

London Mathematical Society
Lecture Note Series 378

Probability and Mathematical Genetics

Papers in Honour of Sir John Kingman

Edited by

N. H. Bingham and C. M. Goldie

*The London
Mathematical
Society*



CAMBRIDGE

London Mathematic

es series: 378

Pro Mathematical Genetics

Edited by

N. H. BINGHAM

Imperial College London

C. M. GOLDIE

University of Sussex



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press,
New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521145770

© Cambridge University Press 2010

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2010

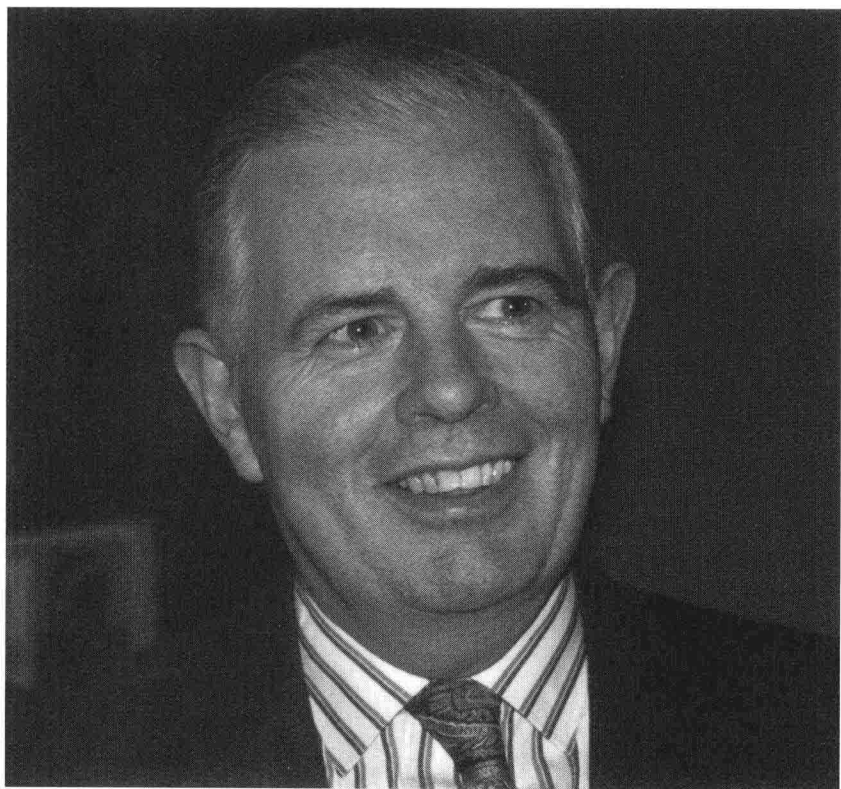
Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-14577-0 Paperback

Additional resources for this publication at www.prob.org.uk

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party Internet websites referred to
in this publication, and does not guarantee that any content on such
websites is, or will remain, accurate or appropriate.



Contributors

- David J. Aldous *University of California at Berkeley*
Graeme K. Ambler *University of Cambridge*
Andrew D. Barbour *University of Zürich*
John D. Biggins *University of Sheffield*
Nicholas H. Bingham *Imperial College London*
Rosie Cornish *University of Bristol*
Daryl J. Daley *Australian National University and University of Melbourne*
Aurore Delaigle *University of Melbourne and University of Bristol*
Peter Donnelly *University of Oxford*
Warren J. Ewens *University of Pennsylvania*
Alexander V. Gnedin *University of Utrecht*
Charles M. Goldie *University of Sussex*
Peter J. Green *University of Bristol*
David R. Grey *University of Sheffield*
Robert C. Griffiths *University of Oxford*
Geoffrey Grimmett *University of Cambridge*
Peter G. Hall *University of Melbourne and University of California at Davis*
Chris Haulk *University of California at Berkeley*
Peter H. Haynes *University of Cambridge*
Viet Ha Hoang *Nanyang Technological University, Singapore*
Frank P. Kelly *University of Cambridge*
Wilfrid S. Kendall *University of Warwick*
Sir John [J. F. C.] Kingman *University of Bristol*
Stephen Leslie *University of Oxford*
James R. Norris *University of Cambridge*
Adam J. Ostaszewski *London School of Economics*
Jim Pitman *University of California at Berkeley*
Carol L. Robinson *Loughborough University*

