



Vereniging voor Ordinatie en Classificatie

Voorzitter: Ron Wehrens, Radboud University Nijmegen, Analytical Chemistry, Toernooiveld 1, 6525 ED Nijmegen (r.wehrens@science.ru.nl)
Secretaris: Hugo Duivenvoorden, Erasmus Universiteit, Fac. Geneeskunde en Gezondheidswetenschappen, Postbus 1738, 3000 DR Rotterdam (h.duivenvoorden@erasmusmc.nl)
Penningmeester: Berrie Zielman, Algemene rekenkamer, Directie Beleid en Communicatie, Afd. Statistiek, Lange Voorhout 8, 2414 ED Den Haag (a.Zielman@rekenkamer.nl) Postbankrekening 161723 t.n.v. Vereniging voor Ordinatie en Classificatie, Louise Henriettestraat 163, 2595 TP Den Haag.
Redactie: Eva Ceulemans, Centrum voor methodologie van het pedagogisch onderzoek, Katholieke Universiteit Leuven, Vesaliusstraat 2, B-3000 Leuven, België (Eva.Ceulemans@ped.kuleuven.be)
VOC-home page: <http://www.voc.ac>

Spring Meeting of the VOC

May 28, 2010

Erasmus Universiteit Rotterdam,
Campus Woudestein, J-Gebouw (J1-41)

10.00	Registration and Coffee
10.15	Jelke Bethlehem: About the quality of surveys
11.00	Coffee & VOC annual member meeting
11.30	Berrie Zielman: Effectiveness of a policy measure for reducing violence in nightlife
12.00	Lunch
13.00	Jean Pierre Verhaeghe: How to make educational practitioners understand the concept of "value added"?
13.35	Ruud Hoogendoorn
14.10	Tea
14.25	Jeroen van Oostrum: Applying mathematical models to surgical patient planning
15.00	Elise Dusseldorp: Treatment Interaction Trees (TINT): A tool to identify disordinal treatment-subgroup interactions
15.45	Drinks

In this issue:

Program Spring Meeting	1
Registration details	1
From the president	2
News from the IFCS	2
Abstracts of the VOC Spring Meeting	2
Publications	5
Book Review	6
Financieel overzicht over het jaar 2009	9
Jaarverslag van de Secretaris over 2009	10
Notulen Ledenvergadering VOC 17 april 2009	10
Agenda voor de ledenvergadering van de VOC	11
Personalialia	11
Meeting	11
Route description	12

Registration details for the Spring Meeting:

Those who would like to participate are welcome and are kindly requested to register at the VOC website (<http://www.voc.ac/html/registration.htm>), or by sending an e-mail to meeting@voc.ac. Registration deadline: May 21nd.

From the President

The VOC spring meeting will this year be held on May 28, in Rotterdam. A bit later than usual, perhaps, but we did not want to be too close to the Statistical Day of the VVS (April 1st), where the VOC also hosted a session, chaired by Mark de Rooij. The theme of this year's spring meeting is "Statistics and Policy", with a keynote address by prof. Bethlehem (UvA). The full programme can be found on the (brand-new) VOC website and elsewhere in this newsletter.

Traditionally, this is also the occasion of the annual members meeting of the VOC. This is your chance to speak up and tell the board what you think - "live"! The reports of the secretary and the treasurer are included in this newsletter. And then there is the opportunity to vote for new board members: two positions will be open. After serving many years in the board, Eva Ceulemans has decided it is time for others to take up her tasks - the newsletter has been in safe hands for the last years. Moreover, she has always been very much involved in organizational issues, an example being last year's jubilee conference in Wageningen. The other open position is that of the chairman. Because of my relocation to Italy, where I am now working at the Fondazione Edmund Mach, it is no longer possible for me to be actively involved in the VOC, something that I regret because I have really enjoyed the atmosphere both within the board and at VOC meetings. Interested candidates can present themselves up to 24 hours before the meeting with the secretary, Hugo Duivenvoorden.

At this point in time it is still uncertain that I will be able to attend the spring meeting - if not, I would like to take this opportunity to wish the VOC the very best: I am sure the need for dissipating statistical know-how and bringing statisticians of very different backgrounds together, two of the main goals of the VOC, is as large as ever and will only grow in the foreseeable future. So hopefully we will meet again, if not at this meeting, then at one of the next.

Ron Wehrens

News from the IFCS

We are happy to announce that Iven Van Mechelen (K.U.Leuven), a member of the VOC, has been elected President-Elect of IFCS. Iven will be President-Elect till the end of 2011, President till the end of 2013 and, finally, Past President till the end of 2015.

Abstracts for the Spring Meeting

Jelke Bethlehem (Statistics Netherlands, Methodology Department): About the quality of surveys

Surveys research is a type of research where data is collected by asking questions to a sample of persons from a population. On the basis of the collected data, conclusions are drawn about the population as a whole. The question is whether this always is a scientifically sound research method.

The presentation gives an historic overview of how surveys indeed became a reliable research instrument. However, not every survey is a good survey. If a survey is not properly designed, wrong conclusions may be drawn. This presentation is about some of the methodological problems and their consequences.

The fast development of the Internet has led to a new form of survey, and this is the web survey. Almost everybody can do a web survey. There are many examples of badly designed web surveys. They suffer from methodological problems like under-coverage and self-selection.

