

The Open University Statistics Department/Group 2007 Report

1. Membership

Professor P.H. Garthwaite (Head of Department to March)
Professor C.P. Farrington (Head of Department from March to September,
Head of Group from October)
Dr C.J. Albers
Dr K. Anaya
Professor F. Critchley
Dr A.E. Faria
Professor J.C. Gower
Dr M. Hocine (from February)
Dr G.E. Iossif
Professor M.C. Jones
Mr I.W. Martin
Dr K.J. McConway
Dr P. Musonda (to July)
Dr C.M. Queen
Dr N.T. Trendafilov
Dr S.K. Vines
Dr H.J. Whitaker

Long-term visitors

Dr P.C. Taylor (University of Hertfordshire)

Short-term visitors

Dr S. Kato (Keio University, Japan)
Professor N. Le Roux (University of Stellenbosch, South Africa)
Professor P.K. Marriott (University of Waterloo, Canada)
Professor P.W. Vos (East Carolina University, U.S.A.)

Full-time research students

Mr S. De
Mr D. Gragn (from October)
Mr D.J. Jenkinson (to September)
Mr E. Mubwandarikwa
Ms A. Noufaily (from October)
Mr S. Unkel

Part-time research students

Ms H. Beaumont (to November)
Mr N. Calleja
Ms A. Gjini
Mr G. Kafatos
Mr M. Merens (from October)
Mr E. Mubwandarikwa

Mr A. Owen
Mr J. Urquhart (joint with Applied Mathematics, from October)

Support staff

Mrs S. Frain, Departmental Secretary (from May to September), Group Secretary (from October)
Mrs W. Hunt, temporary Departmental Secretary (to July)

2. Introduction

2007 was, inevitably, another active year for Statistics at the Open University in both research and teaching. Regrettably, in October the Department of Statistics per se ceased to be. Instead, we are now a Statistics Group within a large Department of Mathematics & Statistics (within a merged Faculty of Mathematics, Computing and Technology). These unnecessary changes might, at best, prove neutral.

One aspect of our ongoing success is stability in academic personnel. Heather Whitaker has a further new appointment as (full-time, permanent) Lecturer in the Department, sorry, Group. In May, we welcomed our new Groupal Secretary, Sarah Frain; Sarah took over from Wendy Hunt who is warmly thanked for admirable support as temporary cover. Kevin McConway failed to take the opportunity to relinquish his role as an Associate Dean in whatever Faculty we happen to be in at the time.

The introduction of [M249](#) *Practical Modern Statistics* (see Section 3) appears to have been as successful as any new course of ours has ever been. It scored outstandingly well in terms of student retention and success and the responses of students when surveyed. M249 is one of five 30 point courses all told, three at second level and two at third, which we run along with a 10 point course at Level 1 and continued contributions to other courses in Faculty and University. The Department's courses attracted a total of almost 2500 students.

In the course of the year, members of the Department published 19 papers in refereed journals (see Section 6), plus contributions to conference proceedings and the like. Pre-publication papers and other research material can be found in our technical report series, at <http://statistics.open.ac.uk/TechnicalReports/TechnicalReportsIntro.htm>. Of course, (too much) time was spent on preparing our RAE submission.

Alvaro Faria (and Paul Garthwaite when here; see Section 12) continued to run a Statistical Advisory Service to support research in other departments at the Open University. A new website explains the service: <http://www.mathematics.open.ac.uk/advisory>.

Further information on the Department is available through our web pages at <http://statistics.open.ac.uk/index.html>

3. Teaching

Presentation

The Department presented five courses of 30 CATS points each, namely, [BM240](#) *Quantitative Methods in Business*, [M248](#) *Analysing Data*, [M249](#) *Practical Modern Statistics*, [M343](#) *Applications of Probability* and [M346](#) *Linear Statistical Modelling*, and one course of 10 CATS points, [SMK184](#) *Chance, Risk and Health*, which runs four times a year.

The Department's hard work on the production of our new second-level course in Statistics, [M249](#), paid off handsomely, as mentioned in Section 2. The course consists of four blocks on the topics of Bayesian statistics, medical statistics, multivariate statistics and time series analysis, respectively. M249 rounds off our undergraduate Diploma in Statistics (by replacing the previous elective element) and also forms a core component of other degrees involving Statistics, along with M343, M346 and one of BM240 or M248.

Each of the Faculty of Mathematics and Computing's first level mathematics courses, [MU120](#) *Open Mathematics* and [MST121](#) *Using Mathematics*, contains substantial amounts of statistical material (roughly one quarter of each) with which Alvaro Faria was heavily involved (and Paul Garthwaite for a while).

