Nieuwsbrief no. 39
Oktober 2007

Vereniging voor Ordinatie en Classificatie

Voorzitter: Ron Wehrens, Radboud University Nijmegen, Analytical Chemistry, Toernooiveld 1, 6525 ED Nijmegen (r.wehrens@science.ru.nl)
Secretaris: Marieke Timmerman, RU Groningen, Heymans Instituut (DPMG), Grote Kruisstraat 2/1, 9712 TS Groningen (m.e.timmerman@ppsw.rug.nl)
Penningmeester: Laurence Frank, Disciplinegroep M&T, Fac. Soc. Wet., Universiteit Utrecht, Postbus 80140, 3508 TC Utrecht (L.E.Frank@fss.uu.nl). Postbankrekening 161723 t.n.v. Vereniging voor Ordinatie en Classificatie, Naarden. Bankrekening nummer 777-5952385-56 Dexia Bank t.n.v. VOC, Naarden
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

VOC Fall Meeting

November 16, 2007

Pieter de la Court gebouw, Leiden

10.15	Registration and Coffee
10.45	Jos Twisk: *The analysis of recurrent event data: An overview*
11.30	Caspar Looman: *Decomposition techniques for Health Expectancy: Linking causes for disability (diseases) to disability prevalence*
12.00	Frans Oort: *Response shift and measurement bias*
12.30	Lunch
14.00	Hein Putter: *Joint analysis of multiple longitudinal outcomes: Application of a latent class model*
14.30	Luc Bonneux: *From bias to politics*
15.00	Tea
15.25	Richard Gill: *Lies, damned lies and legal truths: statistics and data-analysis in the courtroom*
16.10	Drinks

In this issue:

Program VOC Fall Meeting 1
Registration details 1
Van de voorzitter 2
Abstracts of the VOC Fall Meeting 2
Book reviews 5
Personalia 7
Agenda 7
Publications 8
Route description 10

Registration details for the VOC Fall Meeting:

Those who would like to participate are welcome and are kindly requested to register via the VOC website (http://www.voc.ac). Participation is free, however the lunch needs to be paid (10 euro). Registration deadline: November the 12th.

Van de voorzitter

In mijn eerste bijdrage aan de VOC-nieuwsbrief kan ik natuurlijk de gelegenheid niet voorbij laten gaan mijn voorganger, Patrick Groenen, te bedanken voor al het werk dat hij voor de VOC heeft gedaan, eerst in diverse bestuursfuncties en de laatste zes jaar als voorzitter. Hij heeft een bestuur achtergelaten dat draait als een geoliede machine, en ik ben ervan overtuigd dat we de komende jaren veel moois te zien gaan krijgen. Eén van de dingen die het vermelden waard is: volgend jaar houdt onze Duitse zustervereniging, GfKl, haar 32ste jaarlijkse conferentie, en wel van 16 - 18 juli (2008). Deze zal samen met de VOC en de Britse Classification Society worden georganiseerd. Noteert u het vast.

Natuurlijk zijn er al eerder activiteiten, zoals de traditionele najaarsbijeenkomst, dit keer te houden in Leiden. Niet, zoals eerder gemeld, op 6 december; die datum was lang geleden geprikt zodat we van een bezoek van Frank Harrell aan Nederland zouden kunnen profiteren. Helaas bleek het onmogelijk om een en ander te organiseren, gegeven een aantal praktische randvoorwaarden. Daarom zijn we uitgeweken naar 16 november. Het goede nieuws is dat we nu niet één maar twee keynote speakers kunnen verwelkomen: Jos Twisk (VUMC, Amsterdam), die een overzicht zal geven van de analyse van recurrent data, en Richard Gill (Universiteit Leiden), die het geval Lucia de B. nog eens langs de statistische meetlat legt. De andere sprekers zijn Caspar Looman (Erasmus MC, Rotterdam), Frans Oort (Universiteit van Amsterdam), Hein Putter (Leiden UMC) en Luc Bonneux (NIDI, Den Haag). Op de VOC website staan meer details. Ik hoop velen dan te zien!

Ron Wehrens (Voorzitter)

Abstracts for the VOC Fall Meeting

Jos Twisk (Vrije Universiteit in Amsterdam): The analysis of recurrent event data: An overview

The purpose of this presentation is to give an overview of different easily applicable statistical techniques to analyse recurrent event data. These techniques include naive techniques that are mostly used in epidemiological studies and longitudinal techniques such as Cox regression for recurrent events, generalised estimating equations (GEE), and random coefficient analysis. The different techniques are illustrated with a dataset from a randomised controlled trial regarding the treatment of lateral epicondylitis. It is striking to see that the different statistical techniques lead to different results and different conclusions regarding the effectiveness of the different intervention strategies. It is concluded that if one is interested in a particular short term or long term result, simple naive techniques are appropriate. However, if the development of a particular outcome is of interest, statistical techniques that consider the recurrent events and additionally correct for the dependency of the observations are necessary.

