The Open University Statistics Department 2006 Report

Membership

Professor P.H. Garthwaite (Head of Department from July)
Dr K.J. McConway (Head of Department to June)

Dr C.J. Albers

Dr K. Anaya
Professor F. Critchley
Dr A.E. Faria
Professor C.P. Farrington

Professor J.C. Gower
Dr G.E. Iossif
Professor M.C. Jones

Mr I.W. Martin
Dr P. Musonda (from October)
Dr C.M. Queen
Dr N.T. Trendafilov
Dr S.K. Vines
Dr H.J. Whitaker

Long-term visitors

Dr M. Hocine (INSERM U780 Epidemiologie et Biostatistique, Paris)

Dr P.C. Taylor (University of Hertfordshire)

Short–term visitors

Mr E.B. Ferreira (Universidade Federal de Lavras, Brazil)

Professor H. Oja (University of Tampere, Finland)

Professor D.H. Tyler (Rutgers University, USA)

Full–time research students

Mr C.J. Banks (joint with Earth Sciences, to November)

Mr S. De

Mr D.J. Jenkinson

Mr E. Mubwandarikwa (to September)

Mr P. Musonda (to September)

Mr A. Sarkar (joint with Computing, to October)

Mr S. Unkel (from October)

Part–time research students

Ms H. Beaumont

Mr N. Calleja (from October)

Ms A. Gjini

Mr G. Kafatos

Mr E. Mubwandarikwa (from October)

Mr A. Owen

Mrs L. Sun (to February)

Support staff

Mrs V. Spearman, Departmental Secretary (to August)

2. Introduction

2006 was, inevitably, another active year for the Department in both research and teaching. Regrettably, this is the last time that this form of words can be used because at some time in 2007 the Department of Statistics will cease to be: we will be incorporated with the current Department of Mathematics into a single new, large, department. There is little logical reason for this change and few obvious positives to be gained; we can only hope that negatives will be few and minor. This is an enormous pity for a thriving, excellent and happy Department (and for the discipline of Statistics).

One aspect of our success is stability in academic personnel. Relatively small changes are that Heather Whitaker's new appointment (from October) is as half-and-half lecturer and research fellow, while Patrick Musonda became a research fellow on completion of his research studentship. Unfortunately, we lost our popular and efficient secretary Val Spearman to retirement and a new life in Spain! Kevin McConway relinquished his Headship of the Department in mid-term to be crowned Associate Dean (Curriculum) in the Faculty of Mathematics and Computing.

Particular success was achieved this year in obtaining substantial research grants from the Research Councils (see Section 11).

In the course of the year, members of the Department published a book and another 26 papers in refereed journals (including one *Royal Statistical Society* read paper) (see Section 6). Pre-publication papers and other research material can be found in our technical report series, at
http://statistics.open.ac.uk/TechnicalReports/TechnicalReportsIntro.htm

Our teaching continued to flourish. We ran five 30 point courses, three at second level and two at third, a new 10 point course at Level 1, and continued to contribute to other courses in Faculty and University. The Department's courses attracted a total of about 2,400 students. We also completed a major new course production effort (details below).

Alvaro Faria and Paul Garthwaite continued to run a Statistical Advisory Service to support research in other departments at the Open University.

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*, MDST242 *Statistics in Society*, M248 *Analysing Data*, M343 *Applications of Probability* and M346 *Linear Statistical Modelling*, and, for the first time, one course of 10 CATS points, SMK184 *Chance, Risk and Health*, which runs four times a year.

This year was the last in which MDST242 was presented. This venerable course – partially rewritten during the course of its life – has been presented every year from 1983 (a total of 24 times) and has been studied by almost 14,000 students.

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 Paul Garthwaite and Alvaro Faria were involved.

Kevin McConway contributed to the presentation of U205 *Health and Disease*.