Vadim Shcherbakov *Moscow State University*
Bernard W. Silverman *University of Oxford*
Dario Spanó *University of Warwick*
Simon Tavaré *University of Cambridge*
Stanislav Volkov *University of Bristol*
Geoffrey A. Watterson *Monash University*
Peter Whittle *University of Cambridge*
David Williams *Swansea University*
Ruth J. Williams *University of California at San Diego*
Konstantinos C. Zygalakis *University of Oxford*

Preface

John Frank Charles Kingman was born on 28th August 1939, a few days before the outbreak of World War II. This Festschrift is in honour of his seventieth birthday.

John Kingman was born in Beckenham, Kent, the son of the scientist Dr F. E. T. Kingman FRSC and the grandson of a coalminer. He was brought up in north London, where he attended Christ's College, Finchley. He was an undergraduate at Cambridge, where at age 19 at the end of his second year he took a First in Part II of the Mathematical Tripos, following it with a Distinction in the graduate-level Part III a year later, for his degree. He began postgraduate work as a research student under Peter Whittle, but transferred to David Kendall in Oxford when Peter left for Manchester in 1961, returning to Cambridge when Kendall became the first Professor of Mathematical Statistics there in 1962.

John's early work was on *queueing theory*, a subject he had worked on with Whittle, but was also an interest of Kendall's. His lifelong interest in *mathematical genetics* also dates back to this time (1961). His next major interest was in Markov chains, and in a related matter—what happens to Feller's theory of recurrent events in continuous time. His first work here dates from 1962, and led to his landmark 1964 paper on regenerative phenomena, where we meet (Kingman) p -functions. This line of work led on to his celebrated characterisation of those p -functions that are diagonal Markov transition probabilities (1971), and to his book, *Regenerative Phenomena* (1972). Meanwhile, he had produced his work on queues in *heavy traffic* (1965). His work on subadditivity began in 1968, and led to the *Kingman subadditive ergodic theorem* of 1973. His genetic interests led to his book *Mathematics of Genetic Diversity* of 1980, and his famous paper on the (Kingman) coalescent of 1982. Later work includes his book *Poisson Processes* of 1993. Other interests include

Spitzer's identity and its connection with queues, the subject of his *The Algebra of Queues* of 1966.

John began his academic career in Cambridge, as Assistant Lecturer (1962–64) and Lecturer (1964–65), with a fellowship at his undergraduate college, Pembroke (1961–65). He left for a Readership at the University of Sussex, where he was promoted to Professor at the very early age of 26 in 1966, the year in which he published his first book, *Introduction to Measure and Probability*, with S. James Taylor. He left Sussex to be Professor at Oxford from 1969–85. He was elected a Fellow of the Royal Society in 1971 at age 31. He was made a Foreign Associate of the US National Academy of Sciences in 2007.

We all know very good mathematicians who could not run a corner sweetshop, let alone a mathematics department, still less a university. On the other hand, mathematicians who are not very bad at administration are often very good at it. John Kingman is a shining example of the latter category. This led to his secondment, while at Oxford, to chair the Science Board of the Science Research Council (1979–81), and later to serve as Chairman of the Science and Engineering Research Council (1981–85), for which he was knighted in 1985. It led also to John's career change in 1985, when he became Vice-Chancellor of the University of Bristol, serving a remarkable sixteen years until 2001. He then served for 5 years as Director of the Isaac Newton Institute for Mathematical Sciences in Cambridge. In 2000 he became the first chairman of the Statistics Commission, overseeing the Office of National Statistics.

John Kingman is the only person who has been President of both the Royal Statistical Society (1987–89) and the London Mathematical Society (1990–92). He has also served as President of the European Mathematical Society (2003–06). He received the LMS Berwick Prize in 1967, the RSS Guy Medal in silver in 1981, and the RS Royal Medal in 1983 (for his work on queueing theory, regenerative phenomena and mathematical genetics). He holds a number of honorary doctorates. He does not hold a PhD, being Mr Kingman until he was made Professor Kingman at Sussex, later taking a Cambridge ScD.

