ifcs

International Federation of Classification Societies Newsletter

Number 23 June 2002

President: Carlo Lauro

Editors: Paul De Boeck & Krzysztof Jajuga

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, 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.

Contents

From the President

New policies and challenges for IFCS

It is a great honour for me to take over the IFCS chair from Jean Paul, a dear friend of mine, but also from Chikio Hayashi whom I consider one of my masters. I hope to be equal to the moral and scientific heritage they have left me by continuing their work as President of the IFCS for the term 2001/2003. It is my aim to accept and cope with the IFCS new challenges by putting my energies to the Federation's disposal together with the experience I have gained during my lasting story inside the Federation, first as a Founding member in 1985 then as a Council Member until 1993 and also as a member of international scientific committees in most of IFCS conferences. Thus, I feel very pleased to finally serve in this new role where I hope to be expression, at my best, of the continuity and traditions of this Federation while bringing innovative ideas resulting also from previous experiences as IASC President and ISI Vice-President.
The following three aims are the main objectives of my program which I hope to accomplish in co-operation with all the IFCS members:

1. Expand the influence of the Federation
In the recent years new societies joined the Federation (SOCCAD, IPRCS). In order to expand its influence the Federation should:

- establish permanent contacts with International Statistical Societies (e.g. ISI and his branches, Biometric Society, Psychometric Society, ASA, etc.);
- develop contacts with Statistical Societies not affiliated with the Federation (e.g.: Australian Statistical Society, Indian Statistical Institute, Statistical Society of China, etc.) but also with Scientific Societies in which Classification plays a relevant role (Biology, Genetics, Astronomy, Marketing, Management, Sociology, etc.);
- promote the applications of Classification in new fields (e.g. business and finances, etc.) and emerging countries by organizing ad hoc workshops;pay more attention to students, young researchers or professionals by setting up special educational activities (IFCS summer schools, masters and PhDs' on Classification)

2. *Making more participative the life in the Federation*

In order to reach this objective some actions should be undertaken

- by improving communications at Council level and with affiliate Societies;
- by Involving in the IFCS life the past Presidents and outstanding individual members;
- by offering more services to IFCS individual members and non members (active web site, links, internet forums and news group, newsletter dissemination via net, articles, software and members directories). A web site Committee have to be set up;
- favouring the creation of *transversal group* of interest in new research domains

3. *Developing an effective publication policy in accordance with IFCS statute and by-laws*

"….It is intention of the Federation to publish one or more journals. Book(s) and/or newsletters…"

Doug Carroll in his Presidential address (IFCS Newsletter N.12- March 1966) proposed to adopt the Journal of Classification as the official Journal of the Federation. The discussion was still ongoing during the Council meeting in Namur. It is time now to take a decision on this subject matter.

In order to prepare a paper for the Council, a *Publication Committee*, chaired by Krzysztof Jajuga, will be appointed to evaluate the following alternatives:

- adopting the *JoC as official journal of the Federation* and defining IFCS role in it (Editor(s) and Editorial Board appointments, policies, etc.);
- maintain the sponsorship to the JoC and publish a *new journal* as flag of the Federation (eventually an e-journal);

As regard other publications, I would like to promote, by appointing an ad hoc Editorial board, the publication of a scientific *monograph on the "History of Classification"* followed by a poster on the same subject (similar to the one on the History of Mathematics). The worldwide diffusion of such publications and posters will certainly contribute to increase the visibility of the Federation.

As proposed by J.P. Rasson in Namur , in order to improve the management of the IFCS business and procedures ,as well as to achieve the above mentioned goals, some amendments to the Statute and by-laws became necessary:

- number of council members, additional members and group at large,
- roles of the officers,
- voting rules and majority for ordinary business,
- elections calendar and rules,
- use of web site based vote to avoid (privacy problems).

A *Committee of Sage* (say: the old IFCS Presidents) should manage this delicate task.

The Committee on *Long Range Planning of Conferences* will be soon appointed in order to make proposals regarding *IFCS 2004* and *IFCS 2006*, as well as to establish workshops and summer schools.

In order to start the above mentioned initiatives special attention to *Finances improvements* have to be paid- by:

- searching for sponsors,
- renegotiating the Societies fees,
- evaluating the possibility to collect royalties from IFCS conferences, schools and publications in order to support new activities.

A special *Fund Raising Committee*, as the one proposed by Chikio Hayashi, could be usefully serve to this scope.

N.Carlo Lauro
IFCS President

News from the IFCS

The latest round of IFCS elections were completed in January, 2002. I am pleased to report that Krzysztof Jajuga has taken over from Paul De Boeck as the new IFCS Publications Officer. People who attend the IFCS meeting in Krakow will have the opportunity to join me in applauding Paul for his long and excellent service, and to personally thank Krzysztof for agreeing to undertake this job for the IFCS.

Planning for the Krakow Council Meeting:
Carlo Lauro, the IFCS president, invited me and Jean-Paul to visit him in Naples in early May to discuss IFCS business and to prepare for the Krakow meeting. Our discussions were extremely productive, and raised a number of points that are likely to be included on IFCS Council agenda.

