



Vereniging voor Ordinaties en Classificatie

Nieuwsbrief no. 31
oktober 2003

Voorzitter: Patrick Groenen, Econometrisch Instituut, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR Rotterdam.
(groenen@few.eur.nl)

Secretaris: Marieke Timmerman, RU Groningen, Heymans Instituut (DPMG), Grote Kruisstraat 2/1, 9712 TS Groningen
(m.e.timmerman@ppsw.rug.nl)

Penningmeester: Paul Arents, Quest International BV, Sensory Science & Consumer Acceptance, Huizerstraatweg 28, 1411 Naarden
(paul.arents@questintl.com)
Postgirorekening 161723 t.n.v. Vereniging voor Ordinaties en Classificatie, Naarden
Bankrekening nummer 777-5952385-56 Bacob Bank t.n.v. VOC, Naarden

Redactie: Mark de Rooij, Faculteit Sociale Wetenschappen, Universiteit Leiden, Postbus 9555, 2300 RB Leiden.
rooijm@fsw.leidenuniv.nl

VOC-home page: <http://www.voc.ac>

VOC Fall Meeting November 21st 2003,

Item Response Theorie

Leiden University Medical Centre

Lecture Room, Physiology Building

10.30-11.00	Welcome
11.00-11.35	Stef van Buuren (TNO-PG) <i>Verbeteren van vergelijkbaarheid met Response Conversie</i>
11.35-12.10	Frank Rijmen (Leuven University) <i>Mixed models in Item Response Theory</i>
12.10-12.45	Rebecca Holman (Academic Medical Centre, Amsterdam) <i>The AMC Linear Disability Score project: using item response theory to construct and calibrate an item bank to measure the ability to perform activities of daily life</i>
12.45-14.15	Lunch
14.15-14.50	Martijn Berger (Maastricht University) <i>Optimal design in educational testing: a review</i>
14.50-15.25	Andries van der Ark (Tilburg University) <i>Nonparametric item response theory</i>
15.25-16.00	Tea
16.00-17.00	Norman Verhelst (Cito). <i>Item Response Theory and Multiple Choice Questions</i>
17.00-	Drinks

In dit nummer:

Programma VOC Fall Meeting 21 november 2003	1
Abstracts	2
Agenda	4
Boekbespreking	5
Publicaties en rapporten	7
Routebeschrijving + Aanmelden najaarsbijeenkomst	8

Abstracts of the Fall Meeting in Leiden November 21, 2003

Stef van Buuren (TNO-PG)

Verbeteren van vergelijkbaarheid met Response Conversie

De Europese Unie telt binnenkort 25 lidstaten, elk met een eigen informatievoorziening. In opdracht van de Europese Commissie heeft TNO-PG een methode ontwikkeld waarmee de vergelijkbaarheid van informatie uit verschillende lidstaten kan worden verbeterd. Basisgedachte is dat gelijksoortige informatie naar een gemeenschappelijke schaal vertaald kan worden, zodat vergelijkingen tussen landen op de gezamenlijke schaal mogelijk worden. Dit proces gaat in twee stappen. Stap 1 is de constructie van conversiesleutel, stap 2 betreft de werkelijke conversie naar de gezamenlijke schaal. Een doorslaggevend voordeel van de methode is dat ze nieuwe vergelijkende analyses mogelijk maakt op basis van bestaande data. De methode is een min of meer rechtstreekse toepassing van Item Response Theorie. De structuur van de gegevens brengt echter enkele specifieke complicaties met zich mee die zich niet, of in mindere mate, voordoen in meer klassieke toepassingen van IRT. De presentatie geeft voorbeelden van de methode, en gaat in op nog openstaande vragen.

Stef van Buuren is Hoofd Statistiek van TNO Preventie en Gezondheid. Hij houdt zich bezig met o.m. de groei van Nederlandse kinderen, multiële imputatie van incomplete data, vergelijkbaarheid, simulatiesystemen, en publiceert daarover in zowel de toegepaste als de statistische literatuur.