Unfortunately, the media are not able to distinguish the good from the bad. Therefore, this presentation can be seen as a warning not to believe everything that is reported about survey results.

Jelke Bethlehem studied mathematical statistics at the University of Amsterdam. After obtaining his pre-doctoral degree he worked for the Mathematical Centre in Amsterdam. The focus of his work was multivariate statistical analysis and development of statistical software.

In 1978 he joined the Department for Statistical Methods of Statistics Netherlands. His main topics were the treatment of nonresponse in sample surveys, in which he obtained his Ph.D., and disclosure control of published survey data. From 1987 to 1996 he was head of the Statistical Informatics Department, which developed standard software for collecting and processing survey data. He was responsible for the development of the Blaise System for computer-assisted survey data collection.

Currently, he is Senior Advisor of the Methodology Department. He is involved in research projects in the area of survey methodology (nonresponse, web surveys), and he co-ordinated a European research project. He is also part-time professor in Statistical Information Processing at the University of Amsterdam.

Berrie Zielman (the Netherlands court of audit): Effectiveness of a policy measure for reducing violence in nightlife

Increasing cooperation between different parties such as bar owners, municipalities, doormen and police force was proposed as a policy measure. The tasks and duties of the different parties were described in an agreement. The effectiveness of the agreement was studied using police records. From the police databases records of violence in a nightlife area were extracted and aggregated on a monthly basis. Collecting these data for a four year period we obtained a series of count data. To answer the question whether the introduction of an agreement was effective, we used Poisson regression.

Berrie Zielman studied psychology at the university of Leiden, where he also obtained his PhD. The topic of his thesis was the analysis of skew-symmetry. Currently he is a statistical consultant/ researcher at the Netherlands court of audit. At the NCA he reviews statistical procedures used by the government, designs sample plans and performs statistical analyses.

Jean Pierre Verhaeghe (Ugent): How to make educational practitioners understand the concept of “value added”? Implications for statistical analysis and school performance feedback practice

A key element in many school performance feedback practices is the concept of “value added”, also known as “residual gain score”. It refers to the notion that schools should not be compared with other schools based on the raw performance scores of their students. To make fair comparisons, the influence of student and family background characteristics and eventually school context should be filtered out. What is then left – the net school effect – is considered to reflect the “real” impact of the school on students’ learning.

There are two ways to explain the concept of value added in school feedback reports: either as the difference between a school’s observed mean and its expected mean, or as the difference between the school’s adjusted mean and the grand mean for the reference category. From a pure mathematical point of view, both ways are strictly equivalent. Following the output provided by the statistical software they use, many researchers are inclined to use the latter way. Research on school practitioners’ understanding of school performance feedback however shows that the concept of value added is better understood when the first way is used. Since software such as MLWin does not provide school level predictions including fixed effects other than the intercept (expected means and standard errors), putting this finding into school feedback practice becomes quite challenging. The Flemish School Feedback Project has tried to tackle this problem, taking its basic principles into consideration: (1) the school feedback is based on a comparison of new school data with an available

representative data set, and (2) the school feedback should be provided in a fully automated way, which includes automation of all processes involved: data collection, data processing and the production and distribution of school tailored feedback reports. The present paper focuses on school feedback based on one measurement occasion, but the findings can be extended to feedback on more measurement occasions.

Jean Pierre Verhaeghe obtained his PhD in Educational Sciences in 1994. Currently he is a senior researcher at UGent and KULeuven. He coordinates the School Feedback Project (KULeuven - UGent - UAntwerpen). His fields of interest are school effectiveness, growth curve modeling and IRT-analysis.

Jeroen van Oostrum (Erasmus University Medical Center): Applying mathematical models to surgical patient planning

On a daily basis surgeons, nurses, and managers face cancellation of surgery, peak demands on wards, and overtime in operating rooms. Moreover, the lack of an integral planning approach for operating rooms, wards, and intensive care units causes low resource utilization and makes patient flows unpredictable. An ageing population and advances in medicine are putting the available healthcare budget under great pressure. Under these circumstances, hospitals are seeking innovative ways of providing optimal quality at the lowest costs.

I developed during my PhD research instruments for optimizing surgical patient planning on basis of a cyclic and integrated operating room planning approach, called master surgical scheduling. One of the stages in this approach is clustering individual surgical cases into standardized surgical cases that can be scheduled in a repetitive manner. Moreover, I studied additional models to deal for instance efficiently with emergency operations. Application of these instruments enables the simultaneous optimization of the utilization of operating rooms, ward and intensive care units. Moreover, iteratively executing a master schedule of surgical case types provides steady and thus more predictable patient flows in hospitals.

The approach is generic and so can be implemented taking account of specific characteristics of individual hospitals. Prerequisites for successful implementation of logistical models in hospitals comprise sufficient room for last-minute changes as well as keeping the ultimate responsibility for individual patient scheduling with medical specialists. Both are satisfied in the master surgical scheduling approach which has already been successfully implemented in hospitals.

Jeroen van Oostrum (1980) studied applied mathematics at the University of Twente. From 2004 onwards he has worked with applications of Operations Research in health care and performed research within the Expertisecentrum Erasmus Health Care Logistics. Jeroen

obtained his PhD in 2009 by defending his thesis “Applying mathematical model to surgical patient planning”.

