Online

Lifestat 2008

Statistics and Life Sciences: Perspectives and Challenges

March 10 – 13, 2008, Munich

Abstracts of Talks and Posters

First Conference of the Central European Network

54. Biometrisches Kolloquium

25. ROeS Seminar

Joint Biometric Conference of the German Region (DR), the Region Austria-Switzerland (ROeS) and the National Group Poland of the International Biometric Society (IBS)

Editors:
L. A. Hothorn, U. Mansmann, G. Tutz, U. Burger, S. Mejza

Ludwig Maximilians Universität Munich

Lifestat 2008

Statistics and Life Sciences:
Perspectives and Challenges

March 10 - 13, 2008, Munich, Germany

Abstracts of Talks and Posters

First Conference of the Central European Network

54. Biometrisches Kolloquium

25. ROeS Seminar

Joint Biometric Conference of the German Region (DR),
the Region Austria-Switzerland (ROeS) and the National Group Poland of
the International Biometric Society (IBS).

Editors:
L. A. Hothorn (Hannover), U. Mansmann (Munich), G. Tutz (Munich),
H. U. Burger (Basel), S. Mejza (Poznan)

Lifestat 2008
Statistics and Life Sciences: Perspectives and Challenges
Statistik in den Lebenswissenschaften: Perspektive und Herausforderung
La Statistique dans les Sciences de la Vie: perspectives et enjeux
Statistica nelle scienze umane: prospettive e sfide
Statystyka w naukach przyrodniczych: Perspektywy i wyzwania

– Abstracts of Talks and Posters –

First Conference of the Central European Network
54. Biometrisches Kolloquium
25. ROeS Seminar

March 10 - 13, 2008

Ludwig-Maximilians-Universität (LMU) Munich, Germany

Design: Markus Schmidberger
Cover Design: Markus Schmidberger, Uwe Friedrich
Print: AZ Druck und Datentechnik, Kempten (Germany)

ISBN: 978-3-86541-266-9

Foreword

Dear Participant, dear Reader,

Last year the Central European Network of the International Biometric Society (IBS) was established and is currently comprising the German Region, the Austrian-Swiss Region and the Polish group of the IBS. One of the main purposes of the network is to facilitate scientific exchange in the area of biometrics among the member groups and regions of the IBS and to organize joint meetings in which scientific exchange can take place.

The conference LIFESTAT 2008, Statistics and Life Sciences, is the first joint meeting of this network with more than 200 paper and poster submissions from the three regions and groups. Taking place in Munich the German Region took the lead in organizing this event. With the dedicated support of the Austrian-Swiss Region and the Polish group, however, we hope that it is notable already from the program and this abstract book that this conference goes beyond a traditional Biometric Colloquium as held by the German Region every year and is comprising ideas of the two other members of the network as well. We hope that this conference fulfills already some of the objectives of the joint network and really enriches our scientific perspectives in fruitful presentations and discussions with as many colleagues as possible from all member regions and groups of the network.

Finally, it is also an honor for our network, that the president of the International Biometric Society, Andrew Mead (UK), will be present at this first joint conference and will give a presentation on experimental design and analysis approaches for large and complex time-course gene expression microarray studies.

For this conference we organized a number of sessions on interesting and hot topics of biometry. The backbone of the conference program is a series of introductory sessions on topics as Flexible designs, Adaptive seamless design for combining phase II / III clinical studies, Boosting for biomedical data, Non-inferiority trials, Bayesian modeling in biostatistics, Cost efficient designs for biostatisticians, Genome-wide association studies and, finally, Mapping approaches in plants and animals. We are happy that we have been able to gain support in organizing these sessions by many well known colleagues as W. Maurer (Basel), W. Brannath (Vienna), P. Bühlmann (Zurich), A. Munk (Göttingen) together with J. Röhmel (Bremen), H. Dette (Bochum), I. R. König (Lübeck) and G. Freyer (Dummerstorf).

Other sessions were organized by the working groups (AGs, Sektionen) of our society and reflect very well the diversity and the quality of biometry in our regions and groups: the working group on 'Nichtparametrische Methoden', on 'Adaptive und multiple Verfahren', on 'Pharmazeutische Forschung', on 'Landwirtschaftliches Versuchswesen', on 'Populationsgenetik-Genomanalyse', on 'Bayes Methodik', on 'Statistical Computing', on 'Bioinformatik', on 'Ökologie-Umwelt', on 'Statistische Methoden in der Medizin' as well as on 'Räumliche Statistik'. There are also contributed sessions which are reflecting the present work of many active colleagues.

We want to make the LIFESTAT 2008 conference interesting and attractive for all our colleagues and to make it a first success and model for joint network meetings. We will certainly monitor the results of this meeting to find out what we can improve for future meetings.

Last but not least we would like to thank Ulrich Mansmann, Gerhard Tutz and the local organizing committee for all their dedication and the many hours they have put into the preparation of this meeting to make it a success. Only those who have organized such a congress once really know how much work is really connected with it. We would also like to thank the session chairs and all the presenters for their contributions to this conference. Finally we hope that every participant will enjoy the conference and will be able to attend many interesting sessions and to participate in many fruitful discussions.

Munich,
March 2008

Ludwig A. Hothorn
Stanislaw Mejza
H. Ulrich Burger
(conference presidents)

Greeting of the International Biometric Society

Dear Participant, dear Reader,

It is a great pleasure to have been invited, as President of the International Biometric Society, to attend LIFESTAT 2008, the first conference of the IBS Central European Network, in Munich. I look forward to a scientific programme containing talks from the usual diversity of biometrical methodological topics and application areas that is the trademark of all IBS Conferences, and to the many opportunities of meeting with colleagues from the constituent parts of the IBS that comprise the Central European Network. The establishment of the Central European Network, by the German and Austria-Swiss Regions and Polish National Group of the IBS, reflects a continued change in the demography of the society in Europe, and the perceived need to strengthen the underlying organization of the society. This new network follows closely behind the establishment of the Channel Network, by the Belgian, British and Irish, French and Netherlands Regions of the IBS in 2005, as a response to a similar perceived need. I believe that both developments are to the overall benefit of the IBS, and anticipate a further beneficial consolidation of Biometry and the IBS within Europe over the next few years. Indeed, similar consolidation activities are being considered in other parts of the IBS world, with current discussions about both a 'South-East Asia' Network and a 'Spanish-Speaking' Network, linking Europe with South and Central America, with all of these activities following in the successful foot-steps of SUSAN, the Sub-Saharan African Network.

As the IBS develops and moves forward with these plans, it is important to recognise and retain the diversity of the society, as seen through the individual characters of each Region or National Group, whilst benefiting from the

strength provided through the linking of different components of the society within a network. In particular, the Central European Network can provide a focus for the development of the IBS into parts of Eastern Europe which have either had a minimal IBS presence in the past, or, like the Hungarian National Group, have had a variable presence. I look forward to the opportunity during my Presidency of working with the members of the Central European Network to broaden the geographical footprint of the society. These networks also provide opportunities for individual members of the IBS working in particularly specialised fields, to interact with a wider group of biometricians than those within their own Region or National Group. I would encourage all IBS members to take advantage of the great breadth of biometrical knowledge within our society, and to attend meetings and conferences organised by the IBS throughout the world - since being elected President-Elect I have been fortunate enough to be able to attend five IBS conferences in different parts of the world, and to experience the wide diversity of biometry within our society.

Of course, the diversity of the society will be showcased at our next International Biometric Conference in Dublin, 13-18 July. I find myself in the unusual position of presiding at a conference that I am also responsible for organising. Having just reached the deadline for contributed papers and posters, I am confident that the conference will be a scientific success, and I hope that I will see many of you in Dublin in July.

Congratulations to the founders of the Central European Network on this timely development, and to the organisers of this first Central European Network Conference for what I am sure will be a highly successful conference.

Munich,
March 2008

Andrew Mead
(IBS president)

Greeting of the Local Organization

Dear Friends, dear Colleagues,

The LIFESTAT 2008 conference will consolidate the three regions of the International Biometric Society (IBS) - Poland, Germany, Switzerland/Austria - into the Central European Network for the first time.

This new consortium will open up for further integration of other smaller regional groups of IBS (Czech Republic, Slovakia, Romania, Hungary, etc.) and will thus strengthen the potential for development of Biostatistics in Eastern Europe. The conference will foster the exchange of ideas, interests, research and industrial statisticians between Eastern Europe and the German-speaking scientific community.

The organizers chose "Statistics and Life Sciences" as a motto for this first conference of the Central European Network, since molecular life sciences, medicine and epidemiology are the framing, trendsetting subjects, which cannot be appropriately compiled or scientifically interpreted without statistical methodology.

For today's statisticians new contexts emerge from cooperation with the life scientists and stimulated by this dialog, changes in thinking in statistical methods and development take place.

The interaction between scientific and commercial organizations in the life sciences brings to our field a dynamism which stimulates an attractive occupational picture of the biostatistician.

Munich, with its scientific industrial infrastructure, is an ideal venue for discussions about Statistics and Life Sciences. The integration of our colleagues from East Europe in the formative processes and structures is an essential

integration step in the process of professional development of new statistical thinking.

European scientists from life sciences and statistics must position themselves creatively to realize our potential to formulate our own unique vision and to compete successfully in the world and the USA. The strengthening of the European position with our own unique creativity, has become a necessity also as a result of economic developments in India and China.