Production

The final year's hard work on the production of our new second-level course in Statistics, M249, entitled *Practical Modern Statistics,* took place under the chairing of Paddy Farrington and the academic editing of Gillian Iossif. The course is being presented for the first time in 2007. It 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 elective element; see below) and will also form a core component of other degrees involving Statistics.

Kevin McConway contributed 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 and will have their first graduates in 2007. They 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 currently awarded to students who successfully study M248 or BM240, M343 and M346 (or their predecessors) plus one other course from a list of electives. Around 100 students are awarded the Diploma each year.

4.   PhDs awarded

Christopher Banks for his thesis *Sea Ice Thickness and Iceberg Distribution in the Southern Ocean* (joint with Earth Sciences; supervisors: M.A. Brandon, Earth Sciences, and P.H. Garthwaite).

Patrick Musonda for his thesis *The Self-Controlled Case Series Method: Performance and Design in Studies of Vaccine Safety* (supervisors: C.P. Farrington and F. Critchley).

Avik Sarkar for his thesis *Term Burstiness: Evidence, Model and Applications* (joint with Computing; supervisors: A. DeRoeck, Computing, and P.H. Garthwaite).

Linjuan Sun for her thesis *Simple Principal Components* (supervisors: S.K. Vines and F. Critchley).

5.   Research interests

The Department is home to three Research Groups:

Multivariate Statistics (Critchley (chair), Albers, 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, Faria, Farrington, Garthwaite, McConway, Vines, Whitaker).

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

Clearly, these three general groupings cover a huge variety of research projects in Statistics. 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 Book

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R**., Garthwaite, P.H., Jenkinson, D.J.,** Oakley, J.E. & Rakow, T. (2006) *Uncertain Judgements: Eliciting Expert Probabilities.* Wiley, Chichester.

6.2 Publications in refereed academic journals

Al-Awadhi, S.A. & **Garthwaite, P.H.** (2006) Quantifying expert opinion for modelling fauna habitat distributions. *Computational Statistics*, 21, 121-140.

Albers, C.J., Jansen, C.R., Kok, J., Kuipers, O.P. & van Hijum, S.A.F.T. (2006) SIMAGE: simulation of DNA-microarray gene expression data. *BMC Bioinformatics,* 7, Article 205 ('highly accessed')

Basu, S. Basu, A. & **Jones, M.C.** (2006) Robust and efficient parametric estimation for censored survival data. *Annals of the Institute of Statistical Mathematics*, 58, 341-355.

Crawford, J.R. & **Garthwaite, P.H.** (2006) Comparing patients' predicted test scores from a regression equation with their obtained scores: a significance test and point estimate of abnormality with accompanying confidence limits. *Neuropsychology*, 20, 259-271.

Crawford, J.R. & **Garthwaite, P.H.** (2006) Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. *Cognitive Neuropsychology*, 23, 877-904.

Crawford, J.R. & **Garthwaite, P.H.** (2006) Detecting dissociations in single-case studies: Type 1 errors, statistical power and classical versus strong dissociation. *Neuropsychologia*, 44, 2249-2258.

Crawford, J.R., **Garthwaite, P.H.,** Azzalini, A., Howell, D.C. & Laws, K.R. (2006). Testing for a deficit in single-case studies: Effects of departures from normality. *Neuropsychologia*, 44, 666-677.

Farrington, C.P. & **Whitaker, H.J.** (2006) Semiparametric analysis of case series data (with discussion). *Journal of the Royal Statistical Society, Series C*, 55, 553–594.

Gower, J.C. (2006) Statistica data analytica est et aliter. *Statistica Neerlandica*, 60, 124-134.

Gower, J.C. (2006). An application of the modified Leverrier-Faddeev algorithm to the singular value decomposition of block-circulant matrices and the spectral decomposition of symmetric block-circulant matrices. *Computational Statistics and Data Analysis*, 50, 89-106.