Currently Jeroen is head of the Business Intelligence Center of Erasmus University Medical Center. In this position he is on a daily basis involved with operations research and management control challenges. Jeroen is on a part-time basis affiliated to the department of Health care management (Institute Health Policy & Law) at the Erasmus University Rotterdam.

which the combined influences of factors on an outcome can be assessed.

Elise Dusseldorp (TNO Quality of Life and Department of Psychology, K.U.Leuven): Treatment INteraction Trees (TINT): A tool to identify disordinal treatment-subgroup interactions

When two competitive treatments, A and B, are available, some subgroup of patients may display a better outcome with treatment A than with B, whereas for another subgroup the reverse may be true. If this is the case, a disordinal (i.e., a qualitative) treatment-subgroup interaction is present. Such interactions imply that some subgroups of patients should be treated differently, and are therefore most relevant for policy with regard to assignment of patients to programs (i.e., treatment plans). In case of data from randomized clinical trials with many patient characteristics that could interact with treatment in a complex way, a suitable statistical approach to detect disordinal treatment-subgroup interactions is not yet available. In this presentation, we introduce a new method for this purpose, called Treatment INteraction Trees (TINT). TINT results in a binary tree that subdivides the patients into terminal nodes on the basis of patient characteristics; these nodes are further assigned to one of three classes: a first one for which A is better than B, a second one for which B is better than A, and an optional third one for which type of treatment makes no difference. The tree can be used to develop rules for patient-tailored treatment assignment. Results of an application of TINT to real data from the Breast Cancer Recovery Project will be shown. A short demonstration of the software will be given.

Elise Dusseldorp studied psychology at the university of Leiden, where she also obtained her PhD in 2001. Elise is a statistician at the Quality of Life division of TNO (the Netherlands Centre for Applied Scientific Research) in Leiden, and a post-doctoral researcher in the Quantitative Psychology and Individual Differences Research Group of the Catholic University Leuven, Belgium. Her main area of interest in the field of statistics is the modelling of interaction effects in prediction problems. While many statistical models focus on the unique effects of predictive factors on one or more outcome variables, her challenge is to explore the way in

Publications

- Bannink, E. M. N., van Sassen, C., van Buuren, S., de Jong, F. H., Lequin, M., Mulder, P. G. H., & de Muinck Keizer-Schrama, S. M. P. F. (2009). Puberty induction in Turner syndrome: Results of estrogen treatment on development of secondary sexual characteristics, uterine dimensions, and serum hormone levels. *Clinical Endocrinology*, 70(2), 265-273. doi:10.1111/j.1365-2265.2008.03446.x
- Braeken, J., & Tuerlinckx, F. (2009). A mixed model framework for teratology studies. *Biostatistics*, 10, 744-755.
- Braeken, J., & Tuerlinckx, F. (2009). Investigating latent constructs with item response models: A MATLAB IRTM toolbox. *Behavior Research Methods*, 41, 1127-1137.
- Busing, F.M.T.A., & De Rooij, M. (2009). Unfolding incomplete data. *Journal of Classification*, 26, 329-360.
- Ceulemans, E., & Storms, G. (2010). Detecting intra and inter categorical structure in semantic concepts using HICLAS. *Acta Psychologica*, 133, 296-304.
- de Kroon, M. L. A., Renders, C. M., van Wouwe, J. P., van Buuren, S., & Hirasing R. A. (2010, February 11). The Terneuzen birth cohort: BMI changes between 2 and 6 years correlate strongest with adult overweight. *PLoS ONE*, 5(2), e9155.
- De Rooij, M. (2009). Trend vector models for the analysis of change in continuous time for multiple groups. *Computational Statistics and Data Analysis*, 53, 3209-3216.
- De Rooij, M. (2009). Ideal point discriminant analysis with a special emphasis on visualisation. *Psychometrika*, 74, 317-330.
- Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, 16, 1026-1036.
- Eilers, P. (2009). Imtech Contracting Industry park profits from full-service-contracting. *Bwk*, 61(4), 12-13.
- Eilers, P. H. C., Li, B., & Marx, B. D. (2009). Multivariate calibration with single-index signal regression. *Chemometrics and Intelligent Laboratory Systems*, 96(2), 196-202.
- Goegebeur, Y., De Boeck, P., & Molenberghs, G. (2010). Person fit for test speededness: normal curvatures, likelihood ratio tests and empirical Bayes estimates. *Methodology*, 6, 3-16.
- Gravendeel, L. A. M., Kouwenhoven, M. C. M., Gevaert, O., de Rooi, J. J., Stubbs, A. P., Duijm, J. E., ... French, P. J. (2009). Intrinsic Gene Expression Profiles of Gliomas Are a Better Predictor of Survival than Histology. *Cancer Research*, 69(23), 9065-9072.
- Heerdink-Obenhuijsen, N., van Dommelen, P., Kamphuis, M., van Buuren, S., Coenen-van Vroonhoven, E. J. C., & Verkerk, P. H. (2010). *JGZ-Richtlijn. Signalering van en verwijscriteria bij kleine lichaams lengte*. Bilthoven: RIVM, Centrum Jeugdgezondheid.
- Heymans, M. W., Anema, J. R., van Buuren, S., Knol, D. L., van Mechelen, W., & de Vet, H. C. (2009). Return to Work in a Cohort of Low Back Pain Patients: Development and Validation of a Clinical Prediction Rule. *Journal of Occupational Rehabilitation*, 19, 155-165. doi:10.1007/s10926-009-9166-3
- Ip, E. H., Smits, D. J. M., & De Boeck, P. (2009). Locally dependent linear logistic test model with person covariates. *Applied Psychological Measurement*, 33, 555-567.
- Kahraman, N., De Boeck, P., & Janssen, R. (2009). Modeling DIF in complex response data using test design strategies. *International Journal of Testing*, 9, 151-166.
- Lambert, P., & Eilers, P. H. C. (2009). Bayesian density estimation from grouped continuous data. *Computational Statistics & Data Analysis*, 53(4), 1388-1399.
- Marx, B. D., Eilers, P. H. C., Gampe, J. & Rau, R. (2010). Bilinear modulation models for seasonal tables of counts. *Statistics and Computing*, 20(2), 191-202.
- Mook-Kanamori, D. O., Steegers, E. A. P., Eilers, P. H., Raat, H., Hofman, A. & Jaddoe, V. W. V. (2010). Risk Factors and Outcomes Associated With First-Trimester Fetal Growth Restriction. *Jama-journal of the American Medical Association* 303(6), 527-534.
- Oravec, Z., Tuerlinckx, F., & Vandekerckhove, J. (2009). A hierarchical Ornstein-Uhlenbeck model for continuous repeated measurement data. *Psychometrika*, 74, 395-418.
- Polak, M. G., Heiser, W. J., & De Rooij, M. (2009). Two types of single-peaked data: correspondence analysis as an alternative to principal component analysis. *Computational Statistics and Data Analysis*, 53, 3117-3128.
- Schepers, J., & Van Mechelen, I. (2010). Uniqueness of real-valued hierarchical classes models. *Journal of Mathematical Psychology*, 54, 215-221.
- Schnabel, S. K. & Eilers, P. H. C. (2009). An analysis of life expectancy and economic production using expectile frontier zones. *Demographic Research* 21, 5.
- Schnabel, S. K., & Eilers, P. H. C. (2009). Optimal expectile smoothing. *Computational Statistics & Data Analysis*, 53(12), 4168-4177.
- Spinhoven, P., De Rooij, M., Heiser, W. J., Penninx, B., & Smit, J. (2009). The Role of Personality in Comorbidity among Anxiety and Depressive Disorders in Primary Care and Specialty Care: A Cross-Sectional Analysis. *General Hospital Psychiatry*, 31, 470-477.