John Kingman's mathematical work is remarkable for both its breadth and its depth. But what shines out from everything he does, whether his written papers and books or his lectures and seminars, is *lucidity*. Kingman is always clear, and lucid. This even extends to his handwriting—small, neat and beautifully legible. The Wiley typesetters who set his 1972 book worked from his handwritten manuscript, which they said was easier to work from than most authors' typescripts. During his Oxford

years, the secretaries there revered him: they were not used to Chairmen of the Mathematical Institute whose desks were tidy, who handled paperwork promptly, and who would give a decision in real time, rather than procrastinate.

John Kingman has been blessed since his marriage in 1964 in the love and support of his distinguished wife Valerie Cromwell; they have a son and a daughter, who are now acquiring distinction themselves. They now live in retirement in Bristol and London.

While probabilists may regret the loss to probability theory of John's years in administration rather than mathematics, this is offset by the continuing impact of his most important work, whether in queueing theory and heavy traffic, Markov chains and regenerative phenomena (the subject of some of his most recent papers, where he has successfully solved some problems that had remained open since his own work of thirty years ago), subadditive ergodic theory or mathematical genetics and the coalescent. Indeed, the intense concentration of effort on the genetics side associated with the Human Genome Project has thrown in to ever higher relief the fundamental importance of Kingman's work in this area. The editors and contributors to this volume take pleasure in dedicating this book to him, on the occasion of his seventieth birthday.

N. H. Bingham and C. M. Goldie, December 2009.

Acknowledgements All contributions to this collection have been refereed. The editors are most grateful to the referees for their efforts, particularly in those cases where a fast response had to be requested.

The *Bibliography of J. F. C. Kingman* (pp. 1–16) was compiled and arranged by Charles Goldie and Jim Pitman. The editors are grateful to Jim for his collaboration, and also thank John Kingman for providing details of his publications that made the task much easier.

The photograph for the Frontispiece is reproduced by courtesy of the Isaac Newton Institute for Mathematical Sciences, Cambridge.

Contents

<i>List of contributors</i>	page	xiii
<i>Preface</i>		xv
<i>Bibliography of J. F. C. Kingman</i>		1
1 A fragment of autobiography, 1957–1967		
<i>J. F. C. Kingman</i>		17
2 More uses of exchangeability: representations of complex random structures		
<i>David J. Aldous</i>		35
1 Introduction		35
2 Exchangeability		36
3 Using exchangeability to describe complex structures		43
4 Construction of, and convergence to, infinite random combinatorial objects		49
5 Limits of finite deterministic structures		56
6 Miscellaneous comments		59
3 Perfect simulation using dominated coupling from the past with application to area-interaction point processes and wavelet thresholding		
<i>G. K. Ambler and B. W. Silverman</i>		64
1 Introduction		65
2 Perfect simulation		66
3 Area-interaction processes		71
4 Nonparametric regression by wavelet thresholding		77
5 Perfect simulation for wavelet curve estimation		82
6 Conclusion		88
4 Assessing molecular variability in cancer genomes		
<i>A. D. Barbour and S. Tavaré</i>		91

1	Introduction	91
2	Colorectal cancer data	93
3	The Ewens sampling formula	95
4	Analysis of the cancer data	98
5	Poisson approximation	100
6	Conclusion	110
5	Branching out	
	<i>J. D. Biggins</i>	113
1	Introduction	113
2	The basic model	115
3	Spreading out: old results	116
4	Spreading out: first refinements	119
5	Spreading out: recent refinements	120
6	Deterministic theory	122
7	The multitype case	124
8	Anomalous spreading	125
9	Discussion of anomalous spreading	130
6	Kingman, category and combinatorics	
	<i>N. H. Bingham and A. J. Ostaszewski</i>	135
1	Introduction	135
2	Preliminaries	138
3	A bitopological Kingman theorem	143
4	Applications—rational skeletons	151
5	KBD in van der Waerden style	157
6	Applications: additive combinatorics	162
7	Long-range dependence in a Cox process directed by an alternating renewal process	
	<i>D. J. Daley</i>	169
0	Preamble	170
1	Introduction	170
2	Stationary renewal and alternating renewal processes	172
3	Second moments	178
4	An alternating renewal process not satisfying Condition A	180
5	Postlude	181
8	Kernel methods and minimum contrast estimators for empirical deconvolution	
	<i>Aurore Delaigle and Peter Hall</i>	185
1	Introduction	186