Frank Rijmen (Leuven University)

Mixed models in Item Response Theory

Test data are clustered: Each tested person corresponds to a cluster, and the responses of a person on the individual items of the test correspond to the observations within a cluster. Typically, the observations within a cluster tend to be more homogeneous than the observations stemming from different clusters, meaning that the observations within a cluster are not statistically independent. Mixed models are a collection of statistical tools that are well suited for analyzing clustered data. The heterogeneity between clusters can be taken into account by assuming that (some of) the parameters of the model follow a random distribution over the population of clusters. Hence, (some of) the parameters of the model are random variables, and the model is a random effects or mixed model. The random effects represent the cluster-specific effects. Mixed models and related methods were first developed in the context of analysis of variance and regression analysis, leading to the linear mixed model. Other commonly used terms are multilevel models, hierarchical models, and random coefficient

models. More recently, generalized linear and nonlinear mixed models were also developed.

Most parametric IRT models can be conceptualized of as generalized linear and nonlinear mixed models. For example, the linear logistic test model (Fischer, 1973), is a logistic regression model with a random intercept. There are four important assets of this conceptualization. First, the mixed model framework for IRT models relates IRT to a broad statistical literature. Second, applying a common framework to different IRT models can help in our understanding of their differences and commonalities. Third, standard IRT models can readily be adapted and extended, so that a researcher can build his own model, customized to a specific scientific question or data set. Finally, existing and newly formulated models can be estimated using software for generalized linear and nonlinear mixed models.

Frank Rijmen obtained his PhD in psychometrics at the K.U. Leuven in 2002. Now he is a postdoctoral researcher at the same university. His main research interest is in the field of item response theory, with special attention to the relation between parametric IRT models and generalized linear and non-linear mixed models.

Rebecca Holman (Amsterdam Medical Centre)

The AMC Linear Disability Score project: using item response theory to construct and calibrate an item bank to measure the ability to perform activities of daily life

An important aspect of quality of life is the 'disability' status of patients. This is often described in terms of their ability to carry out 'activities of daily life' and measured using questionnaires, which grade each patient on whether they are able to perform certain activities. A questionnaire, which contains many items, can cost patients, clinicians and researchers an excessive amount of time to complete. In the framework of item response theory (IRT) and in conjunction with an item bank, it is possible to use computerized adaptive testing procedures to assess functional status by presenting each individual patient with a smaller selection of items than is possible using traditional methods. This has awakened the interest of medical and epidemiological researchers in the use of IRT as a tool to analyze this type of questionnaire. The AMC Linear Disability Score project aims to develop an item bank containing items, designed to quantify the ability of chronically ill patients to perform 'activities of daily life'. The item bank currently contains approximately 200 items, each describing an activity of daily life connected with self-care, mobility and household management. This paper will describe the stages involved in constructing and calibrating this item bank and the statistical and methodological challenges encountered in the process. In addition, an insight into the applications and uses of the item bank will be given.

Rebecca Holman is employed as a statistician by the Department of Clinical Epidemiology and Biostatistics at the Academic Medical Center, Amsterdam, The Netherlands. Her work concentrates on the statistical and methodological challenges encountered when item response theory techniques are used to analyze data resulting from questionnaires to quantify health status, often in terms of the ability to perform activities of everyday life.

Martijn Berger (Maastricht University)

Optimal design in educational testing: a review

Modern testing in Education and Psychology is mostly based in Item Response Theory (IRT), and several different IRT models have been proposed to analyze the test data. As in any research situation, where data are obtained, the costs of collecting test data may be enormous. Both the required sample size for efficient parameter estimation and the costs of producing test items can be reduced by using optimal design techniques. Optimal design theory was first applied to design issues in testing about 30 years ago.

Basically there are two distinct design issues in testing. The first issue is concerned with the calibration of pre-production test items. Item calibration takes place by administering test items to a selected sample of test takers, simultaneously or sequentially. Since large samples are usually needed to obtain efficient item parameter estimators, it is worthwhile to try to select an optimal sample of test takers, which is preferably much smaller in size. The second design issue is the problem of how to pick the test items for a given test taker or a sample of test takers for efficient trait estimation. Both issues have been tackled by methods originating from optimal design theory.

In this paper a review of the optimal design methods and procedures will be presented. First, IRT models for both dichotomous and polytomous models will be briefly described. Then the optimal design problem for the two design issues (item calibration and trait estimation) will be described together with the problems arising from practical and technical constraints. Optimal design methods have been applied to both computerized adaptive testing and paper and pencil testing in classroom settings. A bibliography will be provided.