Timmerman, M. E., Ceulemans, E., Kiers, H. A. L., & Vichi, M. (2010). Factorial and Reduced K-means reconsidered. *Computational Statistics and Data Analysis*, 54, 1858-1871.

van Buuren S (2010). Effects of Selective Dropout on Infant Growth Standards. In A. Lucas, M. Makrides, & E.E. Ziegler (Eds.): *Importance of Growth for Health and Development. Nestlé Nutr Inst Workshop Ser Pediatr Program, vol 65* pp. 167–179. Vevey: Nestec Ltd./ Basel: S. Karger AG. doi:10.1159/000281161

van Buuren, S. (2010). Item Imputation without specifying Scale Structure. *Methodology*, 6(1), 31-36.

van Buuren, S., Hayes, D. J., Stasinopoulos, D. M., Rigby, R. A., ter Kuile, F. O., & Terlouw, D. J. (2009). Estimating regional centile curves from mixed data sources and countries. *Statistics in Medicine*, 28(23), 2891-2911. doi:10.1002/sim.3667

van Buuren, S., & Ooms, J. C. L. (2009). Stage line diagram: An age-conditional reference diagram for tracking development. *Statistics in Medicine*, 28(11), 1569-1579. doi:10.1002/sim.3567

van den Berg, R. A., Van Mechelen, I., Wilderjans, T. F., Van Deun, K., Kiers, H. A. L., & Smilde, A. K. (2009). Integrating functional genomics data using maximum likelihood based simultaneous component analysis. *BMC Bioinformatics*, 10, 340.

van Dommelen, P., Grote, F. K., Oostdijk, W., de Muinck Keizer-Schrama, S. M., Bouquet, J., Hendriks, J. J., ... Wit, J. M. (2009). Growth monitoring to detect children with cystic fibrosis. *Hormone Research*, 72(4), 218-224.

van Eijk, R., Eilers, P. H. C., Natte, R., Cleton-Jansen, A. M., Morreau, H., van Wezel, T. & Oosting, J. (2010). MLPAinter for MLPA interpretation: an integrated approach for the analysis, visualisation and data management of Multiplex Ligation-dependent Probe Amplification. *Bmc Bioinformatics*, 11, 67.

van Wouwe, J. P., Lanting, C. I., van Dommelen, P., Treffers, P. E., & van Buuren, S (2009). Breastfeeding duration related to practised contraception in the Netherlands. *Acta Paediatrica*, 98(1), 86-90.

Visser, G. H. A., Eilers, P. H. C., Elferink-Stinkens, P. M., Merkus, H. M. W. M. & Wit, J. M. (2009). New Dutch reference curves for birthweight by gestational age. *Early Human Development*, 85(12), 737-744.

