

The Open University

Department of Statistics

Annual Report for 1996

Contents

1. Membership
2. Introduction
3. Teaching
4. Research interests
5. Seminars and conferences at the Open University
6. Conference papers presented
7. Other seminars and talks given
8. Publications
9. Grants awarded
10. Other activities

1. Membership

- Professor D. J. Hand (Head of Department)
- Dr N. Adams
- Dr F. Daly
- Mr J. C. Gower
- Dr S. E. Hitchcock
- Dr G. E. Iossif
- Dr M. C. Jones
- Dr H. G. Li
- Dr A. D. Lunn
- Dr K. J. McConway
- Dr R. Ngouenet
- Dr E. Stanghellini (to June 1996)
- Dr P. C. Taylor
- Dr S. K. Vines
- Mrs J. E. Williams
- Mr K. Yu

Visitors

- Dr M. J. Faddy (University of Queensland)
- Professor C. E. Lunneborg (University of Washington)
- Professor K. Sugawa (Hokusei Gakuen University, Sapporo)

Full-time research students

- Mr M. G. Kelly
- Miss M. Kanaan (from October 1996)

Part-time research students

- Miss M. J. Cottee
- Mr S. Cresswell
- Mr P. Fayers
- Mr G. Kandalajt
- Ms K. Kemsley
- Dr M. King
- Mr D. F. Signorini
- Ms I. Stratton
- Mr K. Yu

Support staff

- Ms E. Ostrowski , Departmental Secretary

2. Introduction

This year was a very productive and successful one for the Department. We were gratified to score a grade 4 in the Research Assessment Exercise, the highest grade in the University's Faculty of Mathematics and Computing. Our research activity continued strongly, particularly in the areas of statistics in finance, classification methods, smoothing and biplots. During 1996, Department members published 2 books and 30 papers, book chapters etc. Dr S. K. Vines took up a new post as advisor on statistical matters to research students across the whole University. In addition, our teaching continued to flourish, with the first presentation of a major revision of one of our second level courses and of one of the Faculty's new first level courses (which has major statistical content), as well as ongoing production of a new third level course in linear modelling and of statistical contributions to another major first level course.

Information on the Department is available through our Web pages at <http://Statistics.open.ac.uk/index.html>

3. Teaching

Presentation

The Department presented for the first time a revised version of [MDST242 *Statistics in Society*](#) (30 CATS points), incorporating a complete rewrite of the middle section of the course. As with the remake of the first part, this stage preserves the successful teaching strategy of the existing course, while bringing the data examples up to date and making them more relevant. This revision involved making two television programmes as well

as several course texts and some audio material. The Department presented three other courses (30 CATS points each), [M246 Elements of Statistics](#), [M343 Applications of Probability](#) and M345 *Statistical Methods*. Altogether the Department's courses attracted a total of about 2,000 students.

Production

The main stage of production got under way for [M346 Linear Statistical Models](#), which will replace M345 *Statistical Methods*, and will be presented for the first time in 1998. The new course M346 takes an even more applied line than its predecessor, following on from [M246 Elements of Statistics](#), and will involve the students in heavy use of *Genstat* on a home computer. As well as linear models *per se*, the new course covers generalized linear models and some aspects of experimental design.

The Department continued to make a major contribution to the production of two courses in the entry suite for the Faculty of Mathematics and Computing. Each of these courses, [MU120 Open Mathematics](#) and [MST121 Using Mathematics](#), contains substantial amounts of statistical material (roughly one quarter of each course). MU120 was presented for the first time in 1996, and MST121 will be presented first in 1997.

Work was completed on a minor revision of [M343 Applications of Probability](#), for first presentation in 1997.

Preliminary planning work was commenced on a new course in Business Statistics, to be developed in conjunction with the Open University Business School.

Other

Members of the Department also contributed to the presentation of M101 *Mathematics: a Foundation Course*, [MU120 Open Mathematics](#), MS284 *An Introduction to Calculus*, [M203 Introduction to Pure Mathematics](#), [U205 Health and Disease](#), D300 *Professional Judgement* and B882 *Creative Management*.