Kevin McConway contributed to the presentation of [U205](#) *Health and Disease* and, from October, to the presentation of the new course [SDK125](#) *Introducing Health Science*.

Production

Karen Vines worked hard on updating (both computationally and presentationally) the content of M346, the revised version to go out to students in 2009.

Chris Jones collaborated with Alan Graham (Mathematics Education) on the production of a unit on Statistics for MU123 *Discovering Mathematics*, the forthcoming replacement for MU120.

Kevin McConway continued to contribute to the production of [SDK125](#) *Introducing Health Science*.

Awards

Our courses are core components of degrees in [Mathematics and Statistics](#) and in [Computing and Statistics](#) which were introduced during 2004. These awards had their first graduates at the end of 2007, but numbers are not yet available. Our courses are also core components of the existing degrees in Mathematical Sciences and in Economics and Mathematical Sciences, while BM240 is part of the BA in Business Studies.

The undergraduate [Diploma in Statistics](#) is now awarded to students who successfully study M248 or BM240, M249, M343 and M346 (or their predecessors). Around 100 students are awarded the Diploma each year.

4. Higher degrees awarded

Helen Beaumont (M.Phil.) for her thesis *Modelling Rates of Cognitive Decline in Patients with a Dementing Illness* (supervisors: A.D. Lunn, Oxford University, and P.H. Garthwaite).

David Jenkinson (Ph.D.) for his thesis *Quantifying Expert Opinion as a Probability Distribution* (joint with Computing; supervisors: P.H. Garthwaite and R.C. Laney, Computing).

5. Research interests

The Department is home to three Research Groups:

Multivariate Statistics (Critchley (chair), Albers, Anaya, Faria, Garthwaite, Gower, Jones, Queen, Taylor, Trendafilov, Unkel, Vines, Constantine (CSIRO, retired), Cook (University of Minnesota), Marriott (University of Waterloo)).

Bayesian Statistics (Queen (chair), Albers, De, Faria, Farrington, Garthwaite, Jenkinson, McConway, Vines, Whitaker).

Medical Statistics (Whitaker (chair), Farrington, Garthwaite, Hocine, Jenkinson, Musonda, Vines).

Clearly, these three general groupings cover a huge variety of research projects in Statistics. Inter alia, they organise informal internal research meetings. Research also continues, of course, on topics not directly covered by these Research Groups, of which distribution theory, kernel smoothing, robust statistics, time series forecasting and influence analysis are but a small selection. Strong ongoing collaborations continue in areas such as medicine, psychology and ecology.

For details, see the publications, talks and other items to follow.

6. Publications

6.1 Publications in refereed academic journals

Anaya-Izquierdo, K. & Marriott, P. (2007) Local mixture models of exponential families. *Bernoulli*, 13, 623-640.

Anaya-Izquierdo, K.A. & Marriott, P.K. (2007) Local mixtures of the exponential distribution. *Annals of the Institute of Statistical Mathematics*, 59, 111-134.

Arnold, G.M., **Gower, J.C.**, Gardner, S. & LeRoux, N.J. (2007) Biplots of free-choice-profile data in generalized Procrustes analysis. *Applied Statistics*, 56, 445-458.

Banks, C.J., Brandon, M.A. & **Garthwaite, P.H.** (2006) Measurement of sea ice draft using upward looking ADCP on an autonomous underwater vehicle. *Annals of Glaciology*, 44, 211-216.

Boente, G., **Critchley, F.**, & Orellana, L. (2007) Influence functions for two families of robust estimators under proportional scatter matrices. *Statistical Methods and Applications*, 15, 295-327.

Crawford, J.R. & **Garthwaite, P.H.** (2007) Using regression equations built from summary data in the assessment of the individual case. *Neuropsychology*, 21, 611-621.

Crawford, J.R. & **Garthwaite, P.H.** (2007) Comparison of a single case to a control or normative sample in neuropsychology: development of a Bayesian approach. *Cognitive Neuropsychology*, 24, 343-372.

Crawford, J.R., **Garthwaite, P.H.** & Gault, C.B. (2007) Estimating the percentage of the population with abnormally low scores (or score differences) on standardized neuropsychological test batteries: a generic method with applications. *Neuropsychology*, 21, 419-430.

Fréville, H., **McConway, K.**, Dodd, M. & Silvertown, J. (2007) Prediction of extinction in plants: interaction of extrinsic threats and life history traits. *Ecology*, 88, 2662-2672.

