Selected Works in Probability and Statistics

For other titles published in this series, go to
www.springer.com/series/8556

Evarist Giné • Vladimir Koltchinskii
Rimas Norvaiša
Editors

Selected Works of R.M. Dudley

Springer

Editors
Evarist Giné
University of Connecticut
Department of Mathematics
Auditorium Road 196
06269-3009 Storrs Connecticut U-9
USA
gine@math.uconn.edu

Vladimir Koltchinskii
Georgia Institute of Technology
School of Mathematics
30332-0160 Atlanta
USA
vlad@math.gatech.edu

Rimas Norvaiša
Institute of Mathematics and Informatics
Akademijos str. 4
LT-08663 Vilnius
Lithuania
norvaisa@ktl.mii.lt

ISBN 978-1-4419-5820-4 e-ISBN 978-1-4419-5821-1
DOI 10.1007/978-1-4419-5821-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010932623

Printed on acid-free paper

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

Preface to the Series

Springer's Selected Works in Probability and Statistics series offers scientists and scholars the opportunity of assembling and commenting upon major classical works in statistics, and honors the work of distinguished scholars in probability and statistics. Each volume contains the original papers, original commentary by experts on the subject's papers, and relevant biographies and bibliographies.

Springer is committed to maintaining the volumes in the series with free access on SpringerLink, as well as to the distribution of print volumes. The full text of the volumes is available on SpringerLink with the exception of a small number of articles for which links to their original publisher is included instead. These publishers have graciously agreed to make the articles freely available on their websites. The goal is maximum dissemination of this material.

The subjects of the volumes have been selected by an editorial board consisting of Anirban DasGupta, Peter Hall, Jim Pitman, Michael Sörensen, and Jon Wellner.

Richard Mansfield Dudley

Preface

It is almost impossible to describe all the research of Richard Mansfield Dudley, as it comprises more than 100 articles and spans over more than 45 years. The three co-editors of this volume have selected what we think are the most influential and representative of his articles and have shared the task of offering a few comments on them. Consistent with the overview of his work that follows, these Selected Works of Richard M. Dudley are divided into six chapters, each preceded by a short note on its content.

Dudley has been extremely influential in the development of Probability and Mathematical Statistics during the second half of the last century (or more exactly, between the nineteen sixties and the present). The two subjects on which he has left the deepest mark so far, in our view, are Gaussian processes and empirical processes. Succinctly, on the first: his research changed the framework of study of Gaussian processes by highlighting the intrinsic metric structure on the parameter space, and provided one of the main tools of their study, the famous metric entropy bound. And on the second: Vapnik and Červonenkis, in connection with their work on machine learning, initiated the modern view of empirical processes as processes indexed by general classes of sets and functions and proved uniform laws of large numbers for them, but it was the results of Dudley on the uniform central limit theorem that a) made this theory so useful and pervasive in asymptotic statistics, and b) initiated a vigorous development by many authors, that in turn made this theory even more useful. We may well say that in the first case Dudley changed the direction of the field in a way that led to the solution of some of its outstanding questions and in the second he created (or co-created with Vapnik and Červonenkis) a whole new field.

There are at least three more quite large fields of research where Richard M. Dudley's achievements are crucial and manifold. When Dudley started his career, the works of Prokhorov, Skorokhod, Varadarajan and LeCam on the general theory of weak convergence were very recent, and Dudley made substantial contributions to this theory in several ways, in particular, to complete the proofs of results of Donsker on the empirical process, and to study metrizability of this convergence in metric spaces. He also wrote, between 1965 and 1973, a series of articles on 'relativistic Markov processes', where he introduced Lorentz invariant diffusions as the only possible analogues in Lorentz space of Brownian motion and (in 3 dimensions) rotationally invariant Lévy processes, gave their properties, including asymptotics, and pointed at possible applications in Cosmology. Empirical processes can be thought of as the 'linear term' in Taylor type developments of statistical functionals, and several researchers in the seventies and eighties used compact differentiation of functionals defined on $D[0,1]$ endowed with the supremum norm; Dudley advocates replacing the sup norm by the p-variation norm and compact differentiability by the more regular Fréchet differentiability, and shows this is possible (and desirable) by using empirical process techniques and by proving the Fréchet differentiability of some of the most usual operators like composition and inverse; this theory is not only quite elegant but it leads to optimality results.

This classification of Dudley's research into large areas leaves out many other remarkable works that touch upon approximation theory, learnability, Wiener functionals, singularity of measures, prediction theory, mathematical statistics, IQ and heredity, ... Some of them are commented upon below.

If we try to think about the common attributes of Dudley's works, three come immediately to mind: *good taste* in the choice of subjects by their relevance and beauty, *rigor* and *scholarship*. These characteristics make of Richard Dudley an excellent adviser, as one of us directly experienced, and this may help explain why he has had as many as thirty-one PhD students, while fairness and generosity are some of his other personal attributes that help further explain why he has been such a successful adviser.

It has been both an honor and a pleasure to select and comment on the works of a scholar of Richard M. Dudley's quality, and even more so as he has had a deep influence on the work (and on more than just the work) of each of us. We are thus grateful to the Editors of the Springer Selected Works in Probability and Statistics for the opportunity to assemble this volume. We also thank Edith Dudley Sylla for correspondence on Dudley's biography. Finally, we wish to acknowledge John Kimmel's and Jim Pitman's support in the preparation of this volume.

Evarist Giné
Vladimir Koltchinskii
Rimas Norvaiša

Acknowledgements

This series of selected works is possible only because of the efforts and cooperation of many people, societies, and publishers. The series editors originated the series and directed its development. The volume editors spent a great deal of time organizing the volumes and compiling the previously published material. The contributors provided comments on the significance of the papers. The societies and publishers who own the copyright to the original material made the volumes possible and affordable by their generous cooperation:

American Institute of Physics
American Mathematical Society
American Statistical Association
Applied Probability Trust
Bernoulli Society
Cambridge University Press
Canadian Mathematical Society
Danish Society for Theoretical Statistics
Elsevier
Finnish Statistical Society
Indian Statistical Institute
Institute of Mathematical Statistics
International Chinese Statistical Association
International Statistical Institute
John Wiley and Sons
New Zealand Statistical Association
Norwegian Statistical Society
Oxford University Press
Princeton University and the Institute for Advanced Studies
Royal Statistical Society
Statistical Society of Australia
Swedish Statistical Society
University of California Press
University of Illinois, Department of Mathematics
University of North Carolina Press

Contents

Part 3: Gaussian Processes

Part 4: Empirical Processes

Part 5: Nonlinear functionals and p-variation

Part 6: Miscellanea

A Biographical Sketch of Richard M. Dudley

Richard Mansfield Dudley was born on July 28, 1938, in Cleveland, Ohio, into a family with a strong university tradition, some of it mathematical. He graduated as valedictorian from Bloom Township High School in Chicago Heights, Illinois in 1955, attended Harvard University from 1955 to 1959, where he graduated Summa cum Laude as Bachelor of Arts, and obtained the PhD in Mathematics from Princeton University in 1962. Richard Dudley had two advisers, Gilbert A. Hunt and Edward Nelson, and wrote a thesis entitled 'Lorentz invariant random distributions'. Professor Dudley's professional career has taken place at the University of California at Berkeley, from 1962 to 1966, and at the Massachusetts Institute of Technology since 1967. He was a Fellow of the A. P. Sloan Foundation in 1966–68 and of the Guggenheim Foundation in 1991.