4. Research interests

The Department is the home of two research groups, in Statistics in Finance (Hand, Adams, Kelly, Li, McConway, Stanghellini) and Classification Methods (Hand, Adams, Gower, Ngouenet, Taylor, Yu). In addition, Dr McConway is a member of the University's Ecology and Conservation research group.

Members of the Department are involved in a variety of research projects, including:

- *Statistical computing* (Daly)
- *Applied multivariate analysis* (Gower)
- *Multidimensional scaling, Procrustes methods* (Gower)
- *Biplots* (Gower, Hand, Ngouenet)
- *Discrimination and classification* (Hand, Adams, Taylor, Yu, Webb (DERA))
- *Measurement theory and metadata* (Hand)
- *Assessing classification rules* (Hand, Adams)
- *Credit risk scoring* (Hand, Adams, Kelly, Li, McConway, Stanghellini)
- *Deconstructing statistical questions* (Hand, Lunneborg)
- *Queues with priorities* (Hitchcock)
- *Stochastic birth-death processes for individuals in clusters* (Hitchcock)
- *Random number generation* (Jones)
- *The local dependence function* (Jones)
- *L-moments* (Jones)
- *Minimum distance estimation* (Jones, Basu (Indian Statistical Institute), Harris (University of Northern Arizona), Hjort (University of Oslo))
- *Semiparametric density and probability estimation* (Jones, Faddy (University of Queensland), Hjort (University of Oslo))
- *Conditional quantile estimation* (Jones, Yu)
- *Bayesian methods in meta-analysis* (Lunn)
- *Medical statistics, clinical trials (categorical data, repeated measures, generalised estimating equations, Bayesian methods)* (Lunn)
- *Performance-based handicapping systems in sport* (Lunn, Cresswell)
- *Robust estimation with circular and spherical data* (Lunn, Fisher (CSIRO))
- *Analysis of spectroscopic data (discrimination, multivariate methods of density reduction)* (Lunn, Kemsley)
- *Bayesian statistics* (Lunn, McConway, Vines)
- *Reliability of software* (Lunn)
- *Statistical methods in plant ecology* (McConway)
- *Comparative methods using phylogenies* (McConway, Hitchcock, Silvertown (Biology, OU))
- *Ecological and genetic correlates of long-term population trends in plants* (McConway, Silvertown (Biology, OU), Miller (Rothamsted Experimental Station))
- *The role of genetic variation in determining ecological amplitude in plants* (McConway, Silvertown (Biology, OU), Miller (Rothamsted Experimental Station), Macnair (University of Exeter))
- *Nonparametric binary regression* (Signorini, Jones)
- *Graphical modelling* (Stanghellini, McConway)

- *Spatial statistics* (Taylor, Kanaan, Mugglestone (Rothamsted Experimental Station), Alexander (Rothamsted Experimental Station))
- *Image analysis* (Taylor, Denham (University of Reading), Pursey (University of Reading))
- *Bayesian analysis of case-control studies* (Vines)
- *Parameterisation of random effects models* (Vines, Gilks (MRC Biostatistics Unit))

5. Seminars and conferences at the Open University

The Department runs an active seminar programme with invited external speakers. During the year we heard the following:

- Professor T. M. F. Smith (University of Southampton) *150 years of survey sampling*
- Professor M. Goldstein (University of Durham) *Pressure matching for hydrocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments*
- Professor G. Murray (University of Glasgow) *Stages in the development and evaluation of a predictive system in severe head injury*
- Dr M. Ramoni (Knowledge Media Institute, The Open University) and Dr P. Sebastiani (City University) *Be aware of your ignorance: robust learning of Bayesian networks from incomplete data*
- Professor S. Nishisato (University of Toronto) *Dual scaling: method, problems and some remedies*
- Dr M. Keeling (University of Cambridge) *Measles epidemics in England and Wales*
- Professor S. Evans (Medicines Control Agency) *The statistical detection of fraud*

In December, the Department again ran the Annual Open University Statistics Conference. This year's topic was *Visualising multivariate data*. The speakers at this meeting were:

