Monographs on Statistics and Applied Probability 89

Algebraic Statistics

Computational Commutative Algebra in Statistics

Giovanni Pistone
Eva Riccomagno
and Henry P. Wynn

CHAPMAN & HALL/CRC

Algebraic Statistics

Computational Commutative Algebra in Statistics

GIOVANNI PISTONE
EVA RICCOMAGNO
HENRY P. WYNN

CHAPMAN & HALL/CRC

Boca Raton London New York Washington, D.C.

MONOGRAPHS ON STATISTICS AND APPLIED PROBABILITY

General Editors

D.R. Cox, V. Isham, N. Keiding, T. Louis, N. Reid, R. Tibshirani, and H. Tong

1 Stochastic Population Models in Ecology and Epidemiology *M.S. Barlett* (1960)
2 Queues *D.R. Cox and W.L. Smith* (1961)
3 Monte Carlo Methods *J.M. Hammersley and D.C. Handscomb* (1964)
4 The Statistical Analysis of Series of Events *D.R. Cox and P.A.W. Lewis* (1966)
5 Population Genetics *W.J. Ewens* (1969)
6 Probability, Statistics and Time *M.S. Barlett* (1975)
7 Statistical Inference *S.D. Silvey* (1975)
8 The Analysis of Contingency Tables *B.S. Everitt* (1977)
9 Multivariate Analysis in Behavioural Research *A.E. Maxwell* (1977)
10 Stochastic Abundance Models *S. Engen* (1978)
11 Some Basic Theory for Statistical Inference *E.J.G. Pitman* (1979)
12 Point Processes *D.R. Cox and V. Isham* (1980)
13 Identification of Outliers *D.M. Hawkins* (1980)
14 Optimal Design *S.D. Silvey* (1980)
15 Finite Mixture Distributions *B.S. Everitt and D.J. Hand* (1981)
16 Classification *A.D. Gordon* (1981)
17 Distribution-Free Statistical Methods, 2nd edition *J.S. Maritz* (1995)
18 Residuals and Influence in Regression *R.D. Cook and S. Weisberg* (1982)
19 Applications of Queueing Theory, 2nd edition *G.F. Newell* (1982)
20 Risk Theory, 3rd edition *R.E. Beard, T. Pentikäinen and E. Pesonen* (1984)
21 Analysis of Survival Data *D.R. Cox and D. Oakes* (1984)
22 An Introduction to Latent Variable Models *B.S. Everitt* (1984)
23 Bandit Problems *D.A. Berry and B. Fristedt* (1985)
24 Stochastic Modelling and Control *M.H.A. Davis and R. Vinter* (1985)
25 The Statistical Analysis of Composition Data *J. Aitchison* (1986)
26 Density Estimation for Statistics and Data Analysis *B.W. Silverman* (1986)
27 Regression Analysis with Applications *G.B. Wetherill* (1986)
28 Sequential Methods in Statistics, 3rd edition
G.B. Wetherill and K.D. Glazebrook (1986)
29 Tensor Methods in Statistics *P. McCullagh* (1987)
30 Transformation and Weighting in Regression
R.J. Carroll and D. Ruppert (1988)
31 Asymptotic Techniques for Use in Statistics
O.E. Bandorff-Nielsen and D.R. Cox (1989)
32 Analysis of Binary Data, 2nd edition *D.R. Cox and E.J. Snell* (1989)

33 Analysis of Infectious Disease Data *N.G. Becker* (1989)
34 Design and Analysis of Cross-Over Trials *B. Jones and M.G. Kenward* (1989)
35 Empirical Bayes Methods, 2nd edition *J.S. Maritz and T. Lwin* (1989)
36 Symmetric Multivariate and Related Distributions
K.T. Fang, S. Kotz and K.W. Ng (1990)
37 Generalized Linear Models, 2nd edition *P. McCullagh and J.A. Nelder* (1989)
38 Cyclic and Computer Generated Designs, 2nd edition
J.A. John and E.R. Williams (1995)
39 Analog Estimation Methods in Econometrics *C.F. Manski* (1988)
40 Subset Selection in Regression *A.J. Miller* (1990)
41 Analysis of Repeated Measures *M.J. Crowder and D.J. Hand* (1990)
42 Statistical Reasoning with Imprecise Probabilities *P. Walley* (1991)
43 Generalized Additive Models *T.J. Hastie and R.J. Tibshirani* (1990)
44 Inspection Errors for Attributes in Quality Control
N.L. Johnson, S. Kotz and X, Wu (1991)
45 The Analysis of Contingency Tables, 2nd edition *B.S. Everitt* (1992)
46 The Analysis of Quantal Response Data *B.J.T. Morgan* (1992)
47 Longitudinal Data with Serial Correlation—A state-space approach
R.H. Jones (1993)
48 Differential Geometry and Statistics *M.K. Murray and J.W. Rice* (1993)
49 Markov Models and Optimization *M.H.A. Davis* (1993)
50 Networks and Chaos—Statistical and probabilistic aspects
O.E. Barndorff-Nielsen, J.L. Jensen and W.S. Kendall (1993)
51 Number-Theoretic Methods in Statistics *K.-T. Fang and Y. Wang* (1994)
52 Inference and Asymptotics *O.E. Barndorff-Nielsen and D.R. Cox* (1994)
53 Practical Risk Theory for Actuaries
C.D. Daykin, T. Pentikäinen and M. Pesonen (1994)
54 Biplots *J.C. Gower and D.J. Hand* (1996)
55 Predictive Inference—An introduction *S. Geisser* (1993)
56 Model-Free Curve Estimation *M.E. Tarter and M.D. Lock* (1993)
57 An Introduction to the Bootstrap *B. Efron and R.J. Tibshirani* (1993)
58 Nonparametric Regression and Generalized Linear Models
P.J. Green and B.W. Silverman (1994)
59 Multidimensional Scaling *T.F. Cox and M.A.A. Cox* (1994)
60 Kernel Smoothing *M.P. Wand and M.C. Jones* (1995)
61 Statistics for Long Memory Processes *J. Beran* (1995)
62 Nonlinear Models for Repeated Measurement Data
M. Davidian and D.M. Giltinan (1995)
63 Measurement Error in Nonlinear Models
R.J. Carroll, D. Rupert and L.A. Stefanski (1995)
64 Analyzing and Modeling Rank Data *J.J. Marden* (1995)
65 Time Series Models—In econometrics, finance and other fields
D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (1996)