The following topics arose:

- We must soon remind committees and officers to prepare their routine reports.
- We must encourage member societies to consider hosting the 2004 and 2006 meetings; the 2004 proposals should be ready for the Krakow meeting in July.
- Carlo has developed possible logos for the IFCS; these must be shown to the Council for consideration.
- It seems desirable for the IFCS president to appoint a nominating committee, whose charge is to ensure that in each IFCS election there are at least two qualified candidates willing to serve.
- The IFCS should examine ways to ensure more regular financial support for our programs, to enable satellite meetings, educational programs, and other activities.
- The IFCS may need to appoint a Publications Committee, with Krzysztof Jajuga as chair, to consider ways in which communications may be improved and modernized.
- The IFCS Council should consider the possibility of fostering new classification societies in other regions of the world, such as India, China, and Australia.
- It may be timely to consider amending the IFCS Constitution and By-Laws to reflect new capabilities of the Internet and the growth of the federation; Carlo is considering the appointment of a committee of previous IFCS presidents to recommend possible changes.

Obviously, the last topic is the most important for the future of the IFCS, and Carlo intends to discuss it more fully at the Krakow meeting.

On a personal note, Jean-Paul and I want to thank Carlo for his hospitality during our visit to Napoli, and to congratulate him upon his good fortune in having wonderful collaborators and students. We think this pre-meeting was an excellent idea, and we were able to accomplish a great deal that could not have been done by email.

David Banks
IFCS Secretary

IFCS – 2002 CRACOW

The Eighth Conference of the International Federation of Classification Societies, IFCS-2002: Data Analysis, Classification and Related Methods, will be held in Cracow, Poland, on July 16-19, 2002.

The opening of the conference is planned for July 16, 9 a.m. and closing is planned for July 19, 4 p.m.

The keynote speakers of the conference are:
Hans-Hermann Bock – Clustering methods: from classical models to new applications
Frank Hampel – Some thoughts about classification
Wojtek J. Krzanowski – Orthogonal components for grouped data – review and applications.

The invited speakers of the conference are:
Hamparsun Bozdogan – A new generation of multivariate mixture-model cluster analysis of normal and nonnormal data using information measure of complexity
Edwin Diday – From data to knowledge: symbolic data analysis, mixture decomposition and spatial pyramidal clustering
Henk A.L. Kiers – Should we use standard errors or cross-validation in component analysis techniques?
Klaus Obermayer – New methods for clustering, visualization and classification of proximity data
Jean-Paul Rasson – Divisive classification and segmentation trees with the Poisson process hypothesis
Yoshiharu Sato – The performance of an autonomous clustering technique
Maurizio Vichi – Clustering and reduction of three-way data

There are also 9 invited sessions (in each 3-5 papers will be given). This is the list of these sessions:
1. Dissimilarity Analysis and Hierarchical Classification
Organizer: Jean Pierre Barthélemy
2. Probability Models for Clustering
Organizer: Hans-Hermann Bock
3. Applications of Classification and Data Analysis in Marketing
Organizers: Daniel Baier, Reinhold Decker
4. Classification and Regression Trees
Organizer: Eugeniusz Gatnar
5. The WEB Mining Challenge
Organizer: Wolfgang Gaul
6. Optimization Methods and Algorithms in Classification and Clustering
Organizers: Patrick Groenen, Hamparsun Bozdogan
7. Bioinformatics and Classification
Organizer: Berthold Lausen
8. Perspectives of Data Science in Japan
Organizer: Noboru Ohsumi
9. Optimization Heuristics in Data Analysis
Organizer: Javier Trejos

In addition, approximately 105 papers will be presented in contributed sessions.

There will be the volume of conference Proceedings (published at the conference) with the selection of 53 papers published by Springer in the series: "Studies in Classification, Data Analysis and Knowledge Organization". This is the following volume:
Krzysztof Jajuga, Andrzej Sokolowski, Hans-Hermann Bock (editors): Classification, Clustring and Data Analysis. Recent Advances and Applications

The organizers provided a social program, including:
the conference dinner, the reception by local government, welcome reception, tourist excursions to Auschwitz and salt mine Wieliczka.
As a tradition, the IFCS Council Meeting will take place as well.

The Chair of Scientific Program Committee: Krzysztof Jajuga, Wroclaw University of Economics, E-mail: jajuga@manager.ae.wroc.pl
The Chair of Local Organizing Committee: Andrzej Sokolowski, Cracow University of Economics, E-mail: sokolows@ae.krakow.pl

E-mail addresses of the conference:

- organizational matters: ifcs2002@ae.krakow.pl
- scientific matters: ifcs2002@credit.ae.wroc.pl

Website of the conference:
http://ifcs2002.ae.krakow.pl

NEWS FROM THE MEMBER SOCIETIES

News from the IPRCS

IPRCS has been closely associated with the annual IMVIP - Irish Machine Vision and Image Processing Conference.
IPRCS, for example, decides on the location of IMVIP each year. In 2002, it was decided to hold IMVIP as a component part of Opto-Ireland, an SPIE regional conference. SPIE is the leading organisation worldwide in optical engineering and and signal processing. Officially, IMVIP is "Co-located conference:
Optical Metrology/Imaging/Machine Vision".