Recent events have led to a strengthening of European contributions in statistics and life sciences to our society and to the international biometrical community. For example, in the spring of 2007 channel Networks were created in which British, French, Dutch and Belgian statisticians from the IBS united into a more focused federation.

The establishment of the Central European Network should facilitate the environment for other European networks including the integration of our East European colleagues in the existing infrastructure and new infrastructures originating from common projects in epidemiology, public health, ecology, agriculture, environment protection and health care.

As we move toward shaping a new European trend in Statistics and Life Sciences, our vision for this first conference of the Central European Network, LIFESTAT 2008, is to consciously foster and strengthen personal working connections with our colleagues from East Europe. This will open new communication channels and facilitate more efficient and better representative organisational structures for our Central European Network and study groups.

We wish you stimulating days in Munich.

Munich,
March 2008

Ulrich Mansmann
Gerhard Tutz
(local organizers)

Organizers and Invited Speakers

Sessions and their organizers:

Systems Biology and Bioinformatics, Diagnostic studies, Genetic Epidemiology (I), Mapping Approaches in Plants and Animals (Invited session organized by G. Freyer (Dummerstorf), AGs Landwirtschaftliches Versuchswesen & Populationsgenetik und Genomanalyse), Adaptive Group Sequential Designs, Biostatistic education at universities of applied sciences, Genome-wide association studies (I) (Invited session organized by I. R. König (Lübeck)), Multiple Testing, Non-Inferiority Trials (I) (Invited session organized by J. Röhmel (Bremen)), Freie Themen (I), Non-Inferiority Trials (II) (Invited session organized by A. Munk (Göttingen)), Freie Themen (II), Event Data Analysis (I), Genetic Epidemiology (II), Flexible designs (Invited session organized by W. Brannath (Vienna)), Agricultural science (I), Event Data Analyses (II), Clinical studies (I), Non-Inferiority Trials (III) (Invited session organized by A. Munk (Göttingen) and J. Röhmel (Bremen)), Genome-wide association studies (II), Advances in Statistical Modelling (I), Agricultural science (II) (Invited session organized by H. P. Piepho (Hohenheim)), Adaptive seamless design for combining phase II / III clinical studies (Invited session organized by W. Maurer (Basle)), Spatial Analysis of Surveillance Data (Invited Session organized by J. Dreesmann (Hannover) AG Räumliche Statistik), Young Statistician Awards (IBS-DR), Structuring high-dimensional data, Clinical studies (II), Advances in Statistical Modelling (II), Boosting for biomedical data (Invited session organized by P. Bühlmann (Zurich)), Advances in Statistical Modelling (III), Hierarchical models (Invited session organized by J. König (Mainz) AG Bayes-Methodik), Bayesian models in biostatistics (I), Bayesian models in biostatistics (II) (Invited session organized by L. Held (Zurich)), Nonparametric and Parametric Multivariate Tests in High Dimensions (Invited session organized by E.

Brunner (Göttingen) AG Nichtparametrische Methoden), Young Statistician Papers (IBS-ROeS), Clinical studies (III), Biological Networks (Invited session organized by A. Benner (Heidelberg) and A. Tresch (Mainz) AGs Statistical Computing und Bioinformatik), Meta-analysis, Additivity tests for mixed models, Cost efficient designs for biostatisticians (Invited session organized by H. Dette (Bochum)), Meta-analysis and Meta-regression (Invited session organized by R. Bender (Köln) AG Statistische Methoden in der Medizin), Statistical Methods in Environmental Monitoring

Invited speakers of the conference:

Fadoua Balabdaoui (Paris, France), Arne Bathke (Lexington, USA), Tim Beissbarth (Heidelberg, Germany), Werner Brannath (Vienna, Austria), Peter Bühlmann (Zurich, Switzerland), Leo Dempfle (Freising, Germany), Holger Dette (Bochum, Germany), David B. Dunson (USA), Rohan Luigi Fernando (Ames, USA), Tim Friede (UK), Leonhard Held (Zurich, Switzerland), Torsten Hothorn (Munich, Germany), Jürg Hüsler (Bern, Switzerland), H. M. James Hung (Silver Spring, USA), Meinhard Kieser (Heidelberg, Germany), Armin Koch (Bonn, Germany), Peter Kraft (Cambridge, USA), Inke R. König (Lübeck, Germany), Jürgen Läuter (Magdeburg, Germany), Michael Branson (Basel, Switzerland), Andrew Mead (Warwick, UK), Matthias Mielke (Göttingen, Germany), Bertram Müller-Myhsok (Munich, Germany), Martin Posch (Vienna, Austria), Michael Proschan (USA), William F. Rosenberger (Fairfax, USA), Joachim Röhmel (Bremen, Germany), Stephen Senn (Glasgow, UK), Christian Sonesson (Mölndal, Sweden), Simon G. Thompson (Cambridge, UK), Sue-Jane Wang (Silver Spring, USA), Gernot Wassmer (Cologne, Germany)

Local organizing committee:

Ursula Gerhardinger, Manuela Hummel, Florian Leitenstorfer, Ulrich Mansmann, Katharina Mildner, Karin Rieger, Markus Schmidberger, Karin Schmidt, Gerhard Tutz, Brigitte Weber, Andrea Willnhammer.

Contents

Online

Online

Part I

Abstracts for talks

Online

The Modular Structure Underlying Analyzes of Gene Set Enrichment

Marit Ackermann[1] and Korbinian Strimmer[2]

[1] Department of Statistics, Technical University of Dortmund, Germany; marit.ackermann@googlemail.com
[2] Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Germany; strimmer@uni-leipzig.de

Abstract: Enrichment analyzes explore whether functional groups of genes are "enriched" for differential expression between two or more phenotypes of interest. An advantage compared to the analysis on the individual gene level is the increase in power because of the smaller number of tests. Consequently, gene set enrichment analysis allows to discover new interesting pathways or other functional gene sets that might have an influence on the process under study.

Correspondingly, in the last few years a large number of procedures for conducting enrichment analysis have been suggested (see, e.g., Goeman and Bühlmann 2007). In this talk we review the most widely used procedures. Furthermore, we identify and discuss a common core and its modular structure. As a result, this not only facilitates a better understanding of the principles underlying gene set enrichment but also allows to choose the optimal analysis strategy for the individual problem at hand. We illustrate our approach with some data examples and a computer simulation.

References:

GOEMAN, J. J. and BÜHLMANN, P. (2007): Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics, 23, 980–987.*

Session:

S38: Biological Networks : Thursday, 13/03/2008, 9:10am - 10:30am

Multivariate analysis of high dimensional data: the one-sample statistic

M. Rauf Ahmad and Edgar Brunner

Department of Medical Statistics, University of Göttingen, Germany;
rahmad@gwdg.de, brunner@ams.med.uni-goettingen.de

Abstract: We consider the one-sample multivariate case, where each of n independent individuals is measured for d treatments. With time courses in such data, the interest usually focuses on the analysis of profile curves (Ahmad *et al*, 2007). When $d < n$, there are classical procedures, like Hotelling's T^2, for the analysis.

Recently, the attention has been focused on the high dimensional case ($d > n$). There have been new procedures introduced in the last few years (Läuter, 2004; Srivastava, 2007). We propose a modified version of the ANOVA-type statistic, based on the Box-approximation. The test statistic and degrees of freedom are defined using bilinear forms. The statistic approximately follows a χ^2-distribution. The estimators are unbiased and consistent, if $n \to \infty$, independent of d. The statistic is robust to normality assumption and does not depend on the covariance structure.

Simulations demonstrate that even for $n = 10$, the statistic maintains the level, independent of d and also has a high power for increasing alternatives and the power increases with increasing d. The statistic is compared with other available competitors and the application of the statistic is also shown through numerical illustration.

References:

AHMAD, M. R., C. WERNER and E. BRUNNER (2007): Analysis of high dimensional repeated measures designs: The one-sample case. *Comp. Stat. and Data An.. Submitted.*

LÄUTER, J. (2004): Two new multivariate tests, in particular for a high dimension. *Acta et comm. univ. tartuensis de math., 8, 179-186.*

SRIVASTAVA, M. S. (2007): Multivariate theory for analyzing high dimensional data. *J. Japan Statist. Soc., 37(1), pp. 53-86.*

Session:

S27: Structuring high-dimensional data : Wednesday, 12/03/2008, 9:10am - 10:30am

Statistical evaluation of mesoscale photochemical pollution simulation results

Lakhdar Aidaoui[1], Stamatis Zoras[2], Abbes Azzi[1], Sofia Papalexiou[3], and Nasreddine Akermi[1]

[1] Faculté de Génie Mécanique, Université des Sciences et de la Technologie d'Oran (USTO), Algeria; laidaoui@gmail.com
[2] Laboratory of atmospheric pollution and environmental physics, Technological Education Institute (TEI) of West Macedonia, Kila, 50100 Kozani, Greece
[3] AIAS Engineering Ltd, D. Solomou 5, Kalamaria, 55134 Thessaloniki, Greece