***Jos Twisk** studied human movement science at the Vrije Universiteit in Amsterdam, and after his graduation in 1990, he started to work at the same faculty, where he joined the research team of the Amsterdam Growth and Health Study. In 1995, he finished his PhD-thesis, which was related to this longitudinal study. In the same year, he moved with the AGHLS from the faculty of human movement science to the EMGO-Institute. After his PhD, he supervised several projects within the AGHSL and participated as a teacher and coordinator in several postdoctoral courses given at the EMGO-Institute. In this period, he specialised himself in the methodological field of longitudinal data analysis and multilevel analysis and wrote two textbooks about it (both published by Cambridge University Press). In 2000, he moved to the department of clinical epidemiology and biostatistics of the VU university medical centre. In 2005, he became head of the department of Methodology and applied biostatistics at the Institute of Health Sciences from the Vrije Universiteit in Amsterdam. He is also head of the expertise centre of applied longitudinal data analysis, which is an interfaculty centre of the Institute of Health Sciences and the Medical Centre of the Vrije Universiteit in Amsterdam. His main activities are statistical and methodological consultancies (both in the clinic and at the university), and teaching.*

Caspar Looman & Wilma Nusselder (Department of Public Health, Erasmus MC Rotterdam): Decomposition techniques for Health Expectancy: Linking causes for disability (diseases) to disability prevalence

Health expectancy is an extension of the concept of life expectancy. Hereby it is possible to compare health of populations not only by mortality, but also by differences in health of the living, for instance, levels of disability. To further investigate these differences it is profitable if bad health (disability) can be linked to causes, i.e., different diseases. For simple life expectancy (based on mortality) causes of death can be used, but for HLE surveys have to be used, which are often only cross-sectional.

After a short introduction to the concepts of life tables and (healthy) life expectancy we will present a method to quantify the attribution of diseases to the prevalence of disability. A regression model with additive hazards performs the trick, but because of the necessary interactions with age we propose a parsimonious model where impacts of different diseases and age patterns are combined in a factorial way.

***Caspar Looman** completed a study for rural engineer at Wageningen University, but changed already during his study to data analysis and statistics and was further educated by Cajo ter Braak. In 1985 he entered the*

Department of Public Health of the Erasmus University Rotterdam as a consulting statistician and still holds the same position.

Frans Oort (University of Amsterdam): Response shift and measurement bias

In research on health-related quality-of-life (HRQL; i.e., self-perceived physical, mental, and social health), we have observed that there are studies in which severe patients report better HRQL than healthy people. Moreover, we have observed that patients who are objectively deteriorating, actually report improving HRQL. In HRQL research such unexpected results are often attributed to *response shifts*, caused by changes in patients' frames of reference.

In an attempt to clarify what response shift is and what it is not, it has been formally defined as a violation of measurement invariance. Measurement invariance is defined as $f(X|T) = g(X|T,V)$, where f and g are distribution functions, X are test scores (e.g., scales of a self-report questionnaire), T are the attributes that we want to measure (e.g., quality of life), and V are possible violators of measurement invariance (e.g., other attributes than T). Variables X, T, and V may be nominal, ordinal, interval or ratio, they may be latent or manifest, and interrelationships may be linear or non-linear. Measurement bias is defined as a violation of measurement invariance. With longitudinal data, time of measurement occasion can be taken as V, and response shift can be defined as a special case of measurement bias.

The measurement bias definition of response shift implies that it can be detected by testing measurement invariance across measurement occasions. The choice of a particular test depends on the nature, measurement levels, and interrelationships of variables X, T, and V. For example, if the X variables are observed and the T variables are latent, then structural equation modelling can be used to detect response shifts. Many more methods are described for research on measurement bias. However, most of this research is geared to item bias (or differential item functioning), with item responses for X and group membership for V.

A possible drawback of the response shift definition is its association with a measurement perspective that may seem shallow to people who want to consider response shift as an explanation for unexpected effects on self-reported health that are much larger than those that can be explained by measurement bias. In addition, advocates of the so-called formative measurement model may object to the reflective measurement model that is used in response shift detection.

So: What do we win by considering response shift as a special case of measurement bias? And what do we loose?

***Frans Oort** studied Psychology at the University of Amsterdam, graduated in 1989, and got his Ph.D. in 1996. He is especially interested in non-standard applications of structural equation modelling (SEM). SEM includes the latent variable modelling of mean and covariance structures. His thesis was about the application of SEM to item response theory and test construction. At Leiden University he studied the application of SEM to three-mode data, such as multi-trait multi-method data, and multivariate longitudinal data. In 1999, he returned to the University of Amsterdam, to work as a statistician at the department of Medical Psychology of the Academic Medical Centre. In 2005 he was appointed as associate professor of Methods and Statistics at the Department of Education of the University of Amsterdam. Current interests include the integration of SEM with multi-level models, generalised linear models, exploratory factor models, and item response models.*

The focus of present research is "unbiased measurement" of psychological attributes in psychological and educational research (e.g., "quality of life"). Many problems in psychometrics, such as item bias, test bias, response shift, culture bias, gender bias, response styles and tendencies, social desirability, etc., can be described as violations of "measurement invariance". This enables a single general approach to these various problems, using SEM to test measurement invariance hypotheses.