Opto-Ireland will be held on 5-6 September 2002 at the Radisson Hotel, Galway, Ireland. Local organiser is Dr Andy Shearer. With just short of 300 submissions to Opto-Ireland, it promises to be a landmark event.

The proceedings will be published in a number of SPIE volumes.

A half-day short course is planned by Fionn Murtagh on "Wavelets and Multiresolution Transforms in the Analysis and Processing of Signals and Images".

Further information is available from the IPRCS page:

www.iprcs.org

News from the VOC

On April 26, the VOC held its Spring Meeting. Our host was the Department for Research and Development of the Internal Revenue Service in Utrecht. The theme of the day was Data Mining and a variety of applications were presented, some of them in the area of tax collection. The presentations were:

- Edith Nijenhuis & Frans J E Vermeulen (Belastingdienst, Utrecht) *Profiles of Companies in the Cleaning Sector*
- Irma Volkers (Belastingdienst, Utrecht) *Early Detection of Insolvency with Data Mining Techniques*
- Marten den Uyl (Sentient Machine Research, Amsterdam) *Data Mining for the End User*
- Machiel Westerdijk (CapGemini, Utrecht) *Data Mining for Business Intelligence*
- Geert Wets (DAM LUC, Diepenbeek) *Detection of Latent Dissatisfaction in Surveys in the Financial Sector*
- Joost Kok (Universiteit Leiden) *Pre-computing and Post-processing of Association Rules*

A more detailed program can be found on our website:
www.voc.ac

On April 26 also the yearly meeting of members was held. Michel Wedel resigned and new members were elected to the board. It now consists of Patrick Groenen (chair), Paul Eilers (secretary), Iven van Mechelen (treasurer), Stef van Buuren (newsletter editor), Paul Arents (treasurer-elect), Ivo van der Lans, Mark de Rooij (editor-elect), Marieke Timmerman (secretary-elect) and Josephine Woltman Elpers.

News from the CSNA

ANNUAL MEETING

Bernard Harris, Meeting Chair; David Dubin, Meeting Webmaster

CSNA 2002 will be held in Madison, Wisconsin, from Thursday, 13 June to Sunday, 16 June, 2002. All sessions are scheduled for the Pyle Center, a University of Wisconsin conference center attractively situated on the shore of Lake Mendota.

Schedule

Two half-day short courses will be offered on Thursday, 13 June:
AM -- Finite Mixture Models, Stanley L. Sclove, University of Illinois at Chicago;
PM -- A Combinatorial Introduction to Cluster Analysis, Melvin Janowitz, DIMACS, Rutgers University.

Technical sessions are scheduled for Friday, Saturday and Sunday, 14, 15 and 16 June, with a number of invited and contributed papers.
In addition, sessions in memory of John van Ryzin and Mark Rorvig are planned.

A welcome reception is scheduled for the evening of Thursday, 13 June, and the meeting banquet is scheduled for the evening of Friday, 14 June.

Accommodations.

A block of dormitory rooms has been reserved in Barnard Hall, which is three blocks from the Pyle Center.

Attractions

Madison has many excellent ethnic restaurants within a ten-minute walk from the Pyle Center. A large number of restaurants in various price ranges may be found in the vicinity of the Pyle Center.

Madison has a number of excellent museums, and a number of attractive resorts and tourist areas are close to Madison.

Transportation

Madison is served by 12 air carriers, many of which are feeder airlines providing service to the hubs of major carriers. There is frequent bus service to the Milwaukee and Chicago airports. Often, there is a substantial fare difference between flying to Chicago or flying to Madison.

Further information

Please see the Website at www.cs-na.org or contact the meeting organizer by e-mail at harris@stat.wisc.edu or regular mail to Prof. Bernard Harris, Dept. of Statistics, University of Wisconsin, Madison, WI 53706-1693, or telephone 608-262-2614, fax 608-262-0032.

JOURNAL OF CLASSIFICATION

Willem Heiser, Editor; Phipps Arabie, Past Editor

Professor Heiser assumed the Editorship effective 1-March-2002, Professor Arabie having passed the torch after 18 years during which the Journal has been very highly rated.
Volume 19, No. 1, will be mailed to subscribers in early June.
Volume 19, No. 2, will be the first issue under Heiser's editorship.
We look forward to a continued high standard of interesting, informative and innovative articles.

BIBLIOGRAPHIC SERVICE

Fionn Murtagh, Editor

The "Classification Literature Automatic Search Service", now in its 30th year, is provided on CD with the first issue each year of the Journal of Classification. From a "profile" of 90 journal articles or books, in 2001 there were 3268 citations to one or more of them.
The CD contains volume 30 for 2001, together with volumes 29, 28, ... 23, i.e., cumulative data from 1994 to date. A graphical user interface provides a search engine, to facilitate quickly finding bibliographic references of interest on the CD.

