

Preface

Information Geometry and Algebraic Statistics are brought together in this volume to suggest that the interaction between them is possible and auspicious.

To meet this aim, we couple expository material with more advanced research topics sometimes within the same chapter, cross-reference the various chapters, include many examples both in the printed volume and in the on-line supplement, held at the Cambridge University press web site [...web address to be input by Cambridge when available...] The on-line part includes proofs instructive but long or repetitive, computer codes and detailed development of special cases.

Chapter 1 gives a brief introduction to both Algebraic Statistics and Information Geometry based on the simplest possible examples and on selected topics that, to the editors, seem most promising for the interlacing between them. Then, the volume splits naturally in two lines. Part I, on contingency tables, and Part II, on designed experiments, are authored by researchers active mainly within Algebraic Statistics, while Part III includes papers on both classical and quantum Information Geometry. The volume rejoins in Part IV which consists of only one chapter by Giovanni Pistone, to whom the volume is dedicated, and provides examples on the interplay between Information Geometry and Algebraic Statistics.

The editors imagine various entry points into the volume according to the reader's own interests. They are indicated with squared boxes in Figure 0.1. Maximum likelihood estimation in models with hidden variables is revisited in an algebraic framework in Chapter 2 (S. E. Fienberg et al.) which is supported by a substantial on-line section, including Chapter 22 (Y. Zhou) where the role of secant varieties for graphical models is detailed. Chapter 3 (A. Slavkovich and S. E. Fienberg) gives old and new geometric characterizations of the joint distribution on $I \times J$ contingency tables and can be used to gain familiarity with algebraic geometric jargon and ideas common in Algebraic Statistics. The next two chapters present fast algorithms for the computation of Markov bases in model selection (Chapter 4 by A. Krampe and S. Kuhnt) and under strictly positive margins (Chapter 5 by Y. Chen et al.). While Chapter 6 (E. Carlini and F. Rapallo) defines a class of algebraic statistical models for category distinguishability in rater agreement problems. The algebraic notion of index of complexity of maximum likelihood equations is used in Chapter 7 (S. Hoşten and S. Sullivant) for bivariate data missing at random. This part of the volume ends with Chapter 8 by S. E. Fienberg and A. Dobra.

Part II considers the two technologies of Algebraic Statistics most employed in design and analysis of experiments. Chapter 12 (R. Fontana and M. P. Rogantin) use the game of sudoku to review polynomial indicator functions and links to Part I via the notion of Markov bases. This link is developed for a special case in Chapter 13 (S. Aoki and A. Takemura). This chapter should appeal to a reader acquainted with the classical theory of experimental design. Chapters 9, 10 and 11 develop in different settings the ideas and techniques outlined in the first part of Chapter 1. Chapter 9 (H. P. Wynn and H. Maruri-Aguilar) argues that algebraic sets can be used as repositories of experimental designs; Chapter 10 (R. Laubenbacher and B. Stigler) presents an application to the identification of biochemical networks from experimental data; and Chapter 11 (E. Riccomagno and R. Notari) considers designs with replicated points.

The Information Geometry part of the volume starts with Chapter 14 (R. F. Streater) which provides a gentle, short although comprehensive, introduction to Information Geometry and its link to the theory of estimation according to Fisher. It keeps as far as possible the analogy between the classical and the quantum case. It extends to the purely quantum case in Chapter 15 (R. F. Streater) which, together with Chapter 16 (A. Jenčová), provides an extension to the quantum case of the statistical manifolds modelled on an Orlicz space. Also Chapter ?? by F. Hansen deals with quantum Information Geometry. A construction of a statistical manifold modelled on a Reproducing Kernel Hilbert Space is presented in Chapter 18 (K. Fukumizu) where the application to the theory of estimation is based on a suitable class of likelihood functions defined point-wise. Chapter 19 (D. Imparato and B. Trivellato) extends the standard non-parametric exponential model by considering its limit developing ideas in Chapter 21. An application of classical information geometry for text analysis is developed by G. Lebanon in Chapter 17.

Chapter 1 includes a glossary of terms from Algebraic Geometry which are recurrent in the volume.