Hein Putter (Leiden University Medical Center): Joint analysis of multiple longitudinal outcomes: Application of a latent class model

We address the problem of joint analysis of two series of longitudinal measurements. The typical way of approaching this problem is as a joint mixed effects model for the two outcomes. Apart from the large number of parameters needed to specify such a model, perhaps the biggest drawback of this approach is the difficulty in interpreting the results of the model, particularly when the main interest is in the relation between the two longitudinal outcomes. Here we propose an alternative approach to this problem. We apply a latent class model to the longitudinal data of the first outcome. We then use the posterior class membership probabilities that follow from this latent class model and multiple imputation to study the relation between the latent classes and the other outcome(s). We apply the method to data from 195 consecutive lung cancer patients in two out-patient clinics of lung diseases in The Hague. At four pre-defined time points, a validated semi-structured interview measuring the level of denial (the DCI) as well as a number of validated questionnaires measuring emotional and physical functioning and QOL were assessed. The aim was to study the relation between denial on the one hand and socio-demographic and illness-related characteristics and a large number of emotional and physical functioning scales on the other hand. Our approach clearly revealed an interesting phenomenon: whereas no difference between classes could be detected for objective measures of health, patients in classes representing higher levels of denial consistently scored significantly higher in subjective measures of health.

***Hein Putter** obtained his PhD in mathematical statistics under the supervision of Willem van Zwet. At present he is working at the department of Medical Statistics and Bioinformatics at the Leiden University Medical Center. His research interests include statistical genetics and survival analysis.*

Luc Bonneux (NIDI): From bias to politics

The bias of a bowl in (English) bowling is the asymmetry which permits the bowl to roll a curved path. In epidemiology, bias is any systematic error which leads the results away from the true values. Statistics may disentangle chance and random noise from true signal, it is powerless against bias, because the error is systematic.

The list of possible biases is very long. They fall apart in three families: bias by selection, information and confounding.

Selection bias means that the study sample is not representative for the study population. For example, are non-smoking partners of smokers representative for the population of non-smokers? The results documenting the fatal consequences of passive smoking critically rely on these populations. If they are not representative, the results are untrue.

Information bias means that the information collected in the study sample is different of the one that would be collected in the study population. For example, in cancer screening trials, the information available of screened and control populations is very different. If that leads to distortion in the registration of the cause of death, cause of death registers are systematically biased. The results of cancer screening trials critically rely on code of death registrations as their main outcome.

Confounding bias means that the relation between study exposure and outcome is systematically distorted by some other exposure, causally related to the study exposure and to the outcome, and accounting for the observed relation. For example, people drinking more alcohol show higher lung cancer rates. But drinkers more often tend to smoke. If alcohol drinkers are separated by smoking status, within smoking strata lung cancer is associated with smoking, but not with drinking.

However, in confounding, your strata may miss information on selected subjects. Epidemiological studies studying the long term health effects of particulate matter compare inhabitants of “dirty” cities with inhabitants of “clean” cities. They meticulously adjust for socio-economic variables. But are socio-economic variables in dirty cities equal to those in healthy cities? And are we measuring the effect of particulate matter, or of specific pollution cocktails in dirty towns?

Biases are the best friends of epidemiologists, as they permit life long debate. They can not be decided by more statistical analysis, or more studies, or more power. If non-smokers exposed to tobacco smoke are not representative of the population non-smokers, if systematic information bias in screening studies leads to differential miscoding, if people living in rusty cities are not comparable to people living in clean cities, multiplying the same study designs only succeed in multiplying the same systematic errors and propagating the wrong results.

The major epidemiological defence has always been randomisation, blinding and distrusting small signals in observational studies. However, political and increasingly financial interests of research organisations and academies has led to scientism, small signals called “scientific results” and used as fascist hammers bludgeoning the political opposition. The cited examples are chosen to highlight the contradictions between scientific fishing in muddy waters and transparent policy making by informed debate. Tobacco smoke is filthy, air pollution is undesirable and the benefits of cancer screening are always small. Whatever study design or analysis.

***Luc Bonneux** studied tropical medicine in Antwerp and worked in a basic health unit in Zaire, now Congo. From 1986 he worked as physician and researcher in the department of microbiology of the institute of tropical medicine in Antwerp and in 1988 he became MSc in epidemiology in the London School of Hygiene and Tropical medicine. From 1989 he worked at the department of public health of Erasmus MC in Rotterdam on the thesis "Degenerative disease in an aging population: Models and conjectures" together with Jan Barendregt. After a period of free-lance epidemiologic projects he works now at the Netherlands Interdisciplinary Demographic Institute.*

Richard Gill (Leiden University): Lies, damned lies and legal truths: statistics and data-analysis in the courtroom

In the legal proceedings against the nurse Lucia de B. it seems that every deadly sin in the statistical book was committed. How does one classify nurses as serial killers or not? How do you order them in degrees of murderousness? Is roster-data (time of medical events, time of shifts on duty) useful for this? Does it make any sense at all to use statistics in court, or would it be better to ban its use entirely?

