ifcs

International Federation of Classification Societies Newsletter

President: David Hand

Number 33   May 2007

The International Federation of Classification Societies, founded in 1985, is composed of:

Associaçáo Portuguesa de Classificaçào e Análise de Dados, British Classification Society, Central American and Carribean Society of Classification and Data Analysis, Classification Society of North America, Gesellschaft für Klassifikation, Irish Pattern Recognition and Classification Society, Japanese Classification Society, Korean Classification Society, Société Francophone de Classification, Società Italiana di Statistica, Statistical Society of Slovenia, Vereniging voor Ordinatie en Classificatie and Section of Classification and Data Analysis of the Polish Statistical Society.

The IFCS is a non-profit, non-political scientific organization, the aims of which are to further classification research. Among other activities, the IFCS organises a biennial conference, and supports the "Journal of Classification" and the journal "Advances in Data Analysis and Classification (ADAC)".

IFCS Homepage:

http://www.classification-society.org/
Designed and maintained by David Dubin.
The website contains among others:
the IFCS Constitution and By-Laws, the IFCS newsletters, and pointers to the websites of the member societies.

Contents

Words from the IFCS President
May 2007

International scientific organisations serve an important role in an increasingly uncertain world. They bring together people from different backgrounds and cultures, who share a common scientific interest, and who can build on that to increase worldwide collaborations and understanding. The IFCS follows the particular model of bringing together national or regional organisations of such people, rather than the individuals oneself, so also serving to integrate people and organisations at a higher level. In continuing to promote this, the IFCS encourages classification societies which are not yet part of the Federation, to join the IFCS.

As part of its activities, the IFCS supports two journals, each with a central interest in classification methods, but with rather different, and indeed complementary, emphases and remits. Details for the *Journal of Classification* can be found at http://www.classification-society.org/csna/joc.html and for *Advances in Data Analysis and Classification* at www.springer.com/journal/11634.

The society also organises a biennial international conference. The next one is due to take place in Dresden on 13-18 March 2009 – and I encourage you to mark the date in your diary now. Details are available on http://www.ifcs2009.de/. The society also encourages further joint meetings between its member societies, such as that of the British Classification Society on *The Analysis and Classification of Large Complex Datasets* which took place in the UK on 2nd Feb 2007, and which was co-sponsored by the Classification Society of North America.

Classification is a huge subject, and it may not always be apparent to people, whether those concerned with developing new methods or those concerned with applying existing methods to new problems, that they are concerned with classification problems. Sometimes I think that the classification community suffers from this, when other organisations other established with the remit of developing or applying particular subtypes of classification tool. The fact is that it is clearly beneficial to put these particular subtypes in the broader context of other classification methods, so that the relative strengths and weaknesses can be identified for application to different problems and areas. This is likely to lead to new ways to tackle problems, as well as to empower people who are not aware of other relevant work which has already been done. With this in mind, I would like to see more interaction between these more specialist organisations and the broader classification societies. As noted in my opening paragraph, raising awareness in this manner can have only beneficial consequences.

David J. Hand
IFCS President

News from the IFCS Scientific Secretary

Dear Colleagues and Friends of the Classification,

At the end of November 2006, the IFCS Council has formally accepted the Statistical Society of Slovenia (SSS) as a Member Society of the IFCS. The Statistical Society of Slovenia, chaired by Andrej Blejec, will be represented in the IFCS Council by Anuska Ferligoj. It is my pleasure to welcome, also on behalf of the other member societies, the Slovenian colleagues into the Federation.

Early this year, the IFCS Election committee, chaired by James Rohlf, has run the election to fill the two vacant positions as "Additional Members" on the IFCS Council. As a result of this election, Claus Weihs (GfKl) and Bernard Fichet (SFC) replace the outgoing members Hans-Hermann Bock and Akinori Okada. While welcoming the incoming members on the Council, I feel to express the gratitude of the whole IFCS community to H.H. Bock and A. Okada for their personal commitment and valuable work done during their terms on the IFCS Council.