66 Local Polynomial Modeling and its Applications *J. Fan and I. Gijbels* (1996)

67 Multivariate Dependencies—Models, analysis and interpretation
D.R. Cox and N. Wermuth (1996)

68 Statistical Inference—Based on the likelihood *A. Azzalini* (1996)

69 Bayes and Empirical Bayes Methods for Data Analysis
B.P. Carlin and T.A Louis (1996)

70 Hidden Markov and Other Models for Discrete-Valued Time Series
I.L. Macdonald and W. Zucchini (1997)

71 Statistical Evidence—A likelihood paradigm *R. Royall* (1997)

72 Analysis of Incomplete Multivariate Data *J.L. Schafer* (1997)

73 Multivariate Models and Dependence Concepts *H. Joe* (1997)

74 Theory of Sample Surveys *M.E. Thompson* (1997)

75 Retrial Queues *G. Falin and J.G.C. Templeton* (1997)

76 Theory of Dispersion Models *B. Jørgensen* (1997)

77 Mixed Poisson Processes *J. Grandell* (1997)

78 Variance Components Estimation—Mixed models, methodologies and applications
P.S.R.S. Rao (1997)

79 Bayesian Methods for Finite Population Sampling
G. Meeden and M. Ghosh (1997)

80 Stochastic Geometry—Likelihood and computation
O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout (1998)

81 Computer-Assisted Analysis of Mixtures and Applications—
Meta-analysis, disease mapping and others *D. Böhning* (1999)

82 Classification, 2nd edition *A.D. Gordon* (1999)

83 Semimartingales and their Statistical Inference *B.L.S. Prakasa Rao* (1999)

84 Statistical Aspects of BSE and vCJD—Models for Epidemics
C.A. Donnelly and N.M. Ferguson (1999)

85 Set-Indexed Martingales *G. Ivanoff and E. Merzbach* (2000)

86 The Theory of the Design of Experiments *D.R. Cox and N. Reid* (2000)

87 Complex Stochastic Systems
O.E. Barndorff-Nielsen, D.R. Cox and C. Klüppelberg (2001)

88 Multidimensional Scaling, 2nd edition *T.F. Cox and M.A.A. Cox* (2001)

89 Algebraic Statistics—Computational Commutative Algebra in Statistics,
G. Pistone, E. Riccomagno and H.P. Wynn (2001)

Algebraic Statistics

Computational Commutative Algebra in Statistics

GIOVANNI PISTONE
EVA RICCOMAGNO
HENRY P. WYNN

CHAPMAN & HALL/CRC

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Pistone, Giovanni.
Algebraic statistics / Giovanni Pistone, Eva Riccomagno, Henry P. Wynn.
p. cm.-- (Monographs on statistics and applied probability ; 89)
Includes bibliographical references and index.
ISBN 1-58488-204-2 (alk. paper)
1. Mathematical statistics. 2. Algebra. I. Riccomagno, Eva. II. Wynn, Henry P. III. Title. IV. Series.

QA276 .P53 2000
519.5—dc21 00-047448

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

International Standard Book Number 1-58488-204-2
Library of Congress Card Number 00-047448
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Contents

List of figures

List of tables

Preface

About thirty-five years ago there was an awakening of interest of researchers in commutative algebra to the algorithmic and computational aspects of their field, marked by the publication of Buckberger's thesis in 1966. His work became the starting point of a new research field, called Computational Commutative Algebra. Currently, computer programs implementing versions of his and related algorithms are readily available both as commercial products and academic prototypes. These are of growing importance in almost every field of applied mathematics because they deal with very basic problems related to systems of polynomial equations. Statisticians, too, should find many useful tools in computational commutative algebra, together with interesting and enriching new perspectives. Just as the introduction of vectors and matrices has greatly improved the mathematics of statistics, these new tools provide a further step forward by offering a constructive methodology for a basic mathematical tool in statistics and probability, that is to say a ring. The mathematical structure of real random variables is precisely a ring, and other rings and ideals appear naturally in distribution theory and modeling. However, the ring of random variables is a ring with lattice operations which are not fully incorporated into the theory we present, at least not yet.

The authors' attention was drawn to the relevance of Gröbner basis theory by a paper on contingency tables by Sturmfels and Diaconis circulated as a manuscript in 1993. With initial help provided by Professor Teo Mora (University of Genova), a first application to design of experiments was published by G. Pistone and H. Wynn in 1996 (*Biometrika*) and this field of application was more fully developed by E. Riccomagno in her Ph.D. thesis work during 1996-97 at the University of Warwick. Subsequent papers in the same direction were published by the authors and a number of coauthors. We are pleased to acknowledge (in alphabetic order) Ron Bates, Massimo Caboara, Roberto Fontana, Beatrice Giglio, Tim Holliday, Maria-Piera Rogantin.

During the few years this monograph was in the making, we have benefitted from many contributions by others, and further related work is in progress. Some of the contents of this book was first exposed at the series of four GROSTAT workshops, which took place in successive years, starting in 1997 at the University of Warwick (UK), the IUT-STID in Nice-Côte

d'Azur in Menton (France), EURANDOM in Eindhoven (NL), and again, in 2000, in Menton. We must thank all the participants and these institutions for their support, in particular Professor Annie Cavarero, director of IUT-STID.

We found keen collaborators at the University of Genova. We should at least mention, together with those above, Professor Lorenzo Robbiano (who also supported GROSTAT IV) and the CoCoA team who have had a major influence on the algebraic and computational aspects of the field. We are very grateful to them all for the early and generous access to their research, for the high level of illumination it provided on the mathematical foundations and the very fast computer code developed under the wings of CoCoA.

We are grateful for many discussions with colleagues and coworkers. A minimal list includes Wilf Kendall, Thomas Richardson, Raffaella Settimi and Jim Smith, in Warwick, and Alessandro Di Bucchianico and Arjeh Cohen, in Eindhoven. Special thanks to Dan Naiman of The Johns Hopkins University for allowing us to draw on recent joint work on tube theory in Chapter 4. Ian Dinwoodie, from Tulane University, helped to strengthen our understanding of the work of Diaconis and Sturmfels on toric ideals, which we reach in the final sections of the book, from our own particular direction. Because a considerable volume of the monograph is based on work in progress, we have, on a few occasions, had to refer to unpublished, although available, technical reports. We thank all the colleagues who helped us by reading different versions of this work, some of them already mentioned, and also Neil Parkin for careful reading of the whole book. We also thank our publishers for their help and considerable patience.

A cocktail of different grants and institutions has funded this research. We should thank the UK Engineering and Physical Sciences Research Council, the Italian Consiglio Nazionale delle Richerche, EURANDOM, and, last but not least, IRMA and the University L. Pasteur of Strasbourg, and Professor Dominique Collombier, who has hosted us during the final revision of the book.

This book is dedicated to our families, with apologies to all for the absences that a triple collaboration must entail.