Vlioger, A. M., Robroch, A., van Buuren, S., Kiers, J., Rijkers, G., Benninga, M. A., & Te Biesebeke, R. (2009, March 31). Tolerance and safety of *Lactobacillus paracasei* ssp. *paracasei* in combination with *Bifidobacterium animalis* ssp. *lactis* in a prebiotic-containing infant formula: a randomised controlled trial. *British Journal of Nutrition*. Advance online publication. doi:10.1017/S0007114509289069

Weisglas-Kuperus, N., Hille, E. T. M., Duivenvoorden, H. J., Finken, M. J. J., Wit, J. M., van Buuren, S., ... Verloove-Vanhorick, S. P. (2009). Intelligence of Very Preterm or Very Low Birth Weight Infants in Young Adulthood. *Arch. Dis. Child. Fetal Neonatal Ed.*, 94(3), F196-200. doi:10.1111/j.1365-2265.2008.03446.x

Wilderjans, T. F., Ceulemans, E., Kiers, H. A. L., & Meers, K. (2009). The LMPCA program: A graphical user interface for fitting the linked-mode PARAFAC-PCA model to coupled real-valued data. *Behavior Research Methods*, 41, 1073-1082.

Book review

Statistical Analysis of Network Data: Methods and Models. E. D. Kolaczyk. Springer. ISBN: 978-0-387-88145-4.

We live in a connected world and increasingly we are becoming aware of this. We are confronted with the interconnectedness of human institutions and processes, for example, in the form of governments and economies (think of the worldwide economic crisis, and the financial problems in Greece), just as we are aware of the interconnectedness of natural systems, for example, in ecology (food webs), and genetics and genomics (gene networks, metabolite and protein networks). A famous network is the one constituted by the seven Königsberg bridges, of which Euler showed in 1735 that it was impossible to walk around the city of Königsberg and passing each bridge only once. This work very likely exemplified the first network analysis. Another famous network is the one investigated by Stanley Milgram, who in 1967 showed that people are on average separated by only 6 others from anyone else on this planet (six degrees of separation).

Some topics that relied on network technology in the near past were the theory of electrical circuits, molecular structure in chemistry, transport and allocation problems, and social networks. However, recently dramatic changes have occurred in science that ask for a fast development of network methodology and model. Firstly, there is a shift in scientific perspective from reductionist approaches to systems-level approaches, accompanied by, secondly, high throughput data collection, storage and management. Especially the developments in biology stimulate the use and creation of network technology. Another strong stimulus for network technology comes from studies on internet traffic.

Many disciplines are nowadays involved in network modeling, but it appears as if a common methodological foundation is lacking. The objective of this book is to provide a first attempt at defining such a common methodological foundation from a statistical point of view. Statistically, a major challenge lies in the relational

nature of networks with measurements being dependent and high dimensional. Networks are nowadays used in a multitude of contexts, each with specific questions. Still, a core set of statistical principles can be derived that supersede the individual applications by imposing a clear statistical taxonomy on the wide range of contexts. The organization of the book follows the classical statistical taxonomy of data description, sampling, and modeling and inference.

The first chapter contains an Introduction and Overview. This chapter describes clearly the motivation of the author to write this book (create a core of statistical network methodology), gives nice examples of networks in different disciplines and explains the structure of the rest of the book.

The second chapter contains a convenient summary of technical material on graphs, probability and statistical inference that is necessary for the rest of the book. The chapter offers a good self check on whether you are ready for the rest of the book. Still, for later chapters more statistical knowledge and skills are required than may be expected after reading chapter 2.

Chapter 3 is on constructing and visualizing networks from data, where the construction is on a descriptive level, i.e., only simple rules are used for whether or not to include an edge between two vertices. Model building and inference follow in later chapters. Although for people interested in networks, a substantial part of the material covered in this chapter will be familiar, it is remarkable how the author, even on this topic, is able to transcend the obvious and produce useful generalizations and new insights.

The next chapter, 4, is dedicated to descriptive statistics of network graphs. Vertex and edge characteristics are described (degree = the number of edges coming in / going out of a vertex; degree distribution across the vertices; degree correlation describing the patterns of association in the network; centrality = importance of a vertex in the whole of the network). The use of these characteristics for interpretation of graphs is well explained. All along the book there is ample attention for explaining why certain statistics and procedures are important and useful and how they can help in answering subject matter questions. Another good feature of the book is that it places network technology well within standard statistical methodology. In this chapter, classical hierarchical cluster analysis and singular value decomposition are integrated with modern network graph description tools to produce tools that were not evident beforehand.

With chapter 5 the inferential part of the book starts: sampling and estimation in network graphs. The author succeeds in showing how classical sampling ideas are also applicable to networks. Various sampling strategies are described and it is shown how the sampling strategy will

affect inference on network characteristics as degrees, vertex totals, totals on vertex pairs, group size, etc.

The next chapter, 6, deals with modeling network graphs. It starts with a discussion of random graph models, building on the seminal work of Erdős and Renyi, who developed a formal probabilistic approach to inference on the characteristics of a graph, by considering the observed graph to be a realization from a (uniform) distribution of graphs sharing the same number of vertices and edges. Of special interest for these graphs is the degree distribution, where the classical random graphs tend to a Poisson distribution of the kind observed for many large-scale real-world networks, with little clustering of vertices. Random graph models are used for the detection of motifs, i.e., small subgraphs occurring more frequently than may be expected on the basis of a random graph. 'Significance' of motifs is derived from simulations of appropriate random networks, the latter being a theme of intense research.