Abstract: In part one of this work the Air Pollution Model TAPM was applied to simulate the pollutants transport and transformation over a medium sized city region, in northwestern Greece during a summer week period of 2006, as well as to assesses the influence of meteorology conditions, topography and the small scale physicochemical phenomena on the performance of mesoscale model simulation results. A quantitative evaluation of the model's performance is presented here in part two. Results of the meteorology and photochemical simulations are compared against observations with the use of appropriate statistical performance measures, which were the index of agreement IOA, Pearson's correlation coefficient r and root mean square error RMSE. The statistical tools used in this study helped also to show clearly the relationship between in one hand the pollutants dispersion or concentrations and meteorology parameters, and on the other hand the liaison between the deferent pollutants (Ozone, NO_2, NO) each one by the other

References:

BRUNELLI, U. and PIAZZA, V. and PIGNATO, L. and SORBELLO, F. and VITABILE, S. (2007): Two-days ahead prediction of daily maximum concentrations of SO2, O3, PM10, NO2, CO in the urban area of Palermo, Italy. *Atmospheric Environment 41. 2967-2995.*

PAPALEXIOU, S. and MOUSSIOPOULOS, N. (2006): Wind flow and photochemical air pollution in Thessaloniki, Greece. Part II: Statistical evaluation of European Zooming Model's simulation results. *Environmental Modelling & Software 21. 1752-1758.*

WILMOTT, C.J. and ACKLESON, S.G. and DAVIS, R.E. and FEDDEMA, J.J. and KLINK, K.M. and LEGATES, D.R. and O'DONNELL, J. and ROWE, C.M. (1985): Statistics for the evaluation and comparison of models. *Journal of Geophysical Research 90, 8995e9005.*

Session:
S43: Statistical Methods in Environmental Monitoring : Thursday, 13/03/2008, 11:00am - 12:40pm

A Simulation Study to Assess the Bias of the Variance Estimator of the Multivariate Nelson–Aalen Estimator in Multistate Models

Arthur Allignol[1,2], Jan Beyersmann[1,2], and Martin Schumacher[2]

[1] Freiburg Centre for Data Analysis and Modelling, University of Freiburg, Eckerstraße 1, 79104, Freiburg, Germany;
arthur.allignol@fdm.uni-freiburg.de
[2] Institute of Medical Biometry and Medical Informatics, University Medical Centre Freiburg, Stefan–Meier–Straße 26, 79104 Freiburg, Germany.

Abstract: The multivariate Nelson–Aalen estimator of cumulative transition hazards is the fundamental nonparametric estimator in event history analysis. Two variance estimators were derived. Klein (1991) investigated the small sample properties of these two variance estimators in the univariate setting, and found through a simulation study that one tends to overestimate the true variance, in contrast to the other estimator which was found to be biased downward, especially when only few individuals are at risk. However, the estimators have never been investigated in the multivariate setting, where risk sets are more likely to be small. Therefore, we conduct a simulation study in this setting, considering several multistate models and transition hazard shapes, to compare the performance of the variance estimators. We propose to improve on variance estimation with the bootstrap and compare it to both variance estimators.

References:

ANDERSEN, P.K., BORGAN, O., GILL, R.D. and KEIDING, N. (1993): Statistical Models Based on Counting Process. *Springer Verlag, New York.*

KLEIN, J.P. (1991): Small Sample Moments of Some Estimators of the Variance of the Kaplan–Meier and Nelson–Aalen estimators. *Scandinavian Journal of Statistics, 18, 333–340.*

Session:
S17: Event Data Analyses (II) : Tuesday, 11/03/2008, 11:00am - 12:40pm

The asymptotic distribution of the likelihood ratio statistic under local and fixed alternative with application in non-inferiority trials

Fadoua Balabdaoui[1,2]

[1] Institut für Mathematische Stochastik, Georg-August Universität Göttingen, Maschmuehlenweg 8-10, D-37073 Göttingen, Germany
[2] Ceremade, Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris CEDEX 16 - France; fadoua@ceremade.dauphine.fr

Abstract: Over the past few years, there has been an increasing need to consider clinical non-inferiority trials with more than two groups. This imposes new challenges when the main goal is to determine the optimal sample sizes of enrolled patients which yield the maximal power. We consider here the general scope of k treatment groups, with sizes $n_1, ..., n_k$ and $k \geq 2$ is an integer. Furthermore, we assume that for each treatment $j \in \{1, ..m_k\}$ that the corresponding measured responses $X_{j1}, ..., X_{jn_j}$ are from a common density $f_j = f(\cdot, \Theta_j)$ with respect to a dominating σ-finite measure.

The Likelihood Ratio (LR) test is considered in two different testing problems. (a) The k-th treatment is not relevantly inferior than any of the other treatments. (b) The k-th treatment is not relevantly inferior than at least one of the other treatments. We will refer to (a) and (b) as the intersection and union case. Under fixed and local alternatives, we establish the limiting distribution of the LR statistic and determine the optimal sample sizes yielding the maximal power of the test. Particular examples including binomial and normal models will be treated.

This is a joint work with Matthias Mielke and Axel Munk.

References:

FEDER, P. I. (1968): On the distribution of the true log likelihood ratio test statistic when the true parameter is 'near' the boundaries of the hypothesis regions. *The Annals of Mathematical Statistics, 6, 2044-2055.*

MUNK, A., MIELKE, M., SKIPKA, G. and FREITAG, G. (2007): Testing noinferiority in three-armed clinical trials based on the likelihood ratio statistics. *The Canadian Journal of Statistics (to appear).*

WHITE, H. (1982): Maximum likelihood estimation of misspecified models. *Econometrica, 50, 1-26.*

WILKS, S.S. (1938): The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics, 9, 60-62.*

Session:

S09: Non-Inferiority Trials (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Nonparametric and Parametric Multivariate Tests

Arne Bathke[1] and Solomon Harrar[2]

[1] Department of Statistics, 875 Patterson Office Tower, University of Kentucky, Lexington, Kentucky 40506-0027, USA; arne@uky.edu
[2] Department of Mathematical Sciences, University of Montana, Missoula, Montana 59812-0864, USA; harrar@mso.umt.edu

Abstract: Multivariate data appear naturally in the life sciences because usually more than one response variable is of interest (multiple endpoints). We look at different ways to analyze multivariate data. Methodological advancements in recent years have made it possible to analyze many multivariate data sets using completely nonparametric methods. We will provide an overview of current statistical research on multivariate methods, illustrated by examples. In particular, we will focus on inference methods that are invariant under monotone transformations of the different variables, and that can be used when the data have a mix of ordinal and quantitative response variables.

References:

BATHKE, A.C. and HARRAR, S.W. (2008): Nonparametric Methods in Multivariate Factorial Designs for Large Number of Factor Levels. *Journal of Statistical Planning and Inference, 138(3), 588-610.*

BATHKE, A.C. and LANKOWSKI, D. (2005): Rank Procedures for a Large Number of Treatments. *Journal of Statistical Planning and Inference 133(2), 223-238.*

HARRAR, S.W. and BATHKE, A.C. (2008): Nonparametric Methods for Unbalanced Multivariate Data and Many Factor Levels. *Journal of Multivariate Analysis, to appear.*

MUNZEL, U. and BRUNNER, E. (2000): Nonparametric Methods in Multivariate Factorial Designs. *Journal of Statistical Planning and Inference, 88(1), 117-132.*

MUNZEL, U. and BRUNNER, E. (2000): Nonparametric Tests in the Unbalanced Multivariate One-Way Design. *Biometrical Journal, 42(7) 837-854.*

Session:

S35: Nonparametric and Parametric Multivariate Tests in High Dimensions : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Predicting Breeding Values of a Parental Population of Self-Pollinated Crops

Andrea M. Bauer and Jens Léon

Institute of Crop Science and Resource Conservation, Chair of Plant Breeding, University of Bonn, D-53115 Bonn, Germany; a.bauer@uni-bonn.de, j.leon@uni-bonn.de

Abstract: Selecting parental lines that are characterized by a superior genotypic value is an essential requirement in breeding self-pollinating crops. Currently, in breeding programs data is obtained from several sources: observations from field experiments, molecular marker data and pedigree information. Though, still plant breeders do not use all these information in their selection decision. It is well-known that a selection based on the individual performance is less efficient than using additionally information of related individuals in the selection process. The prediction of breeding values by best linear unbiased prediction (BLUP) using pedigree information has been found to be superior to commonly-used selection strategies. However, often pedigree data is incomplete or inaccurate. Thus, in our research genetic similarities, which were calculated based on DNA marker data, were considered in the prediction (Bauer et al. 2006). Additionally, we performed a singular value decomposition of genetic similarities if, due to a low number of DNA markers, lines were considered to be genetically identical (Bauer et al. 2008). In a further study, uni- and multivariate predictions of parental breeding values were compared as often agronomically important traits are correlated among each other (Bauer and Léon 2008). In this contribution, results of our research studies will be presented and discussed.