Finally, have a look at the IFCS website (www.classification-society.org) for the latest news on conferences of member societies (SIS CLADAG, CSNA, SFC), journals officially supported by the IFCS (Journal of Classification – JoC, and Advances in Data Analysis and Classification - ADAC) and up-to-date information on the IFCS Council composition with contact details of member societies.

Vincenzo Esposito Vinzi
IFCS Scientific Secretary

News from the GfKl

GfKl has a new board for the next three years. The new president is the old one: Prof. Dr. Claus Weihs (Dortmund). Also the two vice presidents have not changed: Prof. Dr. Reinhold Decker (Bielefeld) and Adj. Prof. Dr. Berthold Lausen

(Erlangen-Nürnberg). In contrast, the treasurer has changed, since Prof. Dr. Hermann Locarek-Junge had resigned the end of 2006. The new treasurer is Prof. Dr. Daniel Baier (Cottbus). The new additional members in the board are Prof. Dr. Hermann Locarek-Junge (Dresden), Prof. Dr. Lars Schmidt-Thieme (Hildesheim), and Prof. Dr. Alfred Ultsch (Marburg).

Much more important, GfKl has its first three honorary members:
Prof. Hans-Hermann Bock, being GfKl's most influential member and its most well-known face, its international representative for so many years,
Prof. Dr. Wolfgang Gaul, having chaired so many program committees and having edited so many proceedings volumes in the Springer Series "Studies in Classification, Data Analysis, and Knowledge Organization", as well as
Prof. Dr. Otto Opitz, so many pupils of him having been / being engaged for the society, e.g. as members of the board or as organizers of annual conferences.
The society thanks these members for their long-lasting dedication.

Last, but not least, GfKl even has an honorary president from now on: Prof. Dr. Hans-Hermann Bock being the secret president all the time, unimportant who was the elected president: Standing ovations for Prof. Bock.

C. Weihs (President of the GfKl)

32nd GfKl Annual Conference

The German Classification Society GfKl will hold its 32nd Annual Conference on "Advances in Data Analysis, Data Handling, and Business Intelligence" at the Helmut-Schmidt-University of Hamburg (Germany) from July 16 – 18, 2008.

The scientific program will include sessions with invited talks and contributed presentations from a broad range of topics. Special emphasis will be laid on interdisciplinary research and the interaction between theory and practice. The local organizers of the conference are Wilfried Seidel, Claudia Fantapié Altobelli, Andreas Fink, Hartmut Hebbel, Detlef Steuer and Ulrich Tüshaus of the Faculty of Economics and Social Sciences, Helmut-Schmidt-University Hamburg.

The first call for the conference will be announced in June 2007 (see www.gfkl.de).

The broad range of relevant topics is illustrated by the following list of intended sections:

Theory and Methods: Exploratory Data Analysis; Clustering and Classification; Pattern Recognition and Machine Learning; Forecasting; Visualization and Scaling Methods; Mixture Analysis; Bayesian, Neural, and Fuzzy Clustering; Graphs, Trees, and Hierarchies; Evaluation of Clustering Algorithms and Data Structures; Data Science and Innovative Tools; Data Cleaning and Pre-Processing; Data, Text, and Web Mining; Computational Intelligence; Information Extraction and Retrieval; Personalization and Intelligent Agents; Tools for Intelligent Data Analysis; Knowledge Representation and Discovery.

Applications: Subject Indexing and Library Science; Marketing and Management Science; e-commerce, Recommender Systems, and Business Intelligence; Targeting in Social Sciences; Banking and Finance; Production, Controlling, and OR; Biostatistics and Bioinformatics; Genome and DNA Analysis; Medical and Health Sciences; Archaeology and Geography; Engineering and Environment; Image and Signal Processing; Administrative Record Census; Linguistics and Statistical Musicology; Education; Psychology.

The Annual GfKl Conference 2008 will be hosted by the Helmut-Schmidt-University, Hamburg. Hamburg is the second largest city of Germany and, according to its inhabitants, one of the most beautiful cities in the world. Hamburg's geographical location is so favourable that it is within easy reach for the rest of the world.