After the random graphs, small-world models are discussed in chapter 6, of which the Watts-Strogatz model is the best known representative. These models are studied extensively in social science and epidemiology. In small-world models, neighbour-to-neighbour exchanges are sufficient to transmit information fast. Still in the same chapter, we find a treatise on models for network growth, with applications to the World Wide Web, scientific citations and many biological and social networks. Data on network growth consist typically of a series of snapshots, complicating the fitting of network growth models. Recursive likelihood approaches have been proposed, but these are difficult to apply for any network not being of very modest size. Solutions are sought in sequential importance sampling.

The above network models were all useful, but difficult to estimate and hard to assess goodness of fit of. In contrast, the exponential random graph models, depart from a statistically well defined framework: exponential family distributions are assumed for the joint distribution of presence and absence of edges in the networks. Actually, these models are graphical models whose properties are discussed in chapter 10. In chapter 6 examples are given of Bernoulli random graphs and Markov random graphs. The likelihoods are again complicated to calculate due to the dependencies induced by the network structure, but approximations by MCMC methods may offer a way out.

Chapter 7 looks at network topology inference, where we try to infer an unobserved portion of a graph from measurements on observed vertices and edges. For example, in link prediction, where we predict a link between two vertices from properties of the vertices. This boils down to a kind of logistic regression. Section 7.3 deserves special mentioning, as it treats the construction of association networks, joining up the dots (vertices) based on a data observed on the dots. This is an activity for which many researchers are getting increasingly interested for application to their own data. The section

runs through correlation networks, partial correlation networks, and Gaussian graphical models and shows how multiple testing issues and penalized regressions have become techniques of choice in inference on association networks. The last section of chapter 7 is on inferring 'interior' components of networks. Here attention focuses on trees, a sub class of graphs, where there are known, observed vertices, the leaves, and we want to infer unobserved vertices giving rise to the leaves. This problem is well known in phylogenetics, where hierarchical and model based clustering methods are used for solution.

Chapter 8 takes the graph to be known and looks at prediction of processes on the graph. An example is the prediction of protein function as a function of the structure of the graph and the properties of its vertices. Techniques that are discussed are nearest neighbor prediction with the neighborhood of the vertex deciding on its function value, Markov random fields and kernel-based regression. The approaches being proposed are known from image analysis, times series and spatial analyses and rely heavily on Bayesian implementations using MCMC methods.

The penultimate chapter, 9, discusses the analysis of network flow data, where we can think of transportation networks, traffic, social and economic processes. Gravity models predict flows from a product of functions on properties of 1) origin vertices, 2) destination vertices,

and 3) a vector of separation (dissimilarity) attributes between origin and destination. In particular cases, these models have the form of log linear models.

The final chapter, 10, gives a rather brief overview of the properties of graphical models, directed and undirected, and shows how a number of models that were discussed in earlier chapters can be treated as graphical models.

It may be apparent of the above description that I am rather enthusiastic about this book. The author has succeeded in writing an accessible, but far from simplified, treatment of current network methodology. The perspective is statistical and the organization of the material logical and didactically sound. The style of the writing is excellent. Because of the width of the book, most techniques are only succinctly described, but at any time ample references allow quick access to further literature. I can recommend this book to anyone with a serious statistical interest in networks.

Fred van Eeuwijk
WUR-Biometris
Wageningen

Financieel overzicht over het jaar 2009

Ontvangsten		Uitgaven	
contributie	1220	betalingsverkeer	62,98
jubileum	8220	kvk	26,14
rente plus rekening	24,45	vocbijeenkomst eur	133,5
sponsor RU nijmegen	500	website voc	119
aned wageningen	252,1	reiskosten c hennig	157,13
<u>lunchgeldten tilburg</u>	<u>100</u>	provisie	5
		geschenk ron wehrens	46,25
		kosten dexia bank	9,67
		<u>classification society</u>	<u>136,18</u>
Totaal	10316	Totaal	695

BALANS

Debet		Credit	
Saldo ING Rekening	10491	Crediteuren	7670
Saldo spaarrekening	2656	<u>Eigen vermogen</u>	<u>8189</u>
<u>Dexia Rekening</u>	<u>2712</u>		
Totaal	15859		15859

Toelichting bij het financieel overzicht 2009

- (1) 61 leden hebben in 2009 contributie betaald.
- (2) Het eigen vermogen is licht gestegen doordat de Jubileum bijeenkomst een kleine winst heeft gegenereerd.
- (3) In de balans is een post crediteuren opgenomen omdat de rekening voor de Wageningse berg bij het sluiten van de boeken nog niet betaald is.
- (4) De Dexia rekening wordt dit jaar opgeheven, omdat deze rekening in het euro tijdperk geen lagere kosten met zich meebrengt. In het verleden konden de Vlaamse leden de contributie naar deze rekening overmaken. Voor het betalingsverkeer tussen Nederland en België worden op dit moment geen kosten in rekening gebracht.
- (5) Een overzicht van de ontwikkeling van het eigen vermogen:

2009	€8189
2008	€6248
2007	€5914
2006	€6869
2005	€6057
2004	€5019
2003	€6795
2002	€6408
2001	€5898
2000	€5731
1999	€4871
1998	€5100

Jaarverslag van de Secretaris over 2009

1. De VOC startte eind 2008 met 106 leden en eindigde eind 2009 met 94 leden. Er waren vijftien opzeggingen, en drie aanmeldingen.