References:

BAUER, A.M., REETZ, T.C. and LÉON, J. (2006): Estimation of breeding values of inbred lines using best linear unbiased prediction (BLUP) and genetic similarities. *Crop Science, 46, 2685–2691.*

BAUER, A.M., REETZ, T.C. and LÉON, J. (2008): Predicting breeding values of spring barley accessions by using the singular value decomposition of genetic similarities. *Plant Breeding (in press).*

BAUER, A.M. and LÉON, J. (2008): Multiple-trait breeding values for parental selection in self-pollinating crops. *Theoretical and Applied Genetics, 116, 235–242.*

Session:

S16: Agricultural science (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Estimating Signaling Networks through Nested Effects Models

Holger Froehlich, Mark Fellmann, Holger Sueltmann, and Tim Beissbarth

DKFZ, INF 580, Heidelberg, Germany; t.beissbarth@dkfz-heidelberg.de

Abstract: In the modern field of systems biology scientists aim to get insights into the architecture and behavior of complex cellular and genomic processes. An important task in this context is the detection of novel interdependencies between gene products. This insight into the molecular networks is an important step towards a better understanding of the functional aspects of a biological system. The advent of RNA interference techniques enables the selective silencing of biologically interesting genes in an efficient way. The combination of targeted interventions using the RNA interference technique with measuring effects on gene expression by DNA microarrays thus enables researchers to gain insights into the signal flow between proteins in a cell based on the observation of downstream effects. For example, in a signaling pathway that activates several transcription factors, blocking an upstream element of the pathway will affect all transcription factor targets, while perturbing one of the downstream transcription factor will only affect its targets, which are a subset of the genes effected by blocking the complete pathway. Markowetz et al. have proposed Nested Effect Models as a statistical framework for scoring networks hypotheses in a Bayesian manner.

We will show extensions of that framework that go in several directions: We show how prior assumptions on the network structure can be incorporated into the scoring scheme by defining appropriate prior distributions on the network structure as well as on hyperparameters. A new approach called module networks is introduced to scale up the original approach, which is limited to around 5 genes, to infer large scale networks. We compare several heuristic approaches for their performance in terms of sensitivity, specificity and speed. Instead of the data discretization step needed in the original framework, we propose the usage of a beta-uniform mixture distribution on the p-value profile, resulting from differential gene expression calculation, to quantify effects. Extensive simulations on artificial data and application of our module network approach to infer the signaling network between 13 genes in the ER-alpha pathway in human MCF-7 breast cancer cells show that our approach gives sensible results. Using a bootstrapping approach this reconstruction is found to be statistically stable. The code for the module network inference method is available in the latest version of the R-package nem, which can be obtained from the Bioconductor homepage.

Session:
S38: Biological Networks : Thursday, 13/03/2008, 9:10am - 10:30am

Time depending dose response modeling

Norbert Benda

Novartis Pharma AG Basel, Statistical Methodology, Switzerland;
norbert.benda@novartis.com

Abstract: The goal of a clinical Phase II dose finding study is to describe the dose response relationship and to find a target dose or dose range that ensures a certain efficacy. In many applications, however, it is useful to consider combinations of dose and treatment duration instead of the dose only. The estimation of a minimum effective dose as a function of time allows for a decision on the optimal treatment duration.

Bretz et al. proposed a methodology that combines formal hypothesis testing for dose response with flexible modeling of the dose response relationship and estimating a target dose.

In this paper an extension of this methodology is proposed that takes into account both, dose and treatment duration, based on a clinical study with repeated binary data. A set of nonlinear mixed effects models is considered. The minimum effective dose as a function of treatment duration is estimated after selection of a model.

Operational characteristics, as precision of the target dose estimation are evaluated under different scenarios to compare different design settings.

The proposed methodology is illustrated with an example on the treatment of psoriasis.

References:

BRETZ, F., PINHEIRO, J. and BRANSON, M. (2005): Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics, 61, 738–748.*

PINHEIRO, J., BORNKAMP, B. and BRETZ, F. (2006): Design and analysis of dose finding studies combining comparison and modeling procedures. *Journal of Biopharmaceutical Statistics, 16, 639-656.*

Session:
S37: Clinical studies (III) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Application of the Population Impact Number (PIN) to Describe the Effect of Interventions in the Entire Community

Ralf Bender and Ulrich Grouven

Institute for Quality and Efficiency in Health Care, Dillenburger Str. 27, D-51105 Cologne, Germany; Ralf.Bender@iqwig.de, Ulrich.Grouven@iqwig.de

Abstract: The risk difference and its inverse, the number needed to treat (NNT), are frequently used to describe the absolute effects of treatments in randomized controlled trials (BENDER, 2005). Recently, the population impact number (PIN) was proposed to describe the effect of interventions in the entire community (HELLER and DOBSON, 2000). However, only methods to perform point estimation of PINs were developed. In a subsequent article, the use of sensitivity analyses was proposed as substitution of confidence intervals (HELLER et al., 2002). In this paper, a simple method to calculate confidence intervals for PINs is proposed, which is based on interval estimation methods for NNTs (BENDER, 2001). The features of NNTs and PINs as effect measures for population-wide interventions are compared and discussed. For illustration, the methods are applied to German quality assurance data to quantify the effect of minimum provider volumes concerning total knee replacement.

References:

BENDER, R. (2001): Calculating confidence intervals for the number needed to treat. *Controlled Clinical Trials, 22, 102-110.*

BENDER, R. (2005): Number needed to treat (NNT). *In: P. Armitage & T. Colton (Eds.): Encyclopedia of Biostatistics, 2nd Ed. Wiley, Chichester, 3752-3761.*

HELLER, R.F. and DOBSON, A.J. (2000): Disease impact number and population impact number: Population perspectives to measures of risk and benefit. *British Medical Journal, 321, 950-952.*

HELLER, R.F., DOBSON, A.J., ATTIA, J. and PAGE, J. (2002): Impact numbers: Measures of risk factor impact on the whole population from case-control and cohort studies. *Journal of Epidemiology and Community Health, 56, 606-610.*

Session:
S12: Freie Themen (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

The impact of time-dependent bias in proportional hazards modelling

Jan Beyersmann[1,2], Martin Wolkewitz[1,2], and Martin Schumacher[2]

[1] Freiburg Centre for Data Analysis and Modelling, University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany; jan.beyersmann@fdm.uni-freiburg.de
[2] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Stefan-Meier-Straße 26, 79104 Freiburg, Germany

Abstract: In the clinical literature, time-dependent exposure status has regularly been analysed as if known at time origin (VAN WALRAVEN et al. 2004). Although statisticians agree that such an analysis yields biased results when analysing the effect on the time until some endpoint of interest, this work is the first to study in detail the bias arising in a proportional hazards analysis. We show that the biased hazard ratio estimate will always be less than the unbiased one; this leads to either an inflated or a damped effect of exposure, depending on the sign of the correct log hazard ratio estimate. We find an explicit formula of the asymptotic bias based on generalized rank estimators (SCHUMACHER et al. 1987), and we investigate the role of censoring, which may prevent an individual from being considered as baseline exposed in the biased analysis. We illustrate our results with data on hospital infection status and different, artificially introduced censoring patterns.

References:

SCHUMACHER M., OLSCHEWSKI M. and SCHMOOR C. (1987): The impact of heterogeneity on the comparison of survival times. *Statistics in Medicine, 6, 773–784.*

VAN WALRAVEN C., DAVIS D., FORSTER A. and WELLS G. (2004): Time-dependent bias was common in survival analyses published in leading clinical journals. *Journal of Clinical Epidemiology, 57, 672–682.*

Session:

S13: Event Data Analysis (I) : Tuesday, 11/03/2008, 9:10am - 10:30am

Mandatory Covariates in High-Dimensional Survival Models Estimated by Boosting

Harald Binder[1,2] and Martin Schumacher[2]

[1] Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany; binderh@fdm.uni-freiburg.de
[2] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; ms@imbi.uni-freiburg.de

Abstract: While there are several techniques for the fitting of predictive high-dimensional survival models by penalized parameter estimation or by boosting, none allows for explicit consideration of mandatory covariates, such as clinical scores. We therefore introduce a new boosting algorithm for censored time-to-event data, that estimates Cox proportional hazards models, employing an offset-based update mechanism in combination with componentwise updates of the linear predictor. This results in sparse model fits, similar to fits obtained from Lasso-like estimation under a L_1 constraint on the parameter vector. However, the offset-based update mechanism also allows for tailored penalization of the covariates under consideration. Specifically, unpenalized mandatory covariates can be introduced. Furthermore, we explore approaches for reducing the bias of parameter estimates, seen for estimation under L_1 constraint and conventional componentwise boosting. Microarray survival data from patients with diffuse large B-cell lymphoma, in combination with the recent, bootstrap-based prediction error curve technique, is used to illustrate the advantages of the new procedure. It is demonstrated that it can be highly beneficial in terms of prediction performance to use an estimation procedure that incorporates mandatory covariates into high-dimensional survival models. The new approach also allows to answer the question, whether improved predictions are obtained by including microarray features in addition to classical clinical criteria.

Session:
S30: Boosting for biomedical data : Wednesday, 12/03/2008, 11:00am - 12:40pm

Knot Removal for Judging the Stability of Additive Spline Model Fits

Harald Binder[1,2] and Willi Sauerbrei[2]

[1] Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany; binderh@fdm.uni-freiburg.de
[2] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; wfs@imbi.uni-freiburg.de

Abstract: Replication stability, i.e., stability of fitted model components with respect to perturbations in the data, is an important aspect for judging the usefulness of statistical models. This is especially important for flexible techniques, such as generalized additive models, that allow for a smooth influence function for each single covariate, typically fitted by spline techniques. However, for the latter only few techniques for stability analysis are available. We therefore investigate existing tools for stability analysis based on bootstrap samples, which were developed for the multivariate fractional polynomial approach, and apply them to the generalized additive model setting. Furthermore, as the focus is on model fits based on B-splines, knot removal techniques are available. It is shown how these can be employed for simplifying the difficult task of stability analysis, providing more insight into the stability of local features fitted in bootstrap samples. For this purpose, the bootstrap result matrix is analyzed via log-linear models. The proposed techniques are illustrated in an application example with children's lung function data.