Martijn P.F. Berger has been involved in optimal design research, not only for testing with IRT models, but also for generalized linear mixed models (GLMM), with emphasis on longitudinal designs. Recent research focuses on maximin procedures, which have been applied to the nominal response model, to the GLMM models and to the robustness of design problem.

Andries van der Ark (Tilburg University)

Nonparametric item response theory

The aim of nonparametric item response theory (NIRT) is to order respondents on a latent variable (called the latent trait) using the ordinal scores on several manifest variables (called items). In this presentation, I will first explain the basic ideas of NIRT for dichotomous item scores. Second, I will present an application where NIRT is used for the construction of a psychological test measuring crying behavior. In this test, the unweighted total score is used as an ordinal estimate of the latent trait. The analyses were conducted using the software program MSP. It is shown how in subsequent analyses quantitative estimates of the latent trait may be obtained. Third, I will discuss some open problems that are encountered when NIRT is extended from dichotomous items to polytomous items.

Andries van der Ark is an assistant professor at the Department of Methodology and Statistics of the Faculty of Social and Behavioral Sciences at Tilburg University. His research interest is the modeling of test and questionnaire data. He is currently working on nonparametric item response theory and the analysis of missing data.

Norman Verhelst (Cito)

Item Response Theory and Multiple-Choice Questions

One of the main controversies between American (U.S) and European (and Australian) students of IRT is the question whether a simple model like the Rasch model is suited to model responses to multiple-choice questions. The mainstream attitude in the U.S. is to reject this model with the argument that it cannot handle correct answers that come about by guessing. It will be argued that it is fairly well possible to apply the Rasch model in the case of multiple choice items, by considering the items as a choice situation with a finite number of alternatives and by applying Luce's choice axiom to these situations and at the same time accounting for individual differences in the scale values resulting from this choice axiom. Some implications for practical applications will also be discussed.

Norman Verhelst was born in 1946. He studied Psychology at the Catholic University of Leuven (Belgium) with emphasis on mathematical psychology and psychometrics. Further career: Teaching Statistics, Methodology and Psychometrics at the universities of Leuven, Nijmegen, Utrecht and Twente. Since 1985 he worked in research and consultancy in Psychometrics at the National Institute for Educational Measurement (Cito), Arnhem, The Netherlands. Main interest: Item Response Theory as applied in National Assessment Programs and International comparative studies (PISA).

