The IMA Volumes in Mathematics and its Applications

Volume 147

Institute for Mathematics and its Applications (IMA)

The Institute for Mathematics and its Applications was established by a grant from the National Science Foundation to the University of Minnesota in 1982. The primary mission of the IMA is to foster research of a truly interdisciplinary nature, establishing links between mathematics of the highest caliber and important scientific and technological problems from other disciplines and industries. To this end, the IMA organizes a wide variety of programs, ranging from short intense workshops in areas of exceptional interest and opportunity to extensive thematic programs lasting a year. IMA Volumes are used to communicate results of these programs that we believe are of particular value to the broader scientific community.

The full list of IMA books can be found at the Web site of the Institute for Mathematics and its Applications:

http://www.ima.umn.edu/springer/volumes.html

Presentation materials from the IMA talks are available at

http://www.ima.umn.edu/talks/

Douglas N. Arnold, Director of the IMA

* * * * * * * * * *

IMA ANNUAL PROGRAMS

1982–1983	Statistical and Continuum Approaches to Phase Transition
1983–1984	Mathematical Models for the Economics of Decentralized Resource Allocation
1984–1985	Continuum Physics and Partial Differential Equations
1985–1986	Stochastic Differential Equations and Their Applications
1986–1987	Scientific Computation
1987–1988	Applied Combinatorics
1988–1989	Nonlinear Waves
1989–1990	Dynamical Systems and Their Applications
1990–1991	Phase Transitions and Free Boundaries
1991–1992	Applied Linear Algebra
1992–1993	Control Theory and its Applications
1993–1994	Emerging Applications of Probability
1994–1995	Waves and Scattering
1995–1996	Mathematical Methods in Material Science
1996–1997	Mathematics of High Performance Computing

(Continued at the back)

Grzegorz A. Rempała Jacek Wesołowski
Authors

Symmetric Functionals on Random Matrices and Random Matchings Problems

Springer

Grzegorz A. Rempała
Department of Mathematics
University of Louisville, KY, USA
Louisville 40292
http://www.louisville.edu/∼garemp01

Jacek Wesołowski
Wydzial Matematyki i Nauk
Informacyjnych
Politechnika Warszawska, Warszawa,
1 Pl. Politechniki
Warszawa 00-661
Poland

ISBN: 978-0-387-75145-0 e-ISBN: 970-0-387-75146-7

Mathematics Subject Classification (2000): 60F05, 60F17, 62G20, 62G10, 05A16

Library of Congress Control Number: 2007938212

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Moim Rodzicom, Helenie oraz Jasiowi i Antosiowi
(Grzegorz Rempała)

Moim Rodzicom
(Jacek Wesołowski)

Foreword

This IMA Volume in Mathematics and its Applications

SYMMETRIC FUNCTIONALS ON RANDOM MATRICES AND RANDOM MATCHINGS PROBLEMS

During the academic year 2003–2004, the Institute for Mathematics and its Applications (IMA) held a thematic program on Probability and Statistics in Complex Systems. The program focused on large complex systems in which stochasticity plays a significant role. Such systems are very diverse, and the IMA program emphasized systems as varied as the human genome, the internet, and world financial markets. Although quite different, these systems have features in common, such as multitudes of interconnecting parts and the availability of large amounts of high-dimensional noisy data. The program emphasized the development and application of common mathematical and computational techniques to model and analyze such systems. About 1,000 mathematicians, statisticians, scientists, and engineers participated in the IMA thematic program, including about 50 who were in residence at the IMA during much or all of the year.

The present volume was born during the 2003–2004 thematic program at the IMA. The two authors were visitors to the IMA during the program year, with the first author resident for the entire ten months. This volume is a result of the authors' interactions at the IMA, and, especially their discussions with the many other program participants in the program and their involvement in the numerous tutorials, workshops, and seminars held during the year. The book treats recent progress in random matrix permanents, random matchings and their asymptotic behavior, an area of stochastic modeling and analysis which has applications to a variety of complex systems and problems of high dimensional data analysis.

