Probability Metrics and the Stability of Stochastic Models

SVETLOZAR T. RACHEV
Department of Statistics and Applied Probability
University of California
Santa Barbara
USA

JOHN WILEY & SONS
Chichester . New York . Brisbane . Toronto . Singapore

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

Jacaranda Wiley Ltd, G.P.O. Box 859, Brisbane,
Queensland 4001, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

John Wiley & Sons (SEA) Pte Ltd, 37 Jalan Pemimpin 05-04,
Block B, Union Industrial Building, Singapore 2057

Library of Congress Cataloging-in-Publication Data:

Rachev, S. T. (Svetlozar Todorov)
Probability metrics / Svetlozar T. Rachev.
p. cm. — (Wiley series in probability and mathematical statistics. Applied probability and statistics section)
Includes bibliographical references and index.
ISBN 0 471 92877 1
1. Limit theorems (Probability theory) 2. Metric spaces.
I. Title. II. Series.
QA273.67.R33 1991
519.2—dc20 90-43733
CIP

British Library Cataloguing in Publication Data:

Rachev, Svetlozar T.
Probability metrics.
1. Probabilities & statistical mathematics
I. Title
519.2

ISBN 0 471 92877 1

Typeset by Techset Composition Ltd, Salisbury, UK.
Printed in Great Britain by Courier International, Tiptree, Essex

To my children

Borjana and Vladimir

Contents

Preface xi

PART I GENERAL TOPICS IN THE THEORY OF PROBABILITY METRICS

Chapter 1 Main Directions in the Theory of Probability Metrics **3**

Chapter 2 Probability Distances and Probability Metrics: Definitions **5**

2.1 Some examples of metrics in probability theory 5
2.2 Metric and semimetric spaces; distance and semidistance spaces 7
2.3 Definitions of probability distance and probability metric 10
2.4 Universally measurable separable metric spaces 12
2.5 The equivalence of the notions of p. (semi-)distance on $\mathscr{P}_2$ and on $\mathfrak{X}$ 17

Chapter 3 Primary, Simple and Compound p. Distances. Minimal and Maximal Distances and Norms **21**

3.1 Primary distances and primary metrics 21
3.2 Simple distances and metrics; co-minimal functionals and minimal norms 25
3.3 Compound distances and moment functions 39

Chapter 4 A Structural Classification of the Probability Distances **51**

4.1 Hausdorff structure of p. semidistances 51
4.2 Λ-structure of p. semidistances 68
4.3 ζ-structure of p. semidistances 71

PART II RELATIONS BETWEEN COMPOUND, SIMPLE AND PRIMARY DISTANCES

Chapter 5 Monge–Kantorovich Mass Transference Problem. Minimal Distances and Minimal Norms **89**

5.1 Statement of Monge–Kantorovich problem 89
5.2 Multi-dimensional Kantorovich theorem 89

5.3 Dual representation of the mimimal norms $\mathring{\mu}_c$; a generalization of the Kantorovich–Rubinstein theorem 107
5.4 Application: explicit representations for a class of minimal norms 115

Chapter 6 Quantitative Relationships between Minimal Distances and Minimal Norms 119

6.1 Kantorovich metric $\hat{\mu}_c$ is equal to Kantorovich–Rubinstein norm $\mathring{\mu}_c$ if and only if the cost function c is a metric 119
6.2 Inequalities between $\hat{\mu}_c$, $\mathring{\mu}_c$ 121
6.3 Convergence, compactness and completeness in $(\mathscr{P}(U), \hat{\mu}_c)$ and $(\mathscr{P}(U), \mathring{\mu}_c)$ 130
6.4 $\mathring{\mu}_c$- and $\hat{\mu}_c$-uniformity 134
6.5 Generalized Kantorovich and Kantorovich–Rubinstein functionals 136

Chapter 7 *K*-minimal Metrics 141

7.1 Definition; general properties 141
7.2 Two examples of K-minimal metrics 146
7.3 K-minimal metrics of given probability metrics; the case $U = R$ 147
7.4 The case: U is a separable metric space 155
7.5 Relations between multi-dimensional Kantorovich and Strassen theorems; convergence of minimal metrics and minimal distances 162