***Richard Gill** studied Statistics at the University of Cambridge. He obtained his Ph.D. in mathematics and natural sciences at the Free University in Amsterdam in 1979. In 2006 he was appointed as full professor of Mathematical Statistics at the University of Leiden. At present, he is a member of the biomathematics Leiden* *node of the NDNS+ cluster (the PLUS for probability and statistics), and advisor/project-coordinator in mathematical statistics at EURANDOM (Eindhoven). He is proud and honoured [and not a little humbled] to be president of the Dutch society for Statistics and Operations Research, VVS-OR. Current interests include quantum statistics, statistics in molecular biology and genetics, causality, graphical models, statistics health and tobacco, statistics and law, statistical and computational learning, missing data, and censoring.*

Book reviews

Modern Multidimensional Scaling: Theory and Applications. Second edition. I. Borg and P.J.F. Groenen. New York: Springer, 2005. (Homepage: http://people.few.eur.nl/groenen/mmds)

Multidimensional Scaling (MDS) is a scaling technique or a family of scaling procedures for analyzing interactions between a set of objects. The relationships between objects are usually summarized by pairwise proximities (similarities or dissimilarities), and MDS aims at finding multidimensional scores for the objects that minimize some form of stress or misfit. A MDS analysis results in a visual expression of the interactions between the objects. The geometric representation or spatial map may reveal relationships between the objects that may be difficult to uncover from the observed proximities.

The book *Modern Multidimensional Scaling* is about MDS. The present book is the second edition; the first edition was released in 1997. Reviews of the 1997 book were presented in, for example,

- Kiers, H. A. L. (1999), *Psychometrika*, *64*, 95-97.
- O'Connell, A. A. (1999), *Journal of the American Statistical Association*, *94*, 338-339.
- Sireci, S. G. (2003), *Journal of Educational Measurement*, *40*, 277-280

The 1997 edition was already considered by some the authoritative reference for MDS at the present time. The book provided a rigorous treatment of the complex, perhaps even obscure, topic MDS. It is a comprehensive text of the MDS theory and may serve as reference text for the experienced applied data-analyst. Furthermore, the book was considered to be a thoughtful combination of theory and practice. The book gives an overview of the different kinds of MDS techniques, but also pays much attention to the practical use of the scaling procedures. The book may therefore be used at different levels. It also serves as an introduction to the fundamentals for readers unfamiliar with MDS. The 1997 edition could already be used as a course text, although it provided no exercises for students.

The 2005 edition of *Modern Multidimensional Scaling* contains 24 chapters, 614 pages, and 176 figures, which is 2 chapters, about 144 pages, and 60 figures more compared to the first edition. Similar to the first edition, the 2005 edition is divided into five parts: I) Fundamentals of MDS, II) MDS Problems and Solving MDS Problems, III) Unfolding, IV) Geometry as a Substantive Model, and V) MDS and Related Methods. An overview of the contents of the chapters and a guide on ``how to read this book" are presented in the preface. The latter feature is especially useful for readers unfamiliar with MDS who may experience the organization of the book a bit cumbersome. The book contains two new chapters, chapter 15 on avoiding trivial solutions in unfolding, and chapter 23 on models for asymmetric data. The remainder of extra pages was used for updates, extensions and some additional sections in existing chapters. The foremost improvement with respect to the 1997 edition is probably the addition of exercises to each chapter. The exercises make the book easier to use as a course text on MDS.

Several major difficulties with unfolding models, degenerate solutions and indeterminacies when allowing for transformations of the data, are discussed in chapter 14. When the book was released in 1997, how to avoid trivial solutions in unfolding was still work in progress. Recent proposals and new developments for avoiding unwanted solutions are discussed in chapter 15 (18 pages), which is the first new chapter of the 2005 edition. Chapter 23 (25 pages), the other new chapter, is devoted to models for visualizing asymmetry and skewed-symmetry data. The chapter is an expansion of a section on asymmetric models into a full chapter. A distinction is made between models that decompose the asymmetric data into a symmetric part and a skew-symmetric part, models that analyze only the skew-symmetry part, and models that fit the asymmetric data directly. In contrast to other chapters in Part IV, chapter 23 is limited to the analysis of two-way asymmetric data.

What should not be ignored is the homepage of the book *Modern Multidimensional Scaling* at

http://people.few.eur.nl/groenen/mmds

Here, one finds all data used for the applications and some of the exercises in the book. The website also provides information on how to obtain the software packages discussed in Appendix A of the book. There are no errata on the website (yet).

