Lecture Notes in Mathematics 1953

Khadiga A. Arwini · Christopher T.J. Dodson

Information Geometry

Near Randomness and Near Independence

Springer

Authors

Khadiga A. Arwini
Al-Fateh University
Faculty of Sciences
Mathematics Department
Box 13496
Tripoli, Libya
arwini2001@yahoo.com

Christopher T.J. Dodson
University of Manchester
School of Mathematics
Manchester M13 9PL, United Kingdom
ctdodson@manchester.ac.uk

ISBN: 978-3-540-69391-8 e-ISBN: 978-3-540-69393-2
DOI: 10.1007/978-3-540-69393-2

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2008930087

Mathematics Subject Classification (2000): 53B50, 60D05, 62B10, 62P35, 74E35, 92D20

Cover design: SPi Publishing Services

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The main motivation for this book lies in the breadth of applications in which a statistical model is used to represent small departures from, for example, a Poisson process. Our approach uses information geometry to provide a common context but we need only rather elementary material from differential geometry, information theory and mathematical statistics. Introductory sections serve together to help those interested from the applications side in making use of our methods and results. We have available *Mathematica* notebooks to perform many of the computations for those who wish to pursue their own calculations or developments.

Some 44 years ago, the second author first encountered, at about the same time, differential geometry via relativity from Weyl's book [209] during undergraduate studies and information theory from Tribus [200, 201] via spatial statistical processes while working on research projects at Wiggins Teape Research and Development Ltd—cf. the Foreword in [196] and [170, 47, 58]. Having started work there as a student laboratory assistant in 1959, this research environment engendered a recognition of the importance of international collaboration, and a lifelong research interest in randomness and near-Poisson statistical geometric processes, persisting at various rates through a career mainly involved with global differential geometry. From correspondence in the 1960s with Gabriel Kron [4, 124, 125] on his Diakoptics, and with Kazuo Kondo who influenced the post-war Japanese schools of differential geometry and supervised Shun-ichi Amari's doctorate [6], it was clear that both had a much wider remit than traditionally pursued elsewhere. Indeed, on moving to Lancaster University in 1969, receipt of the latest *RAAG Memoirs Volume 4 1968* [121] provided one of Amari's early articles on information geometry [7], which subsequently led to his greatly influential 1985 Lecture Note volume [8] and our 1987 *Geometrization of Statistical Theory Workshop* at Lancaster University [10, 59].

Reported in this monograph is a body of results, and computer-algebraic methods that seem to have quite general applicability to statistical models admitting representation through parametric families of probability density

functions. Some illustrations are given from a variety of contexts for geometric characterization of statistical states near to the three important standard basic reference states: (Poisson) randomness, uniformity, independence. The individual applications are somewhat heuristic models from various fields and we incline more to terminology and notation from the applications rather than from formal statistics. However, a common thread is a geometrical representation for statistical perturbations of the basic standard states, and hence results gain qualitative stability. Moreover, the geometry is controlled by a metric structure that owes its heritage through maximum likelihood to information theory so the quantitative features—lengths of curves, geodesics, scalar curvatures etc.—have some respectable authority. We see in the applications simple models for galactic void distributions and galaxy clustering, amino acid clustering along protein chains, cryptographic protection, stochastic fibre networks, coupled geometric features in hydrology and quantum chaotic behaviour. An ambition since the publication by Richard Dawkins of *The Selfish Gene* [51] has been to provide a suitable differential geometric framework for dynamics of natural evolutionary processes, but it remains elusive. On the other hand, in application to the statistics of amino acid spacing sequences along protein chains, we describe in Chapter 7 a stable statistical qualitative property that may have evolutionary significance. Namely, to widely varying extents, all twenty amino acids exhibit greater clustering than expected from Poisson processes. Chapter 11 considers eigenvalue spacings of infinite random matrices and near-Poisson quantum chaotic processes.

The second author has benefited from collaboration (cf. [34]) with the group headed by Andrew Doig of the Manchester Interdisciplinary Biocentre, the University of Manchester, and has had long-standing collaborations with groups headed by Bill Sampson of the School of Materials, the University of Manchester (cf.eg. [73]) and Jacob Scharcanski of the Instituto de Informatica, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil (cf.eg. [76]) on stochastic modelling. We are pleased therefore to have co-authored with these colleagues three chapters: titled respectively, Amino Acid Clustering, Stochastic Fibre Networks, Stochastic Porous Media and Hydrology.

The original draft of the present monograph was prepared as notes for short Workshops given by the second author at Centro de Investigaciones de Matematica (CIMAT), Guanajuato, Mexico in May 2004 and also in the Departamento de Xeometra e Topoloxa, Facultade de Matemáticas, Universidade de Santiago de Compostela, Spain in February 2005.

The authors have benefited at different times from discussions with many people but we mention in particular Shun-ichi Amari, Peter Jupp, Patrick Laycock, Hiroshi Matsuzoe, T. Subba Rao and anonymous referees. However, any overstatements in this monograph will indicate that good advice may have been missed or ignored, but actual errors are due to the authors alone.

Khadiga Arwini, Department of Mathematics, Al-Fateh University, Libya
Kit Dodson, School of Mathematics, the University of Manchester, England

Contents