- Professor A. Unwin (University of Augsburg) *Looking at data directly: linking and active graphics*
- Dr B. Francis (University of Lancaster) *Visualisation tools in the exploration of event histories*
- Mr J. C. Gower (The Open University) *Generalised biplot representations of multivariate data*
- Dr R. Ngouenet (The Open University) *Writing general biplot software in S-Plus*
- Professor W. J. Krzanowski (University of Exeter) *Some simple projections for multivariate selection problems*
- Professor C. M. Bishop (Aston University) *Latent variable models for data visualisation*

6. Conference papers presented

- Third ASUENITIAA Conference, Nantes: (Gower) *Biplots* (Invited talk)
- Fifth Conference of the International Federation of Classification Societies, Tokyo: (Gower) *A unified treatment of biplot methodology* (Invited talk)
- Salamanca Statistics Seminar III Advances in Descriptive Multivariate Analysis: (Gower) *Prediction with nonlinear biplots* (Invited talk)
- Fifth International Workshop on Matrix Methods for Statistics, Shrewsbury: (Gower) *Approximating a matrix of nominal values*
- Conference on Statistics, Science, and Public Policy, Herstmonceaux Castle: (Hand) *The importance of statistics in the gathering and analysis of data* (Invited talk)
- Biometric Section of the International Biometric Society, Basle: (Hand) *Deconstructing statistical questions* (Invited talk)
- Conference on Intelligent Data Management, London: (Hand) *Statistical modelling* (Invited talk)
- Sydney International Statistical Congress: (Hand) *Classification rules matching the solution to the question* (Invited talk)
- COMPSTAT 96, Barcelona: (Hand) *Classification* (Invited talk)
- Neural Computing Applications Forum, Edinburgh: (Hand) *Assessment and generalizability of classification rules* (Invited talk)
- Data Analysis Workshop, Augsburg: (Hand) *Deconstructing statistical questions* (Invited talk)
- Royal Statistical Society Conference, Guildford: (Jones) *Semiparametric density estimation*

7. Other seminars and talks given

J. C. Gower

- *Nonmetric linear biplots* University of Barcelona
- *Nonmetric linear biplots* University of Salamanca

D. J. Hand

- *Measuring the performance of allocation rules* Free University of Berlin
- *Graphical models in credit scoring* Barclays Bank, Coventry
- *Repeated measures analysis* Denmark
- *Statistics and the theory of measurement* Ciba-Geigy, Basle
- *Statistical aspects of credit scoring* Royal Statistical Society Local Group, Exeter
- *Statistics and the theory of measurement* Royal Statistical Society read paper, London

- *Graphical models in credit scoring* Forward Trust Group, Birmingham
- *Statistics and credit scoring* Joint Royal Statistical Society local group/British Computer Society meeting, Sheffield

A. D. Lunn

- *Bayesian methods in the analysis of clinical trials* University of Sydney (series of six seminars)

8. Publications

8.1 Books

- Gower, J. C. and Hand, D. J. (1996) *Biplots*, Chapman & Hall.
- Hand, D. J. and Crowder, M. J. (1996) *Practical longitudinal data analysis*, Chapman & Hall.

8.2 Publications in refereed academic journals

- Channon, S. and Hand, D. J. (1996) Verbal memory following temporal lobectomy, *Data and Statistics*, **1**, 215219.
- Denis, J.-B. and Gower, J. C. (1996) Asymptotic confidence regions for biadditive models: interpreting genotypeenvironment interactions, *Applied Statistics*, **45**, 479493.
- Gower, J. C. and Greenacre, M. J. (1996) Unfolding a symmetric matrix, *Journal of Classification*, **13**, 81105.
- Hand, D. J. (1996) Statistics and the theory of measurement (with discussion), *Journal of the Royal Statistical Society, Series A*, **159**, 445-492.
- Henley, W. E. and Hand, D. J. (1996) A k -nearest neighbour classifier for assessing consumer credit risk, *The Statistician*, **45**, 7795.
- Hjort, N. L. and Jones, M. C. (1996) Locally parametric nonparametric density estimation, *Annals of Statistics*, **24**, 1619-1647.
- Jones, M. C. (1996) Comment on 'Flexible smoothing with Bsplines and penalties' by P. H. C. Eilers and B. D. Marx, *Statistical Science*, **11**, 108109.
- Jones, M. C. (1996) Comment on 'Analysis of gamma-ray bursts distribution' by Zharkov and Zharkov, *Physics Letters A*, **212**, 115-117.
- Jones, M. C. (1996) On close relations of local likelihood density estimation, *Test*, **5**, 345356.
- Jones, M. C. (1996) The local dependence function, *Biometrika*, **83**, 899904.