A scanned copy of John Hartigan's 1975 book "Clustering Algorithms" (formerly published by Wiley) is available on the CD.
A copy of the CD will be available to participants at IFCS 2002, Crackow.

NEWSLETTER

Jennifer Pittman, Editor

The Newsletter is available via www.cs-na.org. Please send comments and information to Jennifer Pittman, e-mail jennifer@stat.duke.edu or regular mail to Jennifer Pittman, Institute of Statistics and Decision Sciences, 201A Old Chemistry Bldg., Box 90251, Duke University, Durham, NC 27708-0251.

Stan Sclove, Secretary

News from the GfKl

On 29. and 30. November the traditional Fall Meeting of the Working Group „Data Analysis und Numerical Classification“ (AG-DANK) will take place at the University of Bonn. The meeting is invited by Prof. Dr. Joachim Buhmann, Department of Computer Science III of the Universitiy of Bonn.

Two foci will be mean field techniques and robustness and stability in classification and data analysis.

Further information is available at:

http://stoch.fmi.uni-passau.de/agdank/Bonn2002

Gunter Ritter

New Publications of the GfKl

1. The Gesellschaft fuer Klassifikation (GfKl) has recently published a new volume in the Springer Series on „Classification, Data Analysis, and Knowledge Organization“, with authors from the international membership of IFCS:

Wolfgang Gaul, Gunter Ritter (eds.):
Classification, Automation, and New Media.
Proc. 24th Annual Conference of the GfKl, Univ. of Passau, March 15-17, 2000. Studies in Classification, Data Analysis, and Knowledge Organization, Vol. 20. Springer-Verlag, Heidelberg, 2002. 535 pages. ISBN 3-540-43233-7

It contains 57 reviewed papers organized in the following sections:

- D*ata Analysis, Statistics and Classification (15)*
- *Pattern Recognition and Automation (3)*
- *Data Mining, Information Processing, and Automation (8)*
- *New Media, Web Mining, and Automation (7)*
- *Applications in Management Science, Finance, and Marketing (11)*
- Applications in Medicine, Biology, Archaeology etc. (13).

2. Another volume is in just being printed by Springer Verlag: the Proceedings of the 8th Conference of the IFCS to be held in Cracow in July 2002:

Krzysztof Jajuga, Andrzej Sokolowski, Hans-Hermann Bock (eds.):
Classification, Clustering, and Data Analysis. Recent Advances and Applications.
Springer-Verlag, Heidelberg-Berlin, 2002. 492 pages. ISBN 3-540-43691-x

It contains 53 refereed papers to be presented in Cracow, with sections dealing with:

- *Clustering and classification, cluster validation, discrimination*
- *Multivariate data analysis, statistical methods, analysis of symbolic data*
- *Regression trees, neural networks and genetic algorithms*
- *Consensus trees and phylogenetics, genome analysis*
- *Applications from medicine, economics, marketing, and psychology.*

3. IFCS members who are well acquainted with the German language may also be interested in the Festschrift which has been issued at the occasion of the 25th birthday of the GfKl, on February 12, 2002:

Hans-Hermann Bock, Peter Ihm (Hrsg.):
25 Jahre Gesellschaft für Klassifikation. Klassifikation und Datenanalyse im Wandel der Zeit.
Shaker Verlag, Aachen, 2001. 184 pages.
ISBN 3-8265-9778-8
It includes not only many facts and reminiscences from the GfKl, but also a section on the origins of the International Federation of Classification Societies.

H.H. Bock

News from the JCS

The 18th Annual Research Meeting of the Japanese Classification Society

The 18th Annual Research Meeting of the Japanese Classification Society was held at the Institute of Statistical Mathematics on December 22, 2001. The program is summarized as the following:

Invited talks

In this meeting, the following invited talks were presented as the special session of ‘Symbolic Data Analysis.”

(1) A Region-Based Fuzzy Pattern Classifier for Symbolic Data
Manabu Ichino (Tokyo Denki University)

This paper presents a region-based fuzzy pattern classifier that is applicable to symbolic data. For symbolic data, each sample pattern is described not only by quantitative features but also by qualitative features. To manipulate these general forms of sample patterns, we introduce a mathematical model so called the Cartesian system model (CSM). Then, we define the mutual neighborhood graph (MNG) as a tool to know the structures between pattern classes. Our region-based fuzzy classifier is designed based on the MNG. We present several numerical examples in order to show the effectiveness of our approach.

(2) SODAS: Symbolic Official Data Analysis System
- A Software for Symbolic Data Analysis -
Nobuo Shimizu (The Institute of Statistical Mathematics)

Symbolic data is a concept which was proposed in order to describe various statistical data generally. We can summarize large data sets in small symbolic data table by using the concept. And also, mathematical expressions of each object at the symbolic data table are called symbolic objects (SOs). The statistical analysis for these data tables and SOs are defined Symbolic Data Analysis (SDA) which is studied in various ways and can be generalized to many existing statistical analysis methods by applying them to symbolic data.
SODAS (Symbolic Official Data Analysis System) is a SDA software package. In this paper, I introduce the characteristics of SODAS and show how to install and to use it.