Agenda

- 08 – 12 December 2003. **Atlantic City, USA.** 59th Annual Deming Conference on Applied Statistics. <http://www.demingconference.com/>
- 18 – 20 December 2003. **Hong-Kong, China.** Bernoulli Society East Asian and Pacific Regional (EAPR) Conference 2003. <http://www.bm.ust.hk/~eapr2003>.
- 09 – 11 March 2004. **Dortmund, Germany.** Classification: the ubiquitous challenge. 28. Annual Conference of the German Classification Society (GfKl), organized jointly with the Dutch / Flemish Classification Society (VOC). <http://www.gfkl.de/gfkl2004/>
- 19 – 21 May 2004. **Gathersburg, USA.** Spring Research Conference on Statistics in Industry and Technology. <http://www-math.cudenver.-edu/SRC2004>
- 19 – 22 May 2004. **Houston USA.** The 2nd Lehmann Symposium. <http://www.stat.rice.-edu/lehmann>
- 20 – 24 May 2004. **Quebec City, Canada.** International Conference on Dependence Modelling: Statistical Theory and Applications in Finance and Insurance (DeMoSTAFI). <http://www.fsa.ulaval.ca/demostafi/>
- 23 – 27 May 2004. **Viña Del Mar, Chile.** ISBA 2004 World Meeting. <http://www.bayesian.org>
- 24 – 28 May 2004. **Cape Town, South Africa.** 9th International Meeting on Statistical Climatology. <http://www.csag.uct.ac.za/IMSC>
- 30 May – 02 June 2004. **Montreal, Canada.** 32nd Annual Meeting of the Statistical Society of Canada, Montréal. <http://www.ssc.ca>
- 06 – 09 June 2004. **La Jolla (San Diego), USA.** ICSA 2004 Applied Statistics Symposium. <http://www.icsa.org>
- 14 – 17 June 2004. **Pacific Grove (Monterey), California, USA.** Annual meeting of the psychometric society. <http://www.-psychometrika.org>.
- 23 – 25 June 2004. **Nantes, France.** International Conference on Statistics in Health Sciences. <http://www.sante.univ-nantes.fr/STAT/>
- 28 June – 03 July 2004. **Lund, Sweden.** International Association for Statistical Education (IASE) 2004. http://hobbes.lite.msu.edu/~IASE_2004_Roundtable
- 04 – 07 July 2004. **Sydney, Australia.** 24th International Symposium on Forecasting. <http://www.isf2004.org>.
- 04 – 08 July 2004. **Florence, Italy.** 19th International Workshop on Statistical Modelling. <http://www.ds.unifi.it/iwsm2004/>
- 11 – 16 July 2004. **Cairns, Australia.** XXIIInd International Biometric Conference in parallel with the Australian Statistical Conference. <http://www.ozacom.com.au/cairns2004>.
- 14-18 July 2004. **Neuchâtel, Switzerland.** Celebrating Statistics: International conference in honour of Sir David Cox on the occasion of his 80th birthday. <http://www.unine.ch-statistics/cox/Welcome.html>
- 15-18 July 2004. **Chicago, USA.** International federation of classification societies. <http://www.classification-society.org/>
- 19 – 23 July 2004. **Aveiro, Portugal.** 3rd International Symposium on Extreme Value Analysis -- Theory and Practice (EVA 2004). <http://www.mat.ua.pt/eva2004/>
- 25-28 July 2004. **Budapest, Hungary.** IEEE International Conference on Fuzzy Systems 2004. <http://www.conferences.hu/budapest-2004>
- 26 – 31 July 2004. **Barcelona, Spain.** 6th World Congress of the Bernoulli Society for Mathematical Statistics & Probability / 67th Annual Meeting of the Institute of Mathematical Statistics. <http://www.imub.ub.-es/events/wc2004/>
- 08 – 12 August, 2004. **Toronto, Canada.** JSM. Statistics as a Unified Discipline. <http://www.amstat.org/meetings/jsm/2004-/index.cfm>
- 19 – 21 August 2004. **Poznan, Poland.** The 13th International Workshop on Matrices and Statistics. <http://matrix04.amu.edu.pl>
- 23 – 27 August 2004. **Prague.** COMPSTAT 2004, 16th Symposium of IASC. <http://compstat2004.cuni.cz/>
- 02 – 04 september 2004. **Gent, Belgium.** 35th European Mathematical Psychology Group Meeting <http://allserv.ugent.be/~yrosseel/empg>

Boekbespreking

Een VOC boekbespreking betekent een boek cadeau!

Beste VOC leden,

Voor de volgende VOC nieuwsbrieven (Februari 2004 en Oktober 2004) zoeken wij op korte termijn naar leden die een boek willen bespreken. Hieronder vindt u enkele suggesties van recente boeken, waarover wij graag een boekbespreking in de VOC nieuwsbrief willen hebben. U kunt ook altijd in overleg zelf een boek aandragen. In alle gevallen geldt, dat u als dank voor de boekbespreking het boek cadeau krijgt.

- 1) Ramsay, J.O. & Silverman, B.W. (2002) Applied Functional Data Analysis: Methods and Case Studies. Springer.
- 2) Cabrera J, McDougall (2001). Statistical consulting. Springer.
- 3) Pepe, M.S. (2003). Statistical evaluation of medical tests for classification and prediction. Oxford.
- 4) Semple, C, Steel M. (2003). Phylogenetics. Oxford.
- 5) Everitt. B. (2002). Cambridge dictionary of statistics. Second Edition. Cambridge Un Press.
- 6) Little, RJA, Rubin DB (2002). Statistical analysis with missing data. Second ed. Wiley.
- 7) Salburg, S (2001). The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century. Freeman & Co.

Bent u geïnteresseerd? Schrijf dan een email naar Josephine Woltman Elpers (j.elpers@eco.rug.nl) onder vermelding van het boek dat u graag wil bespreken.