GIOVANNI PISTONE
EVA RICCOMAGNO
HENRY WYNN

Strasbourg, France, October 2000

Notation

Common symbols

$\mathbb{N}$	positive integer numbers
$\mathbb{Z}$	integer numbers
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
S^*	* excludes the 0 from the set S
S_+	non-negative entries of the set of numbers S: for example $\mathbb{Z}_+ = \{a \in \mathbb{Z} : a \geq 0\} = \{0\} \cup \mathbb{N}$
d superscript	dimension of the cartesian product for example, $\mathbb{Z}^d$ stands for $\underbrace{\mathbb{Z} \times \cdots \times \mathbb{Z}}_{d \text{ times}}$
$\{a\}$	1. component-wise fractional part operator, $a \in \mathbb{R}^d$ 2. the set whose element is a
$\#A$	number of elements in the set A
$[p]$	vector or list p as a column vector
$[a_1 \cdots a_n]$	matrix with the vectors a_i, $i = 1, \ldots, n$ as columns
$[[\ldots], \ldots, [\ldots]]$	matrix as a list of rows
A^t	transpose of A where A is a matrix or a vector
I	identity matrix
$x_1, \ldots, x_d$	factors, variables, indeterminates
d	1. number of independent factors 2. number of variables 3. number of indeterminates
s	number of x_i's if the algebra is emphasised
N	1. sample size 2. number of design points 3. number of support points
k, $\mathcal{K}$	fields of coefficients for example, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}(\theta)$, transcendental extension, $\mathbb{Q}(\sqrt{2})$, algebraic extension

Notation for Gröbner bases

$k[x_1, \ldots, x_s]$	ring of polynomials in $x_1, \ldots, x_s$ and with coefficients in k
$x^\alpha = x_1^{\alpha_1} \ldots x_s^{\alpha_s}$	monomial in $k[x_1, \ldots, x_s]$
$p(x_1, \ldots, x_s)$	polynomial in $k[x_1, \ldots, x_s]$
τ, $\succ$, $\succ_\tau$	term-ordering
$x_{i_1} \succ \ldots \succ x_{i_s}$	initial ordering on the indeterminates
$\tau(x_{i_1} \succ \ldots \succ x_{i_s})$	emphasis on the initial ordering
$\mathrm{LT}_\tau(p(x))$	leading term of the polynomial p with respect to the term-ordering τ
$\mathrm{Ideal}\,(g_1, \ldots, g_h)$ $\langle g_1, \ldots, g_h \rangle$	ideal of $k[x_1, \ldots, x_s]$ generated by $g_1, \ldots, g_h$
$\mathrm{Variety}(I)$	set of zeros of all polynomials in I
$\mathrm{Ideal}(V)$	set of all polynomials vanishing at V
$\mathrm{Variety}(f_1, \ldots, f_l)$	set of common roots of f_i, $i = 1, \ldots, l$
$\mathrm{Rem}(f), \mathrm{Rem}(f, G)$	1. normal form of f with respect to the Gröbner basis G 2. remainder of the division of f with respect to the set of polynomials G

Notation for experimental design

D, D_N	1. experimental design 2. support for a discrete distribution
a, x	design point
$x(i)$, $(x(i)_1, \ldots, x(i)_d)$	ith design point for $i = 1, \ldots, N$
$\mathcal{X}$	design region
$\mathrm{Est}_\tau(D)$	estimable terms with respect to τ and D
$\mathcal{F}$	polynomial regression vector
$Z = [f(x)]_{x \in D, f \in \mathcal{F}}$	design matrix for a model with support $\mathcal{F}$ and a design D; the orderings on D and $\mathcal{F}$ carry over to Z
$Z^t Z$	information matrix
$y = (y_1, \ldots, y_N)$	responses, values at the support points
θ, c, b, a	parameters or coefficients
$k[x_1, \ldots, x_d]/\,\mathrm{Ideal}(D)$ $k[x]/\,\mathrm{Ideal}(D)$	quotient ring
L	list of exponents of a vector space basis of $k[x_1, \ldots, x_d]/\mathrm{Ideal}(D)$
L_0	$L \setminus \{(0, \ldots, 0)\}$
L'	$L' \subseteq L$

Notation for logic and reliability

$\mathcal{B}(\vee, \wedge, ^{-}, 0, 1)$	Boolean algebra
$\vee$	maximum, union
$\wedge$	minimum, intersection
$\emptyset$	empty set
D_{2^d}	2^d full factorial design
$D \setminus D_{2^d}$, $\bar{D}$	complementary set of $D \subset D_{2^d}$
$f_a(x)$	polynomial indicator function of $a \in D_{2^d}$
$f_D(x)$	polynomial indicator function of $D \subset D_{2^d}$
$\mathrm{E}(f)$	mean value of f
$\triangle$	symmetric difference operator

Notation for probability and statistics

D, Ω	support of a probability space
$D^\star$	support of an image probability
A_i	elementary event
A	event
f_A	indicator function of the event A
$\mathcal{L}(D, \mathcal{K})$, $\mathcal{L}(D)$, $\mathcal{L}$	the set of functions from D to $\mathcal{K}$
X	function in $\mathcal{L}(D)$
P	probability
P_0	uniform probability
K	the constant in the exponential model
$K(\Phi)$, $K(\theta)$	cumulant generating function
$\mathrm{E}_0\left(X\right)$	expectation of X with respect to P_0
$\mathrm{E}_P\left(X\right)$	expectation of X with respect to P
m_α	raw moments $\mathrm{E}_0\left(X^\alpha\right)$
θ_α	θ-parameters of a probability
μ_α	μ-parameters $\mathrm{E}_P\left(X^\alpha\right)$
p_i	p-parameters $\mathrm{P}\left(a(i)\right)$
ψ_α	ψ-parameters in exponential models
ζ_α	ζ-parameters: $\zeta_\alpha = \exp(\psi_\alpha)$
R	three-dimensional multi-array where $\mathrm{Rem}\left(X^{\alpha+\beta}\right) = \sum_{\gamma \in L} R(\alpha, \beta, \gamma) X^\gamma$
$R(\beta)$	matrix $[R(\alpha, \beta, \gamma)]_{\gamma, \alpha \in L}$
$r(\delta, \gamma)$	$R(\alpha, \beta, \gamma)$ with $\delta = \alpha + \beta$
$Q(\alpha, \beta)$, $\alpha, \beta \in L$	$\mathrm{E}_0\left(X^{\alpha+\beta}\right) = \sum_{\gamma \in L} r(\alpha + \beta, \gamma) m_\gamma$

CHAPTER 1

Introduction

1.1 Outline

One of the most basic issues in statistical modeling is to set problems up correctly, or at least well. This means, typically, that a sample space needs to be defined together with some distribution on this sample space with some parameters. After that one can decide if the parameters or even the form of the distribution are known, and, given the motivation and resources, enter into full-blown statistical inference. Great care needs to be taken with data capture or, to put it more precisely, with experimental design, if the model is to be properly postulated, tested and used for prediction.

Some of the questions which need to be addressed in carrying out these operations are intrinsically algebraic, or can be recast as algebraic. By algebra here we will typically mean polynomial algebra. It may not at first be obvious that polynomials have a fundamental role to play.

Here is, perhaps, the simplest example possible. Suppose that two people (small enough) stand together on a bathroom scale. Our model is that the measurement is additive, so that if there is no error, and θ_1 and θ_2 are the two weights, the reading should be

$$Y = \theta_1 + \theta_2$$

Without any other information it is not possible to estimate, or compute, the individual weights θ_1 and θ_2. If there is an unknown zero correction θ_0 then $Y = \theta_0 + \theta_1 + \theta_2$ and we are in worse trouble.