Erlangen and Hamburg,
Berthold Lausen and Wilfried Seidel

29th Fall Meeting 2007 of the AG-DANK (GfKl)

The traditional Fall Meeting of the working group AG-DANK 2007 will be held at the Faculty of Economics, Technical University of Dresden. We follow an invitation of Professor Locarek-Junge. This year's focus will lie on the analysis of discrete (categorial) data as well as on economic and financial data. A call for participation will be issued soon. Persons interested in the meeting are asked to contact   ritter@fim.uni-passau.de .

At the general assembly of the AG-DANK, March 7, 2007, at Freiburg, Prof. Gunter Ritter (Passau) and Dr. Christian Hennig (London) were reelected chairman and vice chairman, respectively.

Gunter Ritter

News from the JCS

The 23rd Annual Research Meeting of the Japanese Classification Society

The 23rd Annual Research Meeting of the Japanese Classification Society was held at the RENAISSANCE CENTER of Tama University on November 25-26, 2006. The speakers and the titles of their talks are as follows:

[Japan Korea Session]
The Measurement of Crime Severity

Hiroshi Inoue（Kyushu Institute of Technology)

Tree Structured Prognostic Model for Hepatocellular Carcinoma Patients with Gene Data
Lee, Taerim (Korea National Open University)

Mixture Models for Multivariate Count Data
Kazunori Yamaguchi (Rikkyo University)

On Clustering of Time Varying Data
Tadashi Imaizumi (Tama University)

[Invited Session]
Multidimensional Scaling of Asymmetric Relationships

Akinori Okada（Rikkyo University）

Missing Region Modeling and Multivariate Normal Mixture Model
Detecting Multiple Outliers in Linear Regression Using a Cluster Method Combined with Graphical Visualization
Kim, Sung-Soo (Korea National Open University) and Krzanowski, W. J. (University of Exeter)

[Contributed Session]
Structure of Occupational Evaluations in Korea and Japan
Keiko Genji and Atsuho Nakayama (Rikkyo University)

A Study on One-mode Three-way Overlapping Cluster Analysis
Satoru Yokoyama (Keio University), Atsuho Nakayama (Rikkyo University) and Akinori Okada (Rikkyo University)

Analysis of Adaptability in "Rupin the third" from Conversation
Kei Ogura, Atsuho Nakayama, and Akinori Okada (Rikkyo University)

Three-way Multidimensional Scaling and Residual Analysis
Atsuho Nakayama (Rikkyo University)

A Bayesian Approach for Nonmetric Multidimensional Scaling
Kensuke Okada and Kazuo Shigemasu (The University of Tokyo)

A Study on the Book Mapping using Table of Contents and Indexes
Takao Ishizuka (Asia University)

An application to Internet survey of covariate adjustment by a latent class model
Shun Xu, Fumitake Sakaori and Kazunori Yamaguchi (Rikkyo University)

Analysis of mega hit music compact disc for the diffusion mechanism and prediction by probabilistic model

Kotaro Shibuya (Saison Information Systems CO. LTD.)

Analysis of purchase behavior using panels of households and individuals
Toshinori Saito (Rikkyo University)

Exploring the Possibilities of the Facet Approach in the Context of Japanese Reality
Kazufumi Manabe (Kwansei Gakuin University)

Yoshiro Yamamoto

News from the SoCCCAD

The Central American and Caribbean Society for Classification and Data Analysis (SoCCCAD) announces the XVI INTERNATIONAL SYMPOSIUM ON MATHEMATICAL METHODS APPLIED TO THE SCIENCES (SIMMAC, Simposio Internacional de Metodos Matematicos Aplicados a las Ciencias) and the VIIth SoCCCAD's Meeting for Classification and Data Analysis.

It will take place in San Jose, Costa Rica, on February 19-22, 2008.