References:

BINDER, H. and SAUERBREI, W. (2007): Shape-Preserving Simplification of Additive Spline Models by Knot Removal. *FDM-Preprint 99, University of Freiburg.*

ROYSTON, P. and SAUERBREI, W. (2003): Stability of Multivariable Fractional Polynomial Models With Selection of Variables and Transformations: a Bootstrap Investigation. *Statistics in Medicine, 22, 639–659.*

Session:

S21: Advances in Statistical Modelling (I) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Bayesian Nonparametric Estimation of Monotone Functions

Björn Bornkamp and Katja Ickstadt

Fakultät Statistik, Technische Universität Dortmund, Germany;
{bornkamp},{ickstadt}@statistik.uni-dortmund.de

Abstract: In this talk we consider monotone nonparametric regression in a Bayesian framework. Motivated by traditional nonlinear models for dose-response analysis the monotone function is modelled as a shifted and scaled probability distribution function. Due to its interpretability the model facilitates the incorporation of prior information on different aspects of the curve. The underlying probability distribution function is modelled nonparametrically as a discrete mixture of parametric distribution functions, and a general random probability measure is assumed as a prior for the discrete mixing distribution (Ongaro and Cattaneo, 2004). Additionally we investigate the choice of the base parametric distribution function and show that the two-sided power distribution function (van Dorp and Kotz, 2002), compared to the more common choice of a beta distribution function, is well suited for this purpose. It can be evaluated quickly, allowing a fast calculation of the likelihood, as needed for Markov Chain Monte Carlo algorithms. In addition we can show that any continuous probability distribution function can be approximated by a mixture of two-sided power distribution functions.

We illustrate our methodology on a data set from dose-response analysis and compare it in a simulation study with other recent approaches to monotone nonparametric regression (Dette et al. 2006, Wood 2007).

References:

DETTE, H., NEUMEYER, N. and PILZ, K. (2006): A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli, 12, 469–490.*

ONGARO, A. and CATTANEO, C. (2004): Discrete random probability measures: A general framework for nonparametric Bayesian inference. *Statistics and Probability Letters, 67, 33–45.*

VAN DORP, J. and KOTZ, S. (2002): The standard two-sided power distribution and its properties: with applications in financial engineering. *The American Statistician, 56, 90–99.*

WOOD, S. (2007): `mgcv` R-package.

Session:
S18: Clinical studies (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Additional predictive power of microarray data and hybrid classifiers

Anne-Laure Boulesteix[1], Christine Porzelius[1,2], and Martin Daumer[1]

[1] Sylvia Lawry Centre for Multiple Sclerosis Research, D-81677 Munich, Germany; boulesteix@slcmsr.org, daumer@slcmsr.org
[2] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, D-79104 Freiburg, Germany, cp@fdm.uni-freiburg.de

Abstract: In the context of class prediction, statistical methods are needed for assessing the additional predictive power of high-dimensional microarray data compared to simple clinical parameters alone. While, according to Ntzani and Ioannidis (Lancet 2003, 362:1439–1444), "adjustment for other classic predictors of the disease outcome [is] essential", the step is often performed using sub-optimal methods and not adequately described in medical literature. Such clinical parameters, which include, e.g., age and sex of the patient, disease duration, relapse rate or tumor grade are usually much easier and cheaper to collect than microarray data.

A critical study of breast cancer outcome prediction (Eden et al., Eur J Cancer 2004, 40:1837-1841) suggests that "good old clinical markers have similar power in breast cancer prognosis as microarray [...] profilers". We found similar results when we evaluated the additional predictive value of MRI for the prediction of MS relapses (Daumer et al., Multiple Sclerosis 2006, 12:S46-S47). The question of the additional predictive power of microarray predictors must be addressed fairly, i.e. without favouring microarray data through overfitting mechanisms. Ideally, such methods should also provide an optimal prediction rule making use of all potentialities of both types of data: in particular, they should be able to catch subtypes which are not identified by clinical parameters alone.

We propose a novel two-step approach based on PLS dimension reduction and random forests embedding the idea of pre-validation suggested by Tibshirani and Efron (Stat Appl Genet Mol Biol 2002, 1:1). This approach is fast, flexible and can be used both for assessing the overall significance of the microarray predictors and for building improved hybrid classification rules. It can also be easily generalized to other prediction problems (regression, survival). The new method is illustrated through simulations and an application to breast cancer data.

Session:
S01: Systems Biology and Bioinformatics : Monday, 10/03/2008, 9:10am - 10:30am

Estimation following treatment selection in flexible designs

Werner Brannath

Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria;
werner.brannath@meduniwien.ac.at

Abstract: We discuss the estimation of treatment effects at the end of a flexible clinical trial where treatments are selected and sample sizes are reassessed at an adaptive interim analysis. Parameter estimation at the end of a flexible design has been considered as difficult and yet no satisfactory solution has been proposed. We therefore consider flexible design where single-step adjusted p-values are used for the multiplicity adjustment and where the construction of simultaneous confidence intervals is possible. Like in the fixed size sample case the single-step designs are expected to be somewhat less powerful than the commonly used step-down designs – a price that one seem to have to pay for the construction of simultaneous confidence intervals. We will discuss and quantify the statistical properties of the resulting confidence intervals and will compare the single-step flexible designs to the usual step-down flexible designs in a simulation study.

Session:
S23: Adaptive seamless design for combining phase II / III clinical studies : Tuesday, 11/03/2008, 4:00pm - 6:10pm

On time-varying effects in survival analysis

Anika Buchholz[1,2], Willi Sauerbrei[2], and Patrick Royston[3]

[1] Freiburger Zentrum für Datenanalyse und Modellbildung, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany; ab@fdm.uni-freiburg.de
[2] Institut für Medizinische Biometrie und Medizinische Informatik, Universitätsklinikum Freiburg, 79104 Freiburg, Germany
[3] MRC Clinical Trials Unit, 222 Euston Road, London NW1 2DA, UK

Abstract: The Cox proportional hazards (PH) model is the standard tool for the analysis of survival time data. With long-term survival effects of covariates may vary in time, which means that the PH assumption is violated.

Several new procedures for modelling time-varying effects have been proposed, such as the MFPT approach (Sauerbrei et al. 2007) based on fractional polynomials (FP). Others include a further FP based approach (Berger et al. 2003; Stat Med 22:1163–80), reduced rank models (Perperoglou et al. 2006; Stat Med 25:2831–45), and a semiparametric approach based on cumulative regression functions (Scheike and Martinussen 2004; Scand J Stat 31:51–62).

In the Rotterdam breast cancer series (N=2982, 1518 events, 20 years follow-up) the first two steps of MFPT selected a PH model including 8 variables. Using the different approaches we will investigate for this model whether the effect of any of the 8 variables varies in time.

References:

SAUERBREI, W., ROYSTON, P. and LOOK, M. (2007): A new proposal for multivariable modelling of time-varying effects in survival data based on fractional polynomial time-transformation. *Biometrical Journal, 49, 453–473.*

Session:

S13: Event Data Analysis (I) : Tuesday, 11/03/2008, 9:10am - 10:30am

Boosting for High-Dimensional Data: an Introduction

Peter Bühlmann

ETH Zentrum, 8092 Zürich, Switzerland; buhlmann@stat.math.ethz.ch

Abstract: 13 years ago, Freund and Schapire (1995) proposed in the machine learning community a first practical version of their famous AdaBoost algorithm for binary classification. The method has very quickly attracted much attention from many people, mainly because of its outstanding performance (at that time) on many fairly high-dimensional classification problems. Nowadays, we largely understand when and why boosting algorithms work well. This talk will give an introduction to generic boosting methods which can be widely used in e.g. generalized nonparametric regression, classification or survival models. We will explain a key property that boosting is a computational and efficient regularization method for potentially very high-dimensional data where the number of covariates (e.g. in the 10 thousands) can greatly exceed sample size (e.g. in the dozens). The concepts will be illustrated with a few examples from biology, and other talks in the session will specialize our general view of boosting to various specific applications.

Session:
S30: Boosting for biomedical data : Wednesday, 12/03/2008, 11:00am - 12:40pm

Sample sizes of multi-rater/multi-study experiments when developing a quality assessment tool in toxicology

Iris Burkholder[1], Annette Kopp-Schneider[1], Klaus Schneider[2], Sebastian Hoffmann[3] and Lutz Edler[1]

[1] DKFZ, D-69120 Heidelberg, Germany; i.burkholder@dkfz.de
[2] Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG), D-79098 Freiburg, Germany
[3] European Centre for the Validation of Alternative Methods (ECVAM), IHCP, JRC, European Commission, I-21027 Ispra, Italy

Abstract: Quality assessment of existing toxicological data must be subject to reliable assessment tools e.g. when many chemicals have to be assessed within the REACH program of the EU or when animal studies are supposed to be replaced by alternative in-vitro methods. To evaluate the quality of toxicological studies, for both in-vivo and in-vitro toxicological data, the so-called Klimisch criteria have been widely used. By addressing e.g. documentation, study design, toxicological and statistical methods used they applied a four point classification.

In an ongoing ECVAM project, we are developing a more comprehensive and more detailed assessment tool for toxicological data by identifying about 25 variables related to quality.