At Berkeley, Dudley coincided with Volker Strassen, who introduced him to Sudakov's idea of using metric entropy for the study of sample Gaussian paths and with whom he also shared an interest in probability distances that metrize weak convergence. He also worked with two other Berkeley colleagues, Lucien Le Cam and Jack Feldman, on different aspects of Gaussian processes. While at Berkeley, Dudley published work on Lorentz-invariant Markov processes, singularity of measures on linear spaces, random walks on groups, sequential convergence, prediction theory for non-stationary processes (the subject of the first PhD thesis he supervised), and on weak convergence of probabilities on nonseparable metric spaces (having to do with empirical processes) and Baire measures, among other topics.

Dudley completed his deep and extensive study of Gaussian processes when he was already at M.I.T., and, although he also started working on empirical processes at Berkeley (empirical measures on Euclidean spaces), he developed this theory at M.I.T., in about twenty publications including his landmark 1978 paper 'Central limit theorems for empirical processes' in Annals of Probability, that basically created a new subfield within Mathematical Statistics, the very influential Saint-Flour lecture notes (1984), and the book 'Uniform Central Limit Theorems' (Cambridge Studies in Advanced Mathematics). Empirical process theory is pervasive in modern Mathematical Statistics and has had a deep impact in the field of machine learning. Moreover, Dudley has worked on measurability problems, testing hypotheses, Wiener functionals, and approximation theory (in connection with metric entropy, Gaussian processes and empirical processes), to name a few subjects. During the last twenty years, he has also been interested in *p*-variation and in differentiable functionals, in connection with the delta method for statistical functionals. Most of the work of Richard Dudley is single authored but at the same time he has collaborated with several colleagues and with a few of his thirty-one PhD students. Aside from the already mentioned Strassen, Le Cam and Feldman, the list of collaborators includes S. L. Cook, M. Durst, J. M. González-Barrios, S. Gutmann, D. Haughton, J. Hoffmann-Jørgensen, Y.-C. Huang, M. Kanter, J. Kuelbs, S. R. Kulkarni, J. Llopis, L. Pakula, N. P. Peng, P. Perkins, W. Philipp, A. Quiroz, B. Randol, T. J. Richardson, L. Shepp, S. Sidenko, D. Smith, D. Stroock, F. Topsøe, Z. Wang, R. S. Wenocur, O. Zeitouni, J. Zinn, and the editors of this volume.

Richard Dudley was visiting professor at the University of Aarhus, Denmark, in the spring of 1976, and, while there, he wrote a very nice set of lectures for a graduate course, Probabilities and Metrics. Most of these notes were later adapted and incorporated in the second part of his deservedly successful graduate text Real Analysis and Probability (1989, 2002). This book is special in particular because the author looked for the best and shortest proofs available and each chapter ends with very accurate and complete historical notes. He has also written monographs on two of the subjects on which he has recently worked most: on empirical processes, he has written the book Uniform Central Limit Theorems, which also originated in lecture notes, in this case the above mentioned Saint-Flour lectures, 'A course on empirical Processes', and on *p*-variation and related matters, jointly with Rimas Norvaiša, he has written two sets of lecture notes (Differentiability of Six Operators on Non-Smooth Functions and *p*-variation, and An Introduction to *p*-variation and Young Integrals), and a forthcoming book, Concrete Functional Calculus. Beside the above, his scholarly service to the profession also includes having been the Editor of the Annals of Probability, 1979–1981, after six years of being an Associate Editor, and having served on the Editorial Board of the Wadsworth Advanced Series in Statistics/Probability, 1982–1992. He has talked and talks about his work at many scientific meetings, and in particular he has been an invited speaker at the 1974 International Congress of Mathematicians, at American Mathematical Society, Institute of Mathematical Statistics and Bernoulli Society meetings, at Vilnius Conferences in Probability Theory and Mathematical Statistics, at Saint-Flour, at many Probability in Banach Spaces meetings (some of which he helped organize), etc.

Dudley has participated at different times in very concrete and effective ways in civic activities related to his convictions or to his interests, so, for instance, he was a volunteer news writer and broadcaster one afternoon a week in 1963–1966 for public radio station KPFA in Berkeley, and in 1979 he was the Editor of the Appalachian Mountain Club White Mountain Guide. He has hiked up all the mountains in New England over 4000 feet, besides Mont Blanc and other mountains in the Alps and elsewhere.

Professor Dudley has received several honors, in particular, he has served in the (honorary) Advisory Board of Stochastic Processes and their Applications, 1987–2001 and is a Fellow of: the Institute of Mathematical Statistics, the American Statistical Association, and the American Association for the Advancement of Science; and an elected Member of the International Statistical Institute.

Richard Dudley lives in Newton, Massachusetts, with Liza Martin, his wife of thirty-one years.

Author's Bibliography 1961–2009

[1] R. M. Dudley. Continuity of homomorphisms. *Duke Math. J.*, 28:587–594, 1961.

[2] R. M. Dudley and B. Randol. Implications of pointwise bounds on polynomials. *Duke Math. J.*, 29:455–458, 1962.

[3] R. M. Dudley. *Lorentz invariant random distributions*. PhD thesis, Princeton, 1962. Mathematics.

[4] R. M. Dudley. Random walks on abelian groups. *Proc. Amer. Math. Soc.*, 13:447–450, 1962.

[5] R. M. Dudley. A property of a class of distributions associated with the Minkowski metric. *Illinois J. Math.*, 8:169–174, 1964.

[6] R. M. Dudley. On sequential convergence. *Trans. Amer. Math. Soc.*, 112:483–507, 1964. Corrections *ibid.* 148:623–624, 1970.

🕮 [7] R. M. Dudley. Pathological topologies and random walks on abelian groups. *Proc. Amer. Math. Soc.*, 15:231–238, 1964.

[8] R. M. Dudley. Singular translates of measures on linear spaces. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 3:128–137, 1964.