Contributed Papers
Eleven contributed papers were presented.

(1) Three-Dimensional Space - Turbulence Phenomena -
Akiyoshi Kondo (Takushoku University)

There are 14 imaginary spheres on the Earth. The imaginary sphere will be considered in the following. Consideration is limited to one of the 14 spheres that exist. Of course, the size of the imaginary sphere is assumed to be the same as that of the Earth.
Thus, a regular hexahedron inscribes the sphere in the interior.
Due to change in the regular hexahedron, and due to the rotation of the Earth and various other phenomena, it is considered that various events arise also on the imaginary. The Earth is thought to move 499 meters per second.
`Turbulence` is the phenomenon that has a scale comparable to human beings so that it has both deterministic aspect and random aspect.

(2) The Structure of Disturbance in Total Hip Replacement Patients During Medical Rehabilitation
Kazuhiko Shimizu (Kitasato University), Toshio Ozawa (Kitasato University)

A cross-sectional survey of 284 patients was conducted before and after total hip replacement (THR) for hip osteoarthritis using two test batteries (i.e. SF-36 and the hip joint function rating criteria of the Japanese Orthopaedic Association), and MMT.
Analysis using structural equation modeling (SEM), which permits the assumption of selected latent variables, suggested that functional disturbance-related items, i.e., pain, ROM, and muscular strength, are mutually associated and capable of explaining performance disturbance, and that the function disturbance status on the intact side, as well as on the affected side, must be taken into consideration.
From the viewpoint of physical therapists, the muscular strength of the muscles around the hip joint is an essential factor, and it may be necessary to distinguish the methods of evaluation and functional rehabilitation between primary total hip replacement surgery and revision surgery.
SEM analysis facilitated the understanding of the hierarchical structure of disturbances and enabled the clear understanding of causal relations.

(3) Difference between Japan and US on Relations between "Shikohin" and Stress
Shuhei Matsuki (Tobacco Research Institute)

Based on the questionnaires on coffee, black tea, tobacco and alcoholic beverages, which were implemented in Japan and the United States, the following matters were suggested.
(1) Based on the study of job stress and the "stress reduction" effect, with respect to all items except for tobacco in the United States, the ratio of those who experienced "stress reduction" tends to be higher as the degree of job stress increases.
(2) With respect to consumption conditions and job stress in Japan, except for alcoholic beverages of men, it was observed that the ratio of consumers is higher as the level of job stress increases, particularly for tobacco.

(3) By examining the reasons why the ratios of coffee drinkers and smokers with high job stress in the United States are lower than those with medium stress, we concluded that in the United States those with high job stress are those of high social standing, so one might say that people of such social standing do not drink coffee because of health consciousness and do not smoke because they are afraid that they may be regarded as unsophisticated.

(4) Statistics, Data Analysis, Classification and Data Science

Chikio Hayashi (The Institute of Statistical Mathematics)

Data design, data collection and data quality evaluation are crucial to data analysis if we are to draw out useful relevant information. Analysis of low-information data never bears fruit; however, data analytic methods can be refined. In spite of the importance of this issue in actual data mining and data analysis, I am forced to ask why these problems cannot be discussed at its most essential level. Perhaps it is a matter of the laborious practical work involved or the otherwise plodding pace of research. Indeed, these problems are rarely addressed because in academic circles it is regarded as unsophisticated. In the present talk, I dare to touch these problems with the fundamental concept of data science.

(5) Fatigue Measured by a Free Answer and Items

Kiyoharu Doi (Osaka Shoin Women's University), Noboru Ohsumi (The Institute of Statistical Mathematics)

Stress and fatigue were investigated with a free answer and items targeting 62 male university students. The words from the free answer were classified in "fatigue", "frustration", "aggression", and so on. Some words about "fatigue" were the same as the items of fatigue. But, chi-square values with the words and the items were not significant. Furthermore, the expressions of the words about "aggression" were different from the items of aggression.

(6) Attitudes towards Educational Evaluation and Morality of University Students in the Faculty of Education

Osamu Yoshimura (Okayama University)

Attitudes towards educational evaluation, morality and self-concept of university students were investigated. Analyzing free answers to questions concerning educational evaluation and morality revealed the following:

1) Most of the students think the state of educational evaluation should be improved, but were satisfied with the state.
2) Most of the students have shortsighted morality that would not be extended beyond their family and circle of friends.

Moreover, it was found that the students had somewhat curious personalities, that they felt self-determination, self-esteem and self-efficacy but security.

On the bases of the results, problems of a curriculum of a training school for teachers were discussed.