Hocine, M., Farrington, C.P., Touzé, E., **Whitaker, H.J.,** Fourrier, A., Moreau, T. & Tubert-Bitter, P. (2007) Hepatitis B vaccination and first central nervous system demyelinating events: reanalysis of a case-control study using the self-controlled case series method. *Vaccine*, 25, 5938-5943.

Hocine, M.N., Tubert-Bitter, P., Moreau, T., Chavance, M., Varon, E. & Guillemot, D. (2007) Relative risk-ratio was a useful measure of differential association in cohort and case-series studies. *Journal of Clinical Epidemiology*, 60, 361-365.

Jones, M.C. (2007) Connecting distributions with power tails on the real line, the half line and the interval. *International Statistical Review*, 75, 58-69.

Jones, M.C. (2007) On a class of distributions defined by the relationship between their density and distribution functions. *Communications in Statistics – Theory and Methods*, 36, 1835-1843.

Jones, M.C. & Henderson, D.A. (2007) Kernel-type density estimation on the unit interval. *Biometrika*, 94, 977-984.

Jones, M.C. & Yu, K. (2007) Improved double kernel local linear quantile regression. *Statistical Modelling*, 7, 377-389.

Pewsey, A., Lewis, T. & **Jones, M.C.** (2007) The wrapped t family of circular distributions. *Australian and New Zealand Journal of Statistics*, 49, 79-91.

Queen, C.M., Wright, B.J. & **Albers, C.J.** (2007) Eliciting a directed acyclic graph for a multivariate time series of vehicle counts in a traffic network. *Australian and New Zealand Journal of Statistics*, 49, 221-239.

Trendafilov, N.T. & Jolliffe, I.T. (2007) DALASS: variable selection in discriminant analysis via the LASSO. *Computational Statistics and Data Analysis*, 51, 3718-3736.

Whitaker, H.J., Hocine, M.N. & Farrington, C.P. (2007) On case-crossover methods for environmental time series data. *Environmetrics*, 18, 157-171.

6.2 Refereed conference proceedings and book chapters

Albers, C.J., Critchley, F. & Gower, J.C. (2007) Group average representations in Euclidean distance cones. In *Selected Contributions in Data Analysis and Classification*, eds: P. Brito, P. Bertrand, G. Cucumel, & F. de Carvalho, Springer, pp. 445-454.

Farrington, C.P., Whitaker, H.J. & Hocine, M.N. (2007) Statistical issues in vaccine safety evaluation. In *Encyclopedia of Statistics in Quality and Reliability*, eds: F. Ruggieri, R.S. Kenett & F. Faltin, Wiley, pp. 1888-1894.

Hocine, M.N., Whitaker, H.J. & Farrington, C.P. (2007) Case series analysis. In *Encyclopedia of Statistics in Quality and Reliability*, eds: F. Ruggieri, R.S. Kenett & F. Faltin, Wiley, pp. 1783-1788.

Jolliffe, I.T., Learmonth, J.A., Pierce, G.J., Santos, M.B., **Trendafilov, N.,** Zuur, A.F., Ieno, E.N. & Smith, G.M. (2007) Principal component analysis applied to harbour porpoise fatty acid data. In *Analysing Ecological Data*, eds: A.F. Zuur, E.N. Ieno & G.M. Smith, Springer, pp. 515-528.

Unkel, S. & Trendafilov, N.T. (2007) Noisy independent component analysis as a method of rotating the factor scores. In Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation (ICA2007), eds: M.E. Davies et al., Springer, pp. 810-817.

7. Seminars and conferences at The Open University

The 2007 instalment of our regular series of conferences was **HDMDA, High Dimensional Multivariate Data Analysis**, chief organiser Nickolay Trendafilov (ably assisted by Wendy Hunt). Speakers at this one-day meeting were:

John Gower (Open University) *Quadratic minimisation problems in statistics*

Ludovic Lebart (Centre National de la Recherche Scientifique, Paris) *Resampling techniques for assessing the visualisations of high-dimensional data*

Gilbert Saporta (Conservatoire National des Arts et Métiers, Paris) *Linear techniques for regression and classification on functional data*

Tom Fearn (University College London) *Orthogonal signal pretreatment*

Mark Plumbley (Queen Mary, University of London) *Geometry of sparse representations*

Frank Critchley (Open University) *Principal axis analysis*

The department ran its usual seminar programme with invited speakers. Over the year we heard:

Frank Ball (University of Nottingham) *Statistical inference for epidemics among a population of households*

Adrian Bowman (University of Glasgow) *Statistics with a human face*

Steffen Lauritzen (University of Oxford) *Graphical models with edge and vertex symmetries*

Christl Donnelly (Imperial College London) *TB in cattle and badgers - An epidemiological challenge*

Richard Chandler (University College London) *Space-time modelling using independence and generalised estimating equations*

Tom Snijders (Oxford University) *Statistical modelling of social network dynamics*

We also ran our fourth Annual Research Students' Research Day, where speakers were Steffen Unkel, Swarup De, Neville Calleja and David Jenkinson.