In a standard regression model we write in matrix notation

$$Y = Z\theta + \varepsilon$$

and our ability to estimate the parameter vector θ, under standard theory, is equated with "Z is $N \times p$ full rank" or $\mathrm{Rank}(Z) = p < N$ where θ is a p-vector and N is the number of design points. An example is the one-dimensional polynomial regression

$$Y(x) = \sum_{j=0}^{p-1} \theta_j x^j + \varepsilon_x$$

Then, if the experimental design consists of p distinct points $a(1), \ldots, a(p)$, the square design matrix

$$Z = \left[a(i)^j\right]_{i=1,\ldots,p;j=0,\ldots,p-1}$$

has full rank, and for submodels with fewer than p terms, the Z-matrix also has full rank.

Algebraic methods have been used extensively in the construction of designs with suitable properties. However, particularly in the construction of balanced factorial designs with particular aliasing properties, abstract algebra in the form of group theory has also been used to study the identifiability problem. Most students and professionals in statistics will recall a course on experimental design in which Abelian group theory is used in the form of confounding relations such as

$$I = ABC$$

and unless they are experts in experimental design, they may have remained somewhat mystified thereafter. We return to this example in Section 1.3.

Let us consider a simple example. Here is a heuristic proof that there is a unique quadratic curve through the points $(a(1), y_1)$, $(a(2), y_2)$, $(a(3), y_3)$

$$y_i = r(a(i)), \quad i = 1, 2, 3$$

We can think of $a(1), a(2), a(3)$ as the points of an experimental design at which we have observed y_1, y_2, y_3, respectively, without error. We also assume that $a(1), a(2), a(3)$ are distinct.

Define the polynomial

$$d(x) = (x - a(1))(x - a(2))(x - a(3))$$

whose zeros are the design points. Take any competing polynomial, $p(x)$, through the data that is such that $p(a(i)) = y_i$ (for $i = 1, 2, 3$). Write

$$p(x) = s(x)d(x) + r(x)$$

where $r(x)$ is the remainder when $p(x)$ is divided by $d(x)$. Now we can appeal to algebra and say that, given the polynomial $p(x)$, $r(x)$ is unique. But it is clear from the equation that

$$y_i = p(a(i)) = r(a(i)), \quad (i = 1, 2, 3)$$

since by construction $d(a(i)) = 0$, $i = 1, 2, 3$.

The polynomial p above can be interpreted in two ways: (i) as a continuous function with value y_i at the point $a(i)$ and (ii) as a representation of the function defined only on the design points and again with value y_i at $a(i)$ (for $i = 1, 2, 3$). The first way is very convenient when we do regression analysis and thus we call p an *interpolator*. The other interpretation is more suited for applications in discrete probability.

Here we have tried to solve an identifiability problem directly by exhibiting a minimal degree interpolator rather than check the rank of a Z-matrix.

There is a crucial point to make: *all the operations were carried out with polynomials*.

The same argument applies for polynomial regression of all orders in one dimension. However, a very important issue for this book is that if we are to use this argument for x in higher dimensions, then we need to cope with the fact that representation of points as solutions of equations, the operation of division and the remainders themselves are *not*, in general, unique in higher dimensions. The representation of discrete points as the solution of polynomial equations is to treat them as *zero-dimensional algebraic varieties*. The division operation becomes a *quotient* operation and we have jumped into algebraic geometry. The set of all polynomials which are zero on a variety (in this case, a set of points) has the algebraic structure of an *ideal*. Strictly speaking, the quotient operation uses the ideal, not the variety. The use of Gröbner bases will help throughout.

Elementary probability is not immune from this treatment. Consider a random variable X whose support is $a(1), a(2), a(3)$. What was an experimental design, above, is now a support. Since X lives only on the support, we can write (with probability one)

$$(X - a(1))(X - a(2))(X - a(3)) = 0$$

Expanding we obtain

$$\begin{aligned} X^3 = (a(1) + a(2) + a(3))X^2 - \\ (a(1)a(2) + a(1)a(3) + a(2)a(3))X + a(1)a(2)a(3) \end{aligned}$$

Taking expectation and letting the non-central moments of X be $\mu_0 = 1$, $\mu_1 = \mathrm{E}(X)$, $\mu_2 = \mathrm{E}\left(X^2\right)$, ..., we have

$$\begin{aligned} \mu_3 &= (a(1) + a(2) + a(3))\,\mu_2 \\ &\quad - (a(1)a(2) + a(1)a(3) + a(2)a(3))\,\mu_1 \\ &\quad + a(1)a(2)a(3) \\ \mu_{3+k} &= (a(1) + a(2) + a(3))\,\mu_{2+k} \\ &\quad - (a(1)a(2) + a(1)a(3) + a(2)a(3))\,\mu_{1+k} \\ &\quad + a(1)a(2)a(3)\mu_k \end{aligned} \tag{1.1}$$

We can, in this way, express any higher-order moment as a linear function of μ_0, μ_1, μ_2. This is an example of what we shall call *moment aliasing*.

This small example points to several levels of the use of polynomial algebra in statistics. The first level is to set up the machinery for handling sets of points in many dimensions. These points will be thought of first as an experimental design D and then, when we do probability, as the support of a distribution. Of course, the problem is then different. It is the algebra which is, identical, and to emphasize this, we use the same letter D when the set of points is a support. We will cover at some length all the issues to do with description of varieties, ideals, quotient operations and so on. This occupies Chapters 2 and 3. Chapter 5 studies the algebra of random variables over a finite set of points. This is the second level.

The third level is to interpolate the probability masses for our distribution on the support D. Since the algebra has already told us how to set up interpolators, this is now straightforward, except that probabilities are non-negative and must sum to one. Still at this level we have two basic alternatives: to interpolate the raw probabilities or to interpolate their logarithm. For example, suppose we have a two-state (binary) random variable taking the values in $D = \{0, 1\}$ with probabilities $1-q$ and q, respectively: a Bernoulli random variable. The raw interpolator is

$$p(x) = 1 - q + (2q - 1)x$$

whereas the interpolator of the logarithm, after exponentiation, gives

$$p(x) = \exp\left(\log(1-q) + \log\left(\frac{q}{1-q}\right) x\right)$$

The second of these is the usual exponential family representation of the Bernoulli.

The fourth level of algebraisation, and perhaps the most profound, arises from noticing that when the support D lies at integer grid points, an exponential term such as $e^{\psi_1 x_1}$ can be written $\zeta_1^{x_1}$ where

$$\zeta_1 = e^{\psi_1}$$

Using this trick, we can rewrite models in the exponential form as polynomials. For the Bernoulli, let $\psi_0 = \log(1-q)$ and $\psi_1 = \log\left(\frac{q}{1-q}\right)$. Then, setting $\zeta_0 = e^{\psi_0}$ and $\zeta_1 = e^{\psi_1}$ we have the representation

$$p(x) = \zeta_0 \zeta_1^x$$

This coincides with the familiar form $p(x) = q^x(1-q)^{1-x}$. We shall also discuss this form, which is closely related to the work of Diaconis and Sturmfels (1998) on toric ideals.