Gower, J.C., Gardner, S. & LeRoux, N.J. (2006). A synthesis of canonical variate analysis, generalised canonical correlation and Procrustes analysis. *Computational Statistics and Data Analysis*, 50, 107-134.

Hannachi, A, Jolliffe, I.T., Stephenson, D.B & **Trendafilov, N.T.** (2006) In search of simple structures in climate: Simplifying EOFs. *International Journal of Climatology*, 26, 7-28.

Jones, M.C. (2006) A note on rescalings, reparametrizations and classes of distributions. *Journal of Statistical Planning and Inference*, 136, 3730-3733.

Jones, M.C., Williams-Thorpe, O., Potts, P.J. & Webb, P.C. (2006) Using field-portable XRF to assess geochemical variations within and between dolerite outcrops of Preseli, South Wales. *Geostandards and Geoanalysis Research*, 29, 251-269.

Musonda, P., Farrington C.P. & **Whitaker, H.J.** (2006) Sample sizes for self-controlled case series studies. *Statistics in Medicine*, 25, 2618-2631.

Nowell, D.A.G., **Jones, M.C.** & Pyle, D.M. (2006) Episodic Quaternary volcanism in France and Germany. *Journal of Quaternary Science*, 21, 645-675.

Plummer, M., Best, N., Cowles, K. & **Vines, K.** (2006) CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6, 7-11.

Potts, P.J., Bernadini, F., **Jones, M.C.,** Williams-Thorpe, O. & Webb, P.C. (2006) Effects of weathering on in situ portable X-ray fluorescence analyses of geological outcrops: dolerite and rhyolite outcrops from the Preseli Mountains, South Wales. *X-Ray Spectrometry*, 35, 8-18.

Silvertown, J., Dodd, M., Gowing D., Lawson, C. & **McConway, K.** (2006) Phylogeny and the hierarchical organisation of plant diversity. *Ecology*, 87, S39-S49.

Silvertown, J., **McConway, K.**, Gowing D., Dodd, M., Fay M.F., Joseph J.A. & Dolphin K. (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B – Biological Sciences*, 273, 39-44.

Trendafilov, N.T. (2006) The dynamical system approach to multivariate data analysis. *Journal of Computational and Graphical Statistics*, 15, 628-650.

Trendafilov, N.T. & Jolliffe, I.T. (2006) Projected gradient approach to the numerical solution of the SCoTLASS. *Computational Statistics and Data Analysis*, 50, 242-253.

Whitaker, H.J., Farrington, C. P., Spiessen, B. & **Musonda, P**. (2006) Tutorial in biostatistics: The self-controlled case series method. *Statistics in Medicine,* 25, 1768-1798.

Williams-Thorpe, O., **Jones, M.C.,** Potts, P.J. & Webb, P.C. (2006) Preseli dolerite bluestones: Axe-heads, Stonehenge monoliths, and outcrop sources. *Oxford Journal of Archaeology*, 25, 29-46.

Woodward, A., Hunter, J., Ixer, R., Maltby, M., Potts, P.J., Webb, P.C. Watson, J.S. & **Jones, M.C.** (2005) Ritual in some early Bronze Age gravegoods. *The Archaeological Journal*, 162, 31-64.

Woodward, A., Hunter, J., Ixer, R., Roe, F., Potts, P.J., Webb. P.C., Watson, J.S. & **Jones, M.C.** (2006) Beaker age bracers in England: Sources, function and use. *Antiquity,* 80, 530-543.

6.3 Refereed conference proceedings and book chapters

Gower, J.C. (2006) Divided by a common language. In Multiple Correspondence Analysis and Related Methods, eds: J. Blasius & M. Greenacre, Chapman and Hall, Boca Raton, Florida, pp. 77-105.

Sarkar, A., DeRoeck, A. & **Garthwaite, P. H.** (2005) Term re-occurrence measures for analyzing style. In Proceedings of the SIGIR 2005 Workshop on Stylistic Analysis of Text for Information Access, eds: S. Argamon, J. Karlgren & J.G. Shanahan, ACM Press, pp. 28-36.