Part I (Fundamentals of MDS; chapters 1-6) is an introduction to MDS. This part of the book may be used as an introductory course in MDS. The four purposes and history of MDS are discussed in chapter 1. Using a ruler and compasses, it is demonstrated in the second chapter how a MDS representation may be constructed by hand. This chapter also explains the distinction between ratio and ordinal MDS. Loss functions and model fit (stress) are dealt with in chapter 3, and chapter 4 describes three applications of MDS. Chapter 5 is devoted to the use of MDS in the context of facet theory, whereas chapter 6 is devoted to the important topic of how proximities may be obtained.

Part II (MDS Models and Solving MDS Problems; chapters 7-13) contains the mathematical details of MDS. Chapter 7 covers the matrix algebra for a MDS analysis, which is needed for a more technical understanding of MDS. Some of the chapters later on in the book are difficult to understand without the mathematics in chapter 7. Optimization of a function is explained in the beginning of chapter 8. After this introduction in optimization the majorization algorithm, which is perhaps characteristic for modern MDS scaling procedures, is discussed. Chapter 9 describes theory and applications of several metric and nonmetric models. Chapter 10 on confirmatory MDS is used to show how prior knowledge may be incorporated in a MDS analysis. Fit measures and their relations are the topic of chapter 11, and classical scaling, the first practical available method for MDS, is discussed in chapter 12. Chapter 13 is an important section on the difficulties that may arise when applying MDS techniques. The chapter covers degenerate solutions, how these may be avoided, artificial solutions and the problem of local minima.

Part III (Unfolding; chapters 14-16) concerns the analysis of proximities between the elements of two sets. Chapter 14 provides a general introduction to unfolding models. A majorizing algorithm for unfolding and the difficulties that arise with earlier unfolding models are also discussed in this chapter. Recent proposals for avoiding unwanted solutions are discussed in chapter 15. Chapter 16 covers a variety of unfolding models.

Part IV (MDS Geometry as a Substantive Model; chapters 17-19) is used to take a closer look at the geometry of MDS. In chapters 17 and 18 MDS is used as a technique for modeling psychological phenomena. The Minkowski distance family is used as a model for psychological distance in chapter 17. The interpretation of scalar products as similarity judgments is investigated in chapter 18. Chapter 19 is a theoretical chapter on whether or not the data can be mapped into a Euclidean space, with or without some form of transformation.

Part V (MDS and Related Methods; chapters 20-23) contains a variety of topics. This part is not limited to analysis of just two-way data, but three-way Procrustean models (chapter 21) and three-way MDS models (chapter 22) are also considered. Multiple MDS solutions may be compared using the Procrustean models discussed in chapters 20 and 21. In chapter 22 it is shown how multiple data sets may be analyzed simultaneously and how the differences among the sets may be explained. Chapter 23 is the second new chapter. It is an expansion of a section on asymmetric models into a full chapter. Finally, chapter 24 contains other geometric approaches to data analysis, such as principal component analysis and correspondence analysis.

Matthijs J. Warrens

Exploring Multivariate Data with the Forward Search. A.C., Atkinson, M., Riani, and A. Cerioli. Springer, 2004.

This book has all ingredients a reader would like to find in a book: sound statistical theory, combined with applications to real data sets, exercises after each chapter, detailed solutions to the exercises (!), and a website with software (in *Splus*) and datasets to download. The software is also available in *R* and *Stata*.

The content of this book is rather technical and the statistical language is of a high level. The authors themselves state: "The book is written to be of use and interest both to professional statisticians and other scientists concerned with data analysis as well as to postgraduate students". They suggest that the book can serve as a textbook for a postgraduate course on "Modern multivariate statistics". The term "modern" in this course needs further explanation, in my view. The book is not modern in the sense that it describes new fancy statistical techniques. On the contrary, most techniques that are described are classical statistical methods, like principal component analysis, discriminant analysis, and cluster analysis. The most 'modern' techniques described in the book are spatial linear models. However, the approach to the use of these techniques is modern. The approach is called "the forward search". It can be seen as a new strategy for robust model fitting. Its purpose is "to identify observations which are different from the majority of the data and to determine the effect of these observations on inferences made about the correct model" (p. 55). Because the book focuses on multivariate normal distributions, Mahalanobis distances play an important role in the forward search. The search starts with fitting a model to a subset m of the data consisting of observations that are not outlying in any univariate or bivariate boxplot. The forward search moves to a subset of size $m + 1$ by selecting the $m+1$ observations with the smallest squared Mahalanobis distances. Then, the model is fitted again to this subset, and the Mahalanobis distances are recomputed. And so on, and so on.

The book is very interesting to those of you who work with the above mentioned classical statistical methods and who want to get more insight in the influence of departures in real data from multivariate normality. Also, a very interesting chapter is the chapter dealing with multivariate transformations to normality. These transformations can improve the stability of model parameter estimates.

The only disadvantage I could detect in the book is the technical description of the results from the forward search applied to the real data sets. I missed at some points some conclusions at a more interpretative level. But probably this tells you more about me, than about the book.

Elise Dusseldorp

Personalia

Eva Ceulemans has been appointed Assistant Professor at the Centre for Methodology of Educational Research, Katholieke Universiteit Leuven.