An introduction to generalized linear models.
Annette J. Dobson, Chapman & Hall/CRC, 2001,
ISBN 1584881658

Zoals de titel laat vermoeden, krijgt de lezer van dit boek een inleiding in 'generalized linear models' (GLM). Wat de titel niet meteen laat vermoeden, is dat dit boek een helder betoeg heeft, bijzonder instructief en bovenal vlot verstaanbaar is (evenwel op voorwaarde dat de lezer weet heeft van de basisconcepten uit de statistiek, de lineaire algebra en de calculus). Je wordt als lezer niet geconfronteerd met uitdrukkingen als 'it can be easily seen' of 'after some algebra', waarbij je zelf alles kunt

uitvloeien en na een uur vaak gefrustreerd moet besluiten dat je toch niet helemaal zeker bent of je het nu juist begrepen hebt. Kortom, dit boek is uiterst geschikt om de jonge promovendus, die niet in de statistiek is opgeleid maar zich wel methodologisch wil bezighouden, onder de neus te schuiven. Ik heb het lezen van dit boek alleszins heel plezierig gevonden.

Ook de onderwerpen aangesneden in de opeenvolgende hoofdstukken zijn didactisch goed gekozen. Na de obligatoire inleiding, worden in hoofdstuk 2 de kernideeën van het statistisch modelleren uitgelegd. De volgende drie hoofdstukken zijn de meest belangrijke van het boek, in hoofdstuk 3 worden de verdelingen van de exponentiële familie geïntroduceerd, in hoofdstuk 4 schattingsmethoden en in hoofdstuk 5 de principes van hypothesetoetsing. Deze drie hoofdstukken verklaarde me duidelijk waarom ik overal hoor en lees dat 'GLM een algemeen denkkader voor uiteenlopende statistische technieken is'. Die algemeenheid wordt nog eens duidelijk geïllustreerd in de daarop volgende hoofdstukken die handelen over multiple lineaire regressie, (co)variantie-analyse, logistische regressie, nominale en ordinale logistische regressie, loglineaire modellen, survival analyse, generalized estimating equations en multi-level analyse. Een indrukwekkend oeuvre, zeker als je weet dat het boek slechts 240 bladzijden dik (beter: dun) is. Dit is niet zo omdat Dobson te kort door de bocht gaat of geen toepassingen bespreekt, maar eerder omdat ze uitdiepingen en technische details achterwege laat. Dit is zeker geen verwijt, het boek is immers als inleiding bedoeld en die uitdiepingen en technische details vind je wel ergens anders, bijvoorbeeld in 'Generalized linear models' van J.A. Nelder & P. McCullagh. Kortom, ik vind dat Dobson echt geslaagd is in haar opzet en het boek maakt waar wat diens titel belooft. Wat ik (als GLM-autodidacticus) wel jammer vind, is dat de oefeningen, gegeven op het einde van elk hoofdstuk, niet besproken worden. Desondanks, vind ik het boek echt een aanrader.

Kaatje Bollaerts

The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
Trevor Hastie, Robert Tibshirani and Jerome Friedman (2001) New York: Springer.
533+xvi pagina's. ISBN 0-387-95284-5

Alleen al het hoofdonderwerp van dit boek zou de interesse van menig VOC-lid moeten wekken, omdat het overeenkomt met een van de hoofdonderwerpen van de VOC: classificatie. Gebruik makend van de terminologie van Machine Learning en Datamining noemen de auteurs het Statistical Learning, waarbij ze een onderscheid maken tussen supervised learning, classificatie voor de beschrijving en predictie van een respons variabele aan de hand van inputs/features/onafhankelijke variabelen, en unsupervised learning, classificatie/clustering voor de beschrijving van de (samenhang tussen) inputs/features

alleen. Machine Learning en Datamining zijn velden die in feite de vruchten plukken van verschillende disciplines, maar tot nog toe een sterke nadruk leggen op hun oorsprong: de computer wetenschappen. Een gevolg is dat wanneer je bijv. als psychometrist een bijeenkomst van de Machine Learning gemeenschap bijwoont (zoals BENELEARN), er zowel een grote overeenkomst is in methoden en technieken als een enorm verschil in terminologie. Daarbij is het evident dat de verschillende disciplines veel van elkaar kunnen leren.

Daarom is het prettig dat er nu een boek is verschenen dat de belangrijkste onderwerpen van het hele veld poot te beschrijven vanuit één gemeenschappelijke terminologie, maar belangrijker nog is dat het boek dit doet vanuit een overkoepelend perspectief. Wie de auteurs kent -- en menig VOC lid heeft ze al eens in Nederland kunnen begroeten, bijv. als gastspreker op een VOC bijeenkomst (J. Friedman), maar ook bij de cursus Modern Regression and Classification (T. Hastie en R. Tibshirani) -- zal het niet verbazen dat dit perspectief statistisch van aard is.