2	Methodology and theory	191
3	Relationship to minimum contrast methods	195
9	The coalescent and its descendants	
	<i>Peter Donnelly and Stephen Leslie</i>	204
1	Introduction	205
2	The coalescent and the Fleming–Viot process	206
3	Inference under the coalescent	213
4	The Li and Stephens model	215
5	Application: modelling population structure	221
10	Kingman and mathematical population genetics	
	<i>Warren J. Ewens and Geoffrey A. Watterson</i>	238
1	Introduction	238
2	Background	239
3	Putting it together	244
4	Robustness	247
5	A convergence result	248
6	Partition structures	249
7	‘Age’ properties and the GEM distribution	251
8	The coalescent	256
9	Other matters	260
11	Characterizations of exchangeable partitions and random discrete distributions by deletion properties	
	<i>Alexander Gnedin, Chris Haulk and Jim Pitman</i>	264
1	Introduction	265
2	Partition structures	266
3	Partially exchangeable partitions	274
4	Exchangeable partitions	278
5	The deletion property without the regularity condition	285
6	Regeneration and τ -deletion	286
12	Applying coupon-collecting theory to computer-aided assessments	
	<i>C. M. Goldie, R. Cornish and C. L. Robinson</i>	299
1	Introduction	299
2	Coupon collecting	300
3	How many tests?	302
4	Asymptotics	303
5	Proofs for §4	305
6	Numerical results	314

7	Discussion	315
13	Colouring and breaking sticks: random distributions and heterogeneous clustering	
	<i>Peter J. Green</i>	319
1	Introduction	319
2	Mixture models and the Dirichlet process	320
3	Applications and generalisations	325
4	Pólya urn schemes and MCMC samplers	329
5	A coloured Dirichlet process	333
14	The associated random walk and martingales in random walks with stationary increments	
	<i>D. R. Grey</i>	345
1	Introduction and definition	345
2	Three examples	349
3	Some remarks on duality and asymptotic independence	355
15	Diffusion processes and coalescent trees	
	<i>R. C. Griffiths and D. Spanó</i>	358
1	Introduction	359
2	A coalescent dual process	361
3	Processes with beta stationary distributions and Jacobi polynomial eigenfunctions	368
4	Subordinated Jacobi diffusion processes	371
5	Subordinated coalescent process	375
16	Three problems for the clairvoyant demon	
	<i>Geoffrey Grimmett</i>	380
1	Introduction	380
2	Site percolation	381
3	Clairvoyant scheduling	383
4	Clairvoyant compatibility	384
5	Clairvoyant embedding	385
6	Dependent percolation	390
7	Percolation of words	393
17	Homogenization for advection-diffusion in a perforated domain	
	<i>P. H. Haynes, V. H. Hoang, J. R. Norris and K. C. Zygalakis</i>	397
1	Introduction	398

2	Review of homogenization for diffusion with periodic drift	400
3	Existence of a volume growth rate for a diffusion sausage with periodic drift	402
4	Estimates for the diffusion sausage	403
5	Asymptotics of the growth rate for small and large cross-sections	405
6	Homogenization of the advection-diffusion equation in a perforated domain	408
7	The case of diffusivity $\varepsilon^2 I$	410
8	Monte Carlo computation of the asymptotic growth rate	411
18	Heavy traffic on a controlled motorway	
	<i>F. P. Kelly and R. J. Williams</i>	416
1	Introduction	416
2	A single queue	418
3	A model of Internet congestion	422
4	A Brownian network model	426
5	A model of a controlled motorway	430
6	Route choices	438
7	Concluding remarks	442
19	Coupling time distribution asymptotics for some couplings of the Lévy stochastic area	
	<i>W. S. Kendall</i>	446
1	Different kinds of couplings	448
2	Reflection coupling	450
3	Coupling more than one feature of the process	451
4	Conclusion	461
20	Queueing with neighbours	
	<i>V. Shcherbakov and S. Volkov</i>	464
1	Introduction	464
2	Results	467
3	Asymmetric interaction	470
4	Symmetric interaction	475
5	Appendix	480
21	Optimal information feed	
	<i>P. Whittle</i>	483
1	Interrogation, transmission and coding	483
2	A tractable infinite-horizon case	486