(7) Consciousness of Women's Attitude toward Independence
- A Comparison of Two Surveys -

Setsuko Takakura (Nagasaki Junshin Catholic University), Mariko Murata (Statistical Information Institute for Consulting and Analysis)

We have conducted two surveys on the topic of women’s attitudes toward independence; in 1994 and in 2001. The subjects of our survey were female graduates from five colleges and eight universities, in the classes of '58, '67, '75, '81, '86 and '91. Some '98 graduates were included in the '01 survey. The sample size was 3,023 in 1994, and 2,500 in 2001, of which 1,407 and 866 were effective respondents, respectively.

The survey items included “image of independence”, “factors contributing to independence”, “motivation”, “factors deterring independence”, “family role consciousness” and other private and social factors.

The comparison of the two surveys revealed that the respondents' mental independence was as important as their economic independence. The two surveys have yielded quite similar results, but their consciousness of independence has become slightly stronger, and the consciousness of male-female equality became higher. They gave shorter responses in the 2001 survey for free responses. This implies that women have come to obtain more clear-cut ideas about independence.

(8) The Classification by Dual Scaling and the Replacement Procedure

Takahiro Nakamura (The Institute of Statistical Mathematics), Yasumasa Baba (The Institute of Statistical Mathematics)

As a method for quantification of the qualitative data, various method has been proposed (Guttman(1941), Hayashi(1952), Benzécri(1973).) Dual scaling (Nishisato, 1980) is the one of these methods.

The method can be applied for various types of qualitative data.
The one of the purposes of these methods is to give measure for qualitative data. We can describe the data structure by the measure and classify the variables and individuals. Even if the data structure can be described, the partition of individuals or variables can't be obtained by the methods mentioned above.
In this paper, we propose the classification method combined with dual scaling and replacement procedure. The purpose of this method is to classify individuals or variables. The solution of dual scaling is obtained by maximizing the correlation ratio. The correlation ratio is related to the between-individuals sum of squares. The method proposed here uses between-groups sum of squares. The replacement procedure is an algorithm used in clustering method based on homogeneity (Nakamura and Baba, 2000). We can obtain the partition with maximum correlation ratio by the algorithm. To illustrate the advantage of the method, we will show examples.

(9) Regression Analysis Based on Observation with Errors

Emiko Baba (Nihon University), Yasumasa Baba (The Institute of Statistical Mathematics)

In ordinal regression analysis observed values of explanatory variables and criteria variables are considered as points.
However we need to note that the points given by observation include some errors and the points given as averages are not true points. In this paper we discussed the following cases.
1) Regression analysis in the case that observed points are given as averages.
2) Regression analysis in the case that observation has errors.
In the latter case it was shown that if observation errors of a criterium variable and the explanatory variables are independent then the procedure of parameter estimation of a regression plane reduces to ordinal regression but it is different for polynomial regression.

(10) Fuzzy Cluster Loadings for Weighted Regression Analysis

Mika Sato-Ilic (University of Tsukuba)

Conventional clustering means classifying the given observations into exclusive subsets (clusters). That is, we can discriminate clearly if an object belongs to a cluster or not. However, such a partition is hardly enough to represent many real situations.
Then a fuzzy clustering method is offered to contract clusters with vague boundaries, this method allows that one object belongs to some overlapping clusters with some grades. In other words, the essence of fuzzy clustering is to consider not only the belonging status to the assumed clusters, but also to consider how much the objects belong to the clusters. So, there is a merit to representing the complex data situations which real data almost always have. The state of fuzzy clustering is represented by a partition matrix whose elements show the grade of belongingness of the objects to the clusters.
Replaced by the representativeness of fuzzy clustering to real complex data, the interpretation of such a fuzzy clustering causes us some confusion, because we sometimes think that objects which have a similar degree of belongingness can together form one more cluster. In order to solve this problem, we propose a measure which detects the meaning of the clusters. This measure is directly related with the loading of variables to the clusters, so we call this measure fuzzy cluster loading.
Related with the interpretation of the fuzzy clustering result, the validity of the fuzzy clustering result has been discussed. In the conventional measures of validity of fuzzy clustering, the partition coefficient or entropy coefficient are well known [1].
However, these measures are essentially based on the idea that clear classification is a better result. Using the idea of within-class dispersion and between-class dispersion, separation coefficients are introduced.[4] Moreover, according to the fuzzy hypervolume, partition density was discussed [3].
And recently, the method of evaluation of fuzzy clustering result, which used the idea of the homogeneity of homogeneity analysis has been proposed [5].
However, these measures show the extraction of relations of the fuzzy clustering result, which is the degree of belongingness of objects to clusters, and observations indirectly.
In order to obtain a direct interpretation of fuzzy clustering results, we propose a model for the fuzzy cluster loading. The estimate of fuzzy cluster loading are obtained by minimizing normalized sum of squared errors between the obtained degree of belongingness of the objects to the clusters and the model. Moreover, we discuss the relationship between the weighted regression analysis [2] and the proposed model.
Using this relationship, we show that we can estimate the fuzzy cluster loading in a similar way as the estimate of regression coefficients in the weighted regression analysis.