Dit is tegelijkertijd de kracht van het boek en de belangrijkste beperking: een lezer zal wellicht niet alle technieken in het boek terugvinden die hij/zij vanuit het eigen veld verwacht, omdat de auteurs bewust de nadruk leggen op technieken die ze zelf belangrijk vinden, waar ze zelf bekend mee zijn, en belangrijke bijdragen aan hebben geleverd, en omdat niet alle methoden zich goed in dit statistisch framework laten vatten. Veel belangrijker is echter dat door het consequent doorgevoerde statistisch perspectief de verschillen en overeenkomsten tussen de technieken helder worden weergegeven. Daarbij voegt het boek ook belangrijk materiaal toe, omdat sommige technieken zonder statistische rationale zijn ontwikkeld, maar wel degelijk een statistische basis te geven zijn.

Na een eerste hoofdstuk, waarin geprobeerd wordt de lezer te motiveren door de hoofdonderwerpen te definiëren aan de hand van een aantal praktische voorbeelden, wordt in het tweede hoofdstuk een overzicht van het hele veld gegeven. Door lineaire regressie als een prototype locale methode te zetten tegenover neirest neighbor als een prototype locale methode, wordt de lezer op natuurlijke wijze geïntroduceerd in begrippen als (verwachte) voorspellingsfouten, waarbij de stabiliteit en mogelijk grote systematische fout van lineaire regressie wordt gezet tegenover het probleem van locale methoden in hoge dimensionaliteit. Model selectie wordt aldus gemotiveerd aan de hand van statistische concepten als de Bias-Variance Tradeoff, en model complexiteit en het aantal effectieve parameters. Deze begrippen komen in vrijwel ieder hoofdstuk terug, evenals de gewoonte de modellen te illustreren aan de hand van een zich steeds verder ontwikkelende toepassingen op dezelfde data sets binnen een hoofdstuk. Aan het eind van ieder hoofdstuk wordt daarbij zeer kort ingegaan op de complexiteit van de beschreven methoden en hun geschiktheid voor

grootschalige problemen, waarna een aantal nuttige referenties wordt gegeven.

H. 3 en 4 geven belangrijk statistisch basis materiaal over lineaire methoden voor Regressie en Classificatie, waarna de lineaire methoden worden uitgebreid d.m.v. basis expansies van de inputs met behulp van splines en wavelets en hun regularisatie/smoothing in H. 5 (helaas geen behandeling van de vergelijkbare mogelijkheden tot expansie van de output variabelen). In H.6 wordt een uitgebreide klasse van locale methoden behandeld onder de noemer van Kernel Methods. De statistische basis wordt daarna afgerond, door uitgebreider in te gaan op Model selectie in H. 7 aan de hand van test sets en "Model Inference and Averaging" in H.8, dat Maximum Likelihood en de Bayesiaanse statistiek koppelt aan de bootstrap.

Afhankelijk van de achtergrond van de lezer, komt daarna voor velen de hoofdmoot van populaire recente methoden: Additieve modellen en classificatie/regressie bomen in H. 9; Boosting van vooral boom-methoden in H. 10; Neurale Netwerken in H. 11; Support Vector Machines en Flexibele Discriminant analyse in H. 12; en Prototype en Neirest Neighbor methoden in H. 13. Persoonlijk vond ik vooral de hoofdstukken over Boosting en Vector Support Machines erg interessant door de inzichten die worden geboden in de redenen waarom deze methoden zo krachtig zijn, maar deze hoofdstukken zullen wellicht ook het meest tot controverse leiden vanwege het sterke persoonlijke standpunt van de auteurs dat er uit spreekt (zie bijv. tabel 10.1 p. 313 met de karakteristieken van een aantal methoden voor "off the shelf" classificatie, waarbij de beargumenteerde superioriteit van boom-methoden en Boosting wordt geïllustreerd met veel groene bolletjes, waar alternatieven voorgesteld door andere auteurs veel rode bolletjes krijgen ;-).