[9] R. M. Dudley. Fourier analysis of sub-stationary processes with a finite moment. *Trans. Amer. Math. Soc.*, 118:360–375, 1965.

[10] R. M. Dudley. Gaussian processes on several parameters. *Ann. Math. Statist.*, 36:771–788, 1965.

[11] R. M. Dudley. Convergence of Baire measures. *Studia Math.*, 27:251–268, 1966.

🕮 [12] R. M. Dudley. Lorentz-invariant Markov processes in relativistic phase space. *Ark. Mat.*, 6:241–268, 1966.

[13] R. M. Dudley. Singularity of measures on linear spaces. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 6:129–132, 1966.

🕮 [14] R. M. Dudley. Weak convergences of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.*, 10:109–126, 1966.

🕮 [15] R. M. Dudley. A note on Lorentz-invariant Markov processes. *Ark. Mat.*, 6:575–581, 1967.

🕮 [16] R. M. Dudley. Measures on non-separable metric spaces. *Illinois J. Math.*, 11:449–453, 1967.

[17] R. M. Dudley. On prediction theory for nonstationary sequences. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, pages Vol. II: Contributions to Probability Theory, Part 1, pp. 223–234. Univ. California Press, Berkeley, Calif., 1967.

[18] R. M. Dudley. Sub-stationary processes. *Pacific J. Math.*, 20:207–215, 1967.

🕮 [19] R. M. Dudley. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis*, 1:290–330, 1967.

🕮 [20] R. M. Dudley. Distances of probability measures and random variables. *Ann. Math. Statist.*, 39:1563–1572, 1968.

🕮 Indicates article found in this volume

[21] R. M. Dudley. The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Statist.*, 40:40–50, 1968.

[22] R. M. Dudley. Fourier transforms of stationary processes are not functions. *J. Math. Mech.*, 19:309–316, 1969.

[23] R. M. Dudley. Random linear functionals. *Trans. Amer. Math. Soc.*, 136:1–24, 1969.

[24] V. Strassen and R. M. Dudley. The central limit theorem and ε-entropy. In *Probability and Information Theory (Proc. Internat. Sympos., McMaster Univ., Hamilton, Ont., 1968)*, pages 224–231. Springer, Berlin, 1969.

[25] R. M. Dudley. Random linear functionals: Some recent results. In C. T. Taam, editor, *Lectures in Modern Analysis and Applications, III*, Lecture Notes in Mathematics, Vol. 170, pages 62–70. Springer, Berlin, 1970.

[26] R. M. Dudley. Small operators between Banach and Hilbert spaces. *Studia Math.*, 38:35–41, 1970.

[27] R. M. Dudley. Some uses of ε-entropy in probability theory. In *Les probabilités sur les structures algébriques (Actes Colloq. Internat. CNRS, No. 186, (Clermont-Ferrand, 1969)*, pages 113–122. Éditions Centre Nat. Recherche Sci., Paris, 1970. With discussion by L. Le Cam, Ju. V. Prohorov, A. Badrikian and R. M. Dudley.

[28] R. M. Dudley. Convergence of sequences of distributions. *Proc. Amer. Math. Soc.*, 27:531–534, 1971.

[29] R. M. Dudley. Non-linear equivalence transformations of Brownian motion. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 20:249–258, 1971.

[30] R. M. Dudley. On measurability over product spaces. *Bull. Amer. Math. Soc.*, 77:271–274, 1971.

[31] R. M. Dudley, J. Feldman, and L. Le Cam. On seminorms and probabilities, and abstract Wiener spaces. *Ann. of Math.*, 93:390–408, 1971. Corrections *ibid.* 104:391, 1976.

[32] R. M. Dudley and L. Pakula. A counter-example on the inner product of measures. *Indiana Univ. Math. J.*, 21:843–845, 1972.

[33] R. M. Dudley. A counterexample on measurable processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 57–66, Berkeley, Calif., 1972. Univ. California Press. Corrections *ibid.* 1:191–192, 1973.

[34] R. M. Dudley. Speeds of metric probability convergence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 22:323–332, 1972.

[35] R. M. Dudley. A note on products of spectral measures. In *Vector and operator valued measures and applications (Proc. Sympos., Alta, Utah, 1972)*, pages 125–126. Academic Press, New York, 1973.

[36] R. M. Dudley. Asymptotics of some relativistic Markov processes. *Proc. Nat. Acad. Sci. U.S.A.*, 70:3551–3555, 1973.

[37] R. M. Dudley. Sample functions of the Gaussian process. *Ann. Probability*, 1:66–103, 1973.

[38] R. M. Dudley. Metric entropy and the central limit theorem in $C(S)$. *Ann. Inst. Fourier (Grenoble)*, 24:49–60, 1974. Colloque International sur les Processus Gaussiens et les Distributions Aléatoires (Colloque Internat. du CNRS, No. 222, Strasbourg, 1973).

[39] R. M. Dudley. Metric entropy of some classes of sets with differentiable boundaries. *J. Approximation Theory*, 10:227–236, 1974.

[40] R. M. Dudley. Recession of some relativistic Markov processes. *Rocky Mountain J. Math.*, 4:401–406, 1974. Papers arising from a Conference on Stochastic Differential Equations (Univ. Alberta, 1972).

[41] R. M. Dudley and M. Kanter. Zero-one laws for stable measures. *Proc. Amer. Math. Soc.*, 45:245–252, 1974. Corrections *ibid.* 88:689–690, 1983.

[42] R. M. Dudley, P. Perkins, and E. Giné. Statistical tests for preferred orientation. *J. Geology*, 83:685–705, 1975.

[43] R. M. Dudley. The Gaussian process and how to approach it. In *Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974), Vol. 2*, pages 143–146. Canad. Math. Congress, Montreal, Que., 1975.

[44] R. M. Dudley. *Probabilities and Metrics: Convergence of Laws on Metric Spaces with a View to Statistical Testing*, volume 45 of *Lecture Notes Series*. Matematisk Institut, Aarhus Universitet, Aarhus, 1976.

[45] F. Topsøe, R. M. Dudley, and J. Hoffmann-Jørgensen. Two examples concerning uniform convergence of measures w.r.t. balls in Banach spaces. In P. Gaenssler and P. Révész, editors, *Empirical distributions and processes (Selected Papers, Meeting on Math. Stochastics, Oberwolfach, 1976)*, pages 141–146. Lecture Notes in Math., Vol. 566. Springer, Berlin, 1976.

[46] R. M. Dudley. Unsolvable problems in mathematics. *Science*, 191:807–808, 1976.

[47] R. M. Dudley. A survey on central limit theorems in Banach spaces. In *Symposia Mathematica, Vol. XXI (Convegno sulle Misure su Gruppi e su Spazi Vettoriali, Convegno sui Gruppi e Anelli Ordinati, INDAM, Rome, 1975)*, pages 11–14. Academic Press, London, 1977.