Like many outcomes of IMA thematic programs, the seed for this volume was planted at the IMA, but it took time to grow and flourish. The final fruit is realized well after the program ends. While all credit and responsibility for the contents of the book reside with the authors, the IMA is delighted to have supplied the fertile ground for this work to take place.

We take this opportunity to thank the National Science Foundation for its support of the IMA.

Series Editors

Douglas N. Arnold, Director of the IMA

Arnd Scheel, Deputy Director of the IMA

Preface

The idea of writing this monograph came about through discussions which we held as participants in the activities of an annual program "Probability and Statistics in Complex Systems" of the Institute for Mathematics and Its Applications at the University of Minnesota (IMA) which was hosted there during the 2003/04 academic year. In the course of interactions with the Institute's visitors and guests, we came to a realization that many of the ideas and techniques developed recently for analyzing asymptotic behavior of random matchings are relatively unknown and could be of interest to a broader community of researchers interested in the theory of random matrices and statistical methods for high dimensional inference. In our IMA discussions it also transpired that many of the tools developed for the analysis of asymptotic behavior of random permanents and the likes may be also useful in more general context of problems emerging in the area of complex stochastic systems. In such systems, often in the context of modeling, statistical hypothesis testing or estimation of the relevant quantities, the distributional properties of the functionals on the entries of random matrices are of concern. From this viewpoint, the interest in the laws of various random matrix functionals useful in statistical analysis contrasts with the interest of a classical theory of random matrices which is primarily concerned with asymptotic distributional laws of eigenvalues and eigenvectors.

The text's content is drawn from the recent literature on questions related to asymptotics for random permanents and random matchings. That material has been augmented with a sizable amount of preliminary material in order to make the text somewhat self-contained. With this supplementary material, the text should be accessible to any mathematics, statistics or engineering graduate student who has taken basic introductory courses in probability theory and mathematical statistics.

The presentation is organized in seven chapters. Chapter 1 gives a general introduction to the topics covered in the text while also providing the reader with some examples of their applications to problems in stochastic complex systems formulated in terms of random matchings. This preliminary chapter makes a connection between random matchings, random permanents and U-statistics. Also a concept of a P- statistic, which connects the three concepts is introduced there. Chapter 2 builds upon these connections and contains a number of results for a general class of random matchings which, like for instance the variance formula for a P-statistic, are fundamental to the developments further in the text. Taken together the material of Chapters 1 and 2 should give the reader the necessary background to approach the topics covered later in the text.

Chapters 3 and 4 deal with random permanents and a problem of describing asymptotic distributions for a "classical" count of perfect matchings in random bipartite graphs. Chapter 3 details a relatively straightforward but computationally tedious approach leading to central limit theorems and laws of large numbers for random permanents. Chapter 4 presents a more general treatment of the subject by means of functional limit theorems and weak convergence of iterative stochastic integrals. The basic facts of the theory of stochastic integration are outlined in the first sections of Chapter 4 as necessary.

In Chapter 5 the results on asymptotics of random permanents are extended to P-statistics, at the same time covering a large class of matchings. The limiting laws are expressed with the help of multiple Wiener-Itô integrals. The basic properties of a multiple Wiener-Itô integral are summarized in the first part of the chapter. Several applications of the asymptotic results to particular counting problems introduced in earlier chapters are presented in detail.

Chapter 6 makes a connection between P-statistics and matchings on one side and the "incomplete" U-statistics on the other. The incomplete permanent design is analyzed first. An overview of the analysis of both asymptotic and finite sample properties of P-statistics in terms of their variance efficiency as compared with the corresponding "complete" statistics is presented. In the second part minimum rectangular designs (much lighter that permanent designs) are introduced and their efficiency is analyzed. Also their relations to the concept of mutual orthogonal Latin squares of classical statistical design theory is discussed there.

Chapter 7 covers some of the recent results on the asymptotic lognormality of sequences of products of increasing sums of independent identically distributed random variables and their U-statistics counterparts. The developments of the chapter lead eventually to a limit theorem for random determinants for Wishart matrices. Here again, similarly as in some of the earlier-discussed limit theorems for random permanents, the lognormal law appears in the limit.