Voor wie is het boek geschikt? De nadruk in het boek (en in de oefeningen achter ieder hoofdstuk!) ligt op de conceptuele beschrijving van de methoden, hoewel het hele boek doorspekt is met voorbeelden van interessante toepassingen, waarvan de data sets op het web beschikbaar gesteld zijn. Daarbij kan de technisch en mathematisch goed onderlegde lezer aan de hand van de algoritme beschrijvingen wel aan de slag, maar de meesten zullen toch gebruik maken van gespecialiseerde software (vaak wordt verwezen naar S-Plus). Wie probeert de voorbeelden na te rekenen zal merken dat de aanwijzingen summier zijn en soms op verassingen stuiten.

Hoewel het boek zeer goed is geschreven, is het een pittig boek; met name de statistische kennis die vooraf vereist is moet wat hoger worden ingeschat dan de auteurs zelf aangeven: een elementaire basis cursus. Daarbij dient de minder goed ingevoerde lezer ook goed nota te nemen van de al uitgebreide lijst met errata die op de website staat (en nog uitbreiding behoeft). Al met al is het boek een must voor de onderzoeker die is ingevoerd in een van de deelgebieden van supervised learning en de

achtergronden van alternatieve methoden wil leren kennen, en kan het boek ook gebruikt worden voor cursussen aan de gevorderde student of promovendus.

Een wens betreft het laatste hoofdstuk over unsupervised learning, een gebied waar de auteurs in sneltrein vaart doorheen razen. Dit motiveren ze door te stellen dat dit gebied minder uitgebreid ontwikkeld is. Dit doet mijns inziens geen recht aan de uitgebreide literatuur die ook op dit gebied voor handen is, maar is te begrijpen, omdat het op veel gebieden ontbreekt aan statistische criteria om de kwaliteit van de methoden te beoordelen -- zoals de auteurs zelf aangeven. Dat mag deels zo zijn, maar ik zou wensen dat er ook zo'n goed en omvattend boek over unsupervised learning wordt geschreven.

Bart-Jan van Os

Publicaties en rapporten

Het doel van deze rubriek is dat VOC leden elkaar op de hoogte brengen van hun recente publicaties en rapporten. Voor de volgende nieuwsbrief vraag ik u een lijstje aan mij te sturen met nieuwe openbare rapporten en publicaties. Het is voor mij het gemakkelijkst als dit per e-mail gebeurt (rooijm@fsw.leidenuniv.nl), maar het mag ook per gewone post.

Bandeira LPC, Sousa JMC & Kaymak U (2003).

Fuzzy clustering in classification using weighted features. *Lecture Notes in Computer Science*, 2715, 560-567.

Bechger TM, Maris, G Verstralen, HHFM & Béguin, AA (2003). Combining Classical Test Theory and Item Response Theory. *Applied Psychological Measurement*, 27, 319-334.

Bonneux L & Looman CWN (2003). High coronary heart disease rates among dutch women of the baby boom, born 1945-1959. *European Journal of Public Health*, 13, 226-229.

Bro R & Kiers HAL (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics*, 17, 274-286.

Ceulemans E & Van Mechelen I (2003). Uniqueness of N-way N-mode hierarchical classes models. *Journal of Mathematical Psychology*, 47, 259-264.

Cornelissen AMG, van den Berg J, Koops WJ & Kaymak U (2003). Elicitation of expert knowledge for fuzzy evaluation of agricultural

production systems. *Agriculture, Ecosystems and Environment*, 95, 1-18.

Currie I, Durban M & Eilers PHC (2003). Using P-splines to extrapolate two-dimensional Poisson data. *Proceedings of the 18th International Workshop on Statistical Modelling*. Leuven.

Eilers PHC (2003). Modern regression techniques in chemometrics. *Proceedings of the XXVemes Journees de Statistiques*. Lyon.

Eilers PHC (2003). Mixture models for background estimation. *Proceedings of the 18th International Workshop on Statistical Modelling*. Leuven.

Eilers PHC (2003). A perfect smoother. *Analytical Chemistry*, 75, 3631-3636.

Eilers PHC & Marx BD (2003). Multivariate calibration with temperature interaction using two-dimensional penalized signal regression. *Chemometrics and Intelligent Laboratory Instruments*, 66, 159-174.

Eilers PHC & Schimek MG (2003). Generalized additive models in particulate matter studies: statistical and computational perspectives. *Proceedings of the 54th ISI Session*. Berlin.

Elzinga CH (2003). Sequence Similarity: A Non-Aligning Technique. *Sociological Methods & Research*, 31, 3-29.