Note that we have been a little lazy with the notation here. All the forms of $p(x)$ have a different structure but agree numerically on D.

Much of the real usefulness of algebra in statistics comes from the interplay between these different parametrisations. We shall also need another parametrisation in terms of moments. This is made harder by the fact that, typically, statistical models or submodels are obtained by imposing restrictions on the parameters. We shall define an *algebraic statistical model* as one which adopts one of these representations and for which the restrictions on the parameters are themselves polynomial. However, and this is the most complex issue in the book, the forms of these submodels may be different depending on the parametrisation. Only sometimes can they be perfectly linked. An important example is the independence condition, which forces factorisation of the raw polynomial interpolators, maps to additivity inside the exponential representation and factorisation in the ζ and q^x forms. Conditional independence, as used in Bayes networks, also has this multiple representation. Chapters 5 and 6 discuss all these issues.

The book can be seen from different angles and we are grateful to a reviewer for making us more aware of this. The ambitious angle, and more relevant to researchers in statistics, is to rewrite the foundations of discrete probability and statistics in the language of algebraic geometry. We have only partly succeeded in doing this. There is still much to be done, particularly in sorting out fundamental issues arising from submodels discussed in the last chapter, both theoretically and computationally. This effort must surely draw on the important work of Andrews and Stafford (2000) on general application of computer algebra to statistics.

The more modest objective in which we hope to have succeeded is to enlarge the kitbag of tools available to the statistician. The Gröbner basis method in experimental design can now be used routinely, and is by the authors, to investigate the identifiability of experimental design/model combinations in real applications. The use of the methods in statistical modeling should also proceed rapidly. After the seminal work by Diaconis and Sturmfels (1998), there have been advances in using Gröbner basis methods for Monte Carlo style sampling on contingency tables, notably by Dinwoodie (1998). Promising ongoing work on the use of Gröbner basis methods in Bayes networks is being carried out by J. Q. Smith and R. Settimi. We also include in Section 4.5 work by the authors and other collaborators on reliability on binary (two-level) factorial design.

1.2 Computer Algebra

Several packages for symbolic computation and Gröbner basis computation are available: CoCoA, Maple, Mathematica and GB, to mention a few. We have used mostly Maple and CoCoA. Some points need to be made about these packages.

The package CoCoA (COmputations in COmmutative Algebra, freely available at `http://cocoa.dima.unige.it`) is specially developed for research in algebraic geometry and commutative algebra. Thus it is faster than most other software in computing Gröbner bases, although at times not intuitive, and it allows more refined computations. The interface needs further development and the use of unknown constants is not implemented. Nevertheless in some cases ad hoc tricks can be used to force some indeterminates to play the role of unknown constants. An example is the case of complex numbers for which an indeterminate i is introduced to represent the complex unit. For details see Caboara and Riccomagno (1998).

Robbiano and other members of the CoCoA team are very active in the research area described in Chapters 2 and 3 of this book. They concentrate mainly on links to algebraic geometry with forays into statistics (Robbiano and Rogantin (1998), Caboara and Robbiano (1997)), while the authors are led by applications in statistics with some expeditions into the mathematics and computation.

Maple (University of Waterloo, Canada `http://www.maplesoft.com`) is a general purpose package for symbolic computations. It is quite fast, simple to use and with a good online help. It has a very good interface, allows the use of unknown constants or free parameters, but it is slower than CoCoA for the specialized application described here. Maple V-5 includes the package `Groebner` for doing Gröbner basis computation, and allows the use of unknown constants and user-defined term-orderings.

Sometimes our examples will be over the set of integers, $\mathbb{Z}$, which is not a field. Gröbner basis theory has a counterpart for polynomials with integer coefficients, but it is more expensive. For example, in CoCoA, when the ring $\mathbb{Z}[x_1, x_2]$ is input, a message appears warning that *G-basis-related computations could fail to terminate or can be wrong.* However, $\mathbb{Z}$ is embedded in $\mathbb{Q}$, and one can work with rational coefficients and multiply everything out to obtain integers. On other occasions one has to work with a finite set of coefficients, say $\mathbb{Z}_p$. For p, a prime integer, $\mathbb{Z}_p$ forms a field and the algebraic theory of Gröbner bases is similar to that over rational numbers. In other cases, such as the trigonometric case (see Section 3.14), difficulties arise from the fact that the sine and cosine of rational values are typically irrational numbers and thus the coefficient field is not embeddable in $\mathbb{Q}$. Ad hoc procedures have been considered based on simple algebraic extensions of rational numbers.

As mentioned, the authors prefer to use Maple and CoCoA. Lists of software that include routines to compute Gröbner bases are maintained at `http://SymbolicNet.mcs.kent.edu/` and `http://anillos.ugr.es/`. We should mention: Matematica for its popularity, REDUCE written in LISP and whose main characteristics are code stability, full source code availability and portability, and AXIOM, which takes an object-oriented approach to computer algebra and its overall structure is strongly typed and hierarchical. Among the freely available software there is GROEBNER (at `ftp.risc.uni-linz.ac.at`) developed at RISC-Linz by W. Windsteiger and B. Buchberger, Macaulay2 (`http://www.math.uiuc.edu/Macaulay2/`) developed by D. Grayson and M. Stillman to support research in algebraic geometry and in commutative algebra. The package SINGULAR (`http://www.singular.uni-kl.de/`) is advertised as the most powerful and efficient systems for polynomial computations with a kernel written in C++.

Next we anticipate some notions from Chapter 2. Historically a first application of Gröbner bases is as polynomial system solver in that it can rewrite a system of polynomial equations in an equivalent form which is easier to solve. Equivalent means with the same set of solutions. For example, if the system has a finite number of solutions, there is a Gröbner basis including a polynomial in only one indeterminate, a polynomial in that indeterminate and another one, and so on. In this way the system can be solved by backward substitution. The great advantage of Gröbner bases with respect to, say, numerical methods for solving systems of polynomial equations, is that it can also be used when the system has infinitely many solutions. All the solutions are returned but in a parametric, or implicit, form, which sometimes seems even more complicated than the original. This is why it is generally recommended to couple Gröbner basis with numerical methods when used as system solver.

In this book we are concerned with two slightly different algebraic aspects which use the same Gröbner basis techniques. 1. We know the solutions (so to speak) and are interested in determining the set of polynomials interpolating them. Then, Gröbner basis methods return a basis of the set of functions defined over the solutions. 2. We have a system of polynomial equations and would like to check whether there are some algebraic relations. That is, we need to rewrite the system in a different form. The operations we allow are sums of elements in the polynomial set considered and products with any polynomial. This leads to the definition of a polynomial ideal for which we refer to the main text.

1.2.1 A quick introduction to Gröbner bases

A polynomial, in one or more variables, is a linear combination of monomials. Thus $1 + 2x_1 + 3x_2 + 4x_1x_2$ is a polynomial and $1, x_1, x_2, x_1x_2$ are monomials.