Moreover, the extended two models for 3-way data are proposed. One defines the fuzzy cluster loading as a measure for each time (or situation) which shows how each cluster can be explained by each variable at each time. The other defines the fuzzy cluster loading as a measure over all times (or situations). So, the difference of purpose between these two models is the capturing of properties of the clusters that exist in each time, or properties of the clusters over all of the times. Numerical examples for the models of fuzzy cluster loadings and estimate of the fuzzy cluster loading using the weighted regression analysis are shown in the paper.

[References]
[1] J.C. Bezdek, J. Keller, R. Krisnapuram, N.R. Pal (1999). *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Kluwer Academic Publishers.
[2] C. Brunsdon, S. Fotheringham and M. Charlton (1998). Geographically weighted regression-modelling spatial non-stationarity, *Journal of the Royal Statistical Society*, Vol. 47, Part 3, pp. 431-443.
[3] I. Gath and A.B. Geva (1989). Unsupervised optimal fuzzy clustering, *IEEE Trans. Patt. Anal. and Machine Intell*., Vol. 11, pp. 773-781.
[4] R. Gunderson (1978). Applications of fuzzy ISODATA algorithms to star-tracker printing systems, *Proc. Triennial World IFAC Congress*, pp. 1319-1323.
[5] M. Sato-Ilic (2000). On Evaluation of Clustering using Homogeneity Analysis, *IEEE International Conference on Systems, Man and Cybernetics*, pp. 3588-3593.

(11) On a Clustering Process of an Autonomous Clustering Method

Yoshiharu Sato (Hokkaido University)

In an autonomous clustering method, this paper has discussed the clustering process based on the concrete examples. First we investigated the forming process of clusters for the grid data. The characteristic feature, in this case, is that the clusters are inclined to merge from marginal of the configuration. This feature is also found the case of the data which are given by the normal random numbers with different two mean values. Moreover, using a suitable value of the control parameter of the action rule, the points located at the center between two clusters are never to assign to one of the clusters. They are inclined to generate isolated clusters by themselves. This seems interesting point which shows the different property with usual k-means method. Comparing with the traditional clustering methods, the value of the control parameter will be given. This remains to be further problem.

Other Meeting Information

(1) The Japanese Classification Society, the Behaviormetric Society of Japan and Tama University Renaissance Center opened the Joint Tutorial Seminar on "Dual Scaling" on November 10, 2001 at Tama University Renaissance Center (Tokyo). In this seminar, Professor Shizuhiko Nishisato gave a special lecture on the topic.

(2) A JCS Symposium is held in Toyo Eiwa University Graduate School (Tokyo) on September 11, 2002. The 19th Annual Research Meeting is in the planning stage to be opened at the end of this year. The detailed announcement will be appeared on the JCS Web-page in the future (browse http://wwwsoc.nii.ac.jp/jcs/).

News from the SFC

1. Elections:

Bernard FICHET (University of Marseille - France) has been reelected as President of the SFC Bruno LECLERC (CAMS – CNRS, Paris) has been reelected as Treasurer of the SFC

2. 9th Annual Meeting of the SFC -Toulouse, September 16-18, 2002

This francophone meeting is jointly organized in Toulouse (the pink town), by 2 Universities and 3 laboratories: the team GRIMM (statistics and graph theory) of the Mirail University of Social and Human Sciences, and the TCI (images processing) and LSP (statistics and probability) laboratories of the Sabatier University of Sciences.
The various and usual thematic sessions of Classification more specially include the topics: neural networks, graph theory, image processings, remote sensing, multivariate modelling, pattern recognition, datamining.

Main speakers: A. Appriou (Fr), J-P. Barthelemy (Fr), A. Ciampi (Canada), F. Critchley (England), A. Degenne (Fr), T. Denoeux (Fr), G. Govaert (Fr), S. Holmes (USA), D. Ladiray (Luxembourg), J-P. Rasson (Belgium), F. Rossi (Fr), T. Snijders (Netherlands), J. Ten Berge (Netherlands), S. Thiria (Fr), B. Victorri (Fr).

Three grants with support of AUPELF: A. Lazraq (Maroc), D. Mizere (Congo), R. Romero Alvarez (Mexique).
The Simon Regnier award will be granted to a young researcher for its applied or theoretical results in classification.

Location: Mirail University.

The deadline for submitting the 4-pages papers is June 1. More information on the internet:

http://www.irit.fr/SFC2002/

Contact: sfc@cict.fr

CONFERENCE ANNOUNCEMENTS

26th Annual Conference of the German Classification Society (GfKl)

Between Data Science and Everyday Web Practice

In the week following IFCS 2002, the German Classification Society will hold its 26th Annual Conference from July 22 to July 24, 2002, at the University of Mannheim. More than 170 papers were selected for presentation. Also, several internationally renownend scientists will contribute plenary and semi-plenary talks. In this year, the organization of the conference was strongly supported by CLADAG, the Classification and Data Analysis Group of Societa Italiana di Statistica.