TOPICS:

- Data Analysis, Multivariate Statistics, Clustering and Classification
- Probability, Stochastic Processes, Financial Mathematics, Stochastic Control
- Optimization, Operations Research, Approximation,
- Numerical Analysis, Dynamical Systems, Differential Equations
- Modeling, Biomathematics
- Applications in the above topics

PARTICIPATION:

- Short courses (3 hours)

- Long (40 min) and short (20 min) communications

For submitting a short course or a communication, please visit our Website for instructions:
http://simmac.emate.ucr.ac.cr/

Abstracts deadline: October 1, 2007
Notification to authors: December 1, 2007
Final version deadline: January 4, 2008
Languages: English and Spanish.
Send abstracts to: simmac@emate.ucr.ac.cr
The preliminary program will be available by the end of January, 2008.

Please visit our Website for prices and deadlines. We have special fares for Central Americans and students. On the Website you can find the Scientific Committee, Organizing Committee, how to reach the place of the Symposium, hotel information, excursions, tourist information, and other useful information.

Chairman: Javier Trejos,
jtrejos@cariari.ucr.ac.cr

Javier Trejos

News from the VOC

The VOC had a successful joint spring meeting with the Belgium-Netherlands Association for Artificial Intelligence on April 27, 2007 in the historical Faculty Club in Utrecht. Six speakers covered nicely different aspect of Data Mining. Patrick Groenen (Erasmus University Rotterdam) discussed minimization for support vector machines (svm) by iterative majorization, and meanwhile clearly explained the ideas behind svm. Christophe Croux (Catholic University of Leuven) presented an influence function approach
to robust discrimination. Bernard de Baets (Ghent University) discussed monotone distribution classifiers. Koen Vanhoof (University of Hasselt) talked on aggregation operators' measures. Lambert Schomaker

(University of Groningen) offered a nice overview of the possibilities and impossibilities of writer identification and handwritten manuscript retrieval. Paul Eilers (Utrecht University and Leiden University) closed the meeting with a nice example of statistical classification for reliable high-volume genetic measurements.

The next VOC meeting on classification in medical contexts will take place on Thursday December 6, 2007. Our keynote speaker will be Frank Harrell. The up-to-date information on the activities of the VOC is available via www.voc.ac.

Marieke Timmerman

News from the BCS

1) The BCS web site is at http://thames.cs.rhul.ac.uk/bcs

President is Fionn Murtagh, Past-President Nick Fieller, Secretary Jo Padmore, Treasurer Clive Moncrieff, and Committee Members are Christian Hennig, Boris Mirkin and Trevor Ringrose.

2) **British Classification Society and CSNA workshop, Brunel University, London, England, Friday February 2, 2007.**

Talks were given by Xiaohui Liu (local organiser, Brunel University) on clustering and data analysis aspects of microarray data analysis; Bob Nichol (Portsmouth, formerly Carnegie-Mellon), on data analysis and statistical problems in cosmology; Mel Janowitz (DIMACS) on formal concept analysis and its relation to clustering; Boris Mirkin (Birkbeck, University of London), on neighborhood similarity and use of domain knowledge; and Peter Willett (Sheffield University) on virtual screening techniques for chemical databases.

In the afternoon there was a wide ranging discussion on the BCS in the future. The benefits to members of a society like the BCS include:

- Championing multidisciplinarity.
- Possible involvement in consultations by government and research funding agencies.
- Among young researchers in particular, conferences and meetings are very important.
- Access to publications can be of importance, notwithstanding online availability of practically all the journal literature.
- New products and services were mentioned, including near future perspectives for CSNA Service on the web.

Interest was expressed in possibly restoring the Royal Statistical Society's multivariate studies group. There was discussion on having as an "umbrella" organisation a resurrected Classification Society, federating again the BCS and CSNA, and acknowledging that other classification societies were constituted on a language basis (Dutch, German, French, and maybe others in the future).

3) **Joint half-day meeting of the British Classification Society (BCS) and the Statistical Computing Section of the Royal Statistical Society, on "Representation and Generalisation in Pattern Recognition", Royal Statistical Society, 6 October 2006.**

Speakers:

- Robert P.W. Duin, Delft University of Technology, "Representation related problems in pattern recognition"
- David M.J. Tax, Delft University of Technology, "Outlier detection using ball descriptions with adjustable metric"
- Dick de Ridder, Delft University of Technology, "Integrative bioinformatics"
- Elzbieta Pekalska, University of Manchester and Delft University of Technology, "Statistical learning with general proximity measures".