And at an internal research morning in March, speakers were Paul Garthwaite, Chris Jones, Nickolay Trendafilov and Karen Vines.

8. Conference papers and posters presented

C.J. Albers

International Meeting of the Psychometric Society (IMPS2007), Tokyo, Japan *Minimising translated quadratic forms subject to quadratic constraints, with applications to missing value estimation in Procrustes analysis*

F. Critchley

High Dimensional Multivariate Data Analysis (HDMDA), Open University *Principal axis analysis*

C.P. Farrington

28th Annual Conference of the International Society for Clinical Biostatistics, Alexandropolis, Greece *Bias in case-crossover analyses of environmental time series data*

P.H. Garthwaite

SAMSI Workshop on Risk Analysis, Extreme Events and Decision Theory, Research Triangle Park, North Carolina, U.S.A. *Elicitation of expert opinion as a prior distribution*

Spring Bayes Conference, Coolangatta, Queensland, Australia *Selection of prior weights for weighted model averaging*

J.C. Gower

Correspondence Analysis and Related Methods, Rotterdam, Netherlands *Enhanced biplots*

High Dimensional Multivariate Data Analysis (HDMDA), Open University *Quadratic minimisation problems in statistics*

International Meeting of the Psychometric Society (IMPS2007), Tokyo, Japan *Matrices do matter* [Major Invited Talk]

International Meeting of the Psychometric Society (IMPS2007), Tokyo, Japan *Asymmetry analysis: the place of models*

International Meeting of the Psychometric Society (IMPS2007), Tokyo, Japan *Generalised Procrustes analysis with optimal scaling: exploring transactional data and customer satisfaction with power supplier*

M. Hocine

Vaccine Safety Evaluation: Post-marketing Surveillance Conference, Bethesda, Maryland, U.S.A. *Sequential methodology in case series studies*

28th Annual Conference of the International Society for Clinical Biostatistics, Alexandropolis, Greece *Hepatitis B vaccination and multiple sclerosis: case-control and case series methods*

23rd International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Quebec, Canada *Prospective surveillance of vaccine safety using the self-controlled case series method* [poster]

M.C. Jones

International Indian Statistical Association Joint Statistical Meeting and International Conference on Statistics, Probability and Related Areas, Cochin, India *Jones's Johnson distributions*

P. Musonda

Conference for Applied Statistics for Development in Africa (SADA 07), Cotonou, Benin
Application of the self-controlled case series method in surveillance

N.T. Trendafilov

International Workshop on Computational and Financial Econometrics, Geneva, Switzerland
Continuous-time approach to common principal components estimation

S. Unkel

30th Research Students' Conference in Probability and Statistics (RSC 2007), Durham
Noisy independent component analysis as a method of rotating the factor scores

7th International Conference on Independent Component Analysis and Signal Separation (ICA 2007), London
Noisy independent component analysis as a method of rotating the factor scores

H.J. Whitaker

Society for Academic Primary Care Annual Scientific Meeting, London
Using cases as their own controls: the self-controlled case series method

Royal Statistical Society Conference, Statistics and Policy Making, York
A measure of assortativeness for directly transmitted infections [poster]

9. Other seminars and talks given

C.P. Farrington

Lancaster University *Case series models for censoring events*

(with H.J. Whitaker) POLYMOD (EU Project) Workshop, Rome, Italy
Measures of assortativeness, and their estimation from serological and contact data

Zurich University, Switzerland *The self-controlled case series method*

Zurich University, Switzerland *Case series models for censoring events*

Healthcare Commission, London *Prospective surveillance of vaccine safety*

Université Victor Segalen Bordeaux 2, France
Analyse des séries de cas: principe et applications en pharmacoépidémiologie

P.H. Garthwaite

University of Edinburgh *Selection of weights for weighted model averaging*

University of Durham *Selection of weights for weighted model averaging*

University of New South Wales, Sydney, Australia *The Open University*

University of New South Wales, Sydney, Australia *Selection of prior weights for weighted model averaging*