[48] R. M. Dudley. On second derivatives of convex functions. *Math. Scand.*, 41:159–174, 1977. Acknowledgement of priority, *ibid.* 46:6, 1980.

[49] R. M. Dudley and S. Gutmann. Stopping times with given laws. In *Séminaire de Probabilités, XI (Univ. Strasbourg, Strasbourg, 1975/1976)*, pages 51–58. Lecture Notes in Math., Vol. 581. Springer, Berlin, 1977.

[50] R. M. Dudley. Wiener functionals as Itô integrals. *Ann. Probability*, 5:140–141, 1977.

[51] R. M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.*, 6:899–929, 1978. Corrections *ibid.* 7:909–911, 1979.

[52] R. M. Dudley. Balls in $\mathbf{R}^k$ do not cut all subsets of $k+2$ points. *Adv. in Math.*, 31:306–308, 1979.

[53] R. M. Dudley. Lower layers in R^2 and convex sets in R^3 are not GB classes. In *Probability in Banach spaces, II (Proc. Second Internat. Conf., Oberwolfach, 1978)*, volume 709 of *Lecture Notes in Math.*, pages 97–102. Springer, Berlin, 1979.

[54] J. Hoffmann-Jørgensen, L. A. Shepp, and R. M. Dudley. On the lower tail of Gaussian seminorms. *Ann. Probab.*, 7:319–342, 1979.

[55] R. M. Dudley. On χ^2 tests of composite hypotheses. In *Probability theory (Papers, VIIth Semester, Stefan Banach Internat. Math. Center, Warsaw, 1976)*, volume 5 of *Banach Center Publ.*, pages 75–87. PWN, Warsaw, 1979.

[56] J. Kuelbs and R. M. Dudley. Log log laws for empirical measures. *Ann. Probab.*, 8:405–418, 1980.

[57] R. M. Dudley. Donsker classes of functions. In M. Csörgő et al., editors, *Statistics and related topics (Ottawa, Ont., 1980)*, pages 341–352. North-Holland, Amsterdam, 1981.

[58] M. Durst and R. M. Dudley. Empirical processes, Vapnik-Chervonenkis classes and Poisson processes. *Probab. Math. Statist.*, 1:109–115, 1981.

[59] R. M. Dudley. Some recent results on empirical processes. In *Probability in Banach spaces, III (Medford, Mass., 1980)*, volume 860 of *Lecture Notes in Math.*, pages 107–123. Springer, Berlin, 1981.

[60] R. S. Wenocur and R. M. Dudley. Some special Vapnik-Chervonenkis classes. *Discrete Math.*, 33:313–318, 1981.

[61] R. M. Dudley. Vapnik-Červonenkis Donsker classes of functions. In *Aspects Statistiques et Aspects Physiques des Processus Gaussiens: Saint-Flour, 1980*, volume 307 of *Colloq. Internat. CNRS*, pages 251–269. CNRS, Paris, 1981.

[62] R. M. Dudley. Empirical and Poisson processes on classes of sets or functions too large for central limit theorems. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 61:355–368, 1982.

[63] R. M. Dudley, P. Humblet, and M. Goldring. Appendix to: Single photon transduction in *Limulus* photoreceptors and the Borsellino-Fuortes model, by M. A. Goldring and J. E. Lisman. *IEEE Trans. Systems, Man, Cybernet*, SMC-13:727–731, 1983.

[64] R. M. Dudley. Heritability under genotype-environment interaction and dependence. *Bull. Inst. Math. Statist.*, 12:152, 1983. Abstract.

[65] R. M. Dudley and W. Philipp. Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 62:509–552, 1983.

[66] R. M. Dudley. A course on empirical processes. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 1–142. Springer, Berlin, 1984.

[67] R. M. Dudley. Discussion of Special Invited Paper by E. Giné and J. Zinn "Some limit theorems for empirical processes". *Ann. Probab.*, 12:991–992, 1984.

[68] R. M. Dudley. An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. In A. Beck, R. Dudley, M. Hahn, J. Kuelbs, and M. Marcus, editors, *Probability in Banach spaces, V (Medford, Mass., 1984)*, volume 1153 of *Lecture Notes in Math.*, pages 141–178. Springer, Berlin, 1985.

[69] R. M. Dudley. Manifolds. In *Encyclopedia of Statistical Sciences*, volume 5, pages 198–201. Wiley, New York, 1985.

[70] R. M. Dudley. The structure of some Vapnik-Červonenkis classes. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983)*, Wadsworth Statist./Probab. Ser., pages 495–508, Belmont, CA, 1985. Wadsworth.

[71] R. Dudley and D. W. Stroock. Slepian's inequality and commuting semigroups. In *Séminaire de Probabilités, XXI*, volume 1247 of *Lecture Notes in Math.*, pages 574–578. Springer, Berlin, 1987.

[72] R. M. Dudley. Some inequalities for continued fractions. *Math. Comp.*, 49:585–593, 1987.

[73] R. M. Dudley. Some universal Donsker classes of functions. In Yu. V. Prohorov, V. A. Statulevičius, V. V. Sazonov, and B. Grigelionis, editors, *Probability theory and mathematical statistics, Vol. I (Vilnius, 1985)*, pages 433–438. VNU Sci. Press, Utrecht, 1987.

[74] R. M. Dudley. Universal Donsker classes and metric entropy. *Ann. Probab.*, 15:1306–1326, 1987.

[75] R. M. Dudley. Comment on D. Pollard, "Asymptotics via empirical processes". *Statistical Science*, 4:354, 1989.

[76] R. M. Dudley. *Real Analysis and Probability*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989.

[77] R. M. Dudley, S. L. Cook, J. Llopis, and N. P. Peng. A. N. Kolmogorov and statistics: a citation bibliography. *Ann. Statist.*, 18:1017–1031, 1990.

[78] R. M. Dudley. Interaction and dependence prevent estimation (Commentary). *Behavioral and Brain Sciences*, 13:132–133, 1990.

[79] R. M. Dudley. Nonlinear functionals of empirical measures and the bootstrap. In E. Eberlein, J. Kuelbs, and M. B. Marcus, editors, *Probability in Banach Spaces, 7 (Oberwolfach, 1988)*, volume 21 of *Progr. Probab.*, pages 63–82. Birkhäuser Boston, Boston, MA, 1990.

[80] R. M. Dudley. Nonmetric compact spaces and nonmeasurable processes. *Proc. Amer. Math. Soc.*, 108:1001–1005, 1990.

[81] R. M. Dudley. Program verification. *Notices Amer. Math. Soc.*, 37:123–124, 1990.