4) A near future BCS meeting will concern itself with evaluating research and scholarly output, a theme of very considerable importance right now in the UK (where the periodic Research Assessment Exercise, currently underway nationwide, is to give way to a new evaluation approach based on so-called metrics).

Analysis Methodologies for Post-RAE Scientometrics, and AGM
Friday 6 July 2007, International Building room IN244 Royal Holloway, University of London, Egham (Surrey)

Motivation and background are as follows.

Multivariate data analysis, and related areas - classification and clustering, psychometrics, low-dimensional or other interpretation-friendly visualizations derived from heterogeneous measurements, measurement theory and data coding/encoding, multivariate correlation, underlying variables, multidimensional features, etc. etc. - are all bread and butter methodologies within the remit of the British Classification Society. The selection of appropriate and/or best data analysis methodologies are a result of a number of issues: the overriding goals of course, but also the availability of well formatted, and ease of access to such, data. The meeting will focus on the early stages of the analysis pipeline.

There is a strong movement towards openness and accesibility in the products of research and scholarly work, cf.

1) Harnad, S. (2007) Open Access Scientometrics and the UK Research Assessment Exercise. In Proceedings of 11th Annual Meeting of the International Society for Scientometrics and Informetrics (in press), Madrid, Spain. http://eprints.ecs.soton.ac.uk/13804

2) Brody, T., Harnad, S. and Carr, L. (2006) Earlier Web Usage Statistics as Predictors of Later Citation Impact. Journal of the American Association for Information Science and Technology (JASIST) 57(8) pp. 1060-1072.

3) Carr, L., Hitchcock, S., Oppenheim, C., McDonald, J. W., Champion, T. and Harnad, S. (2006) Extending journal-based research impact assessment to book-based disciplines.
The BCS meeting will be addressed by Drs Les Carr and Tim Brody, of the Intelligence, Agents, Media group, Electronics and Computer Science, University of Southampton.

Discussion will follow.

An aim of this meeting is to discuss and plan the writing of a BCS white paper on data analysis methodologies in the context of what can be considered as open, objective and universal in a metrics context of scholarly and applied research. A part of this discussion is the extent to which unified (perhaps discipline-based) perspectives are desirable over and above the support of a plurality of perspectives.

The "Analysis Methodologies for Post-RAE Scientometrics" meeting will be directly followed by the AGM of the BCS.

Fionn Murtagh

News from the CLAD
 "Associação Portuguesa de Classificação e Análise de Dados"

In the year of 2007 most CLAD activities are centred on the organization or on co-sponsorship of scientific meetings and conferences

- The CLAD annual meeting, JOCLAD2007, held at the University of Porto (www.JOCLAD2007.isep.ipp.pt) was a very successful conference.

- CLAD co-sponsors the next meeting of the International Statistical Institute, ISI2007, that will take place in Lisbon (http://www.isi2007.com.pt/) and supports the ISI2007 Invited Paper Meeting IPM 90 on Analysis of Complex Data in Health Sciences.

- CLAD and LEAD (http://lead.paginas. sapo.pt/) co-sponsor the two following ISI2007 satellite meetings:

  o ISBIS2007 - International Symposium on Business and Industrial Statistics (http://www.isbis2007.uac.pt/intro.php), that will be held at the University of Azores, and

- o ICSMRA2007 - International Conference on Statistical Methods for Risk Analysis, that will be held in the Faculty of Medicine at the University of Lisbon (http://www.fm.ul.pt/ICSMRA2007).

Helena Bacelar-Nicolau
hbacelar@fpce.ul.pt
cladlead@fpce.ul.pt

News from SFC

SFC 2007

The SFC Conference will be held at the Ecole nationale supérieure des télécommunications, in Paris (France), on September 5-7, 2007. The organizers of the conference are Irène Charon, Georges Hébrail and Olivier Hudry. Besides the usual topics of this conference, the focus in 2007 will be on:

- Graphs and classification
- Bioinformatics
- Data stream clustering
- Social networks, Web
- Textual analysis
- Symbolic data analysis.