- Jones, M. C. and Foster, P. J. (1996) A simple nonnegative boundary correction method for kernel density estimation, *Statistica Sinica*, **6**, 10051013.
- Jones, M. C. and Hössjer, O. (1996) From basic to reduced bias kernel density estimators: links via Taylor series approximations, *Journal of Nonparametric Statistics*, **7**, 2334.
- Jones, M. C. and Lunn, A. D. (1996) Transformations and random variate generation: generalised ratioofuniform methods, *Journal of Statistical Computation and Simulation*, **55**, 4955.
- Jones, M. C., Marron, J. S. and Sheather, S. J. (1996) A brief survey of bandwidth selection for density estimation, *Journal of the American Statistical Association*, **91**, 401407.
- Jones, M. C., Marron, J. S. and Sheather, S. J. (1996) Progress in data-based bandwidth selection for kernel density estimation, *Computational Statistics*, **11**, 337381.
- Oliver, J. J. and Hand, D. J. (1996) Averaging over decision trees, *Journal of Classification*, **13**, 281297.
- Samiuddin, S., al-Harbey, A. H., Jones, M. C. and Maatouk, T. A. H., (1996) Simple smoothed histograms, *Pakistan Journal of Statistics*, **12**, 113129.
- Taylor, P. C. and Jones, M. C. (1996) Splitting criteria for regression trees, *Journal of Statistical Computation and Simulation*, **55**, 267-285.
- Vines, S. K., Gilks, W. R. and Wild, P. (1996) Fitting Bayesian multiple random effects models, *Statistics and Computing*, **6**, 337-346.
- Williams-Thorpe, O., Jones, M. C., Tindle, A. G. and Thorpe, R. S. (1996) Magnetic susceptibility variations at Mons Claudianus and in Roman columns a method of sourcing to within a single quarry, *Archaeometry*, **38**, 1541.

8.3 Refereed conference contributions

- Gower, J. C. (1996) A unified treatment of biplot methodology. In *Tutorial Session of the Fifth Conference of the International Federation of Classification Societies*, Japan Marketing Association, Tokyo.
- Gower, J. C. (1996) Biplots. On *Proceedings of the Third ASU-ENITIAA Conference: Food Industry*, Nantes.
- Hand, D. J. (1996) Classification and computers: shifting the focus. In *COMPSTAT Proceedings in Computational Statistics, 1996*, ed. A. Prat, Physica-Verlag, Heidelberg, 77-88.
- Pursey, G., Taylor, P. C. and Denham, M. C. (1996) Image analysis of fungal spores. In *Proceedings in Image Fusion and Shape Variability Techniques*, eds. K. V. Mardia, C. A. Gill and I. L. Dryden, Leeds University Press, 214215.

8.4 Book chapters

- Gower, J. C. (1996) Multivariate and multidimensional analysis. In *Advances in Biometry*, eds. P. A. Armitage and H. A. David, John Wiley & Sons, New York.
- Hand, D. J. (1996) Elementary statistics. In *Research Methods*, ed. A. Greenfield, Edward Arnold, London, 181188.
- Hand, D. J. (1996) Further statistical methods. In *Research Methods*, ed. A. Greenfield, Edward Arnold, London, 189194.
- Jones, M. C. (1996) Comments on 'A personal view of smoothing and statistics' by J. S. Marron, 'Smoothing by local regression: principles and methods' by W. S. Cleveland and C. Loader, and 'Variance properties of local polynomials and ensuing modifications' by B. Seifert and T. Gasser. In *Statistical Theory and Computational Aspects of Smoothing*, eds. W. Härdle and M. G. Schimek, Physica-Verlag, Heidelberg, 8587.
- Lunneborg, C. E. and Hand, D. J. (1996) Data analysis by computer. In *Research Methods*, ed. A. Greenfield, Edward Arnold, London, 195203.