On September the 1st, **Elise Dusseldorp** started working as a statistician at TNO, Quality of Life.

Agenda

December 17 - 21. **Rouen, France.** Nonconvex Programming: Local and Global Approaches. http://ncp07.insa-rouen.fr/

January 11 - 12. **Gainesville, Florida, USA.** University of Florida Tenth Annual Winter Workshop: Bayesian Model Selection & Objective Methods. www.stat.ufl.edu/symposium/2008/index.html

January 17 - 18. **Philadelphia, Pennsylvania, USA.** The 7th International Conference on Health Policy Statistics: Striving for Consensus on Methods. http://www.amstat.org/meetings/ichps/2008/index.cfm

January 30 - February 01. **Houston, Texas, USA.** 2008 Bayesian Biostatistics. www.mdanderson.org/departments/biostats/

March 04 - 07. **Aachen, Germany.** 8th German Open Conference on Probability and Statistics (GOCPS 2008). http://gocps2008.rwth-aachen.de.

April 24 - 26. **Atlanta, Georgia, USA.** SIAM International Conference on Data Mining (SDM'08). www.siam.org/meetings/sdm08

May 16 - 18. **Memphis, Tennessee, USA.** International Conference on Interdisciplinary Mathematical and Statistical Techniques, IMST 2008 / FIM XVI. www.msci.memphis.edu/IMST2008-FIMXVI/

May 25 - 29. **Ottawa, Ontario, Canada.** Joint meeting of the Statistical Society of Canada (SSC) and the Société française de statistique (SFdS). www.ssc.ca/

June 05 - 07. **St. Louis, USA.** The 2008 CSNA meeting.

June 08 - 11. **Protaras, Cyprus.** International Workshop on Recent Advances in Time Series Analysis. www.ucy.ac.cy/~rats2008/

June 26 - 28. **Coimbra, Portugal.** Workshop on Nonparametric Inference – WNI2008. http://www.mat.uc.pt/~wni2008.

June 29 - July 2. **Durham, Newhampshire, USA.** International meeting of the psychometric society. http://www.psychometrika.org/meeting/2008/IMPS2008_trifold.pdf

July 08 - 12. **Oviedo, Spain.** III European Congress of Methodology, held in conjunction with the Society for Multivariate Analysis in the Behavioural and Social Sciences (SMABS). www.methodology.cop.es/

July 13 - 18. **Dublin, Ireland.** XXIVth International Biometric Conference. http://www.cpregistrations.com/ibc/2008/

July 14 - 16. **Liverpool, UK.** 6th Conference of the International Test Commission: The Impact of Testing on People and Society: Enhancing the Value of Test Use. http://www.itc2008.com/

July 16 - 18. **Hamburg, Germany.** GfKl 2008: 32nd Annual Conference of the German Classification Society. Joint Conference with the British Classification Society (BCS) and the Dutch/Flemish Classification Society (VOC).
http://gfkl2008.hsu-hh.de/

July 21 - 25. **Hamilton Island, Australia.** 9th World Meeting of the International Society for Bayesian Analysis. www.maths.qut.edu.au/asba/docs/isba08/

July 24 - 26. **Vienna, Austria.** Current Trends and Challenges in Model Selection and Related Areas. www.univie.ac.at/workshop_modelselection/

August 03 - 07. **Denver, Colaroda, USA.** 2008 Joint Statistical Meetings. http://www.amstat.org/meetings/jsm/2008/index.cfm

August 07 - 09. **Orlando, Florida, USA.** 14th ISSAT International Conference on Reliability and Quality in Design. www.issatconferences.org

August 12 - 14. **Dortmund, Germany.** useR! 2008 - the R User Conference. www.R-Project.org/useR-2008

August 17 - 21. **Copenhagen, Denmark.** 29th Annual Conference of ISCB - the International Society for Clinical Biostatistics. www.iscb2008.info

August 24 - 29. **Porto, Portugal.** COMPSTAT 2008: International Conference on Computational Statistics. http://www.fep.up.pt/compstat08/

Publications

Blom, M.B.J., Jonker, J., Dusseldorp, E., Spinhoven, P., Hoencamp, E., & Haffmans, P.M.J. (2007). Combination treatment for acute depression is superior only when psychotherapy is added to medication. *Psychotherapy and Psychosomatics, 76,* 289-297.

Commandeur, J.J.F., & Koopman, S.J. (2007). *An introduction to state space time series analysis.* Oxford: Oxford University Press.

De Jong, K., Nugter, M.A., Polak, M.G., Wagenborg, J.E.A., Spinhoven, P., & Heiser, W.J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology & Psychotherapy, 14,* 288-301.

Delwarde, A., Denuit, M., & Eilers, P. (2007). Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: a penalized log-likelihood approach. *Statistical Modelling, 7,* 29-48.

Eilers, P.H.C. (2007). Data exploration in meta-analysis with smooth latent distributions. *Statistics in Medicine, 26,* 3358-3368.

Eilers, P.H.C., & Borgdorff, M.W. (2007). Non-parametric log-concave mixtures. *Computational Statistics & Data Analysis, 51,* 5444-5451.

