



Nieuwsbrief no. 38  
April 2007

# Vereniging voor Ordinaties en Classificatie

Voorzitter: Patrick Groenen, Econometrisch Instituut, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR Rotterdam ([groenen@few.eur.nl](mailto: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](mailto:m.e.timmerman@ppsw.rug.nl))  
Penningmeester: Paul Arents, Quest International, Sensory Sciences & Consumer Acceptance, Huizerstraatweg 28, 1411 GP Naarden ([paul.arents@questintl.com](mailto:paul.arents@questintl.com)). Postbankrekening 161723 t.n.v. Vereniging voor Ordinaties en Classificatie, Naarden. Bankrekening nummer 777-5952385-56 Dexia Bank t.n.v. VOC, Naarden  
Redactie: Eva Ceulemans, Onderzoeksgroep kwantitatieve en persoonlijkheidspsychologie, Katholieke Universiteit Leuven, Tiensestraat 102, B-3000 Leuven, België ([Eva.Ceulemans@psy.kuleuven.be](mailto:Eva.Ceulemans@psy.kuleuven.be))  
VOC-home page: <http://www.voc.ac>

## Joint VOC and BNVKI meeting on data mining

April 27, 2007

Faculty Club, Room Kanunikkenzaal,  
Achter de Dom 7, Utrecht

10.00	Registration and Coffee
10.40	Patrick Groenen: <i>Minimization for Support Vector Machines by Iterative Majorization</i>
11.20	Cristophe Croux: <i>Robust Discrimination: an influence function approach</i>
12.00	Bernard de Baets: <i>Monotone distribution classifiers</i>
12.40	Lunch
13.50	Koen Vanhoof: <i>Aggregation operators' measures</i>
14.30	Lambert Schomaker: <i>Machine learning or pattern recognition? On classification methods in writer identification and handwritten manuscript retrieval</i>
15.10	Tea
15.40	Thorsten Joachims: <i>Efficient training of SVMs for structured outputs</i>
16.30	VOC annual member meeting
16.45	Drinks

## In dit nummer:

Programma Voorjaarsbijeenkomst	1
Aanmelden Voorjaarsbijeenkomst	1
Van de voorzitter	2
Abstracts van de Voorjaarsbijeenkomst	2
Boekbesprekingen	4
Personalia	6
Agenda	6
Publicaties en Rapporten	7
Financieel overzicht 2006	9
Agenda ledenvergadering VOC 27 april 2007	10
Notulen ledenvergadering april 2006	11
Verslag van de secretaris over 2006	11
Routebeschrijving	12

## Registration details for the Joint VOC and BNVKI meeting on data mining:

Those who would like to participate are welcome and are kindly requested to register preferably via the VOC website (<http://www.voc.ac>) or send an e-mail to the VOC secretary Marieke Timmerman ([m.e.timmerman@rug.nl](mailto:m.e.timmerman@rug.nl)). Participation is free, however the lunch needs to be paid (17 euro).

## Van de voorzitter

Na zes jaar is dit mijn laatste 'van de voorzitter' bericht. Ik ben blij dat ik de VOC in diverse hoedanigheden heb kunnen dienen. De VOC is voor mij een belangrijke beroepsvereniging waar sprekers op de VOC-bijeenkomsten vaak interessante en uitdagende onderwerpen aandragen. Hoogtepunten uit de afgelopen zes jaar zijn voor mij de gezamenlijke organisatie met de SCF van de IFCS-2000 meeting in Namur, België, de gezamenlijke bijeenkomst met de Duitse zustervereniging, de GfKl, Gesellschaft für Klassifikation, in Dortmund in 2004, en de jubileumbijeenkomst in het najaar van 2004.

Ik verlaat het bestuur met een gerust hart. Het bestuur heeft zich vernieuwd met diverse jonge leden van wie we nog veel zullen horen. Verder ben ik blij dat Ron Wehrens zich bereid heeft verklaard het voorzitterschap van mij te willen overnemen. Ik ben vol vertrouwen dat de vereniging bij hem in uitstekende handen is.

Met mij neemt de penningmeester Paul Arents ook afscheid van het bestuur. Paul heeft zich consciëntieus ingezet voor de VOC en ik wil hem dan ook hartelijk danken namens de VOC voor zijn inzet en de goede staat waarin hij de financiën van de VOC achterlaat.

Tenslotte wil ik aandacht vragen voor de voorjaarsbijeenkomst. Deze bijeenkomst hebben we georganiseerd in samenwerking met de Belgisch Nederlandse Vereniging Voor Kunstmatige Intelligentie (BNVKI). De VOC wil graag meer samenwerking en informatieuitwisseling met verenigingen die overeenkomstige interessegebieden hebben als die van de VOC. De BNVKI is de eerste vereniging waarmee de VOC zo'n gezamenlijke bijeenkomst organiseert rond het algemene thema data mining. Het programma ziet er interessant uit en we hopen dan ook dat deze bijeenkomst een eerste aanzet is tot meer samenwerking. Ik hoop dan ook veel VOC-ers op deze en toekomstige bijeenkomsten te ontmoeten.

Patrick Groenen (Voorzitter)

---

## Abstracts for the Joint VOC and BNVKI meeting on data mining

### Patrick Groenen, Georgi Nalbantov, Cor Bioch (Erasmus University Rotterdam): Minimization for Support Vector Machines by Iterative Majorization

Over the last few years there has been increasing interest in support vector machines (SVM) for the two-group classification problem. Currently, SVM classification belongs to the best prediction methods available. A nice property of the standard SVM is that it can be expressed

as a minimization problem of a quadratic loss function. However, as the solution is often obtained by switching to the dual problem, it is not easy to understand what the method is doing. In this presentation, we stick to the original loss function and discuss an intuitive explanation of what makes an SVM working. It turns out that an SVM is closely related to multiple regression with optimal scaling. In addition, we present a new iterative algorithm to solve the SVM that is based on iterative majorization. We discuss the advantages and disadvantages of this algorithm and compare its performance against standard methods for SVM.