New South Wales Branch of the Statistical Society of Australia, Sydney, Australia *Elicitation methods for quantifying expert opinion*

J.C. Gower

Western Cape Chapter of the South African Statistical Association. Stellenbosch, South Africa *A matrix miscellany*

25th Birthday Meeting of the Herts and Beds Local Group of the Royal Statistical Society, Rothamsted Experimental Station, Harpenden *Science, Statistics, computing and mathematics*

M.C. Jones

Highlands Local Group of the Royal Statistical Society, St Andrews *Don't mention Stonehenge! The statistician's side of the story of a ten year collaboration between archaeology, earth sciences and statistics*

Newcastle University *How parametric distributions assist with some nonparametric smoothing problems*

K.J. McConway

Junior Statisticians' Group, Department of Work and Pensions, London *Statistics and the media*

H.J. Whitaker

(with C.P. Farrington) POLYMOD (EU Project) Workshop, Rome, Italy *Measures of assortativeness, and their estimation from serological and contact data*

University College London *Self-controlled case series method* [STATA practical]

10. Editorial roles

F. Critchley

- Associate Editor: *Journal de la Société Française de Statistique*
- Editorial Board Member: *Studies in Classification, Data Analysis, and Knowledge Organization* (book series, Springer)

C.P. Farrington

- Editorial Board Member: *Epidemiology and Infection*

J.C. Gower

- Associate Editor: *The Mathematical Scientist*

M.C. Jones

- Associate Editor: *Annals of the Institute of Statistical Mathematics*
- Associate Editor: *Communications in Statistics*
- Associate Editor: *Journal of Statistical Planning and Inference*
- Associate Editor: *Pakistan Journal of Statistics*
- Associate Editor: *Sankhya*
- Associate Editor: *Statistica Sinica*
- Associate Editor: *Statistical Methodology*

11. Grants Awarded

K.J. McConway

Department of the Environment, Food and Rural Affairs grant for project entitled "Tools, Training and research for managing Eco-Hydrology of Cape Flora". £325,540. Co-investigator with colleagues from the Open University's Department of Biological Sciences, the South African National Biodiversity Institute and Cape Nature

12. Other activities

F. Critchley

- research visitor at each of: Instituto Superior Técnico, Lisbon, Portugal; Université Libre de Bruxelles, Belgium; and Université de Liège, Belgium

C.P. Farrington

- elected onto the RSS Council
- member of the WHO Ad-Hoc Committee on EPI Serology in Relation to Intermittent Preventive Treatment in Infants against Malaria
- member of the Wellcome Trust's Study Design Expert Group (from October)
- external examiner for the MSc in Modern Epidemiology at Imperial College London
- examined PhDs at Imperial College London and University College London
- member of the Organising Committee of the Conference for Applied Statistics for Development in Africa (SADA 07), Cotonou, Benin

P.H. Garthwaite

- visiting academic in the Department of Mathematics and Statistics, University of New South Wales, Sydney, Australia (from April)
- external examiner for undergraduate statistics courses at the University of Mauritius, yes, Mauritius
- member of the MRC Panel of Experts
- examined a PhD for the University of Durham

J.C. Gower

- research visitor at University of Stellenbosch, South Africa
- examined PhD at the University of Leiden, Netherlands
- member of the Sir Ronald Aylmer Fisher Memorial Committee of Great Britain
- foreign member of the Advisory Board of IOPS (Interuniversitaire Onderzoekschool voor Psychometrie en Sociometrie)
- organiser of session on Procrustes Problems at the International Meeting of the Psychometric Society, Tokyo, Japan

M.C. Jones

- academic reviewer of subject guides for courses on Distribution Theory and Statistical Inference for London School of Economics external students
- examined PhDs for Imperial College London and the University of Sheffield

K.J. McConway

- Associate Dean (Curriculum) of the Faculty of Mathematics and Computing, transmogrifying into Associate Dean (Curriculum and Awards) of the Faculty of Mathematics, Computing and Technology in October
- external examiner for the MSc in Statistics at the University of Kent
- academic consultant on two more series of the BBC Radio4/OU series 'More or Less', including a contribution on air to one programme talking about UNICEF reports on children and young people. He also wrote material for the programme-related website at <http://www.open2.net/moreorless/>

C.M. Queen

- external examiner for a PhD at the University of Sunderland

N.T. Trendafilov

- organiser and chair of session on Matrix Computations and Statistics at the meeting of the ERCIM Working Group on Computing & Statistics within the International Workshop on Computational and Financial Econometrics in Geneva, Switzerland