[82] R. M. Dudley. IQ and heredity (letter). *Science*, 252:191, 1991.

[83] A. J. Quiroz and R. M. Dudley. Some new tests for multivariate normality. *Probab. Theory Related Fields*, 87:521–546, 1991.

[84] R. M. Dudley, E. Giné, and J. Zinn. Uniform and universal Glivenko-Cantelli classes. *J. Theoret. Probab.*, 4:485–510, 1991.

[85] R. M. Dudley. Why are adoptees so similar in IQ? (Commentary.). *Behavioral and Brain Sciences*, 14:336, 1991.

[86] D. L. Smith and R. M. Dudley. Exponential bounds in Vapnik-Červonenkis classes of index 1. In R. M. Dudley, M. G. Hahn, and J. Kuelbs, editors, *Probability in Banach spaces, 8 (Brunswick, ME, 1991)*, volume 30 of *Progr. Probab.*, pages 451–465. Birkhäuser Boston, Boston, MA, 1992.

[87] R. M. Dudley. Fréchet differentiability, p-variation and uniform Donsker classes. *Ann. Probab.*, 20:1968–1982, 1992.

[88] R. M. Dudley. Nonlinear functionals of empirical measures. In R. M. Dudley, M. G. Hahn, and J. Kuelbs, editors, *Probability in Banach spaces, 8 (Brunswick, ME, 1991)*, volume 30 of *Progr. Probab.*, pages 403–410. Birkhäuser Boston, Boston, MA, 1992.

[89] J. M. González-Barrios and R. M. Dudley. Metric entropy conditions for an operator to be of trace class. *Proc. Amer. Math. Soc.*, 118:175–180, 1993.

[90] R. M. Dudley, S. R. Kulkarni, T. Richardson, and O. Zeitouni. A metric entropy bound is not sufficient for learnability. *IEEE Trans. Inform. Theory*, 40:883–885, 1994.

[91] R. M. Dudley and V. I. Koltchinskii. Envelope moment conditions and Donsker classes. *Teor. Ĭmovīr. Mat. Stat.*, 51:39–49, 1994.

[92] R. M. Dudley. Metric marginal problems for set-valued or non-measurable variables. *Probab. Theory Related Fields*, 100:175–189, 1994.

[93] R. M. Dudley and J. González-Barrios. On extensions of Mercer's theorem. In M. E. Caballero and L. G. Gorostiza, editors, *III Simposio de Probabilidad y Procesos Estocásticos*, volume 11 of *Aportaciones Matemáticas*, pages 91–97, 1994.

[94] R. M. Dudley. The order of the remainder in derivatives of composition and inverse operators for p-variation norms. *Ann. Statist.*, 22:1–20, 1994.

[95] R. M. Dudley. Empirical processes and p-variation. In D. Pollard, E. Torgersen, and G. L. Yang, editors, *Festschrift for Lucien Le Cam*, pages 219–233. Springer, New York, 1997.

[96] R. M. Dudley and D. Haughton. Information criteria for multiple data sets and restricted parameters. *Statist. Sinica*, 7:265–284, 1997.

[97] R. M. Dudley and R. Norvaiša. *An Introduction to p-Variation and Young Integrals, with Emphasis on Sample Functions of Stochastic Processes*, volume 1 of *MaPhySto Lecture Notes*. Aarhus, Denmark, 1998.

[98] R. M. Dudley and J. M. González-Barrios. Conditions for integral and other operators to be of trace class. *Bol. Soc. Mat. Mexicana*, 4:105–114, 1998.

[99] R. M. Dudley. Consistency of M-estimators and one-sided bracketing. In E. Eberlein, M. Hahn, and M. Talagrand, editors, *High dimensional probability (Oberwolfach, 1996)*, volume 43 of *Progr. Probab.*, pages 33–58. Birkhäuser, Basel, 1998.

[100] R. M. Dudley and R. Norvaiša. *Differentiability of Six Operators on Nonsmooth Functions and p-Variation*, volume 1703 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. With the collaboration of Jinghua Qian.

[101] R. M. Dudley. *Notes on Empirical Processes*, volume 4 of *MaPhySto Lecture Notes*. Aarhus, Denmark, 1999.

[102] R. M. Dudley. *Uniform Central Limit Theorems*, volume 63 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[103] V. Koltchinskii and R. M. Dudley. On spatial quantiles. In V. Korolyuk, N. Portenko, and H. Syta, editors, *Skorokhod's Ideas in Probability Theory*, pages 195–210. Inst. Math. Nat. Acad. Sci. Ukraine, Kyiv, 2000.

[104] Y.-C. Huang and R. M. Dudley. Speed of convergence of classical empirical processes in p-variation norm. *Ann. Probab.*, 29:1625–1636, 2001.

[105] R. M. Dudley and D. Haughton. Asymptotic normality with small relative errors of posterior probabilities of half-spaces. *Ann. Statist.*, 30:1311–1344, 2002.

[106] R. M. Dudley and D. M. Haughton. One-sided hypotheses in a multinomial model. In C. Huber-Carol, N. Balakrishnan, M. S. Nikulin, and M. Mesbah, editors, *Goodness-of-fit Tests and Model Validity (Paris, 2000)*, Stat. Ind. Technol., pages 387–399. Birkhäuser, Boston, MA, 2002.

[107] R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. 2d. Edition.

[108] R. M. Dudley. Statistical nearly universal Glivenko-Cantelli classes. In J. Hoffmann-Jørgensen, M. B. Marcus, and J. A. Wellner, editors, *High Dimensional Probability, III (Sandbjerg, 2002)*, volume 55 of *Progr. Probab.*, pages 295–312. Birkhäuser, Basel, 2003.

[109] R. M. Dudley. Some facts about functionals of location and scatter. In E. Giné, V. Koltchinskii, W. Li, and J. Zinn, editors, *High Dimensional Probability*, volume 51 of *IMS Lecture Notes Monogr. Ser.*, pages 207–219. Inst. Math. Statist., Beachwood, OH, 2006.

[110] R. M. Dudley, S. Sidenko, and Z. Wang. Differentiability of t-functionals of location and scatter. *Ann. Statist.*, 37:939–960, 2009.

[111] R. M. Dudley and R. Norvaiša. Concrete Functional Calculus, to appear. Research monograph in preparation, about 750 pp., under contract for publication.

PhD students of Richard M. Dudley

Chandrakant Deo, ‘Prediction Theory of Non-stationary Processes,’ University of California, Berkeley, 1965.

José Abreu Leon, ‘Smoothing, Filtering and Prediction of Generalized Stochastic Processes,’ Massachusetts Institute of Technology, 1970.

Lewis Pakula, ‘Covariances of Generalized Stochastic Processes,’ Massachusetts Institute of Technology, 1972.