2. Bestuur

Het bestuur van de VOC had in 2009 de volgende samenstelling:

Ron Wehrens	Voorzitter
Hugo Duivenvoorden	Secretaris
Berrie Zielman	Penningmeester
Eva Ceulemans	Redacteur Nieuwsbrief
Michel van der Velden	Webmaster
Mark de Rooij	Lid

Het bestuur vergaderde in 2009 vier maal, waarvan eenmaal telefonisch. De belangrijkste onderwerpen waren de Bijeenkomsten, en lopende zaken als de Nieuwsbrief en de website.

3. Activiteiten

In 2009 hadden we twee goed bezochte bijeenkomsten. De Voorjaarsbijeenkomst (17/4/09, Tilburg University) had een gevarieerd thema. De sprekers van die dag waren achtereenvolgens: Geert Molenberghs ('A latent variable mixture model as a basis for sensitivity analysis in incomplete longitudinal data'), Gerrit Gort ('Codominant scoring of AFLP: an application of normal mixture models'), Andries van der Ark ('A new reliability coefficient base on latent class analysis'), Carlos Hernandez ('Timing and speed of new product price landings'), Tomoki Tokuda ('Bayesian mixture modelling with variable selection'), en Christian Hennig: ('Merging normal mixture components').

De Jubileumbijeenkomst over 2 dagen (12-13/11/09, 40 deelnemers, Wageningse Berg, Wageningen) was een doorslaand succes, met bijdragen uit uiteenlopende disciplines. De sprekers op de eerste dag waren: Trudy Dehue ('Other times, other suffering: on the changing classification of despair'), Age Smilde ('From metabolomics data to biological networks and back'), Christian Steglich ('Modelling interdependent actors: cross-sectional and longitudinal approaches in social network analysis'), Cajo ter Braak ('Spectral decomposition and fuzzy clustering of network data with an application in genetics'), en Marcel Reinders ('Characterisation and inference of biological networks'). Op de tweede dag waren de sprekers: Iven van Mechelen ('Classification models to retrieve the sequential process basis of person-in-context behavior'), Marian Hickendorff ('Latent variable modeling of strategy choice and strategy accuracy in primary school mathematics'), Ingmar Visser ('Classification through time: Markov models for time series data'), Francis Tuerlinckx ('Some applications of stochastic differential equations in

psychological research'), Han van der Maas ('Sudden change and types'), Ritsert Jansen ('Gene and QTL networks') en Mark van der Laan ('Targeted maximum likelihood machine learning: Applications to causal effect/variable importance assessment and prediction with censored data').

De lezing van Korbinian Strimmer ('Gene sets, false discovery rates, and high-dimensional prediction') ging wegens ziekmelding niet door.

4. Publiciteit.

De Nieuwsbrief verscheen 2 maal. De Bijeenkomsten werden ook aangekondigd voor niet-leden, onder andere via de VVS-site en via mailing lists (IOPS, Stoch-Ned, Bio-MVA).

Notulen Ledenvergadering 17 april 2009

Verslag: Hugo Duivenvoorden

1. Opening en ingekomen stukken.

De voorzitter opent de vergadering. Er zijn geen ingekomen stukken. De voorzitter deelt mede dat de post vanaf nu per email verzonden zal worden: efficiënt en kosteneffectief; bovendien mogelijkheid tot email adres controle.

2. Notulen jaarvergadering 18 april 2008

Na een korte toelichting door Hugo worden de notulen goedgekeurd.

3. Jaarverslag 2008 van de secretaris

Het verslag wordt aanvaard.

4. Financieel verslag 2008

Het verslag van de penningmeester over 2008 wordt aanvaard, na een toelichting over tegenvallende inkomsten in 2008. Dat dient te worden toegeschreven aan de verlate inning van de contributie. Paul Eilers heeft de boeken gecontroleerd. Hij heeft aangegeven dat de boeken keurig in orde zijn. De vergadering kan zich daarin vinden. De nieuwe kascommissie bestaat uit Cees Elzinga en Paul Eilers.

5. Bestuurssamenstelling

Laurence Frank treedt reglementair af. Zij wordt heel hartelijk bedankt voor haar bijdragen aan het bestuurlijk functioneren van VOC. Michel heeft zich herkiesbaar gesteld als lid van het bestuur. Er zijn geen tegenkandidaten; Michel wordt bij acclamatie herkozen. Berry Zielman, die Laurence opvolgt, krijgt de functie van penningmeester.

6, 7, 8. Wvttk, Rondvraag en Sluiting

Verder niets ter tafel gekomen; voorzitter sluit de vergadering.