Chapter 8 Relations between Minimal and Maximal Distances 167

8.1 Duality theorems and explicit representations for $\check{\mu}_c$ and $\overset{(s)}{\mu}_c$ 167
8.2 Convergence of measures with respect to minimal distances and minimal norms 174

Chapter 9 Moment Problems Related to the Theory of Probability Metrics. Relations between Compound and Primary Distances 185

9.1 Primary minimal distances; moment problems with one fixed pair of marginal moments 185
9.2 Moment problems with two fixed pairs of marginal moments; moment problems with fixed linear combinations of moments 191

PART III APPLICATIONS OF MINIMAL p. DISTANCES

Chapter 10 Uniformity in Weak and Vague Convergence 201

10.1 ζ-metrics and uniformity classes 201
10.2 Metrization of the vague convergence 208

Chapter 11 Glivenko–Cantelli Theorem and Bernstein–Kantorovich Invariance Principle **211**

11.1 Fortet–Mourier, Varadarajan and Wellner Theorems 211
11.2 Functional central limit theorem and Bernstein–Kantorovich invariance principle 218

Chapter 12 Stability of Queuing Systems **223**

12.1 Stability of $G|G|1|\infty$-system 223
12.2 Stability of $GI|GI|1|\infty$-system 226
12.3 Approximation of a random queue by means of deterministic queueing models 229

Chapter 13 Optimal Quality Usage **241**

13.1 Optimality of quality usage and the Monge–Kantorovich problem 241
13.2 Estimates of the minimal total losses $\tau_{\phi}(\phi^*)$ 248

PART IV IDEAL METRICS

Chapter 14 Ideal Metrics with Respect to Summation Scheme for i.i.d. Random Variables **257**

14.1 Robustness of χ^2-test of exponentiality 257
14.2 Ideal metrics for sums of independent random variables 264
14.3 Rates of convergence in the CLT in terms of metrics with uniform structure 274

Chapter 15 Ideal Metrics and Rate of Convergence in the CLT for Random Motions **283**

15.1 Ideal metrics in the space of random motions 283
15.2 Rates of convergence in the integral and local CLTs for random motions 288

Chapter 16 Applications of Ideal Metrics for Sums of i.i.d. Random Variables to the Problems of Stability and Approximation in Risk Theory **299**

16.1 The problem of stability in risk theory 299
16.2 The problem of continuity 301
16.3 Stability of the input characteristics 307

Chapter 17 How Close are the Individual and Collective Models in the Risk Theory? 313

17.1 Stop-loss distances as measures of closeness between individual and collective models 313
17.2 Approximation by compound Poisson distributions 326

Chapter 18 Ideal Metric with Respect to Maxima Scheme of i.i.d. Random Elements 337

18.1 Rate of convergence of maxima of random vectors via ideal metrics 337
18.2 Ideal metrics for the problem of rate of convergence to max-stable processes 351
18.3 Doubly ideal metrics 373

Chapter 19 Ideal Metrics and Stability of Characterizations of Probability Distributions 387

19.1 Characterization of an exponential class of distributions $\{F_p, 0 < p \leqslant \infty\}$ and its stability 389
19.2 Stability in de Finnetti's theorem 400
19.3 Characterization and stability of environmental processes 406

Bibliographical Notes 423

References 435

Table of Probability Semidistances 463

Index of Symbols and Abbreviations 479

Index 487

Preface

The study of limit theorems and a number of other questions in probability theory makes it necessary to introduce functionals, defined either on classes of probability distributions or on classes of random elements, and evaluating their nearness in one or another probabilistic sense. Thus various metrics have appeared, among which are the well known Kolmogorov (uniform) metric, L^p metrics, the Prokhorov metric, the metric of convergence in probability (Ky Fan metric) and others. The use of metrics in many problems in probability theory is connected with the following fundamental question:

'Is the proposed stochastic model a satisfactory approximation to the real model, and if so, within what limits?' To answer this question, an investigation of the qualitative and quantitative stability of the stochastic model is required. Analysis of quantitative stability assumes the use of metrics as measures of comparability. The main idea of the *method of metric distances* (MMD), developed by V. M. Zolotarev and his students to solve stability problems, is reduced to the following two aspects.