Evarist Giné-Masdéu, ‘Invariant Tests for Uniformity on Compact Riemannian Manifolds Based on Sobolev Norms,’ Massachusetts Institute of Technology, 1973.

Ruben Klein, ‘Topics on Gaussian Sample Functions,’ Massachusetts Institute of Technology, 1974 .

Michael Zuker, ‘Speeds of Convergence of Random Probability Measures,’ Massachusetts Institute of Technology, 1974.

Marjorie Hahn, ‘Central Limit Theorems for $D[0,1]$-Valued Random Variables,’ Massachusetts Institute of Technology, 1975.

Donald Cohn, ‘Topics in Liftings and Stochastic Processes,’ Harvard University, 1975.

Eric Slud, ‘Inequalities for Binomial, Normal, and Hypergeometric Tail Probabilities,’ Massachusetts Institute of Technology, 1976.

Samuel Gutmann, ‘Non-Stationary Markov Transitions,’ Massachusetts Institute of Technology, 1977.

Steven Pincus, ‘Strong Laws of Large Numbers for Products of Random Matrices,’ Massachusetts Institute of Technology, 1980.

Mark Durst, ‘Donsker Classes, Vapnik-Chervonenkis Classes, and Chi-Squared Tests of Fit with Random Cells,’ Massachusetts Institute of Technology, 1980.

Zakhar Maymin, ‘Minimax Estimation on Subsets of Parameters,’ Massachusetts Institute of Technology, 1981.

Joseph Yukich, ‘Convergence in Empirical Probability Measures, Massachusetts Institute of Technology,’ 1982.

Kenneth Alexander, ‘Some Limit Theorems and Inequalities for Weighted and Non-Identically Distributed Empirical Processes,’ Massachusetts Institute of Technology, 1982.

Rae Michael Shortt, ‘Existence of Laws with Given Marginals and Specified Support,’ Massachusetts Institute of Technology, 1982.

David Marcus, ‘Non-Stable Laws with All Projections Stable and Relationships Between Donsker Classes and Sobolev Spaces,’ Massachusetts Institute of Technology, 1983.

Richard Dante DeBlassie, ‘Hitting Times of Brownian Motion’ , Massachusetts Institute of Technology, 1984.

Dominique Haughton, ‘On the choice of a model to fit data from an exponential family,’ Massachusetts Institute of Technology, 1984.

Joseph Fu, ‘Tubular Neighborhoods of Planar Sets’ (with J. Almgren), Massachusetts Institute of Technology, 1984.

Daphne Smith, ‘Vapnik-Červonenkis Classes and the Supremum Distribution of a Gaussian Process, Massachusetts Institute of Technology,’ 1985.

Adolfo Quiroz Salazar, ‘On Donsker Classes of Functions and their Application to Tests for Goodness of Fit,’ Massachusetts Institute of Technology, 1986.

Robert Holt, ‘Computation of Gamma Tail Probabilities,’ Massachusetts Institute of Technology, 1986.

Michael Schmidt, ‘Optimal Rates of Convergence for Nonparametric Regression Function Estimators,’ Massachusetts Institute of Technology, 1988.

José González-Barrios, ‘On Von Mises Functionals with Emphasis on Trace Class Kernels,’ Massachusetts Institute of Technology, 1990.

Evangelos Tabakis, ‘Asymptotic and Computational Problems in Single-Link Clustering,’ Massachusetts Institute of Technology, 1992 .

Yen-Chin Huang, ‘Empirical Distribution Function Statistics, Speed of Convergence, and p-Variation,’ Massachusetts Institute of Technology, 1994.

Jinghua Qian, ‘The p-Variation of Partial Sum Processes and the Empirical Process,’ Tufts University 1997.

Li He, ‘Modeling and Prediction of Sunspot Cycles,’ Massachusetts Institute of Technology, 2001.

Martynas Manstavičius, ‘The p-variation of Strong Markov Processes,’ Tufts University, 2003.

Xia Hua, ‘Testing regression models with residuals as data’, Massachusetts Institute of Technology, 2010

Part 1

Convergence in Law

Introduction

The papers in this chapter deal with important properties of weak convergence of probability measures on metric spaces. Most of them are motivated by, and applied to, the question of establishing convergence in law of empirical processes, a basic topic of statistics.

The problem considered in the first paper comes from the fact that, due to certain set-theoretic assumptions, a finite, countably additive measure defined on all Borel sets of a metric space is concentrated in a separable subspace whereas, on the other hand, almost all sample paths of empirical processes are not elements of a separable subset of a metric space. Dudley extended the notion of weak convergence to countably additive probability measures defined on σ-algebras of a metric space not necessarily related to its metric topology. Namely, the weak* convergence of measures β_n defined on σ-algebras $\mathcal{B}_n$ of subsets of a metric space S to a Borel measure β_0 on S means that

$$\lim_{n\to\infty} \int^* f\, d\beta_n = \lim_{n\to\infty} \int_* f\, d\beta_n = \int f\, d\beta_0 \tag{1}$$

for every bounded continuous real-valued function f on S, where $\int^*$ and $\int_*$ are upper and lower integrals, respectively. Results were obtained for this convergence in case each $\mathcal{B}_n$ includes the smallest σ-algebra $\mathcal{U}$ generated by all open balls of S. A probability measure on $\mathcal{U}$ does not need to be confined to a separable subspace of S. Dudley proved the weak* convergence of measures α_n, $n\geq 1$, on a suitable function space J with the uniform metric when α_n is the probability distribution of normalized empirical distribution functions induced by a sequence of independent identically distributed $\mathbf{R}^k$-valued random variables. The suggested solution of the problem corrected the main results of M. D. Donsker (Mem. Amer. Math. Soc., 1951, No. 6 and Ann. Math. Statist., 1952, **23**, 277-281) for real-valued random variables and generalized them to random variables with values in a Euclidean space.

The second paper continued the subject of the first one by giving a more general definition of weak* convergence. Let $\mathcal{U}$ be the σ-algebra on a metric space S generated by open balls as before. Let $M(S, \mathcal{U})$ be the set of all finite, countably additive, real-valued set functions on $\mathcal{U}$, and let $C(S, \mathcal{U})$ be the closed linear subspace of $\mathcal{U}$-measurable elements of the Banach space $C(S)$ of all bounded, continuous, real-valued functions on S with the supremum norm. Then any $\mu\in M(S, \mathcal{U})$ defines a bounded linear functional $f\to\int f\, d\mu$ on $C(S, \mathcal{U})$ and we have the weak* topology of pointwise convergence on $C(S, \mathcal{U})$. Let (β_n) be a sequence of nonnegative elements of $M(S, \mathcal{U})$ and let β_0 be a nonnegative element of $M(S, \mathcal{U})$ concentrated in a separable subspace of S. Under the hypothesis that the metric space S is complete Dudley proved that $\beta_n\to\beta_0$ for the weak* topology if and only if (1) holds for every f in $C(S)$.