8.5 Major book reviews

- Jones, M. C. (1996) Review of 'Nonparametric Regression and Generalized Linear Models: a Roughness Penalty Approach' by P. J. Green and B. W. Silverman, *Statistics and Computing*, **6**, 179181.

8.6 Editorships

J. C. Gower

- Associate editor: *Psychometrika*
- Associate editor: *Journal of Classification*
- Associate editor: *The Mathematical Scientist*
- Statistical editor: *Journal of Agricultural Science*

D. J. Hand

- Editor-in-Chief: *Statistics and Computing*
- Joint Editor: *Royal Statistical Society Book Series*
- Editorial Board member: *Journal of Applied Statistics*
- Editorial Board member: *Statistical Methods in Medical Research*
- Editorial Board member: *Data Mining and Knowledge Discovery*
- Editorial Board member: *Studies in Classification, Data Analysis, and Knowledge Acquisition* (book series, Springer-Verlag)
- Editorial Advisory Committee member: *Research in Official Statistics*

M. C. Jones

- Joint editor: *Journal of the Royal Statistical Society, Series B*
- Associate editor: *Statistics and Computing*

C. E. Lunneborg

- Software reviews editor: *Statistics and Computing*

9. Grants awarded

D. J. Hand

- Dynamic graphical methods for credit scoring, £15,000 from Abbey National plc.

P. C. Taylor

- The spatial analysis of storms in Africa, £15,570 from the Open University Research Committee.

10. Other activities

N. Adams, E. Stanghellini and S. K. Vines

- were all awarded PhD degrees, by Liverpool John Moores University, the University of Florence and the University of Cambridge respectively.

M. J. Cottee

- was awarded the degree of PhD (Open) for her work on graphical representations of statistical structures, under the supervision of Professor Hand.

F. Daly

- completed a successful term as Sub-Dean of the Open University's Faculty of Mathematics and Computing, including several periods as Acting Dean.

J. C. Gower

- is the British (Royal Statistical Society) representative on the international board of ECAS (European Courses in Advanced Statistics).
- is Honorary Secretary of the Sir Ronald Aylmer Fisher Memorial Committee of Great Britain.

- is a Foreign Member of the Advisory Board of IOPS (Interuniversitaire Onderzoekschool voor Psychometrie en Sociometrie).

D. J. Hand

- was a member of the programme committees for the following conferences: Eighth International Conference on Scientific and Statistical Database Management, Stockholm (1996); Information, Statistics and Induction in Science, Monash University (1996); Second International Conference on Knowledge Discovery and Data Mining, Portland (1996); Second IASC World Conference, Pasadena (1997); Sixth International Workshop on AI and Statistics, Florida (1997); First European Symposium on Principles of Data Mining and Data Discovery, Trondheim (1997); Third International Conference on Knowledge Discovery and Data Mining, Newport Beach (1997).
- is a member of the international advisory board for Softstat 97, Heidelberg.
- is Additional Council Member of the International Federation of Classification Societies.
- is external examiner for the BSc and MSc in Statistics at University College London.
- examined PhDs at the London School of Economics and the University of Reading.

M. C. Jones

- was awarded the rank of Fellow of the American Statistical Association.
- was *ex officio* member of the Editorial Policy Board, the Programme Committee and the Research Section Committee of the Royal Statistical Society.
- continued as External Examiner for the MSc in Computational Statistics at the University of Bath.
- examined a PhD for the University of Glasgow.

A. D. Lunn

- was on the Organizing Committee of the Sydney International Statistical Congress (1996).