On the set of integer numbers there is one natural total order, the one we all know. The set of monomials in one indeterminate, x, inherits such an order, thus x is lower than x^3 and $1 = x^0$ is lower than x^α for all α positive integers. We do not consider negative integers.

In more than one dimension the uniqueness of a natural way of ordering points on the (non-negative) integer grid is lost. The same is valid for monomials in more than one indeterminate. In Chapter 2 monomial orderings (also called term-orderings) are properly defined. For the moment we only observe that a term-ordering corresponds to a total order on the integer grid and is compatible with cancellation of monomials. There are orderings on the integer grid that do not correspond to any term-ordering.

The most common term-ordering is the lexicographic ordering. In three dimensions x, y and z, first fix z larger (in the ordering) than y and y larger than x. We write $z \succ y \succ x$ and talk of initial ordering. All monomials of the type x^α are lower than any monomial involving y and/or z and the monomials x^α are ordered according to the one-dimensional ordering. Next come the monomials with the y indeterminate at first degree, that is $x^\alpha y$, which are again ordered according to the one-dimensional ordering. After

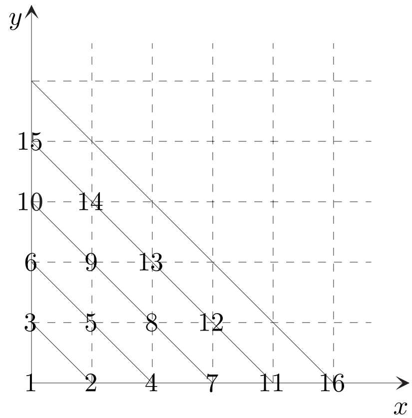

Figure 1.1 *Example of degree reverse lexicographic term-ordering in two dimensions.*

that we have the monomials $x^\alpha y^2$. After all the monomials $x^\alpha y^\beta$, for α and β non-negative integers, it is the turn of the monomials including the z indeterminate.

The degree reverse lexicographic term-ordering is a term-ordering often used. An example in two dimensions is given in Figure 1.1. Monomials on a line parallel to $y = -x$ are ordered in a linear fashion according to the ordering in one dimension and going in the direction bottom to top, that is x^α is smaller than y^α. Monomials on lines closer to the origin are smaller than monomials on lines far away. In higher dimensions, hyper-planes play the role of lines. For a definition see Section 2.3.

Once a term-ordering is chosen, the largest term of a polynomial is well defined and is called its *leading term*.

Consider the system of polynomials

$$\begin{cases} yx - z \\ x^2 - z \end{cases} \tag{1.2}$$

The associated system of equations is obtained by equating to zero the two polynomials. A quick computation shows that there are two sets of solutions

$$\begin{cases} x = 0 \\ y = y \\ z = 0 \end{cases} \quad \text{and} \quad \begin{cases} x = y \\ y = y \\ z = y^2 \end{cases}$$

The following systems of polynomial equations have the same solutions, that is they are algebraically equivalent,

$$\begin{cases} (y-x)x = 0 \\ z - x^2 = 0 \end{cases} \qquad \begin{cases} yx - z = 0 \\ z(y-x) = 0 \\ z - x^2 = 0 \end{cases}$$

The corresponding two sets of polynomials are two different Gröbner bases of the ideal generated by Equation (1.2) with respect to two different term-orderings. That is the lexicographic ordering with initial ordering $z \succ y \succ x$ and the degree reverse lexicographic term-ordering with the same initial ordering, respectively. The leading terms are $\{yx, z\}$ and $\{yx, zy, x^2\}$.

Looking at the solutions of the systems, one is tempted to say that an equivalent set of polynomials is

$$\begin{cases} x - y \\ z - y^2 \end{cases} \tag{1.3}$$

But it cannot be retrieved from the polynomials in (1.2) using sums and products of polynomials. That is, this last system is not algebraically equivalent to the others. The solution $(0,0,0)$ is clearly given in (1.2) while in (1.3) it is deduced from the solution $x = y$, $z = y^2$ for $y = 0$. This phenomenon is referred to as the multiplicity of a solution.

Roughly speaking, Gröbner basis computation allows us to rewrite the system (1.2) without losing or adding solutions, by having the correct set of leading terms. Namely, a polynomial set G is a Gröbner basis for a set of polynomials F and with respect to a term-ordering if the set of polynomials generated by the leading terms of F is equal to the analogous set generated by the leading terms of G. The elements of the set generated by the polynomials $\{f_1, \ldots, f_s\}$ are the polynomials $\sum_{i=1}^{s} h_i f_i$, where the h_i's are generic polynomials. Note the role of a term-ordering in the definition of Gröbner bases. The set of polynomials $F = \{f_1 = yx - z, f_2 = x^2y - z\}$ does not form a Gröbner basis with respect to the lexicographic term-ordering with initial ordering $z \succ y \succ x$. Call this term-ordering τ. Indeed yx cannot be obtained from the leading terms of f_1 and f_2, which is z for both f_1 and f_2, but it is the leading term of $f_1 + f_2$. The (reduced) Gröbner basis of F with respect to τ is given above. There is an algorithm to compute Gröbner bases given a set of polynomials and a term-ordering which is described in Section 2.12.3.

Having the right leading terms also helps in the division of polynomials. Namely the division of a polynomial by a Gröbner basis has a unique remainder, while in general this is not true. The division of a polynomial f by a set F is a way of rewriting f as a polynomial combination of elements of F in such a way that we are left with a reminder whose leading term is not divisible by the leading terms of the polynomials in F. For example consider $f = z$. The division of f by f_1 and f_2, with respect to τ, gives the reminder yx if we divide first by f_1, indeed $f = (-1)f_1 + xy$. But if we first

Table 1.1 *The 2^{3-1} fractional factorial design.*

A	B	C
1	1	1
1	−1	−1
−1	1	−1
−1	−1	1

divide by f_2 and then by f_1 we obtain $f = (-1)f_2 + x^2$ where now x^2 is the remainder. Fortunately when we divide f with respect to the Gröbner basis G, we do not need to consider with respect to which polynomial we divide first, the reminder will always be the same, z itself in this example.

1.3 An example: the 2^{3-1} fractional factorial design

In this section we outline the ideas and techniques presented in this book on an example which we shall return to in the main text as well. Consider the four points of the 2^{3-1} fractional factorial design with levels ± 1 in Table 1.1 (see Box, Hunter and Hunter (1978) and Cox and Reid (2000)). It is defined by the confounding relation $ABC = I$ where A, B and C are the factors and I is the identity. When we refer to the factors in the classical framework, for example when using the mathematics of group theory, we use capital letters. We use small letters a, b and c for factors in our polynomial representation. Moreover some computer algebra software require that indeterminates, the algebraic equivalent of factors, are a single, small letter.