7.   Seminars and conferences at the Open University

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

Hannu Oja (University of Tampere, Finland) *Scatter matrices and independent component analysis (ICA)*

Dave Tyler (Rutgers University, USA) *Invariant coordinate selection (ICS): a robust statistical perspective on independent components analysis (ICA)*

Rodney Holder (St. Edmund's College, Cambridge) *The fine-tuning of the universe: some issues in probability theory*

Christian Hennig (University College London) *Two methods for clusterwise cluster validation*

Ed Godolphin (Royal Holloway, London) *Choosing cross-over designs when few subjects are available*

Alastair Young (Imperial College London) *Conditional properties of the bootstrap*

John Gower (Open University) *A matrix miscellany*

Baibing Li (Loughborough University) *Cluster analysis for gene expression data using the clustering-function-based method*

We also ran our third Annual Research Students' Research Day, where speakers were Swarup De, Eric Ferreira, Gorge Kafatos and Ardiana Gjini.

And at an internal research morning, speakers were Casper Albers, Nickolay Trendafilov,

8.   Conference papers and posters presented

C.J. Albers

Valencia/ISBA 8th World Meeting on Bayesian Statistics, Benidorm, Spain *Modelling vehicle counts in traffic networks* (poster)

K. Anaya

Robust Classification and Discrimination With High Dimensional Data (ROBCLA 2006), Florence, Italy *Optimal simple principal components*

International Conference on Robust Statistics (ICORS 2006), Lisbon, Portugal *Robustification using geometry*

F. Critchley

Robust Classification and Discrimination With High Dimensional Data (ROBCLA 2006), Florence, Italy *Principal axis analysis*

S. De

Research Students' Conference 2006, Glasgow *A class of Bayesian space-time models for environmental forecasting* (poster)

StatGIS06 workshop, Klagenfurt, Austria *A Bayesian model for Chernobyl's radioactive deposition in Bavaria* (poster)

A.E. Faria

Valencia/ISBA 8th World Meeting on Bayesian Statistics, Benidorm, Spain, *The geometric combination of Bayesian forecasting models* (poster)

C.P. Farrington

INSERM Workshop, La Londe-Les-Maures, France *Analyse des séries de cas: principe et applications*

J.C. Gower

10th Jubilee Conference of the International Federation of Classification Societies, Ljubljana, Slovenia, *Similarity in retrospect*

COMPSTAT 06, Rome, Italy *A matrix miscellany*

M.C. Jones

International Conference on Mathematical and Statistical Modeling in Honor of Enrique Castillo, Ciudad Real, Spain *Three families of distributions based on simple initial distributions … and my two favourite simple initial distributions* (Plenary)

ICMS Workshop on Quantile Regression, LMS Method and Robust Statistics in the 21st Century, Edinburgh *Parametric families of distributions and their interaction with the workshop title* (Keynote)

International Conference on Multivariate Statistical Methods in the 21st Century; the Legacy of Prof. S.N. Roy, Kolkata, India *Linking a parametric family of distributions with nonparametric quantile estimation and regression*

H.J. Whitaker

23rd International Biometric Conference. Montreal, Canada *Infections with varying contact rates: application to varicella*

INSERM Workshop, Villejuif, France *Practical analysis of case series data*

9. Other seminars and talks given

F. Critchley

Universidad Carlos III, Madrid, Spain *Principal axis analysis*

C.P. Farrington

American University of Beirut, Lebanon *Serological surveys in infectious disease epidemiology*

Imperial College London *Estimation of contact surfaces from serological survey data*

University of Newcastle *Analysis of epidemiological data using cases only*

Royal Statistical Society Ordinary Meeting, London *Semiparametric analysis of case series data* (with H. J. Whitaker)