In this contribution, we will describe the process of developing an experimental design for a multi-rater - multi-study experiment based on the width of the 95% confidence interval of the kappa agreement coefficient. A graphical procedure for optimization the number of raters and the number of studies per rater is used. Endpoints of statistical analysis focus on the agreement coefficient kappa for multiple raters and are applied both to individual variables outcomes as well as summary scores.

References:

KLIMISCH, H.J. et al. (1997): Regulatory Toxicology and Pharmacology, 25, 1-5.

REACH: Registration, Evaluation, Authorisation and Restriction of Chemical substances. http://ec.europa.eu/environment/chemicals/reach/reach intro.htm.

Session:

S02: Diagnostic studies : Monday, 10/03/2008, 9:10am - 10:30am

Nonlinear Principal Component Analysis: an Application

Paulo Canas Rodrigues

Nova University of Lisbon and Agricultural University of Poznan, Poland;
paulocanas@fct.unl.pt

Abstract: Principal Component Analysis (PCA) (e.g. Pearson, 1901; Hotelling, 1933; Jolliffe, 2002) is one of the most popular techniques to analyze multivariate data. This technique looks for a few linear combinations which can be used to summarize the original data, losing in the process as little information as possible. The linear combinations (or principal components) are also uncorrelated.

The most immediate objective of PCA is to verify if exists a small number of the first principal components that explains a higher proportion of variance of the original data set. If it happens, a few principal components can be used to represent the original data set without a big loss of information. This procedure corresponds to the dimension reduction of the original data set.

Another point of PCA is the visualization of variables underlying the original structure (the principal components) which have physic meant and allows, therefore, to help to see the initial structure of another point of view.

PCA assumes that all the variables in the data set are numerical, or can be treated as numerical. But, when we have variables with different measurement levels we shouldn't use this method and we get in the Nonlinear Multivariate Analysis (Gifi, 1991; De Geer, 1993a; De Geer, 1993b) or, in particular the Nonlinear Principal Component Analysis (NLPCA), a better technique when we aim to reduce the dimensionality and find patterns of this kind of data.

In this work we present a comparison between the PCA and the NLPCA in a practical point of view. We present a data set with 233 patients with Heart Failure and a set of 17 variables in different measurement levels, from a Portuguese Central Hospital in the year of 2001, with a follow up of five years, in maximum.

We use the SPSS software (Meulman, 2004) to obtain the results for the NLPCA and we present a comparison between PCA and NLPCA, which produced very different results because of the wrong results obtained from the traditional method.

References:

DE GEER, J.P. (1993a): *Multivariate Analysis of Categorical Data, Vol. 2: Theory (Advanced Quantitative Techniques in the Social Sciences). SAGE Publications - Leiden University. The Netherlands.*

DE GEER, J.P. (1993b): *Multivariate Analysis of Categorical Data, Vol. 3: Applications (Advanced Quantitative Techniques in the Social Sciences).* SAGE Publications - Leiden University. The Netherlands.

GIFI, A. (1991): *Nonlinear Multivariate Analysis.* John Wiley & Songs - University of Leiden. The Netherlands.

HOTELLING, H. (1933): Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol., 24, 417–441.*

JOLLIFFE, I.T. (2002): *Principal Component Analysis*. Springer Verlag - New York Inc.

MEULMAN, J.J. and HEISER W.J. (2004): *SPSS Categories 13.0*. SPSS Inc - Chicago.

PEARSON, K. (1901): On lines and planes of closest fit to systems of points in space. *Phil. Mag. (6), 2, 559–572*.

Session:
S27: Structuring high-dimensional data : Wednesday, 12/03/2008, 9:10am - 10:30am

Performance of bagging and boosting depends on diversity of base classifiers

Małgorzata Ćwiklińska-Jurkowska[1]

Collegium Medicum, Nicolaus Copernicus University, Jagiellońska 13, 85-067 Bydgoszcz, Poland; mjurkowska@cm.umk.pl

Abstract: Diversity is an important factor for building exact family of classifiers. When a data set is highly dimensional, has small learning set compared to the number of variables, it is often difficult to build a good single classifying function. Such classifier is biased and has large variance and consequently – poor performance. In order to improve weak classifier by stabilizing its decision, the technique of regularization or noise injection has been developed. Another approach is combining classifiers into power decision rule. Different character of constituent classifiers is usually the reason of diversity of the ensemble.

Diversity of classifiers is important and desired feature for performance of combined classifiers. Bagging and boosting are ensemble of classifiers, for which diversity comes from resampling training sets. It is interesting if there is strong dependency of bagging or boosting performance on ensemble diversity for different base classifiers. Various measures of diversity and similarity of the ensemble constructed by resampling classifiers were applied. Standard medical data sets from uci repository were used. Kuncheva et al. (2002) examined this problem for linear classifiers. In the presented work the different, stable and nonstable, classifiers are examined.

The apparent (resubstitution) error, the leave one-out, cross validation and test errors generally are smaller for ensembles build on diverse classifiers, however for some data sets improvement by bagging or boosting is not substantial in comparison with some advanced constituent classifiers such as support vector machines.

Resamling ensemble methods are not beneficial for SVM. Support vector machines and boosting are both based on maximizing margins and in both methods idea is focused on objects difficult for classification

However for bagging ensemble in the case of big data sets the relationship can be quite opposite than it could be expected: bagging has performance negatively dependent on diversity. It can be explained by the fact that for such datasets the ensemble of very diverse classifiers contain probably worse classifiers. So examining diversity in predicting performance can be more useful for boosting than bagging.

References:

KUNCHEVA L.I., SKURCHINA M. and DUIN R.P.W. (2002): An experimental study on diversity for bagging and boosting with linear classifiers. *Information Fusion, 3(2), 245–258.*

SKURICHINA M., DUIN R. (2000): Boosting in Linear Discriminant Analysis. In: J. Kittler, F. Roli (eds.): *Multiple Classifier Systems. Springer-Verlag, Berlin, 190–199.*

Session:
S26b: Boosting for high-dimensional biomedical data : Wednesday, 12/03/2008, 9:50am - 10:30am

Discrimination between non-genotoxic and genotoxic chemical compounds based on cDNA microrray data

Małgorzata Ćwiklińska-Jurkowska[1], Tomasz Burzykowski[2], and Magdalena Wietlicka-Piszcz[1]

[1] Collegium Medicum, Nicolaus Copernicus University, Jagiellońska 13, 85-067 Bydgoszcz, Poland; mjurkowska@cm.umk.pl, mpiszcz@logonet.com.pl
[2] Center for Statistics, Hasselt University, Agoralaan – building D, 3590 Diepenbeek, Belgium; tomasz.burzykowski@uhasselt.be

Abstract: The aim of the study was to find a subset of 596 genes, suspected as possibly active after exposing to genotoxic chemical compounds, that would allow for a discrimination between genotoxic and non-genotoxic compounds. cDNA microarray measurements, representing responses for 20 different genotoxic and non-genotoxic chemicals observed in a dye-swap experiment were analyzed. The raw microrray data were obtained in the ImaGene format. The intensity values were corrected for background using a convolution of the normal and exponential distributions, where the normal part represents the background and the exponential represents the signal intensities [Smyth, 2005]. The within-array normalization of background-corrected intensities for chemical compounds vs. the control (solvents) was carried out. The linear model was fitted to the log ratios of intensities [Smyth, 2004]. The aim of the procedure was to adjust for the dye effect and to choose genes differentially expressed in GTX and NGTX groups.

The subsets of significant genes were obtained based on the linear model by using different methods for multiple-testing-adjustment of p-values. The subsets were subsequently used in different classification methods. In particular, classical LDA, diagonal LDA and diagonal QDA, SVM with different kernels and regularization constants, random forests, bagging and double-bagging trees with LDA and singular LDA, boosting with different modifications and bagboosting, were considered [Gentle&Hardle, 2005]. The misclassification errors were estimated using a test data set of 12 microarrays and cross-validation and compared. Baggbosting method with a limited subset of 15 genes was found to yield the smallest misclassification rates.

References:

SMYTH, G. K. (2005): Limma: linear models for microarray data. In: R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.): *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer, New York, 397–420.

SMYTH, G. K. and SPEED, T. P. (2003): Normalization of cDNA microarray data. *Methods, 31, 265–273.*

SMYTH G. K. (2004): Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology, vol. 3, no. 1, article 3.*

SMYTH G. K., MICHAUD. J. and SCOTT H. (2005): The use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics, 21(9), 2067–2075.*

GENTLE J.E., HARDLE W. and MORI Y. (2004): *Handbook of Computational Statistics: Concepts and Methods.* Springer, Berlin.

Session:

S26b: Boosting for high-dimensional biomedical data : Wednesday, 12/03/2008, 9:50am - 10:30am

Logistic Models for Extreme Risk Data

Andrzej Dabrowski

Department of Mathematics, Wrocław University of Environmental and Life Sciences, Poland; aludwikdabr@gmail.com

Abstract: In the classical binary response model the probability π of the occurrence of an event, conditional on a vector of covariates x and parameters vector β is expressed as $g(\pi) = x'\beta$. The function g links the linear predictor to the probability and determines the shape of the quantal response. Extreme risk data are described by models with link functions accounting for symmetrically distributed heavy tails in the latent variable model for binary response. Moreover the distributions are skewed. The approach of this paper is to compare different models for extreme risk data. The non canonical link functions, depending on the data, such as Gosset and Pregibon link functions, based on Box-Cox transformations will be studied. The performance of the alternative link functions are measured by assessing the accuracy of the estimated success probabilities. The deviance information criterion measure is used alternatively for guiding the choice of links The effectiveness of unknown parameters estimation by classical least square regression and non-standard methods, such as least absolute deviations, lasso regression are compared for the models of extreme risk data with different link functions.

