom Ubaldo has been honoured to work closely for thirty years, is responsible for the introduction into Italy of the Cambridge legacy, mainly mediated by the work of Carnap. His deep interest in a probabilistic view of the world led him to extend this approach to statistical physics. Paolo Viarengo has taken part in this endeavour for the last ten years. We met them many times during the period in which this book was written, we discussed many topics of common interest with them, and most of the already published material contained in this book was joint work with them! Both Ubaldo and Enrico like to recall the periods in which they studied under the supervision of A.C. Levi many years ago. Finally, we thank Lindsay Barnes and Joanna Endell-Cooper at Cambridge University Press as well as Richard Smith, who carefully edited our text, for their excellent work on the manuscript.

Enrico developed his ever increasing interest in probability theory and stochastic processes while working on continuous-time random walks with Rudolf Gorenflo and Francesco Mainardi and, later, with Guido Germano and René L. Schilling.

While working on this book Enrico enjoyed several discussions on related problems with Pier Luigi Ferrari, Fabio Rapallo and John Angle. Michele Manzini and Mauro Politi were very helpful in some technical aspects.

For Enrico, the idea of writing this book was born between 2005 and 2006. In 2005, he discussed the problem of allocating n objects into g boxes with Tomaso Aste and Tiziana di Matteo while visiting the Australian National University. In the same year, he attended a Thematic Institute (TI) sponsored by the EU EXYSTENCE network in Ancona, from May 2 to May 21. Mauro Gallegati and Edoardo Gaffeo (among the TI organizers) independently developed an interest in the possibility of applying the concept of statistical equilibrium to economics. The TI was devoted to *Complexity, Heterogeneity and Interactions in Economics and Finance*. A group of students, among whom were Marco Raberto, Eric Guerci, Alessandra Tedeschi and Giulia De Masi, followed Enrico's lectures on *Statistical equilibrium in physics and economics* where, among other things, the ideas of Costantini and Garibaldi were presented and critically discussed. During the Econophysics Colloquium in November 2006 in Tokyo, the idea of writing a textbook on these topics was suggested by Thomas Lux and Taisei Kaizoji, who encouraged Enrico to pursue this enterprise. Later, during the Econophysics Colloquium 2007 (again in Ancona), Joseph McCauley was kind enough to suggest Enrico as a potential author to Cambridge University Press. Out of this contact, the proposal for this book was written jointly by Ubaldo and Enrico and the joint proposal was later accepted. In October 2008, Giulio Bottazzi (who has some nice papers on the application of the Pólya distribution to the spatial structure of economic activities) invited Enrico for a short introductory course on probability for economists where some of the material in this book was tested on the field.

Last, but not least, Enrico wishes to acknowledge his wife, Tiziana Gaggino, for her patient support during the months of hard work. The formal deadline for this book (31 December 2009) was also the day of our marriage.