*Patrick Groenen is full professor in statistics at the Econometric Institute, Erasmus University Rotterdam. He has published several papers on multivariate analysis methods and multidimensional scaling in the international literature. He is also co-author of a textbook on multidimensional scaling. His research interests include visualization, multidimensional scaling, (nonlinear) multivariate analysis, support vector machines, and majorization. Until this meeting, he is chairman of the VOC.*

### Christophe Croux (Catholic University Leuven): Robust Discrimination: an influence function approach

A discriminant rule allows to classify an observation to a certain group, depending on the characteristics of the observations. Classification rules are constructed from a so-called training sample, where outliers might be present. Hence, robust discriminant rules have been developed. The performance of a classification rule is typically measured by the error rate, being the percentage of incorrectly predicted observations. We study the sensitivity of the error rate with respect to the observations in the training sample by means of an influence functions approach. Such an approach allows to quantify the robustness of classifiers. Besides a classification rule being robust, we would also like it to be efficient, where efficiency is measured by the closeness of the classifiers' error rate to the lowest possible error rate one could get, the Bayes error rate. We show that, in the setting of discriminant analysis, the second order influence function of the error rate can be used to compute this classification efficiency.

Reference: Croux, C., Filzmoser, P., & Joossens, K. (to appear). Classification Efficiencies for Robust Linear Discriminant Analysis. *Statistica Sinica*.

*Christophe Croux is Professor of Statistics and Econometrics at the Catholic University Leuven (Belgium). His research interests are robust statistics, multivariate data analysis, classification, computational statistics, and applied time series analysis. He serves on the editorial board of Computational Statistics and Data Analysis and the Journal of the American Statistical Association.*

### **Bernard De Baets (Ghent University): Monotone distribution classifiers**

We present a rigorous framework for supervised ranking, a specific type of supervised classification:

- (1) objects are assigned labels belonging to a totally ordered set of labels;
- (2) objects are not described in terms of attributes, but in terms of (true) criteria;
- (3) the labels are assigned in a monotone way: objects with equal or higher scores on all criteria do not receive a lower overall score, or, in other words, are not assigned a lower label.

The purpose of supervised ranking is then to produce such a monotone classifier on the basis of a learning sample. Real-world data sets of this kind are usually pervaded with two undesirable phenomena: doubt and reversed preference. We focus on distribution classifiers. The monotonicity constraint then naturally leads to the notion of stochastic dominance. We confine ourselves to an ordinal setting and present a general framework from which several instance-based supervised ranking algorithms can be derived, such as the Ordinal Stochastic Dominance Learner.

***Bernard De Baets** (1966) holds an MSc and PhD in Mathematics (1988, 1995), and a Postgraduate degree in Knowledge Technology (1991). He is a Government of Canada Award holder (1988-89) and is Honorary Professor (2006) of Budapest Tech (Hungary). He is a Professor in Applied Mathematics (1999) at Ghent University, where he is leading the research unit Knowledge-based Systems (KERMIT) at the Faculty of Bioscience Engineering. The activities of KERMIT concern the principles and practice of the extraction, representation and management of knowledge by means of intelligent techniques. He is co-editor-in-chief of "Fuzzy Sets and Systems", area-editor of "4OR" and is on the editorial board of 9 other journals. He coordinates EUROFUSE (the EURO Working Group on Fuzzy Sets), is a member of the Board of Directors of EUSFLAT, of the Administrative Board of the Belgian OR Society, of the Technical Committee on Artificial Intelligence and Expert Systems of IASTED, and of the Executive Board of the International Computational Intelligence Society.*

### **Koen Vanhoof (Universiteit Hasselt): Aggregation Operators' Measures**

Over the last ten years, a whole range of aggregation operators (AGOPs) have been developed and extensively studied. Past research on the applications of AGOPs has mainly been focused on the AGOP's domain representation power or decision-making strength. However, as this study shows, certain AGOP's measures which play a role in the AGOP's behaviour (i.e., behavioural parameters) can be of great importance to practitioners, especially when these behavioural parameters seem to be proxies for domain-specific

concepts which is difficult to measure directly or to derive statistically. In such cases, these AGOP's behavioural parameters become much more than just another mathematical measure. The claim that a behavioural parameter can be interpreted as a proxy for a domain concept has to be validated both theoretically and empirically. Theoretical validation implies a close match between the AGOP's mathematical and behavioural properties on the one hand and the theoretical domain's information fusion process on other hand. Empirical validation is present when the behavioural parameter's empirical results correspond with existing domain knowledge. In case studies both theoretical and empirical validation is provided to support the basic assumption that aggregation operator measures enable us to obtain superior consumer information with substantial managerial relevance.

***Koen Vanhoof** attained a master in Physics in 1982, a master in Computer Science in 1985 and a PH.D. in Computer Science in 1988 at the Katholieke Universiteit Leuven. His major research interests lie in the areas of data mining, statistics, knowledge engineering and modelling, computational intelligence methods, decision support systems and soft computing applications to information management, marketing and finance. He has authored and/or co-authored over 20 peer-reviewed journal articles and about 6 book chapters and 60 conference papers on his research topics. He is co-editor of the International Journal of Information Theory and Applications. He has been appointed as a guest professor at Jagilionski University (Cracow, Poland), the University of Antwerp (Antwerp, Belgium), the University of Maastricht (Maastricht, the Netherlands), the University of Economics (Sofia, Bulgaria), the Technical University (St. Petersburg, Russian Federation) and the Academy of Economics (Wroclaw, Poland). He is Head of the Data mining research group at Hasselt University in Belgium.*

### **Lambert Schomaker (University of Groningen): Machine learning or pattern recognition? On classification methods in writer identification and handwritten manuscript retrieval**

The automatic recognition of handwriting and writer identification constitute a considerable challenge to science and engineering. As in speech recognition, the promising academic results in handwriting recognition are reduced drastically when methods are applied in the real world. In many application domains, the academic results cannot be replicated due to the lack of labeled data for supervised training. Consequently, there is a continuous use for unsupervised clustering methods and an increased use of plain techniques such as nearest-neighbour search and non-probabilistic pattern matching. Examples will be presented from the field of handwritten historical manuscript retrieval and a branch of behavioral

biometrics: writer identification. After an extensive study, using very large data sets, we have found that in the end more is gained from human time spent on feature design, i.e., traditional pattern recognition, than from the time spent in comparing the experimental results of machine-learning methods. Furthermore, many traditional methods appear ill-suited to practically handle the massive data sets (dozens of Gigabytes), with large numbers of instances and large numbers of feature dimensions. Our results in handwriting biometrics belong to the best achievable results to date, with data sets of up to 900 writers. Results in handwritten manuscript retrieval are more recent and very promising. Here, the challenge is to bootstrap a search engine for any new historical handwritten collection from a 'zero-knowledge' starting point, without any class label or any reliable segmentation into individual words. In brief: what to do when statistics are not (yet) present?