Elzinga, C.H., & Liefbroer, A.C. (2007). De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. *European Journal of Population*, *23,* 225-250.

Elzinga, C.H., Hoogendoorn, A.W., & Dijkstra, W. (2007). Linked markov sources: Modeling outcome-dependent social processes. *Sociological Methods & Research*, *36*, 26-47.

Essink-Bot, M.L., Stuifbergen, M.C., Meerding, W.J., Looman, C.W.N., & Bonsel, G.J. (2007). Individual differences in the use of the response scale determine valuations of hypothetical health states: an empirical study. *BMC Health Services Research, 7,* 62.

Frank, L.E. & Heiser, W.J. (2007). Estimating standard errors in Feature Network Models. *British Journal of Mathematical and Statistical Psychology, 60,* 1-28.

Groenen, P.J.F., Kaymak, U., & Van Rosmalen, J. (2007). Fuzzy clustering with Minkowski distance functions. In: J. Valente de Oliveira, & W. Pedrycz (Eds.), *Advances in fuzzy clustering and its applications* (pp. 53-68). Chichester: Wiley. (Appeared also as Econometric Institute Report EI 2006-24).

Groenen, P.J.F., Nalbantov, G., & Bioch, J.C. (2007). Nonlinear support vector machines through iterative majorization and I-splines. In: R. Decker, & H-.J. Lenz (Eds.), *Advances in data analysis* (pp. 149-162). Berlin: Springer. (Appeared also as Econometric Institute Report EI 2006-25).

Haans, J.J., Paridaans, N.P., Eilers, P.H., Doornbos, J., De Roos, A., & Masclee, A.A. (2007). Effect of sildenafil on gastric function assessed with magnetic resonance imaging. *Gastroenterology, 132,* A375-A375.

Heij, C., Groenen, P.J.F., & van Dijk, D.J. (2007). Forecast comparison of principal component regression and principal covariate regression. *Computational Statistics and Data Analysis*, *51*, 3612-3625. (Appeared also as Econometric Institute Report EI 2005-28).

Heiser, W.J., & Meulman, J.J. (2007). Noise or signal? The dilemma of individual differences. In: A. in 't Groen, H.J. de Jonge, E. Klasen, H. Papma, & P. Van Sloten (Eds.), *Knowledge in Ferment: Dilemmas in Science, Scholarship and Society* (pp. 71-83). Leiden: Leiden University Press.

Hidegkuti, I., & De Boeck, P. (2007). Between-group differences and taxometrics. *Psychological Reports*, *100*, 211-230.

Linting, M., Meulman, J.J., Groenen, P.J.F., & Van der Kooij, J.J. (2007). Nonlinear principal components analysis: Introduction and application. *Psychological Methods, 12,* 336-358.

Linting, M., Groenen, P.J.F., & Meulman, J.J. (2005). Stability of nonlinear principal components analysis: An empirical study. *Psychological Methods, 12,* 359-379.

Lips, E.H., de Graaf, E.J., Tollenaar, R.A.E.M., van Eijk, R., Oosting, J., Szuhai, K., Karsten, T., Nanya, Y., Ogawa, S., de Velde, C.J.V., Eilers, P.H.C., Van Wezel, T., & Morreau, H. (2007). Single nucleotide polymorphism array analysis of chromosomal instability patterns discriminates rectal adenomas from carcinomas. *Journal of Pathology, 212,* 269-277.

Nijsten, T., Looman, C.W.N., & Stern, R.S. (2007). A twenty year study of the clinical severity of psoriasis. *Archives of Dermatology, 143,* 1113-1121.

Slingerland, A.S., van Lenthe, F.J., Jukema, J.W., Kamphuis, C.B.M., Looman, C., Giskes, K., Huisman, M., Venkat Narayan, K.M., Mackenbach, J.P., & Brug, J. (2007). Ageing, retirement and changes in physical activity: prospective cohort findings from the GLOBE-study. *American Journal of Epidemiology, 165,* 1356-1363.

Ten Berge, J.M.F., & Sočan, G. (2007). The set of feasible solutions for reliability and factor analysis. In: S.Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 303-320). Amsterdam: Elsevier.

Torres, A., & van de Velden, M. (2007). Perceptual mapping of multiple variable batteries by plotting supplementary variables in correspondence analysis of rating data. *Food Quality and Preference, 18,* 121-129**.**

Van Blankenstein, M., Looman, C.W.N., Kruijshaar, M.E., Siersema, P.D., Kuipers, E.J., & Bytzer, P. (2007). Modelling a population with Barrett's oesophagus from oesophageal adenocarcinoma incidence data. *Scandinavian Journal of Gastroenterology, 42,* 308-317.

Van Deun, K., Heiser, W.J., & Delbeke, L. (2007). Multidimensional unfolding by nonmetric multidimensional scaling of Spearman distances in the extended permutation polytope. *Multivariate Behavioral Research*, *42,* 103-132.