22 A dynamical-system picture of a simple branching-process phase transition***David Williams*** 491

1 Introduction 491

2 Wiener–Hopferization 493

3 How does ODE theory see the phase transition? 496

4 Proof of Theorem 1.1 and more 499

Index 509

Bibliography of J. F. C. Kingman

Compiled by Charles M. Goldie and Jim Pitman^a

- 1 Books authored
- 2 Books edited
- 3 Mathematical articles
- 4 Abstracts
- 5 Book reviews authored
- 6 Discussion contributions
- 7 Other contributions
- 8 Interviews and biographies

1 Books authored

- [B1] Kingman, J. F. C., and Taylor, S. J. 1966. *Introduction to Measure and Probability*. Cambridge: Cambridge University Press.
- [B2] Kingman, J. F. C. 1966. *On the Algebra of Queues*. Methuen's Supplementary Review Series in Applied Probability, vol. 6. London: Methuen & Co. Ltd. Reprint of [M36].
- [B3] Kingman, J. F. C. 1972. *Regenerative Phenomena*. Wiley Series in Probability and Mathematical Statistics. London: John Wiley & Sons.
- [B4] Kingman, J. F. C. 1980. *Mathematics of Genetic Diversity*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 34. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).
- [B5] Kingman, J. F. C. 1993. *Poisson Processes*. Oxford Studies in Probability, vol. 3. Oxford: Oxford University Press.

^a This bibliography was prepared using the BibServer system developed by Jim Pitman with the assistance of NSF Award 0835851, Bibliographic Knowledge Network.

- [B6] Kingman, J. F. C. 2002. *Procesy Poissona*. Wydawnictwo Naukowe PWN. Polish translation of [B5] by Bobrowski, A.

2 Books edited

- [E1] Ibragimov, I. A., and Linnik, Yu. V. 1971. *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff Publishing. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov. Translation from the Russian edited by J. F. C. Kingman.
- [E2] Bodmer, W. F., and Kingman, J. F. C. (eds). 1983. *Mathematical Genetics: Proceedings of a Royal Society Meeting held in London, April 20, 1983*. London: Royal Society.
- [E3] Kingman, J. F. C., and Reuter, G. E. H. (eds). 1983. *Probability, Statistics and Analysis*. London Math. Soc. Lecture Note Ser., vol. 79. Cambridge: Cambridge University Press. Papers dedicated to David G. Kendall on the occasion of his sixty-fifth birthday.
- [E4] Kendall, D. G., with the assistance of Kingman, J. F. C. and Williams, D. (ed). 1986. *Analytic and Geometric Stochastics: Papers in Honour of G. E. H. Reuter*. Adv. in Appl. Probab., vol. 18 (Supplement). Sheffield: Applied Probability Trust.

3 Mathematical articles

- [M1] Kingman, J. F. C. 1960. A note on the axial symmetry of the disturbance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.*, **258**, 87–89.
- [M2] Yih, Chia-Shun (with an appendix by Kingman, J. F. C.). 1960. Instability of a rotating liquid film with a free surface. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.*, **258**, 63–89. Expanded version of [M1].
- [M3] Kingman, J. F. C. 1961. A convexity property of positive matrices. *Quart. J. Math. Oxford Ser. (2)*, **12**, 283–284.
- [M4] Kingman, J. F. C. 1961. A mathematical problem in population genetics. *Proc. Cambridge Philos. Soc.*, **57**, 574–582.
- [M5] Kingman, J. F. C. 1961. On an inequality in partial averages. *Quart. J. Math. Oxford Ser. (2)*, **12**, 78–80.
- [M6] Kingman, J. F. C. 1961. The ergodic behaviour of random walks. *Biometrika*, **48**, 391–396.
- [M7] Kingman, J. F. C. 1961. The single server queue in heavy traffic. *Proc. Cambridge Philos. Soc.*, **57**, 902–904.
- [M8] Kingman, J. F. C. 1961. Two similar queues in parallel. *Ann. Math. Statist.*, **32**, 1314–1323.
- [M9] Kingman, J. F. C. 1962. On queues in heavy traffic. *J. Roy. Statist. Soc. Ser. B*, **24**, 383–392.