The invited speakers are: Hans Bock (Aachen), Joao Gama (Porto), Pascale Kuntz (Nantes), Ludovic Lebart (Paris), Jacques van Helden (Bruxelles).
About 50 talks have been submitted to the conference. A selection of full papers presented at the conference will be published as a special issue of "Mathématiques et Sciences humaines - Mathematics and Social Sciences". The final program as well as all the other details will be soon available from the conference Web site:

http://www.enst.fr/SFC2007

Olivier Hudry, on behalf of the organizers

One-day Statistical Workshop

Friday, June 29th, 2007
Lisieux, France

Cluster analysis is an important part of Data Mining and Machine Learning. Several problems are difficult to be solved as the determination of the number of clusters, the selection of the relevant variables and the stability and relevance of the clusters. The purpose of this workshop is to address these topics from a theoretical point of view. Mixture Models of parametric probability distributions provide a statistically mature approach to answer such questions. But, there is a recent interest for making use of nonparametric stability-based methods to evaluate the clusters and to help in the choice of the parameters of a clustering algorithm. The workshop is aiming to present those two different approaches and discuss their connections.

Program of the workshop. The workshop program includes six invited speakers and a Poster Session which will be held after lunch. Invited speakers are:

- G. Celeux (France): Clustering mixture models: general aspects and model selection criteria.
- U. von Luxburg (Germany): Statistical aspects of clustering stability.
- G. Govaert (France): Some specific aspects of clustering mixture models: block clustering and variable selection.
- P. Bertrand (France): On using criteria of cohesion and isolation to evaluate clustering stability.
- B. Murphy (Ireland): Parsimonious Gaussian Mixture Models.
- G. Valentini (Italy): Stability based methods for the assessment of clustering solutions.

Registration. The deadline for registration is June 15, 2007.

Further information is available on the workshop website:
http://www.iutcaen.unicaen.fr/stid/colloque/main.html

Patrice Bertrand

EGC 2008

Knowledge extraction and management (Extraction et Gestion des Connaissances, EGC'2008) - January 29 - February 1, 2008, Sophia Antipolis, France http://www-sop.inria.fr/axis/egc08

Honorary chair : Osmar Zaïane, Univ. of Alberta, Canada

We are pleased to announce that the 8th French Conference "Extraction et Gestion des Connaissances" (EGC'2008) will take place in the INRIA Sophia Antipolis, close to Nice, France, from January 29 to February 1, 2008. This conference is devoted to theories, methods and applications of Data Mining, Machine Learning, Knowledge Discovery in Data Bases, Knowledge Management and engineering. Its aim is to gather all researchers and company managers in this area in order to present work of high
quality and exchange new ideas. This year, the honorary chair EGC'2008 invites you to propose your contribution in several ways: tutorials, workshops, oral presentations, posters, and software demonstrations. Two special scientific prices of 1500 euros each will be awarded to the best accademic and industrial papers. Papers should be submitted in French or in English. Two formats of submission exist: short papers (up to 6 pages) and long papers (up to 12 pages), please refer to our web site for additional information (http://www-sop.inria.fr/axis/egc08).

The main submission deadlines are:
- Tutorials and workshops: September 20, 2007
- Paper abstract: September 21, 2007
- Full paper: September 28, 2007
- Software demonstrations: October 20, 2007

The proceedings of EGC'2008 will appear in a special issue of the French journal RNTI (Cépaduès publisher). Each paper is read by three reviewers (see the program committee on our web site).

Fabrice Guillet and Brigitte Trousse, (Program chair and local chair).

CLA 2007

The fifth international conference on Concept Lattices and their Applications (CLA 2007: http://www.lirmm.fr/cla07/index.html) will be held in Montpellier, France, October 24-27, 2007.

Selected papers accepted to CLA 2007 will be invited to further revise and extend their work in a special issue of International Journal of General Systems (http://www.tandf.co.uk/journals/titles/03081079.asp) on Concept Lattices and Applications. Selected papers will be competitively chosen by the Program Chairs and the CLA steering committee based on the quality of their content and presentation.