Agenda voor de ledenvergadering van de VOC op 28 mei 2010

1. Opening
2. Notulen Ledenvergadering 17 april 2009
Deze zijn elders in de Nieuwsbrief opgenomen.
3. Jaarverslag van de secretaris over 2009
Dit is elders in deze Nieuwsbrief opgenomen.
4. Financieel verslag
 - Jaarverslag van de penningmeester (zie deze Nieuwsbrief).
 - Verslag van de kascommissie (Cees Elzinga en Paul Eilers)
 - Benoeming van de nieuwe kascommissie.
5. Bestuurssamenstelling
Het bestuur bestaat uit de volgende leden (achter de functie staat de resterende duur van hun termijn):

Ron Wehrens	Voorzitter (0)
Hugo Duivenvoorden	Secretaris (1)
Eva Ceulemans	Redacteur Nieuwsbrief (0)
Mark de Rooij	Lid (1)
Michel van de Velden	Webredacteur (2)
Berrie Zielman	Penningmeester (2)

Ron Wehrens heeft per 1 januari 2010 het voorzitterschap van VOC opgezegd, en wel vanwege het aanvaarden van een baan in Italië. Het bestuur stelt voor om Ron te laten opvolgen door Jeroen Vermunt. Ook Eva, wier bestuurstermijn afloopt, zal het bestuur verlaten. Het bestuur stelt voor om Katrijn van Deun te benoemen in het bestuur als opvolger van Eva. Tegenkandidaten voor zowel Jeroen als Katrijn kunnen worden aangemeld bij de secretaris (h.duivenvoorden@erasmusmc.nl) tot 24 uur vóór de ledenvergadering. De leden mogen tijdens de vergadering zich hierover uitspreken.

6. Wvttk/Rondvraag
7. Sluiting

Personalialia

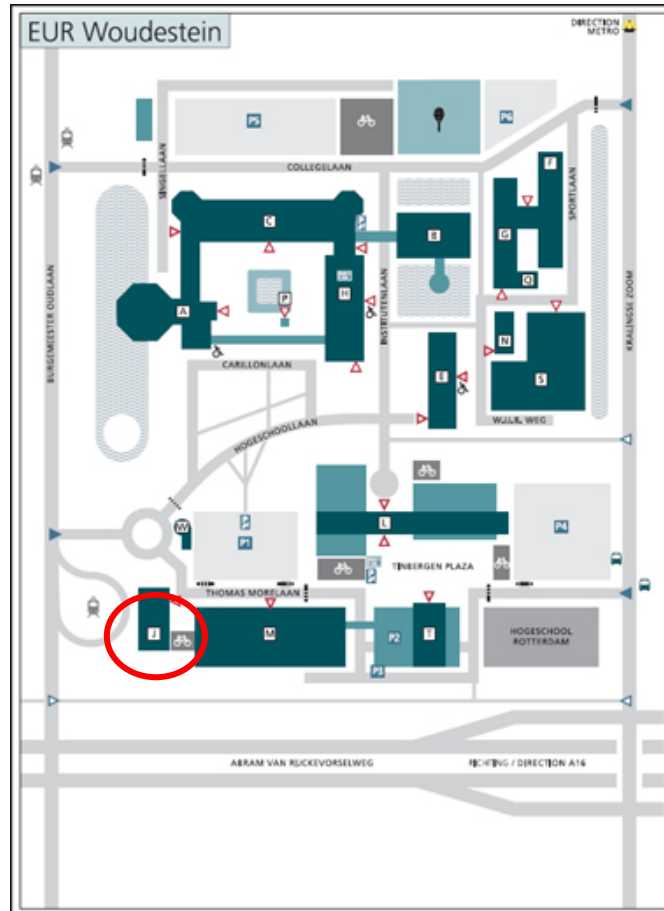
Wegens ruimtegebrek gratis op te halen: de (vrijwel complete) jaargangen 1991 tot en met 2008 van "Computational Statistics & Data Analysis". Alles moet in één keer mee; er kunnen geen afleveringen worden uitgezocht. Bel Niels H. Veldhuijzen, Cito, Arnhem: 026-3521416, of stuur een briefje naar: "niels.veldhuijzen@cito.nl".

Meeting

The Dutch Chemometrics Society invites you to the 26th Annual Symposium on Chemometrics, on May 20, 2010 in Utrecht, The Netherlands. This year's edition will be centred on two themes, namely Visualisation and the use of Prior Knowledge. More information can be found at the following website: www.chemometrie.nl (events and workshops)

Route description

The Fall Meeting takes place at the Erasmus Universiteit Rotterdam, Campus Woudestein, J-building J1-41, Burgemeester Oudlaan 50, 3062 PA Rotterdam.



Route description to campus Woudestein by Car

- The Woudestein campus is situated very near the Van Brienoordbrug ('brug' = bridge). The Van Brienoordbrug is crossed by the A16 motorway. This is where you have to get off the A16.
- If you come from the south (A16, direction Breda and Antwerp): keep following the signs Den Haag until you see the signs Capelle a/d IJssel and Rotterdam Centrum. Keep following the direction Rotterdam Centrum and you will see the direction to the university.
- If you come from the north (A13): keep following the direction Dordrecht/A16 until you see the sign Capelle a/d IJssel and Rotterdam Centrum. Keep following the direction Rotterdam Centrum and automatically you will see the direction of the university.
- If you come from the east (A12/A20): same route as coming from the north (Dordrecht/A16 and Rotterdam Centrum).

Route description to campus Woudestein by Public transport from NS Railway Station

- From Rotterdam Central Station

- with tram 21 in direction of De Esch. Get off at stop Woudestein.
- with tram 7 in direction Woudestein or Burg. Oudlaan. Get off at terminus.
- with metro in direction of Spijkenisse/Slinge. Change at station Beurs on metro in direction of Capelle a/d IJssel, Ommoord or Nesselande. Get off at stop Kralingse Zoom.

- From station Rotterdam Alexander

- with metro in direction of Schiedam Centrum/Spijkenisse. Get off at stop Kralingse Zoom.