References:

AGRESTI, A., (1990): *Categorical Data Analysis*, Wiley, New York.

BIRKES, D. and DODGE, Y. (1993): *Alternative Methods of Regression*, Wiley, New York.

CZADO, C. and MUNK, A. (2000): Noncanonical links in generalized linear models - when is the effort justified? *J. of Statistical Planning and Inference, vol 87, no 2, 317-345.*

FOKIANOS, K., PENG, A. and QIN, J.(2000): A Generalized Moments Specification Test for the Logistic Link. *The Canadian Journal of Statistics, 27, 735-750, 1999 with an addendum in Vol 28, 446, 2000.*

PREGIBON, D. (1981): Logistic regression diagnostics. *Ann. Statistics, 9, 705-724.*

Session:

S31: Advances in Statistical Modelling (III) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Variance Components - Estimation Methods and Applications

Leo Dempfle

Department of Animal Science, TU München, D-85350 Freising-Weihenstephan, Germany; Leo.dempfle@t-online.de

Abstract: Estimation of variance components has a long history dating back to the 19th century. However, for a long time methods were only available for balanced designs. Methods for dealing with unbalanced designs came much later and only in the fifties of last centuries methods were developed which could generally deal with unbalanced data in mixed models. These procedures were very useful at the time but they were still of an ad hoc nature fulfilling few optimality criteria (mainly unbiasedness). Then within a rather short time methods fulfilling several optimal properties (locally best quadratic unbiased) and Maximum Likelihood based procedures as well as Bayesian Methods were developed.

If we look at the most simple problem of estimating the variance in a simple random sample one can, depending on the desired criteria, choose among several estimators: Best unbiased, minimum mean squared error, Maximum likelihood and Bayesian procedure applying 'some' prior distribution. If there are two or more variance components only with balanced designs does there exist a best unbiased method. In the unbalanced case there are crossing risk functions. In addition there is the problem of conflicting optimality criteria. By definition a variance component is positive but unbiased estimators can, with one exception, not always be positive. In general unbiasedness might not be a very important criterion but as will be discussed non-negative estimators can have serious disadvantages in practical applications.

For the application inadmissible estimators (Anova-related methods) should be phased out, unbiasedness vs non-negativity must be carefully judged and good use of any available knowledge should be made in order to choose a method having good properties for the situation at hand. In Genetics eg we can sometime give a lower bound to the ratio of two variance components. Any prior knowledge is also extremely useful for designing efficient experiments for estimating variance components.

References:

SEARLE, S.R., CASELLA, G. and McCULLOCH, C.E (1992): *Variance Components*. Wiley, New York

Session:
S22: Agricultural science (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Optimal designs for dose finding studies

Frank Bretz[1], Holger Dette[2], Andrey Pepelyshev[3], and José Pinheiro[4]

[1] Biostatistics, Novartis Pharma AG, 4002 Basel, Switzerland;
frank.bretz@novartis.com
[2] Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany;
holger.dette@ruhr-uni-bochum.de
[3] St. Petersburg State Univ., Department of Mathematics, St. Petersburg, Russia;
andrey@ap7236.spb.edu
[4] Biostatistics, Novartis Pharmaceuticals, East Hanover NJ, USA;
jose.pinheiro@novartis.com

Abstract: Identifying the right dose is a key goal in the development of any medicinal drug. Its importance cannot be understated: selecting too high a dose can result in an unacceptable toxicity profile while selecting a dose that is too low increases the likelihood that the compound provides insufficient evidence of effectiveness. Hence, dose finding studies are of crucial importance in drug development and need to be planned carefully.

The primary goal of a typical clinical dose finding study is to estimate the minimum effective dose (MED), i.e., the smallest dose, which shows a clinically relevant and a statistically significant effect. Let a clinical relevance threshold Δ and the dose range $[x_1, x_k]$ under investigation be given, where x_1 is often the placebo and x_k the maximum tolerated dose. A common problem is then to derive an optimum design with respect to the number of different dose levels k, the location of the dose levels $x_2, \ldots, x_{k-1}$ and the proportions of patients allocated with the dose levels $x_1, \ldots, x_k$.

In this talk we determine local optimal designs for the MED estimation problem by minimizing the asymptotic variance of the MED estimate for a particular dose response model. We study the finite sample properties of these designs via simulation. Moreover, we investigate the sensitivity of the local optimal designs with respect to mis-specification of the true parameter values and the underlying regression model. Finally, local optimal designs are constructed, which are robust with respect to a class of models commonly used in practice.

Session:
S41: Cost efficient designs for biostatisticians : Thursday, 13/03/2008, 11:00am - 12:40pm

How to link call rate and Hardy-Weinberg equilibrium as measures of genome-wide SNP data quality - results from the KORA 500K project

Helmut Finner[1], Klaus Strassburger[1], Iris M. Heid[3], Christian Herder[2], Wolfgang Rathmann[1], Guido Giani[1], Thorsten Dickhaus[1], Peter Lichtner[5,6], Thomas Meitinger[5,6], H-Erich Wichmann[3,4], Thomas Illig[3], and Christian Gieger[3,4]

[1] German Diabetes Center, Institute of Biometrics and Epidemiology, Leibniz Center at the Heinrich-Heine-University, Düsseldorf, Germany; dickhaus@ddz.uni-duesseldorf.de
[2] German Diabetes Center, Institute of Clinical Diabetology, Leibniz Center at the Heinrich-Heine-University, Düsseldorf, Germany
[3] GSF - Research Center for Environment and Health, Institute for Epidemiology, Neuherberg, Germany
[4] University of Munich, IBE, Chair of Epidemiology, Munich, Germany
[5] GSF - Research Center for Environment and Health, Institute of Human Genetics, Neuherberg, Germany
[6] Technical University of Munich, University Hospital Rechts der Isar, Institute of Human Genetics (IHG), Munich, Germany

Abstract: Based on SNP data from the KORA 500K project, a genome-wide association (GWA) study in a population-based German survey, we study the impact of the call rate (CR) on Hardy-Weinberg equilibrium (HWE) and the data quality. The applied statistical tools are based on (realized randomized) p-values and the false discovery rate (FDR). It turns out that all SNPs with complete genotype information (CR=100%) are nearly perfect in HWE militating in favor of the population being in HWE. The proportion of SNPs not in HWE increases linearly with decreasing CR. This finding has important implications for the analysis of GWA studies. Usage of a single threshold for HWE p-values as quality criterion cannot be recommended. Instead, we propose a stratified analysis with different thresholds depending on CRs. Finally, we employ empirical Bayes methods to compute posterior probabilities for the validity of the HWE hypothesis for a SNP given its p-value.

Session:
S20: Genome-wide association studies (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

TagSNP selection in 19^{th} human chromosome based on HapMap data

Adrian Drozdz[1] and Joanna Szyda[1,2]

[1] Wroclaw University of Envirnomental and Life Sciences, Institute of Animal Genetics, ul. Kozuchowska 7, 51-631 Wroclaw, Poland; enhncr@gmail.com
[2] Institute of Natural Sciences, pl. Grunwaldzki 24, 50-365 Wroclaw, Poland; szyda@ar.wroc.pl

Abstract: The amount of SNPs collected in HapMap Project approaches near to 4 000 000. This bulk of data needs to be reduced to get the smallest possible set of SNPs easy to genotype automatically, but which best covers genes of interest or the whole genome. Decreasing the amount of SNPs, called tagging, is the crucial way to improve time and cost of analysis. Using publicly available tools for choosing tagSNPs it is possible to obtain reduction in the number of SNP. The percent of reduction depends on program and data used. In this analysis three programs were used (FESTA, TAGster and Haploview). The material comprised the p arm of the 19th human chromosome covering 25 Mbp and 21 362 SNPs for the CEU population. The results show the differences in tagging effectiveness and in time of analysis. TAGster is the best SNP tagger. FESTA is the fastest.

References:

BARRETT J.C., FRY B., MALLER J. and DALY M.J. (2005): Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics, 21(2), 263-265.*

LAI E. (2001): Application of SNP Technologies in Medicine: Lessons Learned and Future Challenges. *Genome Research, 11 (issue 6), 927-929.*

QIN Z.S., GOPALAKRISHNAN S. and ABECASIS G.R. (2006): An efficient comprehensive search algorithm for tagSNP selection using linkage disequilibrium criteria. *Bioinformatics, 22(2), 220-225.*

THE INTERNATIONAL HAPMAP CONSORTIUM (2003): The International HapMap Project. *Nature, 426, 789-796.*

XU Z., KAPLAN N. and TAYLOR J. (2007): TAGster: Efficient Selection of LD tag SNPs in Single or Multiple Populations. *Bioinformatics, published online on September 12.*

Session:

S03: Genetic Epidemiology (I) : Monday, 10/03/2008, 9:10am - 10:30am

Bayesian nonparametric modeling in biostatistics

David B. Dunson

Biostatistics Branch, National Institute of Environmental Health Sciences, U.S. National Institutes of Health, USA; dunson1@niehs.nih.gov

Abstract: In recent years, there has been a dramatic increase in the use of Bayesian nonparametric methods in biomedical applications. This increase is attributable to the flexibility of these methods and to the availability of efficient and easy to implement computational algorithms. Bayesian nonparametric methods allow selected distributions within a hierarchical model to be unknown through use of a prior distribution, with the most commonly used prior being the Dirichlet process (DP). The DP is useful in building more flexible models that are robust to deviations from modeling assumptions and in applications, such as density estimation, regression, clustering and meta analysis. Recently developed methods generalize the DP to allow unknown distributions to change nonparametrically with predictors.

