Preface

Algebraic statistics is a rapidly developing field, where ideas from statistics and algebra meet and stimulate new research directions. Statistics has been relying on classical asymptotic theory as a basis for statistical inferences. This classical basis is still very useful. However, when the validity of asymptotic theory is in doubt, for example, when the sample size is small, statisticians rely more and more on various computational methods. Similarly, algebra has long been considered as the purest field of mathematics, far apart from practical computations. However, due mainly to the development of Gröbner basis technology, algebra is now becoming a field where computations for practical applications are feasible. It is an interesting trend, because historically algebra was invented to speed up various calculations.

These two trends meet in the field of algebraic statistics. Algebraic algorithms are now very useful and essential for some practical statistical computations such as Markov chain Monte Carlo tests for discrete exponential families, which is the main topic of this book. On the other hand algebraic structures and computational needs of statistical models provide new challenging problems to algebraists. Some algebraic structures are naturally motivated from statistical modeling, but not necessarily from pure mathematical considerations.

Algebraic statistics has two origins. One origin is the work by Pistone and Wynn in 1996 on the use of Gröbner bases for studying confounding relations in factorial designs of experiments. Another origin is the work by Diaconis and Sturmfels in 1998 on the use of Gröbner bases for constructing a connected Markov chain for performing conditional tests of a discrete exponential family. These two works opened up the whole new field of algebraic statistics. In this book we take up the second topic. We give a detailed treatment of results following the seminal work of Diaconis and Sturmfels. We also briefly consider the first topic in Chap. 15 of this book.

As a general reference to the first origin of algebraic statistics we mention Pistone et al. [118]. For the second origin we mention Drton et al. [55], Pachter and Sturmfels [116], and our review paper [15]. For Japanese people the following two books are very useful: Hibi [86], and JST CREST Hibi team [93]. The Markov bases

database (`http://markov-bases.de/`) provides very useful online material for studying Markov bases.

Algebraic statistics gave us some exciting opportunities for research and collaboration. In particular we enjoyed working with Takayuki Hibi and Hidefumi Ohsugi, who are the leading researchers on Gröbner bases in Japan. Since 2008 Takayuki Hibi has a project, "Harmony of Gröbner Bases and the Modern Industrial Society," in the mathematics program of the Japan Science and Technology Agency. Algebraic statistics offers a rare ground where algebraists and statisticians can talk about the same problems, albeit often with different terminologies. This book is intended for statisticians with minimal backgrounds in algebra. As we ourselves learned algebraic notions through working on statistical problems, we hope that this book with many practical statistical problems is useful for statisticians to start working on algebraic statistics.

In preparing this book we very much benefited from comments of Takayuki Hibi, Hidehiko Kamiya, Kei Kobayashi, Satoshi Kuriki, Mitsunori Ogawa, Hidefumi Ohsugi, Toshio Sakata, Tomonari Sei, Kentaro Tanaka, and Ruriko Yoshida.

Finally we acknowledge great editorial help from John Kimmel.

Kagoshima, Japan   Satoshi Aoki
Niigata, Japan   Hisayuki Hara
Tokyo, Japan   Akimichi Takemura