Gower JC & De Rooij M (2003). A comparison of the multidimensional scaling of triadic and dyadic distances. *Journal of Classification*, 20, 115-136

Kaymak U & Sousa JM (2003). Weighted constraint aggregation in fuzzy optimization. *Constraints*, 8, 29-46.

Kaymak U (2003). Data and cluster weighting in target selection based on fuzzy clustering. *Lecture Notes in Computer Science*, 2715, 568-575.

Kiers HAL & der Kinderen A (2003). A fast method for choosing the numbers of components in Tucker3 analysis. *British Journal of Mathematical and Statistical Psychology*, 56, 119-125.

Kiers HAL (2003). Some alternatives to PLS. In: *Società Italiana di Statistica, Analisi Statistica Multivariata per le Scienze Economico-Sociali, le Scienze Naturali e la Tecnologia*, 171-182. Napoli: RCE Edizioni.

Marx BD & Eilers PHC (2003). Smooth regression coefficient surfaces. *Proceedings of the 18th International Workshop on Statistical Modelling*. Leuven.

Otto SJ, Fracheboud J, Looman CWN, Broeders MJM, Boer R, Hendriks JHCL, Verbeek ALM & de Koning HJ (2003). Initiation of population based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *The Lancet*, 361, 1411-17.

Pas MMC van de, Weima S, Looman CWN & Broekmans FJM (2003). The use of fixed distance embryo transfer after IVF/ICSI equalizes the success rate among physicians. *Human Reproduction*, 4, 774-780.

Rooij IAJ van, Bancsi LFJMM, Broekmans FJM, Looman CWN, Habbema JDF & te Velde ER (2003). Women older than 40 years of age and those with elevated follicle-stimulating hormone levels differ in poor response rate and embryo quality in in vitro fertilization. *Fertility and Sterility*, 79, 482-488.

Scheffer GJ, Broekmans FJM, Looman CWN, Blankenstein M, Fauser BCJM, de Jong FH & te Velde ER (2003). The number of antral follicles in normal women with proven fertility is the best reflection of reproductive age. *Human Reproduction*, 18, 700-706.

Stokdijk M, Eilers PHC, Nagels J, et al (2003). External rotation in the glenohumeral joint during elevation of the arm. *Clinical Biomechanics*, 18, 296-302.

Timmerman ME & Kiers HAL (2003). Four simultaneous component models of multivariate time series from more than one subject to model intraindividual and interindividual differences. *Psychometrika*, 86, 105-121.

Werf MJ van der, de Vlas SJ, Brooker S, Looman CWN, Nagelkerke NJD, Habbema JDF & Engels D (2003). Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. *Acta Tropica*, 86, 125-139.

Zonneveld P van, Scheffer GJ, Broekmans FJM, Blankenstein MA, de Jong FH, Looman CWN, Habbema JDF & te Velde ER (2003). Do cycle disturbances explain the age-related decline of female fertility? Cycle characteristics of women above 40 compared to a reference

population of young women. *Human Reproduction*, 18, 495-501.

Routebeschrijving

Per trein

Verlaat het Leiden Centraal Station aan de LUMC kant, ongeveer 7 minuten lopen. Loop langs de ingang van het LUMC ziekenhuis, vervolgens langs de parkeergarage (Albinusdreef). Steek de Sandiforddreef over en neem het bruggetje over de wetering. De bezoekersingang van het laboratorium is schuin tegenover dit bruggetje.

Per bus

Alle bussen naar Leiden Centraal Station en via station lopen als boven, of buslijn 30 Wassenaarseweg

Met de auto

A44 afslag Oegstgeest, bij verkeerslichten rechtsaf: Rhijngeesterstraatweg, vervolg Endegeesterstraat, bocht naar links: Wassenaarseweg, rotonde: rechtdoor, rechtsaf bij nr 62, LUMC-Preklinische Laboratoria. A4 richting station, vervolgens viaduct onderdoor richting ziekenhuizen/Leidse Hout. Dit is de Rijnsburgerweg. Tweede afslag links (verkeerslichten) Wassenaarseweg, linksaf bij nr 62, LUMC-Preklinische Laboratoria.

Volg bordjes gebouwingang 9.



Attendance is free and open to anyone interested, but registration is mandatory (with the secretary, Marieke Timmerman: M.E.Timmerman@ppsw.rug.nl). The cost of the lunch is 9 euro. When registering, please indicate whether you will join the lunch or not.