The rows in Table 1.1 are solutions of the following system of polynomial equations, which defines the 2^{3-1}design

$$\begin{cases} a^2 - 1 = 0 \\ b^2 - 1 = 0 \\ c^2 - 1 = 0 \\ abc - 1 = 0 \end{cases} \tag{1.4}$$

The aliasing table in Table 1.2 is obtained by multiplying $ABC = I$ by A, B and C, respectively. Now, the system of polynomial equations originated by substituting small letters in Table 1.2 has still the same set of solutions as the system in (1.4). For the polynomials in the system so obtained, namely $abc - 1$, $bc - a$, $ac - b$, $ab - c$, the first polynomial is larger than the other three polynomials as its highest term is divided by the second-order terms of the other three polynomials. In this sense it is redundant

Table 1.2 *Aliasing table for the* 2^{3-1} *design.*

ABC	$=$	I
BC	$=$	A
AC	$=$	B
AB	$=$	C

and it can be substituted by the three polynomials a^2-1, b^2-1 and c^2-1 which are of smaller order. The set of zeros of the system of polynomial equations obtained equating to zero these new three polynomials is the 2^3 full factorial design.

The final set of equations so obtained forms a Gröbner basis

$$\begin{cases} a^2-1 \\ b^2-1 \\ c^2-1 \\ bc-a \\ ac-b \\ ab-c \end{cases} \tag{1.5}$$

General methods to compute Gröbner bases from a set of polynomials are given in Chapter 2.

In the classical theory, one would look at the aliasing table in Table 1.2 and deduce that the interaction AB is aliased to the linear factor C. That is the effects of AB and C are confounded and both AB and C cannot be terms in the same linear regression model. In more mathematical terminology one says that AB and C are linearly dependent functions over the 2^{3-1} design. The approach presented in this book develops this observation. The theory of Gröbner basis automatises the process of finding a vector space basis of the set of functions defined over the 2^{3-1} design. From this vector space basis it is easy to check whether two terms are confounded. This saturated set of independent terms is formed by monomials, that is factors and interactions. It will be the basis with the terms smallest in some sense which will be clear when in Chapter 2 the concept of term-ordering is explained.

We show the process for determining this vector space basis for the 2^{3-1} design. Consider the Gröbner basis in Equation (1.5) and consider the largest terms of each of its polynomials, they are

$$\mathrm{LT} = \left\{a^2, \quad b^2, \quad c^2, \quad ab, \quad ac, \quad bc\right\}$$

The formalization of this process requires again the definition of term-ordering. For the moment it is sufficient to say that, for example, in $ab-c$

the term ab is larger than c because it represents a second-order interaction. In some cases to be considered later it will be possible that a linear term is larger than an interaction.

Now consider all the terms that are not divided by the monomials in LT. They are listed below and they are four, exactly the number of points in the 2^{3-1} design:

$$1, \quad a, \quad b, \quad c$$

The theory of Gröbner bases states that this is a set of linearly independent functions over the 2^{3-1} design. They can be used to build a linear regression model.

In particular all the functions over the 2^{3-1} design can be represented as linear combinations of those four monomials, and a function f is written as

$$f(x) = \theta_0 + \theta_1 a + \theta_2 b + \theta_3 c$$

where x ranges over the points in the 2^{3-1} design. Now probabilities are functions and thus can be represented in this way, and the θ coefficients are chosen so that $\sum_{x \in 2^{3-1}} f(x) = 1$. For example, the probability that assigns mass $1/2$ to the point $(1,1,1)$, mass $1/4$ to the point $(-1,1,-1)$, and equal mass $1/8$ to the other two points is the function

$$1/4 + 1/16a + 1/8b + 1/16c$$

The uniform probability is given by the constant function $1/4$.

Random variables are again linear functions of $1, a, b, c$, for example $Y = A + B + C$. The expectation of Y with respect to the uniform probability can now be computed with linear operations as

$$\mathrm{E}_0\left(Y\right) = \sum_{x \in 2^{3-1}} Y(x) = \sum_{(a,b,c) \in 2^{3-1}} (a+b+c) = 0$$

Analogously, the second-order moment is

$$\mathrm{E}_0\left(Y^2\right) = \sum_{x \in 2^{3-1}} Y(x)^2 = \sum_{(a,b,c) \in 2^{3-1}} (a+b+c)^2 = 12$$

As mentioned previously the relation (1.1) further simplifies the computation of higher-order moments.

We conclude this section by computing the image probability of Y. Let us start with the computation of the image support. Thus adjoin the polynomial for Y, using small letter y, to the equations of the Gröbner basis of the 2^{3-1} design

$$\begin{cases} a^2 - 1 \\ b^2 - 1 \\ c^2 - 1 \\ bc - a \\ ac - b \\ ab - c \\ y - (a + b + c) \end{cases} \tag{1.6}$$

The aim is to find a polynomial involving only y and not the indeterminates a, b and c. That is to check whether y is algebraically independent from a, b and c. The square of the last polynomial in (1.6) above gives

$$y^2 + (a + b + c)^2 - 2y(a + b + c)$$

and thus, using again the definition of y,

$$y^2 - (a + b + c)^2 = y^2 - a^2 - b^2 - c^2 - 2bc - 2ac - 2ab$$

Now $a^2 = b^2 = c^2 = 1$ and $bc = a$, $ac = b$ and $ab = c$, giving

$$y^2 - 2y - 3 = (y + 1)(y - 3) = 0$$

This is the description of the image of Y. In Chapter 5 this process is automatised by considering the Gröbner basis of the polynomials above with respect to a so-called elimination term-ordering.

The image probability of Y takes values on the set $D^* = \{-1, 3\}$ and its density with respect to the uniform distribution has the form of a polynomial supported on $\{1, y\}$. Thus in generic form we can write

$$p_Y = \theta_0 + \theta_1 Y$$

Because the support of p_Y is $\{1, y\}$, the density p_Y is fully known if the first two moments $\mathrm{E}\left(Y^\alpha\right)$, $\alpha = 0, 1$ are known. By using the conditions $Y^2 = 2Y + 3$, $\mathrm{E}\left(Y\right) = 0$, and $\mathrm{E}_*\left(Y\right) = \dfrac{-1+3}{2} = 1$ (the expectation with respect to the uniform on D^*), we obtain the system

$$\begin{cases} 1 = \mathrm{E}_*\left(\theta_0 + \theta_1 Y\right) = \theta_0 + \theta_1 \\ 0 = \mathrm{E}_*\left(\theta_0 Y + \theta_1 Y^2\right) = \mathrm{E}_*\left(\theta_0 Y + \theta_1(2Y + 3)\right) = \theta_0 + 5\theta_1 \end{cases}$$

which gives $p_Y = \dfrac{5}{4} - \dfrac{1}{4} Y$.

The polynomial setup presented here can be used to discuss many probabilistic and statistical concepts. Much of this can be found in the main text but still much work is left for the authors and the interested reader.

References

Abbott J., Bigatti A., Kreutzer M., and Robbiano L. (2000) Computing ideals of points. *J. Symb. Comput.* **30**, 341–356.