*Prof. dr. Lambert Schomaker (19-2-1957) is the Research Director of the Department of Artificial Intelligence at Groningen University. His research concerns pattern-recognition problems in handwriting recognition, writer identification, handwritten-manuscript retrieval and related topics. Recent work involves large-scale historical handwriting retrieval on high-performance computers. He is member of the IEEE Computer Society, is an active member within the IAPR and member of the BNVKI.*

### **Thorsten Joachims (Cornell University, USA): Efficient Training of SVMs for Structured Outputs**

This talk explores a large-margin approach to predicting multivariate objects like trees, clusterings, or alignments. Such problems arise, for example, when a natural language parser needs to predict the correct parse tree for a given sentence, when one needs to determine the co-reference relationships of noun-phrases in a document, or when predicting the alignment between two proteins. In particular, the talk will show how training such complex prediction rules can be formulated as a convex program. This leads to a Support Vector Machine (SVM) that generalizes conventional classification SVMs to a large range of structured outputs and multivariate loss functions. While the resulting optimization problems are convex quadratic, they have an exponential (or infinite) number of constraints. Nevertheless, the talk will show how cutting-plane methods can be used to solve the optimization problems efficiently. A by-product is a linear-time training algorithm also for linear binary classification SVMs. The algorithm is implemented in the SVM-Struct software and empirical results will be given for several application examples.

*Thorsten Joachims is an Associate Professor in the Department of Computer Science at Cornell University. In 2001, he finished his dissertation with the title "The Maximum-Margin Approach to Learning Text Classifiers:*

*Methods, Theory, and Algorithms", advised by Prof. Katharina Morik at the University of Dortmund. From there he also received his Diplom in Computer Science in 1997 with a thesis on WebWatcher, a browsing assistant for the Web. From 1994 to 1996 he was a visiting scientist at Carnegie Mellon University with Prof. Tom Mitchell. His research interests center on a synthesis of theory and system building in the field of machine learning, with a focus on Support Vector Machines and machine learning with text. He authored the SVM-Light algorithm and software for support vector learning.*

---

## **Boekbesprekingen**

### **An Introduction to Support Vector Machines (and other kernel-based learning methods). N. Cristianini and J. Shawe-Taylor. Cambridge University Press, 2000.**

Om met het goede te beginnen: dit boek behandelt op systematische en overzichtelijke wijze alle bouwstenen van een van de belangrijkste ontwikkelingen uit de Machine Learning: Support Vector Machines. Na de inleiding introduceert het tweede hoofdstuk classificatie via lineaire separatie met een vlak en het margin concept (error functie), alsmede de twee vormen van het perceptron algoritme om zo'n separatievlak te vinden. Het derde hoofdstuk beschrijft de Kernel-Induced Feature Space: de niet-lineaire mapping van de data in een hoogdimensionele ruimte waar de data veel beter gesepareerd kunnen worden. De truck is dat de kernels zo zijn geformuleerd dat de noodzakelijke berekeningen kunnen worden uitgevoerd in termen van een eenvoudige functie van de oorspronkelijke data in plaats van de hoogdimensionele feature space. Het vierde hoofdstuk beschrijft het statistische concept achter de SVM's, terwijl het vijfde hoofdstuk de optimalisatie van de achterliggende convexe kwadratische programmeerproblemen beschrijft.

De lezer beschikt nu over alle bouwstenen om SVM's te introduceren en krijgt vervolgens ook voldoende aanwijzingen voor implementatie en pseudocode om zelf aan de slag te gaan. Het laatste hoofdstuk geeft voorbeelden van toepassingen in verschillende gebieden.

Tegenover de systematische aanpak en het inhoudelijk goede overzicht staat dat de praktisch ingestelde lezer weinig houvast aan toepassingen kan ontleen en dat ik de helderheid van de presentatie zeker niet bewonderenswaardig vind voor zowel beginner als expert, zoals op de omslag gesuggereerd. Ik heb me daarbij afgevraagd of dat nu voornamelijk komt omdat ik uit een andere discipline kom, of omdat ik het boek onwillekeurig vergelijk met The Elements of Statistical Learning (een erg hoge standaard), of omdat de schrijfstijl

en notatie van de schrijvers toch wel wat te wensen overlaat. Vermoedelijk wat van alles.

In mijn eigen ogen is de uitleg van de notatie onvolledig, gebruiken ze te vaak woorden om te verwijzen zonder ook het correcte symbool te noemen en staan belangrijke wiskundige definities wel eens een paar pagina's verder dan wenselijk. Mede daardoor produceren de schrijvers met grote regelmaat cryptische zinnen - ook op gebieden die ik wel beheers - die slechts na verschillende keren herlezen voor de expert zijn te duiden, en waarvan ik vermoed dat ze de beginner het bos in sturen. Dat zoiets in de inleiding op verschillende plaatsen gebeurt is beter te accepteren dan bij de uitleg van de grondbegrippen. Op het gevaar af zelf heel cryptisch te worden zal ik een voorbeeld geven. Op p12 worden eerst definities gegeven rond de 'functional margin' die prima te volgen zijn. Daarna stappen de schrijvers over op de 'geometric margin' met de zin: "In both definitions if we replace functional margin by geometric margin we obtain the equivalent quantity for the normalised linear function (onhandige notatie), which therefor measures the Euclidean distances of the points from the decision boundary in the input space". Het probleem is hier dat er voorafgaand niet twee, maar eigenlijk drie definities staan, 'quantity' in enkelvoud staat, 'equivalent' niet in stricte zin moet worden gelezen, en 'therefor' ook daarna op geen enkele manier wordt onderbouwd. Op deze manier worden ook veel interessante zijsprongen naar verwante gebieden gemaakt, die mij voortdurend de neiging gaven er maar een ander boek bij te pakken om de opmerkingen te verhelderen.