Important dates :
- June 10: Abstract due,
- June 22: Paper due,
- August 10: Notification to authors,
- September 14 : Final version due.

Jean Diatta

News from CSNA
the Classification Society of North America

Stanley L. Sclove, Secretary

CSNA ANNUAL MEETING

The CSNA 2007 annual meeting will be held at the University of Illinois at Urbana/Champaign, June 7-10, overlapping with the meeting of the Digital Humanities group. The meeting is sponsored by the National Center for Supercomputing Applications, the UIUC College of Education and the Graduate School of Library and Information Science. The organizers of the meeting are David Dubin and Carolyn Anderson. The keynote speaker will be Michael Kurtz. Dr. Kurtz will be speaking on The Astronomical Information Network. Further information on the annual meeting is available at http://www.classification-society.org/csna/csna07.html.

ELECTIONS

A Nominating Committee for the CSNA 2007 elections has been appointed, consisting of Fionn Murtagh, Chair, Samantha Bates Prins, Jim Rohlf, Padhraic Smyth, and Stan Sclove. The present terms of office extend through 31-December of the year. The Nominating Committee will present a tentative slate by the Fall. There is a 30-day period during which eligible voters may nominate additions to the slate. At least five eligible voters are needed to nominate someone to the slate. Then, the (possibly augmented) slate is announced, and there is a voting period of thirty days. The election will be online, set up by IFCS/CSNA
Webmaster David Dubin.

Journal of Classification

Issue No. 1, Vol. 24 (2007) of the Journal of Classification is going to pressand will be mailed to subscribers shortly. The editor is Willem Heiser of Leiden University. The journal will be mailed with the bibliographic CD, Classification Literature Automated Search Service, now containing several out-of-print books on classification and clustering. This bibliographic service has been edited by Fionn Murtagh. Murtagh, President of the British Classification Society, will become President of CSNA in January, whereupon Michael J. Kurtz of the Smithsonian Astrophysical Observatory will become editor of the bibliographic Service. Further details about the CSNA bibliographic Service are given below.

CSNA Bibliographic Service

Fionn Murtagh, Editor, CSNA Service

The Classification Literature Automated Search Service CD will see the addition of two further tomes to the scanned books. These are: RR Sokal and PHA Sneath, Principles of Numerical Taxonomy, San Francisco, Freeman, 1963, 359 pp. and PHA Sneath and RR Sokal, Numerical Taxonomy: the Principles and Practice of Numerical Classification, San Francisco, Freeman, 1973, 573 pp.

Short commentaries by the authors on the writing of these books are available online at the following addresses. These short commentaries are well worth a look.

http://www.garfield.library.upenn.edu/classics1982/A1982PJ14400001.pdf and

http://www.garfield.library.upenn.edu/classics1987/A1987F272800001.pdf

In another item of news from CSNA Service, after a very long time as editor, Fionn Murtagh is preparing to pass on the baton. Michael Kurtz is editor-elect now. Michael is one on the principals of ADS, the Smithsonian/NASA Astrophysics Data System (see http://adswww.harvard.edu). ADS currently contains more than 5.1 million records in the areas of Astronomy and Astrophysics, Physics, and ArXiv e-prints, and a further 5.3 million links to records maintained externally.

I am currently having all of the 22 years of the journal Les Cahiers de l'Analyse des Données digitized. There were about 500 pages per year in this journal, produced by J.-P. Benzécri, between 1977 and 1997.

Journal 'Advances in Data Analysis and Classification'

As scheduled and announced in the IFCS Newsletters 31 and 32 (2006) the first issue of the journal 'Advances in Data Analysis and Classification' (ADAC) has been published in February 2007 by Springer Verlag. It contains the following scientific articles:

A. Douzal-Chouakria, P.N. Nagabhushan:
Adaptive dissimilarity index for measuring time series proximity

S.J. Frame, S.R. Jammalamadaka:
Generalized mixture models, semi-supervised learning, and unknown class inference

A. Samé, L. Oukhellou, E. Côme, P. Aknin:

Mixture-model-based signal denoising

J. Diatta:
Galois closed entity sets and k-balls of quasi-ultrametric multi-way dissimilarities

G. González-Rodríguez, Á. Blanco, N. Corral, A. Colubi:
Least squares estimation of linear regression models for convex compact random sets.