Problem 1 (Choice of ideal metrics). Find the most appropriate (ideal) metrics for the stability problem under consideration. Then, solve the problem in terms of these ideal metrics.

Problem 2 (Comparison of metrics). If it is required to write the solution of the stability problem in terms of other metrics, one must solve the problem of comparison of these metrics with the chosen (ideal) metrics.

Unlike Problem 1, Problem 2 does not depend on the specific stochastic model under consideration. Thus, the independent solution of the second problem allows its re-use in any particular situation. In addition, it enables us to use a variety of metric relationships without making any effort in different kinds of stability problems. Moreover, following the stated two-stage approach, we get a clear comprehension of the specific regularities which form the stability effect.

In probability theory, metrics have been used for a long time, although one usually exploits a very limited class of metrics. Also, some ideas of the MMD have been used for a long time in approximation theory and functional analysis. In view of the variety of stability problems, there are no regular selection rules determining the 'ideal' metric for the given problem. Therefore, the development of the MMD demands the creation of a *theory of probability metrics* (TPM).

The term 'probability metric' means simply a semimetric in a space of random variables (taking values in some separable metric space). In probability theory, sample spaces are usually not fixed and one is interested in those metrics whose values depend on the joint distributions of the pairs of random variables being considered. Each such metric can be considered just as given by a function defined on the set of probability measures on the Cartesian square of the sample space. Complications connected with the question of existence of pairs of random variables on a given space with given probability laws can be easily avoided. Although such a function is not a metric on a space of probability distributions (it is not a function of pairs of measures), small values of it say that the measure is concentrated near the diagonal. Therefore its marginal distributions are close to each other. Fixing these marginal distributions, one can find the infimum of the values of our function on the class of all measures with the given marginals. Such an infimum is a metric on the class of probability distributions and in some concrete cases (for example, for the L_1 distance in the space of random variables—Kantorovich's theorem; for the Ky Fan metric—Strassen–Dudley's theorem; for the indicator metric—Dobrushin's theorem) were found earlier (giving, respectively, the Kantorovich (or Wasserstein) metric, the Prokhorov metric and the total variation distance).

The necessary classification of the set of probability metrics (*p. metrics*) is naturally carried out from the point of view of metric structure and generating topologies. That is why the following two research directions arise.

Direction 1. Description of the basic structures of p. metrics.

Direction 2. Analysis of the topologies in the space of probability measures, generated by different types of p. metrics. This analysis can be carried out with the help of convergence criteria for different metrics.

At the same time, more specialized research directions arise. Namely,

Direction 3. Characterization of the ideal metrics for the given problem.

Direction 4. Investigations of the main relationships between different type of p. metrics.

In this book, all four directions are considered as well as applications to different problems of probability theory. Much attention is paid to the possibility of giving equivalent definitions of p. metrics (for example, in direct and dual terms, in terms of the Hausdorff metric for sets, etc). Indeed, in concrete applications of p. metrics, the use of different equivalent variants of the definitions in different steps of the proof is often a decisive factor.

One of the main classes of metrics considered is the class of *minimal metrics*, the idea of which goes back to the work of Kantorovich in the 1940s on the transportation problems in linear programming. Such metrics have been found independently by many authors in several parts of probability theory (Markov processes, statistical physics, etc.). They are connected with the widely known method of 'coupling.' Then it is natural to evaluate the distance between variables with metrics of the indicated type. Distances for such metrics are hard to compute, but it is easy to give upper bounds, attained by at least one joint probability distribution. Another useful class of metrics studied in this book is the class of 'ideal' metrics having satisfied the following properties: (1) $\mu(P_c, Q_c) \leqslant |c|^r \mu(P, Q)$ for all $c \in [-C, C], c \neq 0$, where $P_c(A) := P((1/c)A)$ for any Borel set A on a Banach space U, and (2) $\mu(P_1 * Q, P_2 * Q) \leqslant \mu(P_1, P_2)$, where $*$ denotes the convolution. This class is convenient for the study of functionals of sums of independent random variables, giving nearest bounds of the distance to limit distributions.