Uiteindelijk raakte ik wat meer gewend aan de schrijfstijl, en heb ik veel geleerd over een onderwerp dat tot een van de belangrijkste ontwikkelingen in predictive modelling kan worden gerekend. Dat maakt het boek zeker de spreekwoordelijke moeite van het lezen waard.

Bart Jan van Os

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

Probably you are familiar with the "XYZ tour" T-shirt. The back of the shirt shows a long list of cities and dates, telling you where and when a famous pop group has given (or is to give) performances. I don't know much about their singing qualities, but Geert Molenberghs and Geert Verbeke deserve a world tour T-shirt as the Simon and Garfunkel of Mixed Models. Not only do they write about them, but they also presented very popular courses in many places.

This book is a sequel to *Linear Mixed Models for Longitudinal Data* by the same authors (Springer, 2000) which only considered normally distributed variables. Normal distribution theory is a cozy world, because

whether you look at hierarchical components or at marginal variables, distributions always are normal. You might have to do some hefty matrix manipulations, but you can be sure that those matrices completely specify the distribution. All this beauty is gone when the observations are counts or binary observations.

The book can be approached in several ways. It is a handbook of analyses, because many real-life data sets are provided, with SAS code for model estimation. The data and the analyses are presented in detail, using tables and graphs, including model selection and interpretation. Re-running the analyses one-self can be a great way to get acquainted with the practical aspects of modeling non-normal longitudinal data. The book also presents a careful overview of model types (like marginal, conditional and subject-specific), their meaning and their advantages and drawbacks. In that sense it is a handbook on the foundations of mixed models. The presentation is not overly technical and will be accessible to many readers.

Between the extremes of foundations and applications, the book contains a massive amount of technical material, which is far from easy reading in many places. It is clear that the authors know what they are talking about, but the models inherently are complicated in many places and frequently detailed derivations are skipped.

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

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

It is clear that I like the book. It is one to return to over and over again. And if you want to see the authors on stage: July 1, in Barcelona

(<http://mat.uab.es/~iws2007/sc.html>).

Paul Eilers

## Personalia

**Paul Eilers** is per 1 februari 2007 Universitair Hoofddocent op twee plaatsen: bij Methoden en technieken, Faculteit Sociale Wetenschappen, Universiteit Utrecht (80%) en bij de Groep Datatheorie van de Faculteit Sociale Wetenschappen van de Universiteit Leiden (20%).

## Agenda

- May 14 – 15. **Amsterdam, the Netherlands.** Annual Belgian-Dutch Benelearn Conference. <http://staff.science.uva.nl/~katrenko/benelearn07/>
- May 21 – 23. **Ames, Iowa, USA.** 14th Annual Spring Research Conference on Statistics in Industry and Technology. [www.stat.iastate.edu/SRC07](http://www.stat.iastate.edu/SRC07)
- May 31. **Wageningen, the Netherlands.** Chemometrics for Food, Food for Chemometrics. [www.chemometrie.nl](http://www.chemometrie.nl)
- June 07 – 10. **Urbana, Illinois, USA.** The 2007 Meeting of the Classification Society of North America.
- June 09 – 12. **Rome, Italy.** Sixth International Workshop on Objective Bayesian Analysis. <http://3w.eco.uniroma1.it/OB07/>
- June 10 – 14. **Hang Zhou, China.** International Conference on Bioinformatics. [www.stat.umn.edu/~statconf](http://www.stat.umn.edu/~statconf)
- June 10 – 15. **Ascona, Switzerland.** Statistics for Biomolecular Data Integration and Modeling. [www.stat.math.ethz.ch/talks/Ascona\\_07/](http://www.stat.math.ethz.ch/talks/Ascona_07/)
- June 11 – 23. **Cison Di Valmarino (Treviso), Italy.** 2007 Edition of the Summer School On Modern Methods in Epidemiology and Biostatistics. [www.bioepi.org](http://www.bioepi.org)
- June 18 – 21, **Montréal, Québec, Canada.** ICES-III: Third International Conference on Establishment Surveys. <http://www.amstat.org/meetings/ices/2007/index.cfm>
- June 22. **Rotterdam, the Netherlands.** Tweejaarlijkse bijeenkomst van het Nederlands Forum voor Medische Besliskunde. [www.hetcongresbureau.nl](http://www.hetcongresbureau.nl)
- June 25 – 27. **Rotterdam, the Netherlands.** CARME 2007: Correspondence analysis and related methods. <http://www.carme-n.org/carme2007/>
- June 25 – 29. **Fleurance, France.** New directions in Monte Carlo methods. [www.adapmc07.enst.fr](http://www.adapmc07.enst.fr)
- July 02 – 04. **London, UK.** International Conference of Computational Statistics and Data Engineering 2007. [www.iaeng.org/WCE2007/ICCSDE2007.html](http://www.iaeng.org/WCE2007/ICCSDE2007.html)
- July 02 – 06. **Barcelona, Spain.** 22th International Workshop on Statistical Modelling. [mat.uab.es/~iws2007/](http://mat.uab.es/~iws2007/)
- July 04 – 06. **Leeds, UK.** LASR 2007 - Systems Biology & Statistical Bioinformatics. [www.maths.leeds.ac.uk/statistics/workshop/](http://www.maths.leeds.ac.uk/statistics/workshop/)
- July 09 – 13. **Tokyo, Japan.** IMPS2007: International Meeting of the Psychometric Society. <http://www.ech.co.jp/imps2007/>
- July 09 – 20. **Columbus, Ohio, USA.** Summer Program in Applied Biostatistical & Epidemiological Methods. [www.biostatistics.osu.edu](http://www.biostatistics.osu.edu)
- July 12 – 14. **Johnson City, Tennessee, USA.** Sixth International Conference on Lattice Path Combinatorics and Applications. [www.etsu.edu/math/godbole/lattice/index.htm](http://www.etsu.edu/math/godbole/lattice/index.htm)
- July 29 – August 02. **Salt Lake City, Utah, USA.** 2007 joint statistical meetings. <http://www.amstat.org/meetings/jsm/2007/>
- August 02 – 04. **Seattle, Washington, USA.** 13th ISSAT International Conference on Reliability & Quality in Design. [www.issatconferences.org](http://www.issatconferences.org)
- August 13 – 15. **Kunming, China.** International Conference on the Frontiers of Statistics: High Dimensional Data Analysis. [c2s2.yale.edu/kunming](http://c2s2.yale.edu/kunming)
- August 22 – 29. **Lisboa, Portugal.** 56th Session of the International Statistical Institute. <http://www.isi2007.com.pt/isi2007/index.php>
- August 30 – September 01. **Aveiro, Portugal.** IASC 07 - Statistics for Data Mining, Learning and Knowledge Extraction. <http://www2.mat.ua.pt/iasc07/>
- September 10 – 12. **Florence, Italy.** Latent Variable Modeling. [www.ds.unifi.it/mplus/](http://www.ds.unifi.it/mplus/)