Van Deun, K., Marchal, K., Heiser, W., Engelen, K., & Van Mechelen, I. (2007). Joint mapping of genes and conditions via multidimensional unfolding analysis. *BMC Bioinformatics*, *8*, 181.

Van Herk, H., & van de Velden, M. (2007). Insight into the relative merits of rating and ranking in a cross-national context using three-way correspondence analysis. *Food Quality and Preference*, *18,* 1096-1105.

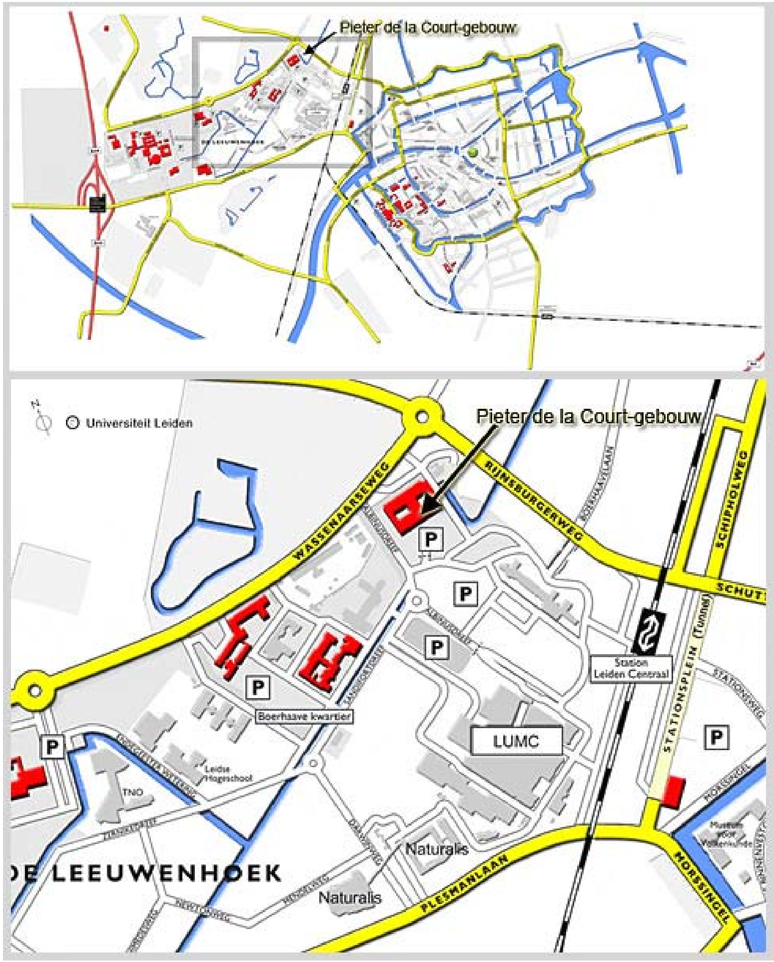
Van Mechelen, I., & Schepers, J. (2007). A unifying model involving a categorical and/or dimensional reduction for multimode data. *Computational Statistics & Data Analysis*, *52*, 537-549.

Warrens, M.J., De Gruijter, D.N.M., & Heiser, W.J. (2007). A systematic comparison between classical optimal scaling and the two-parameter IRT model. *Applied Psychological Measurement*, *31,* 106-120.

Warrens, M.J., & Heiser, W.J. (2007). Robinson Cubes. In: P. Brito, G. Cucumel, P. Bertrand, & F. de Carvalho (Eds.). *Selected contributions in data analysis and classification* (pp. 515-523). Heidelberg: Springer Verlag.

Route description

The VOC Fall Meeting takes place in Room CH10 of the temporary building that is called "het chalet". This building is reached from the Pieter de la Court building by crossing the carpark.



By car:

1. Coming from the A44:
   You leave the A44 at exit 8 (this exit is called Leiden-Valkenburg-Katwijk-Noordwijk from the direction The Hague, and from the direction Amsterdam it is called Leiden-Utrecht). You drive to Leiden Centrum/Naturalis following the Plesmanlaan.
2. Coming from the A4:
   You leave the A4 at exit 7 (Zoeterwoude dorp) and you follow the N206 in the direction Katwijk. At the Plesmanlaan you turn right to Leiden Centrum/Naturalis.

At the Plesmanlaan you take the first option to turn left. You are now following the Einsteinweg. You need to follow the Einsteinweg till the roundabout, where you turn right following the Max Planckweg. You need to follow the Max Planckweg till the Wassenaarseweg, where you turn right. Follow this road till the Pieter de la Court building of the Faculty of Social and Behavioural Sciences (big, yellow, square building along the Wassenaarseweg (zipcode 2333 AK), right before De Posthof restaurant and the roundabout at the Rijnsburgerweg). The entrance to the carpark is on the right-hand side of the building.

By train:

When you arrive by train (Leiden central station), you need to take the 'hospital exit' (not the centre) at the railway station. You now arrive at the premises of the academic hospital. Further on you can already see the yellow Pieter de la Court Building.