The third paper in this chapter compares various metrics on the set of all probability measures of a metric space, and relates the weak* convergence of probability measures with almost surely convergent realizations. Let $\mathcal{P}(S)$ be the set of all Borel probability measures on a separable metric space S, endowed with the weak* topology. For S complete and $\mu, \nu \in \mathcal{P}(S)$, V. Strassen (Ann. Math. Statist., 1965, **36**, 423-439) proved that the Prokhorov distance $\rho(\mu, \nu)$ is the minimum distance in probability between random variables distributed according to μ and ν. Dudley generalized this result without assuming completeness of S and by using the finite combinatorial "marriage lemma". Useful bounds for the Prokhorov and the bounded Lipschitz metrics are given in this paper. Also Dudley proved that if $\beta_n\to\beta_0$ in $\mathcal{P}(S)$ then there exist random variables X_n with distributions β_n such that $X_n\to X_0$ almost surely. This was proved by A. V. Skorokhod (Theor. Prob. Appl., 1956, **1**, 261-290) to hold if S is complete. Later M. J. Wichura (Ann. Math. Statist., 1970, **41**, 284-291) and P. J. Fernandez (Bol. Soc. Brasil. Math., 1974, **5**, 51-61) proved another extension of Skorokhod's result when the metric space S may be non-separable. In this case probability measures β_n, $n\geq 1$, are defined on the σ-algebra $\mathcal{U}$ generated by all open balls of the metric space S, β_0 is a Borel probability as before, and β_n converges to β_0 in the sense Dudley defined.

Further improvement on almost surely convergent realizations of probability measures was made in the last paper of this chapter. Let $(\Omega_n, \mathcal{A}_n, P_n)$ be probability spaces for $n=0,1,2,\ldots$, and let X_n a

E. Giné et al. (eds.), *Selected Works of R.M. Dudley*, Selected Works in Probability and Statistics,
DOI 10.1007/978-1-4419-5821-1_1, © Springer Science+Business Media, LLC 2010

function from Ω_n into S, where X_0 takes values in some separable subset of S and is measurable for the Borel sets on its range. Following J. Hoffmann-Jørgensen (Various Publication Series no. 39, Matematisk Institut, Aarhus Universitet, 1991) one says that X_n converges to X_0 in law if

$$\lim_{n\to\infty} \int^* f(X_n)\, dP_n = \int f(X_0)\, dP_0$$

for every bounded continuous real-valued function f on S, where the upper integral and integral are taken over Ω_n, not S as in (1), so that the laws of X_n for $n \geq 1$ need not be defined on any particular σ-algebra of the metric space S. Dudley proved that X_n converges in law to X_0 if and only if one can redefine each X_n on a new probability space in such a way that the new sequence of S-valued random variables converges almost surely (in fact, almost uniformly). He uses this theorem in particular to show that the empirical process based on a sample from P and indexed by a class of functions converges in law in the sense of Hoffmann-Jørgensen to the corresponding 'P-Brownian bridge' if and only if the class of functions is functional P-Donsker, as previously defined by Dudley (Lect. Notes in Math., 1984, **1097**, 1-142) and by Dudley and Philipp (Z. Wahrsch. verw. Geb., 1983, **62**, 509-552).

Reprinted from Illinois Journal of Mathematics
Vol. 10, No. 1, March 1966
Printed in U.S.A.

WEAK CONVERGENCE OF PROBABILITIES ON NONSEPARABLE METRIC SPACES AND EMPIRICAL MEASURES ON EUCLIDEAN SPACES

BY
R. M. Dudley[1]

1. Introduction

It is known that under certain mild set-theoretic assumptions, a finite, countably additive measure defined on all Borel sets of a metric space is concentrated in a separable subspace (Marczewski and Sikorski [8]). However, there are interesting probability measures on metric spaces not concentrated in separable subspaces. In this paper, we consider countably additive probability measures on the smallest σ-field containing the open balls of a metric space. This σ-field is the Borel field for a separable space, but is smaller in general. A probability measure on it need not be confined to a separable subspace.

A sequence of such measures will be said to converge weak* to a Borel measure if the upper and lower integrals of each bounded continuous real function converge. Some abstract results on this convergence, similar to those in Prokhorov [9] for separable metric spaces, will be given in §2.

The rest of the paper deals with "empirical measures" on Euclidean spaces, whose study motivated the abstract results and provides an application of them. Two of the main results of Donsker [3], [4] for measures on the real line will be generalized to arbitrary Euclidean spaces. At the same time, his results are corrected by replacing some integrals, which may not be defined, by upper and lower integrals.

I discovered after writing most of the rest of this paper that a generalization of Donsker's work to multidimensional spaces was proved several years ago by L. LeCam, who is now revising a paper embodying his results for the Illinois Journal of Mathematics. I shall try to explain what seem to be the main differences between our approaches.

While my abstract results in §2 are for metric spaces and guided by those in Prokhorov [9] for the separable case, LeCam uses a more elaborate abstract apparatus involving the second dual spaces of topological linear spaces and nonmetric topologies; the place of upper and lower integrals is taken by integrals with respect to finitely additive extensions of a measure.

With regard to the more concrete equicontinuity properties of empirical distribution functions, my approach in §4 below uses a sort of Markov property for the random empirical measure μ_n, namely given $\mu_n(E)$ for a set E, the

Received September 28, 1964.

[1] Partially supported by a National Science Foundation Grant and by the Office of Naval Research.

E. Giné et al. (eds.), *Selected Works of R.M. Dudley*, Selected Works in Probability and Statistics, 5
DOI 10.1007/978-1-4419-5821-1_2, © Springer Science+Business Media, LLC 2010

values of μ_n on subsets of E are independent of its values on sets disjoint from E. LeCam instead lets n be a random variable $n(\ )$ with a Poisson distribution and obtains a random measure with independent values on disjoint sets.

Acknowledgments. I am grateful both to Prof. LeCam for providing the above information and to V. Strassen for earlier conversations about the Skorokhod topology, etc.