- September 12 - 14. **Macerata, Italy**. CLADAG 2007: Sixth scientific meeting of the Classification and Data Analysis Group of the Italian Statistical Society. <http://cladag2007.unimc.it/>
- October, 09. **Utrecht, the Netherlands**. Tweede Nederlandse Demografiedag. <http://www.nvdemografie.nl/activiteiten/congres/2007/ndd2007.html>
- October 12 – 14. **Greensboro, North Carolina, USA**. International Conference on Advances in Interdisciplinary Statistics and Combinatorics. [www.uncg.edu/mat/aisc](http://www.uncg.edu/mat/aisc)
- October 24 – 26. **San Francisco, California, USA**. International Conference on Machine Learning and Data Analysis 2007. [www.iaeng.org/WCECS2007/ICMLDA2007.html](http://www.iaeng.org/WCECS2007/ICMLDA2007.html)
- November 05 – 06. **Utrecht, the Netherlands**. The 19th Belgian-Dutch Conference on Artificial Intelligence (BNAIC 2007). <http://www.cs.uu.nl/bnaic2007/>
- 
- Publicaties en rapporten**
- Balázs, K., Hidegkuti, I., & De Boeck, P. (2006). Detecting heterogeneity in logistic regression models. *Applied Psychological Measurement*, 30, 322-344.
- Boersma, S.N., Maes, S., Joekes, K., & Dusseldorp, E. (2006). Goal processes in relation to goal attainment: Predicting health-related quality of life in MI-patients. *Journal of Health Psychology*, 11, 931-945.
- Bollaerts, K., Eilers, P.H.C., & Van Mechelen, I. (2006). Simple and multiple P-splines regression with shape constraints. *British Journal of Mathematical and Statistical Psychology*, 59, 451-469.
- Borghini, E., de Onis, M., Garza, C., Van den Broeck, J., Frongillo, E.A., Grummer-Strawn, L., van Buuren, S., Pan, H., Molinari, L., Martorell, R., Onyango, A., & Martines, J. (2006). Methods for constructing the WHO child growth references: Recommendations of a statistical advisory group. *Statistics in Medicine*, 25, 247-265.
- Elffers, H., Verboon, P., Huisman, W. (eds.) (2006). *Managing and Maintaining Compliance*. Boom Juridische Uitgeverij: Den Haag.
- Coppi, R., Gil, M.A., & Kiers, H.A.L. (2006). The fuzzy approach to statistical analysis. *Statistics and Data Analysis*, 51, 1-14.
- de Zwart, I.M., Haans, J.J.L., Verbeek, P., Eilers, P.H.C., de Roos, A., & Masclee, A.A.M. (2007). Gastric accommodation and motility are influenced by the barostat device: assessment with magnetic resonance imaging. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 292, G208-G214.
- Giordani, P. & Kiers, H.A.L. (2006). A comparison of three methods for principal component analysis of fuzzy interval data. *Computational Statistics and Data Analysis*, 51, 379-397.
- Goegebeur, Y., De Boeck, P., Molenberghs, G., & del Pino, G. (2006). A local-influence-based diagnostic approach to a speeded item response theory model. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 55, 5, 647-676.
- Gonzalez, J., Tuerlinckx, F., De Boeck, P., & Cools, R. (2006). Numerical integration in logistic-normal models. *Computational Statistics and Data Analysis*, 51, 1535-1548.
- Hanafi, M., & Kiers, H.A.L. (2006). Analysis of K sets of data, with differential emphasis on agreement between and within sets. *Computational Statistics and Data Analysis*, 51, 1491-1508.
- Hehenkamp, W.J.K., Looman, C.W.N., Themmen, A.P., de Jong, F.H., Te Velde, E.R., & Broekmans, F.J.M. (2006). Anti-Müllerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. *Journal of Clinical Endocrinology & Metabolism*, 91, 4057-63.
- Jacobusse, G.J., van Buuren, S., Verkerk P.H. (2006). An interval scale for development of children aged 0-2 years. *Statistics in Medicine*, 25, 2272-2283.
- Klinkert, E.R., te Velde, E.R., Weima, S., van Zandvoort, P.M., Hanssen, R.G.J.M., Nilsson, P.R., de Jong, F.H., Looman, C.W.N., & Broekmans, F.J.M. (2006). FSH Receptor genotype is associated with pregnancy but not with ovarian response in IVF. *Reproductive Biomedicine Online*, 13, 687-695.