The editors thank the authors for providing interesting papers, the many referees who helped with the peer-reviewing, our publisher CUP and the ever patient and capable Diana Gillooly, Mauro Gasparini for suggesting the subtitle of the volume. This is also the title of a conference held in Sestri Levante on May 19-20, 2008, where some chapters in this volume were first presented. Some chapters were also presented at the opening workshop of the 2008-09 SAMSI Program on Algebraic Methods in Systems Biology and Statistics, 14-17 September, 2008.

This volume is dedicated to Giovanni Pistone on the occasion of his sixty-fifth birthday. We are grateful for his discreet and constant support.

Roma, Genova and London, 10 October 2008
Paolo, Eva, Maria-Piera and Henry

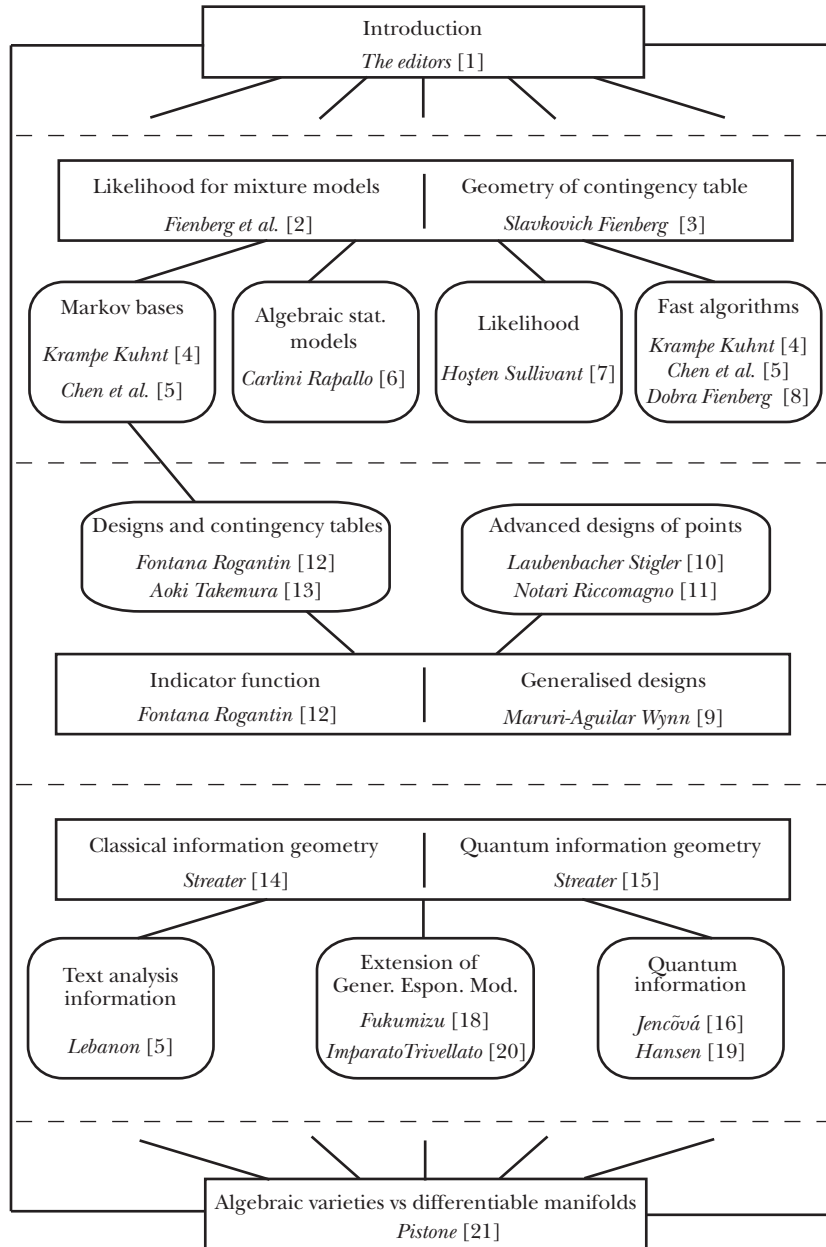


Fig. 0.1. Layout of the volume.