The conference will take place in the lecture halls of the University of Mannheim. The University is located in the Mannheimer Schloss which is within 5-10 minutes walking distance from the Central Railway Station (Hauptbahnhof) and the hotels in the city.
There is a direct intercity connection from Frankfurt Rhein/Main International Airport to Mannheim. Trip length is approximately 30 minutes. A map of Mannheim can be found at
http://www.gfkl.de/gfkl2002/
(menu item Conference Site).

The preliminary conference program can be found at
http://www.gfkl.de/gfkl2002/
(menu item Program).

For members of GfKl and its associated societies there is a reduced registration fee of 70 EUR if they register before June 15, 2002. The normal fee is 100 EUR. It is possible to register online via
http://www.gfkl.de/gfkl2002/
(menu item Online Registration).

Information on hotels near the University is given at
http://www.gfkl.de/gfkl2002/
(menu item Hotels Near Conf. Site).

On Sunday evening (July 21, 2002) there is the possiblity of an informal meeting of conference participants during the get-together in the Rheinterrassen, a beer garden and restaurant near the University at the banks of the river rhine. The location of the Rheinterrassen is indicated on the map at
http://www.gfkl.de/gfkl2002/
(menu item Conference Site).

On Monday (July 22) at 6:00 pm we will have the reception of the University's president in the Rittersaal of the Schloss; on Tuesday (July 23) we will start the excursion to Heidelberg (cocktails on the terrace of Heidelberger Schloss and subsequent dinner in the Schlossweinstuben) at 5:45 pm.
If there is a large enough group of interested participants, there is the possibiltiy to drive to the entrance of Heidelberg's Philosophenweg and to reach the castle by foot via Philosophenweg, Alte Brücke, and Altstadt.

I hope to meet many participants of IFCS member societies in Mannheim.

Martin Schader
(Local Organizer GfKl 2002)

10th Annual meeting of the SFC EGC 2003

The 10th Annual meeting of the SFC will take place on september 10-12, 2003 at Neuchâtel (Switzerland). It is organized by the statistics group of the University of Neuchâtel. The chair of the scientific program commitee is Prof. Yadolah DODGE. The chair of the local organizing committee is Dr G. Melfi.

André HARDY, Secretary of the SFC
Email: Andre.Hardy@fundp.ac.be
http://www.sfc.fundp.ac.be

EGC 2003

Journées francophones d'Extraction et de Gestion des Connaissances
(French-speaking community conference on Knowledge Extraction and Managment)
January 22 – 24, 2003, Lyon, France

URL: http://www.univ-lyon3.fr/EGC2003
email: egc2003@bat710.univ-lyon1.fr

Contact person: Pr. Mohand-Said Hacid (http://bat710.univ-lyon1.fr/~mshacid)

The EGC series of conferences is a forum for the exchange of the latest research results involving french-speaking researchers and industrials from the areas of machine learning, statistics and data analysis, information systems and databases, and knowledge engineering. The previous events were held in Montpellier (http://www.lirmm.fr/EGC2002/) and Nantes.
(http://www.sciences.univ-nantes.fr/irin/EGC2001/)

EGC provides a medium for exchanging scientific research and technological achievements accomplished by the french-speaking community. The conference addresses issues involving solutions of problems regarding data and knowledge representation, pre-processing of huge amount of data, algorithms for knowledge extraction (from databases, text, videos, web data, etc.), intelligent vizualisation, datawarehouses, semantic web, web intelligence, business intelligence, etc.

Open Mailing List GfKl

In December 2001 the GfKl (German Classification Society) has introduced an e-mailing list in order to improve communication between its members and interested non-members. In the meantime there are 65 participants on this list. The GfKl would like to declare this list open to all members of other classification societies. In particular, the list is willing to announce conferences and workshops of the classification societies. Moreover, you might ask for support for the solution of a classification or data analysis problem. Also, you might want to discuss political decisions or public events relevant to the classification societies. Note however, that the languages of the list are German and English.

If you want to join the list, please e-mail the administrator under the address pr@gfkl.de identifying the society you are member of. Then you will be subscribed on the list, and further instructions for its usage will be given.

Prof. Dr. C. Weihs
Officer for public relations of the GfKl

IFCS Newsletters on Internet:

http://edfu.lis.uiuc.edu/~class/ifcs/newsletter.html

The newsletters are available as PDF files, to be read with AcrobatReader. Click on Acrobat Reader to download.

IFCS Homepage:

http://edfu.lis.uiuc.edu/~class/ifcs/
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.

Contributions for the coming issue of the IFCS newsletter can be sent to:

jajuga@manager.ae.wroc.pl

Text files are by preference in ascii or word, with pc format. For graphical materials, by preference GIF is used.

Editor:
Krzysztof Jajuga
Wroclaw University of Economics
Komandorska 118/120
53-345 Wroclaw
Poland
tel: +48 71 3680340 fax: +48 71 3680322

The present newsletter is made in collaboration with the new editor, Krzysztof Jajuga. I want to congratulate him as the new Publication Officer and I wish him all the best in this job.

Paul De Boeck