- Kok, J.D., Looman, C.W.N., Weima, S., & te Velde, E.R. (2006). A high number of oocytes obtained after ovarian hyperstimulation for IVF or ICSI is not associated with decreased pregnancy outcome. *Fertility and Sterility*, *85*, 918-924.
- Loudon, S.E., Fronius, M., Looman, C.W.N., Awan, M., Simonsz, B., van der Maas, P.J., & Simonsz, H.J. (2006). Predictors and a remedy for noncompliance with amblyopia therapy in children measured with the occlusion dose monitor. *Investigative Ophthalmology & Visual Science*, *47*, 4393-4400.
- Missotten, G.S.O., Notting, I.C., Schlingemann, R.O., Zijlmans, H.J., Lau, C., Eilers, P.H.C., Keunen, J.E.E., & Jager, M.J. (2006). Vascular endothelial growth factor A in eyes with uveal melanoma. *Archives of ophthalmology*, *124*, 1428-1434.
- Oosting, J., Lips, E.H., van Eijk, R., Eilers, P.H., Szuhai, K., Wijmenga, C., Morreau, H., van Wezel, T. (2007). High-resolution copy number analysis of paraffin-embedded archival tissue using SNP BeadArrays. *Genome Research*, *17*, 368-376.
- Polinder, S., van Beeck, E.F., Essink-Bot, M.L., Toet, H., Looman, C.W.N., Mulder, S. & Meeding, W.J. (2007). Functional outcome at 2.5, 5, 9 and 24 months after injury in the Netherlands. *Journal of Trauma*, *62*, 133-141.
- Razzaghi-Azar, M., Sadigh, N., Moghimi, A., Montazer, M., Golnari, P., Shulami, L.Z., van Buuren, S., Sadeghi, H.M., Kazemi, A.Z., & Nezhad, S.M.F. (2006). Age of the onset of puberty in girls and boys of Tehran, Iran. *Annals of Human Biology*, *33*, 628-633.
- San Martín, E., del Pino, G., & De Boeck, P. (2006). IRT models for ability-based guessing. *Applied Psychological Measurement*, *30*, 183-203.
- Schepers, J., Van Mechelen, I., & Ceulemans, E. (2006). Three-mode partitioning. *Computational Statistics and Data Analysis*, *51*, 1623-1642.
- Tuerlinckx, F., Rijmen, F., Verbeke, G., & De Boeck, P. (2006). Statistical inference in generalized linear mixed models: A review. *British Journal of Mathematical and Statistical Psychology*, *59*, 225-255.
- van Buuren, S., Boshuizen, H.C., & Reijneveld, S.A. (2006). Towards targeted hypertension screening guidelines. *Medical Decision Making*, *26*, 145-153.
- van Buuren, S., Brand, J.P.L., Groothuis-Oudshoorn, K., & Rubin, D.B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, *76*, 1049-1064.
- van den Hurk, K., van Dommelen, P., de Wilde, J.A., Verkerk, P.H., van Buuren, S., & HiraSing, R.A. (2006). Snelle toename gewicht bij kinderen tussen de 4 en 15 jaar. *Tijdschrift voor Jeugdgezondheidszorg*, *38*, 83-84.
- Zijlmans, H.J.M.A.A., Fleuren, G.J., Baelde, H.J., Eilers, P.H.C., Kenter, G.G., & Gorter, A. (2007). Role of tumor-derived proinflammatory cytokines GM-CSF, TNF-alpha, and IL-12 in the migration and differentiation of antigen-presenting cells in cervical carcinoma. *Cancer*, *109*, 556-565.



## Financieel overzicht over het jaar 2006

### Exploitatie-overzicht (Nederlands en Vlaams)

#### Inkomsten

contributies VOC 2002 (1)	15
contributies VOC 2003 (1)	45
contributies VOC 2004 (1)	60
contributies VOC 2005 (1)	120
contributies VOC 2006 (1)	600
contributies VOC 2007 (1)	1000
lunch ontvangsten voorjaarscongres 2006 (3)	270
lunch ontvangsten najaarscongres 2006 (4)	320
rente dexiarekening	0
rente plusrekening	21
contributie zonder persoonsnaam (28/11/2006) (6)	40
<i>totaal</i>	2491

#### Uitgaven

najaarsbijeenkomst sprekers verblijf 2005 (2)	200
najaarsbijeenkomst sprekers kado 2005 (2)	134
voorjaarsbijeenkomst sprekers verblijf 2006 (3)	123
voorjaarsbijeenkomst sprekers kado 2006 (3)	182
voorjaarsbijeenkomst 2006 (3)	606
najaarsbijeenkomst sprekers verblijf 2006 (4)	138
najaarsbijeenkomst sprekers kado 2006 (4)	84
najaarsbijeenkomst 2006 (4)	908
bank- en transactiekosten 2005	1
bank- en transactiekosten 2006	34
website 2006 via Denit	114
kamer van koophandel 2006	21
IFCS 2005 (5)	83
IFCS 2006 (5)	83
correctie teveel betaald lid (1-2-2006) (6)	64
<i>totaal</i>	2774

### Balans per 31/12/2006

#### Debet

saldo Ned. girorekening 31/12/2006 (7)	2600
saldo Belg. girorekening 31/12/2006 (7)	2663
saldo plusrekening 31/12/2006 (7)	2605
<i>totaal</i>	7869

#### Credit

voortuitbetaalde contributies VOC 2007 (8)	1000
eigen vermogen (9)	6869
<i>totaal</i>	7869

## Voetnoten bij financieel overzicht 2006 en balans

1. De VOC telde op 31/12/2006 107 leden hetgeen netto een daling is van 2 tov van vorig jaar. 86 Leden hebben betaald voor 2006.
2. De kosten van de najaarsbijeenkomst 2005 werden gedeeltelijk verrekend in 2006. Het betrof een teruggave van de hotel/taxi kosten van Gower en de sprekerskado kosten.
3. De lunch voor de voorjaarsbijeenkomst in Rotterdam werd betaald door de deelnemers. Kosten voor zaalhuur, drankjes, sprekerskados en verblijf van keynote spreker Hand werden door de VOC betaald.
4. De lunch voor de najaarsbijeenkomst in Leiden werd betaald door de deelnemers. Kosten voor zaalhuur, drankjes, sprekerskados en sprekersreiskosten werden door de VOC betaald.
5. De bijdrage aan de IFCS is in 2006 uitgevoerd en betreft de jaren 2005-2006.
6. Er is een onduidelijke overboeking geweest naar de rekening van de VOC. Het gaat waarschijnlijk om 2 contributies (40 euro). Verder is een teveel gestort bedrag (uit 2005) terugbetaald (63.50 euro).
7. Op 31/12/2006 bedroeg het saldo van Nederlandse giro- en spaarrekening, en de Belgische rekening €7869. Een jaar daarvoor was dit €8152.
8. In 2006 is er 1000 euro contributie betaald voor het jaar 2007.
9. Het eigen vermogen was op 31/12/2006 toegenomen van €6057 eind 2005 tot €6869.

jaar	eigen vermogen
2006	€ 6869
2005	€ 6057
2004	€ 5019
2003	€ 6795
2002	€ 6408
2001	€ 5898
2000	€ 5731
1999	€ 4871
1998	€ 5100

Paul Arents  
Peningmeester

## Agenda voor de ledenvergadering van de VOC op 27 april 2007

### 1. Opening

### 2. Notulen Ledenvergadering 28 april 2006

Deze zijn elders in de Nieuwsbrief opgenomen.