2. Weak*-convergence in nonseparable metric spaces

Suppose β is a measure on a σ-field $\mathcal{S}_1$ in a space S and let F be any real-valued function on S. Then the *lower integral*

$$\int_* F(s)\, d\beta(s)$$

is defined as the supremum of all integrals

$$\int f(s)\, d\beta(s)$$

where $f \leq F$ on S, f is $\mathcal{S}_1$-measurable, and the integral of f is defined. Similarly, the upper integral

$$\int^* F\, d\beta$$

is defined as the infimum of $\int f\, d\beta$ for $f \geq F$ on S and $\int f\, d\beta$ defined. Clearly

$$\int_* F\, d\beta \leq \int^* F\, d\beta$$

for any F. If χ_A is the indicator function of a set A, let

$$\beta_*(A) = \int_* \chi_A\, d\beta, \qquad \beta^*(A) = \int^* \chi_A\, d\beta;$$

β_* and β^* are clearly the usual inner and outer measures for β.

A *Baire* set in a topological space is a member of the smallest σ-field with respect to which all continuous functions are measurable. In a metric space, the Baire sets are precisely the Borel sets. Now if β_n are measures on a topological space S (not necessarily defined on all Baire sets), and β is a Baire measure on S, we say

$$\beta_n \to \beta \quad (\text{weak}^*)$$

if for every bounded continuous real function F on S,

$$\lim_{n\to\infty} \int^* F\, d\beta_n = \lim_{n\to\infty} \int_* F\, d\beta_n = \int F\, d\beta.$$

Let (S, ρ) be a metric space, and let $\mathcal{S}$ be the σ-field generated by the balls

$$[y \in S : \rho(x, y) < \varepsilon]$$

for any $x \in S$ and $\varepsilon > 0$. Let $\mathcal{B}$ be the Borel σ-field generated by the open sets of S. In the cases of interest here, S will be non-separable and $\mathcal{B}$ strictly larger than $\mathcal{S}$.

Let $(\mathcal{C}, \| \ \|)$ be the Banach space of bounded real continuous functions on S with supremum norm. For any subset A of S and $\varepsilon > 0$ let

$$A^{\varepsilon} = [x \in S : \rho(x, y) < \varepsilon \quad \text{for some} \quad y \in A].$$

We need the following well-known fact:

LEMMA 1. *If F is a continuous real-valued function on a metric space (S, ρ), K is a compact subset of S, and $\varepsilon > 0$, then there is a $\delta > 0$ such that if $x \in K$, $y \in S$, and $\rho(x, y) < \delta$ then*

$$| F(x) - F(y)| < \varepsilon.$$

Proof. If the conclusion is false there are $x_n \in K$ and $y_n \in S$, $n = 1, 2, \cdots$, with $\rho(x_n, y_n) < 1/n$ and

$$| F(x_n) - F(y_n)| \geq \varepsilon.$$

A subsequence of the x_n converges to an $x \in K$ at which F is not continuous, a contradiction which completes the proof.

We call a set $\mathcal{K}$ of measures on S weak*-*precompact* if any sequence $[\beta_n]$ of *distinct* elements of $\mathcal{K}$ has a subsequence which is weak*-convergent (to a Borel measure on S).

THEOREM 1. *If (S, ρ) is a metric space and $\mathcal{K}$ is a set of probability measures each defined at least on $\mathcal{S}$ in S, then $\mathcal{K}$ is weak*-precompact if for every $\varepsilon > 0$ there is a compact set $K \subset S$ such that for every $\delta > 0$,*

$$\beta(K^{\delta}) \geq 1 - \varepsilon$$

for all but finitely many $\beta \in \mathcal{K}$.

Proof. Note that K^{δ} is a countable union of open balls and hence is in $\mathcal{S}$. For each positive integer N, let K_N be a compact set in S such that for any $\delta > 0$,

$$\beta(K_N^{\delta}) \geq 1 - 1/N$$

for all but finitely many $\beta \in \mathcal{K}$. Let $\{F_n\}_{n=1}^{\infty}$ be a countable set of continuous functions on S with $\| F_n \| \leq 1$ for all n, uniformly dense on K_N for each N in the continuous functions F with $\| F \| \leq 1$ (such F_n exist since $\mathcal{C}(K_N)$ is a separable Banach space for each N and we can use the Tietze extension theorem).

Given N and n, let $\delta > 0$ be such that $\rho(x, y) < \delta$ and $x \in K_N$ imply

$$| F_n(x) - F_n(y)| < 1/N.$$

Let $x_1, \cdots, x_r$ be points of K_N such that for each $x \in K_N$,

$$\rho(x_j, x) < \delta$$

for some j. For $j = 1, \cdots, r$ let A_j be the set of all $x \in S$ such that $\rho(x_i, x) \geq \delta$ for $i < j$ and $\rho(x_j, x) < \delta$. Then the sets A_j are disjoint and belong to $\mathcal{S}$; if A is their union,

$$K_N \subset A \subset K_N^\delta .$$

Since A is open, $K_N^\gamma \subset A$ for some $\gamma > 0$, so that $\beta(A) > 1 - 1/N$ for all but finitely many $\beta \in \mathcal{K}$. Let $\varepsilon = 1/N$,

$$\begin{aligned} G(x) &= F_n(x_j) - \varepsilon, & x \in A_j, j = 1, \cdots, r \\ &= -1, & x \notin A \\ H(x) &= F_n(x_j) + \varepsilon, & x \in A_j, j = 1, \cdots, r \\ &= 1, & x \notin A. \end{aligned}$$

Then $G \leq F_n \leq H$, G and H are $\mathcal{S}$-measurable, and

$$\int (H - G)\, d\beta \leq 2\varepsilon + 2\varepsilon = 4\varepsilon$$

for any $\beta \in \mathcal{K}$ with $\beta(A) > 1 - \varepsilon$.

If $\{\beta_m\}$ is any sequence of distinct elements of $\mathcal{K}$, we can find a subsequence $\{\beta_{m_r}\}$ such that

$$\int_* F_n \, d\beta_{m_r}$$

is convergent for a given n, so that

$$\limsup \int^* F_n \, d\beta_{m_r} - \lim \int_* F_n \, d\beta_{m_r} \leq 4\varepsilon.$$

Taking further subsequences and diagonalizing, we can assume this holds for all n. Letting ε tend to zero through some sequence and diagonalizing again, we get a subsequence $\{\gamma_q\}$ of $\{\beta_m\}$ such that

$$\liminf_{q \to \infty} \int_* F_n \, d\gamma_q = \limsup_{q \to \infty} \int^* F_n \, d\gamma_q$$

for all n, so that lim inf and lim sup can be replaced by lim.

Now let F be a bounded continuous function on S with $\| F \| \leq 1$. Given N, choose n so that

$$|(F - F_n)(x)| < 1/N$$

for all $x \in K_N$; then this will hold for all $x \in K_N^\delta$ for some $\delta > 0$. Then except for finitely many $\beta \in \mathcal{K}$,

$$\int_* F \, d\beta \geq \int_* F_n \, d\beta - 3/N,$$

$$\int^* F \, d\beta \leq \int^* F_n \, d\beta + 3/N.$$