Hasselt University, Belgium *Estimation of contact surfaces from serological survey data*

London School of Hygiene and Tropical Medicine *Estimation of contact surfaces from serological survey data*

INSERM U780 Epidemiologie et Biostatistique, Paris, France *Estimation de surfaces de contact à partir de données sérologiques*

P. Garthwaite

Royal Statistical Society Highlands Local Group, Aberdeen *Some applications of Bayesian methods*

J.C. Gower

Imperial College London *Data-mining mining-data*

M.C. Jones

Heriot-Watt University, Edinburgh *Parametric families of distributions and, inter alia, nonparametric quantile estimation and regression*

S.K. Vines

University of Auckland, New Zealand *Simple principal components and beyond …*

H.J. Whitaker

University College London The *self-controlled case series method*

Royal Statistical Society Ordinary Meeting, London *Semiparametric analysis of case series data* (with C.P. Farrington)

MRC Biostatistics Unit, Cambridge *Case series analyses of censoring events*

10. Editorships

F. Critchley

Editorial Board Member: *Studies in Classification, Data Analysis, and Knowledge Organization* (book series, Springer-Verlag).

C.P. Farrington

Associate Editor: *Journal of the Royal Statistical Society, Series B* (to July)

Editorial Board Member: *Epidemiology and Infection*

J.C. Gower

Associate Editor: *Journal of Classification*

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: *Sankhya*

Associate Editor: *Statistica Sinica*

Associate Editor: *Statistical Methodology*

11. #9; Grants Awarded

F. Critchley

Engineering and Physical Sciences Research Council grant for £284,843, '*A Computational Information Geometry Approach to Sensitivity Analysis in Statistical Science*' (post-doctoral researcher: K. Anaya)

C.P. Farrington

Engineering and Physical Sciences Research Council grant for £235,764 '*New Statistics for the Case Series Method – Weakening the Assumptions*' (post-doctoral researcher: M. Hocine)

Medical Research Council grant for £70,000 '*Prospective surveillance of vaccine safety by case series analysis*' (post-doctoral researcher: P. Musonda)

12. Other activities

K. Anaya

research visitor, University of Waterloo, Canada

F. Critchley

research visitor, University of Liège, Belgium, and University of Waterloo, Canada

international member of PhD panels at the University of Antwerp, Belgium, and the National University of Mexico

member of steering committee, European Science Foundation SACD network

C.P. Farrington

member of the RSS Research Section Committee

member of the WHO Ad-Hoc Committee on EPI Serology in Relation to Intermittent Preventive Treatment in Infants against Malaria

external assessor for DEFRA's Review of the Bovine TB Research Programme

(with Mona Kanaan) organised a one-day workshop on 'The case series method in epidemiology' at the (American University of Beirut), Lebanon, and was a research visitor for a week

external examiner for the MSc in Modern Epidemiology at Imperial College London

P.H. Garthwaite

external examiner for undergraduate mathematics courses at the University of Southampton

external examiner for the MSc in Statistics at the University of Kent

member of the MRC Panel of Experts

J.C. Gower

member of the Sir Ronald Aylmer Fisher Memorial Committee of Great Britain

·   member of the Council of the International Federation of Classification Societies

·   foreign member of the Advisory Board of IOPS (Interuniversitaire Onderzoekschool voor Psychometrie en Sociometrie)

K.J. McConway

became Associate Dean (Curriculum) of the Faculty of Mathematics and Computing

academic consultant on the BBC Radio4/OU series 'More or Less', including a contribution on air to one programme talking about school performance tables. He also wrote material for the programme-related website at
http://www.open2.net/moreorless/

also appeared on 'Inside Out', a BBC East current affairs TV programme, talking about the National Lottery

S.K. Vines

spent six months on study leave in the Department of Statistics, University of Auckland, New Zealand, and taught an undergraduate course on Multivariate Analysis for them