### 3. Jaarverslag van de secretaris over 2006

Dit is elders in deze Nieuwsbrief opgenomen.

### 4. Financieel verslag

- Jaarverslag van de penningmeester (zie deze Nieuwsbrief).
- Verslag van de kascommissie (Hugo Duivenvoorden en Kees Elzinga)
- Benoeming van de nieuwe kascommissie.

### 5. Bestuurssamenstelling

Het bestuur bestaat nu uit de volgende leden (achter de functie staat de resterende duur van hun termijn):

Patrick Groenen	Voorzitter (0)
Marieke Timmerman	Secretaris (1)
Paul Arents	Penningmeester (0)
Eva Ceulemans	Redacteur Nieuwsbrief (0)
Mark de Rooij	Lid (1)
Michel van de Velden	Lid (2)
Laurence Frank	Lid (2)

Van drie bestuursleden eindigt hun termijn: Patrick Groenen, Paul Arents en Eva Ceulemans. Als nieuw bestuurslid wordt voorgedragen Ron Wehrens voor de functie van voorzitter. Eva Ceulemans is beschikbaar voor herbenoeming. Laurence Frank is bereid het penningmeesterschap op zich te nemen. Tegenkandidaten kunnen aangemeld worden tot 24 uur voor de ledenvergadering, bij de secretaris.

### 6. CSNA en ADAC affiliate membership.

Toelichting: In het verleden gaf de VOC de mogelijkheid om via de VOC 'affiliate member' te worden van de CSNA waarmee automatisch het Journal of Classification ontvangen wordt. In 2005 heeft het VOC bestuur om administratieve redenen deze mogelijkheid afgeschaft. Hierna bleek dat zowel de CSNA als een aantal VOC leden deze afschaffing zeer betreunden. Het bestuur overweegt daarom sterk om het Affiliate membership weer in te voeren.

Naast het CSNA lidmaatschap kan voor VOC leden ook een abonnement op het nieuwe journal (1<sup>e</sup> issue: februari 2007) *Advances in Data Analysis and Classification*

(ADAC) interessant zijn. Op ADAC is een abonnement via de VOC mogelijk.

Voor zowel CSNA en ADAC geldt dat een lidmaatschap/abonnement via de VOC de leden een financieel voordeel geeft; bovendien geeft het de steun van de VOC aan de betreffende organisatie aan.

De lidmaatschap/abonnementsgelden zijn voor 2007 als volgt:

	Individueel lidmaatschap, niet via VOC	Via VOC
CSNA	€80,00	€60,00
ADAC (excl. VAT & Handling)	€148,00	€39,00

Het bestuur hoort graag de mening van de leden over een al dan niet aanbieden van het lidmaatschap op CSNA en/of ADAC via de VOC.

### 7. Oproep van bestuur aan leden tot organisatorische bijdrage aan reguliere bijeenkomsten en Jubileumbijeenkomst

### 8. Wvttk/Rondvraag

### 9. Sluiting

## Verslag ledenvergadering VOC 28-4-2006

Datum: 28-4-2006

Plaats: Erasmus Universiteit Rotterdam

Verslag: Marieke Timmerman

### 1. Opening en ingekomen stukken.

De voorzitter opent de vergadering. Er zijn geen ingekomen stukken.

### 2. Notulen jaarvergadering 15 april 2005

De notulen worden aanvaard.

### 3. Jaarverslag van de secretaris

Het verslag wordt aanvaard.

### 4+5 Financieel verslag en Jaarverslag van de Penningmeester

De penningmeester geeft een korte toelichting op het Financieel overzicht over het jaar 2005. Het verslag wordt aanvaard.

### 6. Verslag van de kascommissie

Hugo Duivenvoorden en Jeroen Vermunt hebben de boeken gecontroleerd. Zij zijn helaas afwezig, maar laten weten de boeken keurig in orde bevonden te hebben, en doen het voorstel tot decharge van de penningmeester. Dat voorstel neemt de vergadering over.

### 7. Benoeming van de nieuwe kascommissie

Nieuwe kascommissie: Hugo Duivenvoorden en Cees Elzinga

### 8. Bestuursamenstelling

Stef van Buuren en Paul Eilers treden reglementair af. Zij worden heel hartelijk bedankt voor hun jarenlange inspanningen voor de VOC. Laurence Frank en Michel van de Velden hebben zich verkiesbaar gesteld. Er zijn geen tegenkandidaten en zij worden bij acclamatie gekozen. Michel van de Velden wordt webredacteur, en Laurence Frank is beoogd penningmeester, bij einde van Paul Arents' termijn in 2007.

### 9. Heffen entreegeld

Er ontstaat een levendige discussie over het al dan niet heffen van entreegeld. Men is zeker niet onverdeeld voor. Er wordt geopperd om geen entreegeld te heffen, doch frequente niet-lid bezoekers uit te nodigen tot lidmaatschap. Het bestuur buigt zich verder over deze kwestie.

### 10, 11, 12. Wvttk, Rondvraag en Sluiting

Er komt verder niets ter tafel en er wordt geen gebruik gemaakt van de rondvraag. De voorzitter sluit de vergadering.

## Jaarverslag van de Secretaris over 2006

### 1. Leden

De VOC startte 2006 met 108 leden en eindigde met 107 leden. Er waren 4 opzeggingen, en 3 aanmeldingen.

### 2. Bestuur

Het bestuur van de VOC had in 2006 de volgende samenstelling:

Patrick Groenen	Voorzitter
Marieke Timmerman	Secretaris
Paul Arents	Penningmeester
Eva Ceulemans	Redacteur Nieuwsbrief
Michel van der Velden	Webmaster
Mark de Rooij	Lid
Laurence Frank	Lid

Het bestuur vergaderde drie maal telefonisch. De belangrijkste onderwerpen waren de Bijeenkomsten, en lopende zaken als de Nieuwsbrief en de website.

### 3. Activiteiten

In 2006 hadden we twee goed bezochte bijeenkomsten. Deze werden breder aangekondigd dan in het verleden.

De Voorjaarsbijeenkomst (28/4/06, 37 deelnemers, Erasmus Universiteit Rotterdam) had als thema 'measurement theory'. De sprekers waren David Hand en Wim van der Linden (beide keynotes) en Paul Krabbe, Christiaan Heij en Gert Jacobusse.