The journal is methodologically oriented and open for a broad spectrum of topics in data analysis, theoretical as well as applied ones: classification and clustering, multivariate statistical methods, factorial analysis, pattern recognition and machine learning, data and text mining, with applications in science and economy, extending also to domains such as e-commerce, internet, finance, telecommunication, and microarray analysis. Details can be found on the journal's website www.springer.com/journal/11634. Paper submissions should be sent to EditorADAC@stochastik.rwth-aachen.de, for format specifications see previous website.

For the next issues of ADAC the following topics are scheduled:
- Learning metrics in random subspaces for semi-supervised clustering
- Variable selection in discriminant analysis based on the location model
- Fast calibrations of the forward search for testing multiple outliers in regression
- Statistics in musicology
- Convex ordering of Lévy processes with applications in finance

Subscriptions for the journal (about 160 Euro/year) can be addressed directly to Springer (see Springer's website). However, individual members of IFCS Member Societies can profit from a largely reduced rate when ordering via their Member Society. Depending on special arrangements between Springer and the Society, the individual price is then between 19.00 Euro (plus VAT; compulsory subscriptions) and 39.00 Euro (plus VAT; mandatory subscriptions). Please contact the treasurer of your Society.

Hans-Hermann Bock, Wolfgang Gaul, Akinori Okada, Maurizio Vichi

Short book review

Models for Discrete Longitudinal Data.
G. Molenberghs and G. Verbeke.
Springer, 2005.

Probably you are familiar with the "XYZ tour" T-shirt. The back of the shirt shows a long list of cities and dates, telling you where and when a famous pop group has given (or is to give) performances. I don't know much about their singing qualities, but Geert Molenberghs and Geert Verbeke deserve a world tour T-shirt as the Simon and Garfunkel of Mixed Models. Not only do they write about them, but they also presented very popular courses in many places.
This book is a sequel to Linear Mixed Models for Longitudinal Data by the same authors (Springer, 2000) which only considered normally distributed variables. Normal distribution theory is a cozy world, because whether you look at hierarchical components or at marginal variables, distributions always are normal. You might have to do some hefty matrix manipulations, but you can be sure that those matrices completely specify the distribution. All this beauty is gone when the observations are counts or binary observations.
The book can be approached in several ways. It is a handbook of analyses, because many real-life data sets are provided, with SAS code for model estimation. The data and the analyses are presented in detail, using tables and graphs, including model selection and interpretation. Re-running the analyses one-self can be a great way to get acquainted with the practical aspects of modeling non-normal longitudinal data.
The book also presents a careful overview of model types (like marginal, conditional and subject-specific), their meaning and their advantages and drawbacks. In that sense it is a handbook on the foundations of mixed models. The presentation is not overly technical and will be accessible to many readers.
Between the extremes of foundations and applications, the book contains a massive amount of technical material, which is far from easy reading in many places. It is clear that the authors know what they are talking about, but the models inherently are complicated in many places and frequently detailed derivations are skipped.

Like in the preceding volume, missing values and sensitivity issues get ample attention. There is a chapter on non-linear mixed models. It looks a bit out of place. Suddenly we are back at constant-variance additive normal noise. In the context of the data that are analyzed there, concentrations in blood of a drug, this looks unrealistic to me. In such a case constant relative precision is more common than a constant absolute precision. This would have been a good opportunity to introduce gamma-distributed responses, possibly combined with quasi-likelihood for flexible (multiplicative plus constant) error models.

I hope that some day the authors (or someone else) will go through the work of translating the SAS code to R. That would make advance discrete longitudinal modeling available to a much larger audience.

It is clear that I like the book. It is one to return to over and over again. And if you want to see the authors on stage: July 1, in Barcelona (http://mat.uab.es/~iwsm2007/sc.html).

Paul Eilers