Svetlozar T. Rachev · Lev B. Klebanov
Stoyan V. Stoyanov · Frank J. Fabozzi

The Methods of Distances in the Theory of Probability and Statistics

Springer

The Methods of Distances in the Theory of Probability and Statistics

Svetlozar T. Rachev • Lev B. Klebanov
Stoyan V. Stoyanov • Frank J. Fabozzi

The Methods of Distances in the Theory of Probability and Statistics

Springer

Svetlozar T. Rachev
Department of Applied
Mathematics and Statistics
College of Business
Stony Brook University
Stony Brook, NY, USA

FinAnalytica
New York, NY, USA

Stoyan V. Stoyanov
EDHEC Business School
EDHEC-Risk Institute
Singapore

Lev B. Klebanov
Department of Probability and Statistics
Charles University
Prague, Czech Republic

Frank J. Fabozzi
EDHEC Business School
EDHEC-Risk Institute
Nice, France

ISBN 978-1-4614-4868-6 ISBN 978-1-4614-4869-3 (eBook)
DOI 10.1007/978-1-4614-4869-3
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012953265

Mathematics Subject Classification (2000): 60F05, 60F17, 60K25, 62G35, 62E20

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

STR
To my grandchildren Iliana, Zoya,
and Zari

LBK
To my wife Marina

SVS
To my wife Petya

FJF
To my wife Donna
and my children Francesco, Patricia,
and Karly

Preface

The development of the theory of probability metrics – a branch of probability theory – began with the study of problems related to limit theorems in probability theory. In general, the applicability of limit theorems stems from the fact that they can be viewed as an approximation to a given stochastic model and, consequently, can be accepted as an approximate substitute. The key question that arises in adopting the approximate model is the magnitude of the error that must be accepted. Because the theory of probability metrics studies the problem of measuring distances between random quantities or stochastic processes, it can be used to address the key question of how good the approximate substitute is for the stochastic model under consideration. Moreover, it provides the fundamental principles for building probability metrics – the means of measuring such distances.

The theory of probability metrics has been applied and has become an important tool for studying a wide range of fields outside of probability theory such as statistics, queueing theory, engineering, physics, chemistry, information theory, economics, and finance. The principal reason is that because distances are not influenced by the particular stochastic model under consideration, the theory of probability metrics provides some universal principles that can be used to deal with certain kinds of large-scale stochastic models found in these fields.

The first driving force behind the development of the theory of probability metrics was Andrei N. Kolmogorov and his group. It was Kolmogorov who stated that every approximation problem has its own distance measure in which the problem can be solved in a most natural way. Kolmogorov also contended that without estimates of the rate of convergence in the central limit theorem (CLT) (and similar limit theorems such as the functional limit theorem and the max-stable limit theorem), limit theorems provide very limited information. An example worked out by Y.V. Prokhorov and his students is as follows. Regardless of how slowly a function $f(n) > 0$, $n = 1, \ldots$, decays to zero, there exists a corresponding distribution function $F(x)$ with finite variance and mean zero, for which the CLT is valid at a rate slower than $f(n)$. In other words, without estimates for convergence in the CLT, such a theorem is meaningless because the convergence to the normal law of the normalized sum of independent, identically distributed random variables with distribution function $F(x)$ can be slower than any given rate $f(n) \to 0$. The problems associated with finding the appropriate rate of convergence invoked a variety of probability distances in which the speed of convergence (i.e., convergence rate) was estimated. This included the works of Yurii V. Prokhorov, Vladimir V. Sazonov, Vladimir M. Zolotarev, Vygantas Paulauskas, Vladimir V. Senatov, and others.

The second driving force in the development of the theory of probability metrics was mass-transportation problems and duality theorems. This started with the work of Gaspard Monge in the eighteenth century and Leonid V. Kantorovich in the 1940s – for which he was awarded the Nobel Prize in Economics in 1975 – on optimal mass transportation, leading to the birth of linear programming. In mathematical terms, Kantorovich's result on mass transportation can be formulated in the following metric way. Given the marginal distributions of two probability measures P and Q on a general (separable) metric space (U, d), what is the minimal expected value – referred to as $\kappa(P, Q)$ or the Kantorovich metric – of a distance $d(X, Y)$ over the set of all probability measures on the product space $U \times U$ with marginal distributions $P_X = P$ and $P_Y = Q$? If the measures P and Q are discrete, then this is the classic transportation problem in linear programming. If U is the real line, then $\kappa(P, Q)$ is known as the Gini statistical index of dissimilarity formulated by Corrado Gini. The Kantorovich problem has been used in many fields of science – most notably statistical physics, information theory, statistics, and probability theory. The fundamental work in this field was done by Leonid V. Kantorovich, Johannes H. B. Kemperman, Hans G. Kellerer, Richard M. Dudley, Ludger Rüschendorf, Volker Strassen, Vladimir L. Levin, and others. Kantorovich-type duality theorems established the main relationships between metrics in the space of random variables and metrics in the space of probability laws/distributions. The unifying work on those two directions was done by V. M. Zolotarev and his students.

In this book, we concentrate on four specialized research directions in the theory of probability metrics, as well as applications to different problems of probability theory. These include:

- Description of the basic structure of probability metrics,
- Analysis of the topologies in the space of probability measures generated by different types of probability metrics,
- Characterization of the ideal metrics for a given problem, and
- Investigation of the main relationships between different types of probability metrics.

Our presentation in this book is provided in a general form, although specific cases are considered as they arise in the process of finding supplementary bounds or in applications to important special cases.

The target audience for this book is graduate students in the areas of functional analysis, geometry, mathematical programming, probability, statistics, stochastic analysis, and measure theory. It may be partially used as a source of material for lectures for students in probability and statistics. As noted earlier in this preface, the theory of probability metrics has been applied to fields outside of probability theory such as engineering, physics, chemistry, information theory, economics, and finance. Specialists in these areas will find the book to be a useful reference to gain a greater understanding of this specialized area and its potential application.

New York, USA — Svetlozar T. Rachev
Prague, Czech Republic — Lev B. Klebanov
Singapore — Stoyan V. Stoyanov
Nice, France — Frank J. Fabozzi

Contents