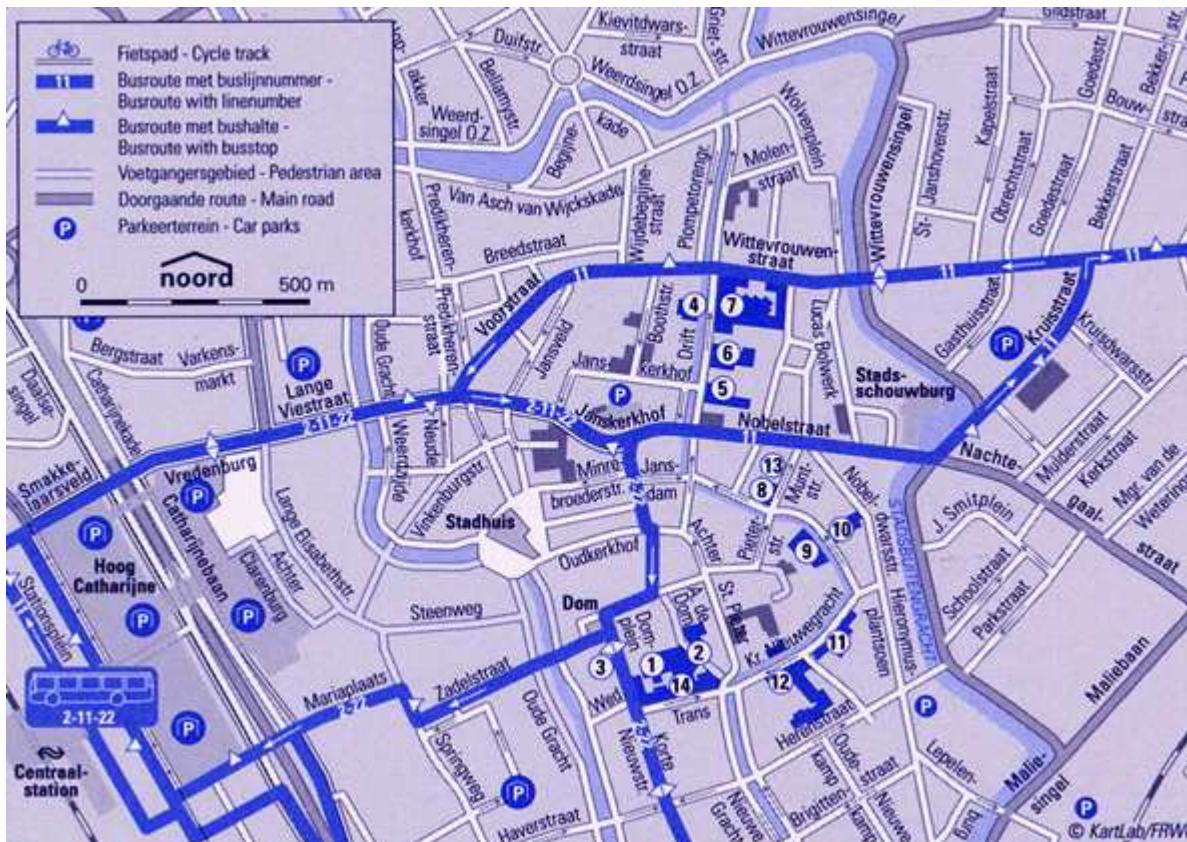
De Najaarsbijeenkomst (3/11/06, 50 deelnemers, Universiteit Leiden) had als thema: 'View on Changes: models for longitudinal data'. De sprekers waren Geert Verbeke en Paul van Geert (beide keynotes) en Reinoud Stoel, Ellen Hamaker, Cees Elzinga en Mark de Rooij.

### 4. Publiciteit.

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

## Routebeschrijving:

Faculty Club, Achter de Dom 7, 3512 JN Utrecht (cijfer 1 op plattegrond)



### Per auto

#### *Komend vanuit de richting A12/A2*

U neemt de afslag Utrecht, vervolgens afslag Utrecht/Jaarbeurs en Jaarbeurs/Centrum. U komt het centrum binnen via de Graadt van Roggweg. U volgt de route P Centrum. Wanneer u de Daalsetunnel ziet (onder het spoor door) gaat u bij de stoplichten rechtsaf. U houdt rechts aan (niet de tunnel in) en gaat bij de volgende stoplichten rechtdoor, onder de Hoog Catharijne Traverse door. Bij P Springweg slaat u linksaf en direct weer rechtsaf. U volgt de weg en voorbij de bocht naar rechts volgt u nog steeds P Springweg. U rijdt vanzelf de parkeergarage in.

U verlaat de parkeergarage richting Oudegracht. Op de Oudegracht gaat u direct de brug over en daarna linksaf. Komend op Het Wed gaat u rechtsaf. U loopt om de Domkerk heen en gaat direct rechts en nogmaals rechts (als de deur van de Pandhof voor de Domkerk geopend is kunt u ook door de Pandhof lopen). U loopt in de straat Achter de Dom. Na ± 150 meter ziet u een torentje met aan uw linkerhand een ingang. Hier is de Faculty Club gehuisvest.

#### *Komend vanuit de richting A27/A28*

U neemt de afslag Rijnsweerd en vervolgens de route P Centrum West. U volgt deze route over de Rubenslaan, Maliesingel en de Catharijnesingel. Halverwege de Catharijnesingel ziet u P Springweg. Volg deze route. Zie verder boven.

### Per openbaar vervoer vanaf Centraal Station

Vanaf Utrecht CS neemt u buslijn 2, halte Domplein of buslijn 3, 5, 9 of 11, halte Janskerkhof. U loopt via de Korte Jansstraat (richting Domtoren) en de Domstraat. Vervolgens rechtsaf en dan linksaf. U komt dan op het Domplein. Zie verder boven.

### Wandelroute vanaf Centraal Station

De wandeling vanaf Utrecht Centraal naar het Domplein neemt ongeveer 15 minuten in beslag. U loopt door Hoog Catharijne en gaat bij V&D met de roltrap naar buiten, u slaat rechtsaf en loopt via de Lange Elisabethstraat en de Steenweg onder de Dom door naar het Domplein, waar u rechts het Academiegebouw ziet liggen.