We would like to express our thanks to several individuals and institutions who helped us in completing this project. We would like to acknowledge the IMA director, Doug Arnold who constantly encouraged our efforts, as well as our many other colleagues, especially André Kézdy, Ofer Zeitouni and Shmuel Friedland, who looked at and commented on the various finished and not-so-finished portions of the text. We would also like to thank Ewa Kubicka and Grzegorz Kubicki for their help with drawing some of the graphs presented in the book. Whereas the idea of writing the current monograph was born at the IMA, the opportunity to do so was also partially provided by other institutions. In particular, the Statistical and Applied Mathematical Sciences Institute in Durham, NC held during 2005/6 academic year a program on "Random Matrices and High Dimensional Inference" and kindly invited the first of the authors to participate in its activities as a long term visitor. The project was also supported by local grants from the Faculty of Mathematics and Information Science of the Warsaw University of Technology, Warszawa, Poland and from the Department of Mathematics at the University of Louisville.

Louisville, KY and Warszawa (Poland)
July 2007

Grzegorz A. Rempała
Jacek Wesołowski

Contents

Foreword VII

Preface IX

1 Basic Concepts 1

1.1 Bipartite Graphs in Complex Stochastic Systems 1

1.2 Perfect Matchings 2

1.3 Permanent Function 4

1.4 U-statistics 6

1.5 The H-decomposition 8

1.6 P-statistics 12

1.7 Examples 14

1.8 Bibliographic Details 16

2 Properties of P-statistics 19

2.1 Preliminaries: Martingales 19

2.2 H-decomposition of a P-statistic 21

2.3 Variance Formula for a P-statistic 27

2.4 Bibliographic Details 32

3 Asymptotics for Random Permanents 35

3.1 Introduction 35

3.2 Preliminaries 37

3.2.1 Limit Theorems for Exchangeable Random Variables 37

3.2.2 Law of Large Numbers for Triangular Arrays 40

3.2.3 More on Elementary Symmetric Polynomials 41

3.3 Limit Theorem for Elementary Symmetric Polynomials 43

3.4 Limit Theorems for Random Permanents 45

3.5 Additional Central Limit Theorems 55

3.6 Strong Laws of Large Numbers 59

3.7 Bibliographic Details 65

4 Weak Convergence of Permanent Processes ... 67
4.1 Introduction ... 67
4.2 Weak Convergence in Metric Spaces ... 68
4.3 The Skorohod Space ... 71
4.4 Permanent Stochastic Process ... 74
4.5 Weak Convergence of Stochastic Integrals and Symmetric Polynomials Processes ... 75
4.6 Convergence of the Component Processes ... 78
4.7 Functional Limit Theorems ... 81
4.8 Bibliographic Details ... 86

5 Weak Convergence of *P*-statistics ... 87
5.1 Multiple Wiener-Itô Integral as a Limit Law for U-statistics ... 88
5.1.1 Multiple Wiener-Itô Integral of a Symmetric Function ... 88
5.1.2 Classical Limit Theorems for U-statistics ... 92
5.1.3 Dynkin-Mandelbaum Theorem ... 94
5.1.4 Limit Theorem for U-statistics of Increasing Order ... 96
5.2 Asymptotics for P-statistics ... 100
5.3 Examples ... 107
5.4 Bibliographic Details ... 120

6 Permanent Designs and Related Topics ... 121
6.1 Incomplete U-statistics ... 121
6.2 Permanent Design ... 125
6.3 Asymptotic properties of USPD ... 129
6.4 Minimal Rectangular Schemes ... 134
6.5 Existence and Construction of MRS ... 140
6.5.1 Strongly Regular Graphs ... 140
6.5.2 MRS and Orthogonal Latin Squares ... 142
6.6 Examples ... 144
6.7 Bibliographic Details ... 147

7 Products of Partial Sums and Wishart Determinants ... 149
7.1 Introduction ... 149
7.2 Products of Partial Sums for Sequences ... 151
7.2.1 Extension to Classical U-statistics ... 157
7.3 Products of Independent Partial Sums ... 160
7.3.1 Extensions ... 165
7.4 Asymptotics for Wishart Determinants ... 167
7.5 Bibliographic Details ... 169

References ... 171

Index ... 177