Adams W.W. and Loustaunau P. (1994) *An Introduction to Gröbner Bases*. American Mathematical Society, Providence, RI.

Amari S.i. (1985) *Differential-geometrical Methods in Statistics*. Springer-Verlag, New York, 2nd printing 1990 corrected ed.

Andrews D.F. and Stafford J.E. (2000) *Symbolic Computation for Statistical Inference*. Oxford University Pres, Oxford.

Anthony M. and Biggs N. (1997) *Computational Learning Theory. An Introduction*. Cambridge University Press, Cambridge.

Barlow R.E. (1998) *Engineering Reliability*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Barndorff-Nielsen O.E. and Cox D.R. (1989) *Asymptotic Techniques for Use in Statistics*. Chapman & Hall, London.

Barndorff-Nielsen O.E. and Cox D.R. (1994) *Inference and Asymptotics*. Chapman & Hall, London.

Becker T. and Weispfenning V. (1993) *Gröbner Bases. A Computational Approach to Commutative Algebra*. Springer-Verlag, New York.

Box G.E.P., Hunter W.G., and Hunter J.S. (1978) *Statistics for Experimenters*. John Wiley & Sons, New York-Chichester-Brisbane.

Buchberger B. (1966) *On Finding a Vector Space Basis of the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal*. PhD thesis, Department of Mathematics, University of Innsbruck.

Caboara M., de Dominicis G., and Robbiano L. (1996) Multigraded Hilbert function and Buchberger algorithm. In Y. Lakshman, (ed.) *Proc. ISSAC'96*, ACM.

Caboara M., Pistone G., Riccomagno E., and Wynn H. (1997) The fan of an experimental design. SCU Research Report 33, Department of Statistics, University of Warwick.

Caboara M. and Riccomagno E. (1998) An algebraic computational approach to the identifiability of Fourier models. *J. Symbolic Comput.* **26**, 245–260.

Caboara M. and Robbiano L. (1997) Families of ideals in Statistics. In Küchlin, (ed.) *Proc. ISSAC '97*, ACM Press, New York.

Collart S., Kalkbrener M., and Mall D. (1997) Converting bases with the Gröbner walk. *J. Symbolic Comput.* **24**, 465–469.

Cox D., Little J., and O'Shea D. (1997) *Ideals, Varieties, and Algorithms*. Springer-Verlag, New York, 2nd ed.

Cox D., Little J., and O'Shea D. (1998) *Using Algebraic Geometry*. Springer-Verlag, New York.

Cox D.R. and Reid N. (2000) *The Theory of the Design of Experiments*. Chapman & Hall, London.

Diaconis P. and Sturmfels B. (1998) Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26**, 363–397.

Dinwoodie I.H. (1998) The Diaconis-Sturmfels algorithm and rules of succession. *Bernoulli* **4**, 401–410.

Dohmen K. (1999) Improved inclusion-exclusion identities and inequalities based on a particular class of abstract tubes. *Electron. J. Probab.* **4**, no. 5, 12 pp. (electronic).

Faugère J.C., Gianni P., Lazard D., and Mora T. (1993) Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Comput.* **16**, 329–344.

Fontana R., Pistone G., and Rogantin M.P. (1997) Algebraic analysis and generation of two-levels designs. *Statistica Applicata* **9**, 15–29.

Fontana R., Pistone G., and Rogantin M.P. (2000) Classification of two-level factorial fractions. *J. Statist. Plann. Inference* **87**, 149–172.

Giglio B., Naiman D.Q., and Wynn H.P. (2000) Gröbner bases, abstract tubes, and inclusion-exclusion reliability bounds. SCU Research Report 25, Department of Statistics, Uuniversity of Warwick.

Giglio B., Riccomagno E., and Wynn H.P. (2000) Gröbner bases strategies in regression. *J. Appl. Statist.* **27**, 923–938.

Halmos P. and Givant S. (1998) *Logic as Algebra*. Mathematical Association of America, Washington, DC.

Holliday T., Pistone G., Riccomagno E., and Wynn H.P. (1999) The application of computational algebraic geometry to the analysis of designed experiments: a case study. *Comput. Statist.* **14**, 213–231.

Kiefer J.C. (1987) *Introduction to Statistical Inference. Edited and with a preface by Gary Lorden*. Springer-Verlag, New York.

Kreuzer M. and Robbiano L. (2000) *Computational Commutative Algebra 1*. Springer, Berlin-Heidelberg.

Lauritzen S.L. (1996) *Graphical Models*. The Clarendon Press Oxford University Press, New York.

Lehmann E.L. (1983) *Theory of Point Estimation*. John Wiley & Sons Inc., New York, a Wiley Publication in Mathematical Statistics.

Lehmann E.L. (1986) *Testing Statistical Hypotheses*. John Wiley & Sons Inc., New York, second ed.

Marinari M., Möller H., and Mora T. (1996) Gröbner duality. Tech. Rep. 312, Università di Genova, Dipartimento di Matematica, preprint.

Marinari M.G., Möller H.M., and Mora T. (1993) Gröbner bases of ideals defined by functionals with an application to ideals of projective points. *Appl. Algebra Engrg. Comm. Comput.* **4**, 103–145.

McCullagh P. and Nelder J.A. (1983) *Generalized Linear Models*. Chapman & Hall, London.

Mora T. (1994) An introduction to commutative and noncommutative Gröbner bases. *Theoret. Comput. Sci.* **134**, 131–173.

Mora T. and Robbiano L. (1988) The Gröbner fan of an ideal. *J. Symbolic Comput.* **6**, 183–208.

Naiman D. and Wynn H.P. (2000) Abstract tubes for simplex and orthant arrangements with applications to reliability bounds. SCU Research Report 24, Department of Statistics, Uuniversity of Warwick.

Pistone G. and Wynn H.P. (1996) Generalised confounding with Gröbner bases. *Biometrika* **83**, 653–666.

Pistone G. and Wynn H.P. (1999) Finitely generated cumulants. *Statist. Sinica* **9**, 1029–1052.

Riccomagno E. (1997) *Algebraic Geometry in Experimental Design and Related Fields*. PhD thesis, Department of Statistics, University of Warwick.

Robbiano L. (1985) Term orderings on the polynomial ring. In *EUROCAL '85, Vol. 2 (Linz, 1985)*, Springer, Berlin, 513–517.

Robbiano L. (1998) Gröbner bases and statistics. In *Gröbner Bases and Applications (Linz, 1998)*, Cambridge Univ. Press, Cambridge, 179–204.

Robbiano L. and Rogantin M.P. (1998) Full factorial designs and distracted fractions. In *Gröbner Bases and Applications (Linz, 1998)*, Cambridge Univ. Press, Cambridge, 473–482.

Sturmfels B. (1996) *Gröbner Bases and Convex Polytopes*. American Mathematical Society, Providence, RI.

Traverso C. (1996) Hilbert functions and the Buchberger algorithm. *J. Symbolic Comput.* **22**, 355–376.