The presentation here is given 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 MMD given herein is illustrated in some concrete problems. First, there are problems of the type of the Glivenko–Cantelli theorem on the convergence of the empirical measures. Originally, this kind of problem was different in view of the various possible natural modes of convergence and was considered as a problem needing an *ad hoc* approach. Here it is shown that from a general point of view such results turn out to be obvious consequences of the SLLN and general properties of metrics. Analogously, we considered a generalization of the Prokhorov theorem on convergence of random polygons to the Wiener process in the case where one considers the question of convergence of distributions of unbounded functionals. Also considered are applications to the rate of convergence for sums and maxima of random variables and convolution of random motions. Special sections are devoted to application of TPM to the solution of stability problems in queueing theory, risk theory, quality usage, and others.

I would like to thank a number of people who have directly contributed to this project. V. M. Zolotarev introduced me to the subject of theory of probability metrics and stability of stochastic models. He has been a constant source of intellectual guidance and inspiration. R. M. Shortt, L. Rüschendorf, L. de Haan, J. Yukich, E. Omey, L. Baxter, P. Todorovich, M. Taksar and J. Beirlant worked with me on papers which formed the foundation of this book. R. M. Dudley's lecture notes from Aarhus University have been a rich source of ideas. In addition, I want to thank H. Robbins, S. Cambanis, G. Simons, H. Kellerer, S. Resnick and G. Samorodnitski for many helpful discussions during the last three years. My warm thanks and appreciation go to Ruth Bahr, Michelle Bebb and Lee Trimble for their expert typing of this manuscript. Finally, I must thank the editorial and production department of Wiley for their support, patience and superb final product.

While trying to correct all the mistakes in my manuscript, I realize that all of them have not been found. I will be very happy to learn of any mistakes found by the reader together with any comments one might have†.

Svetlozar T. Rachev

† Research supported by NATO Scientific Affairs Division Grant CRG 900798 and Grant from UC Regents, University of California, Santa Barbara.

PART I

General Topics in the Theory of Probability Metrics

CHAPTER 1

Main Directions in the Theory of Probability Metric

In the period of the formation of probability theory as a mathematical science, the first limit theorems (the Law of Large Numbers, De Moivre and Laplace Central Limit Theorems) were obtained (cf. Feller 1970, Chap. VI, 4, and Chap. VII, 3). In applications of these limit theorems, a given probability distribution is often approximated by some limiting distribution. In such cases, the convergence rate problem arises. As mentioned in the Preface, the convergence rate problem requires the concept of a metric as a measure of comparability.

Even in the case of probability distributions on the real line, several different types of metrics (Lévy metric, Kolmogorov metric, L_p-metric) are often used to estimate the closeness between distributions. Since the early thirties, the demands of various applications resulted in the creation of new, more complicated probability models. Both the theory of random processes and the theory of distributions on functional spaces were extensively developed. In connection with limit theorems in general spaces, Fortet and Mourier (1953), Prokhorov (1956), Kantorovich and Rubinstein (1958), and Dudley (1966a) suggested a series of new metrics on spaces of distributions. Certainly, the study of metric properties is not confined to probability theory; such investigations have occurred in other mathematical areas. In fact, some metrics on spaces of measures (e.g. the Kantorovich–Rubinstein metric, total variation norm, and L_p-metrics) play an important role in functional analysis (see Dunford and Schwartz 1988, Kantorovich and Akilov 1984). In probability theory, an even greater variety of such metrics arises in completely natural ways. This variety and suitability can be partly explained by the fact that the stochastic problem under consideration often dictates the choice of an appropriate metric. Frequently, this choice will determine the investigation's success.

Questions concerning the bounds within which stochastic models can be applied (as in all probabilistic limit theorems) can only be answered by investigation of qualitative and quantitative stability. Such stability is very often convenient to express in terms of a metric. This was the case with Zolotarev's Method of Metric Distances (MMD) and the Theory of Probability Metrics (TPM) (see Zolotarev 1976a–d, 1977a,b, 1983a,b, Zolotarev and Rachev 1985).