Motivated by practical advantages in biomedical applications, this talk provides a non- technical overview of some interesting recent developments in Bayesian nonparametrics. These developments will be illustrated through a series of applications to genotoxicology and epidemiology studies. The focus will be on providing an intuition for the use of nonparametric Bayes methods and mixture models in solving difficult problems in applied statistics.

Session:
S34: Bayesian models in biostatistics (II) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Geoadditive Latent Variable Models for Analyzing Morbidity and Malnutrition

Ludwig Fahrmeir and Khaled Khatab

Department of Statistics, Ludwig-Maximilians-Universität München, Munich, Germany; ludwig.fahrmeir@stat.uni-muenchen.de

Abstract: Latent variable models (LVM) consist of measurement models for multivariate, mixed discrete and continuous responses and a structural model for the latent variables. Geoadditive LVM extend the usual linear predictor in the structural model by incorporating P-splines and Markov random fields to account for nonlinear and spatial effects (Fahrmeir and Raach, 2007). Inference is fully Bayesian using MCMC simulation. We apply this methodology to data from a Demographic Health Survey (DHS) for Nigeria with z-scores for stunting and underweight and binary disease indicators fever, cough and diarrhea. The latent variables are morbidity and malnutrition. Our analysis based on Khatab (2007) identifies nonlinear covariate effects and geographic effects influencing these latent variables.

References:

FAHRMEIR, L. and RAACH, A. (2007): A Bayesian semiparametric latent variable model for mixed responses. *Psychometrica 72, 327-346.*

KHATAB, K. (2007): Analysis of childhood diseases and malnutrition in developing countries of Africa. *Dissertation, Department of Statistics, Ludwig-Maximilians-Universität München.*

Session:

S24: Spatial Analysis of Surveillance Data : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Strategies for Including Patients Recruited During Interim Analysis of Clinical Trials

Andreas Faldum[1], Gerd Otto[2], Jochen Thies[2], and Gerhard Hommel[1]

[1] Institute of Medical Biostatistics, Epidemiology and Informatics, University of Mainz, Germany; faldum@imbei.uni-mainz.de, hommel@imbei.uni-mainz.de
[2] Transplantation and Hepatobiliopancreatic Surgery, University of Mainz, Germany; otto@transplantation.klinik.uni-mainz.de, thies@transplantation.klinik.uni-mainz.de

Abstract: In clinical trials a periodical check of safety and efficacy data is often needed. For organizational reasons it is rarely desirable to stop a trial during such an interim analysis. Therefore, new study patients are included in the trial while the interim analysis is ongoing. Disregarding the additional information provided by these interim patients would be unsatisfactory, especially for an office of regulatory affairs. Consequently, the rules for group sequential or adaptive decisions must be adjusted to the recruitment of interim patients.

In the context of a prospective randomised multicentre trial to investigate the incidence of ischemic type biliary lesions after liver transplantation, strategies for modifying study designs to consider the analysis of interim patients are proposed. The impact of the proposed design modifications on conditional error function, overall power, average sample size and maximum sample size is demonstrated.

References:

FALDUM, A. and HOMMEL, G. (2007): Strategies for including patients recruited during interim analysis of clinical trials. *Journal of Biopharmaceutical Statistics, 17(6), 1211–1225.*

OTTO, G., MAUER, D., THIES, J., KRONFELD, K., WACHTLIN, D., and FALDUM, A. (2007): Prospective randomised multicentre trial investigating liver preservation with HTK by simple aortic perfusion in comparison to aortic perfusion plus ex situ arterial flushing. *Studienprotokoll.*

WHITEHEAD, J. (1992): Overrunning and underrunning in sequential clinical trials. *Controlled Clinical Trials, 13, 106–121.*

Session:

S05: Adaptive Group Sequential Designs : Monday, 10/03/2008, 2:10pm - 3:30pm

Adjusting Life Expectancy for Population Ageing

Uwe Feldmann

Institute of Medical Biometry, Epidemiology and Medical Informatics, University of Saarland, Germany; uf@med-imbei.uni-saarland.de

Abstract: While it is familiar to adjust the mortality rate for population ageing in order to make this measure comparable over time and across countries, such an age-standardization seems not be available for life expectancy. Time series of demographic parameters were recorded in several federal states of Germany during the decade after its reunion. A paradox seemed to be happened. On one hand, within this decade the new federal states had a gain in life expectancy of more than four years while the gain in the old federal states was only about two years. On the other hand in the new federal states the birth rates sunk dramatically and additionally, there was a considerable emigration rate of people aged between 20 and 30 years. The paper explains why birth rate reduction and emigration of younger people may cause an over-estimation of life expectancy and how to adjust this effect.

References:

FELDMANN, U., MENEHRT, RF. (2006): Assessing the state of population health by age-adjusted life expectancies. *Methods Inf Med, 45, 275–280.*

HALLEEY, E. (1691): An estimate on the decrease of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw. *Philos Transact, Royal Society of London.*

MENEHRT, RF. (2004): Zur gesundheitlichen Entwicklung der saarländischen Bevölkerung im Vergleich zu Sachsen, Baden-Württemberg und Deutschland. *Thesis of the Medical Faculty of the University of Saarland, Germany.*

SANDERSON, WC. and SCHERBOV, S. (2005): Average remaining lifetimes can increase as human populations age. *Nature, 435, 811–813.*

Session:

S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Alternative Approaches for Whole-Genome Analyses

Rohan L. Fernando[1] and Christian Stricker[2]

[1] Department of Animal Science, Iowa State University, Ames, Iowa, USA; rohan@iastate.edu
[2] Applied Genetics Network, Davos, Switzerland; stricker@genetics-network.ch

Abstract: Due to continuing advances in genotyping techniques, high-density, marker data are becoming available for mapping quantitative trait loci (QTL) and for prediction of genotypic values. One advantage of using high-density marker genotypes is the expectation that some of these genotypes that are close to QTL will be in linkage disequilibrium with the QTL (Meuwissen et al. 2001). Thus, estimating the effect of marker genotypes on the quantitative trait provides a simple approach to analyze whole-genome data. The greatest problem with this approach is that the number of marker effects in the model is often much greater than the number individuals with trait phenotypes. Meuwissen et al. (2001) considered four strategies to address this problem. The first was a stepwise regression technique where a marker was added to the model if it significantly improved the fit of the current model. Based on simulation results they concluded that stepwise regression technique was not able to esimate the effects of the markers. In the second, all markers were included in the model as random effects with the same variance for all markers. This variance was set to the genetic variance divided by the number of markers. It has been shown that the accuracy of prediction with this strategy is inversely related to the number of markers included in the model (Fernando et al. 2007). Thus, this second strategy is not suitable for whole-genome analyses. In the third strategy, Bayesian inference was used with unknown, locus-specific variances. A scaled, inverse chi-square distribution was used as the prior. The fourth strategy also used Bayesian inference with unknown, locus-specific variances, but a known proportion π of the markers were assumed to have a null variance, corresponding to markers that are not associated with segregating QTL. The fourth strategy had the best performance. Here, we will examine the consequences of using incorrect values for π. Also, we will consider Bayesian inference with unknown π and alternative prior distributions for π.

References:

MEUWISSEN, T., HAYES, B., and GODDARD, M. (2001): Prediction of total genetic value using genome-wide dense marker maps. *Genetics, 157, 1819–1829.*

FERNANDO, R.L., HABIER, D., STRICKER, C., DEKKERS, J.C.M., TOTIR, L.R. (2007): Genomic selection. *Acta Agriculturae Scandinavica, Section A, Animal* Science. (in press).

Session:

S04: Mapping Approaches in Plants and Animals : Monday, 10/03/2008, 2:10pm - 3:30pm

A new model for rank-based regression analysis

Václav Fidler

Department of Epidemiology, University Medical Center Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands; V.Fidler@epi.umcg.nl

Abstract: We consider a multiple regression problem in which the analysis is to be based on the ranks of the original continuous outcomes. Existing models for rank data will be briefly reviewed and their possible use in a model based rank regression analysis discussed. An important example of a rank regression model is the proportional hazards model of Cox, also known as the forward selection model. It is much used for analysis of survival data but rarely as a general method for nonparametric regression. This may be due to lack of invariance of the model under changing the sign of all observations (reversing the ranking). We propose a new probability model for ranks which does not have this drawback. The model can be used as a starting point for maximum likelihood based inference. We illustrate its use on real and simulated data sets.

Session:
S35: Nonparametric and Parametric Multivariate Tests in High Dimensions : Wednesday, 12/03/2008, 2:10pm - 3:30pm

The Conditional Synergy Index to identify a biological gene-gene-interaction

Ronja Foraita

Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Germany; foraita@bips.uni-bremen.de

Abstract: Understanding the etiology of common human diseases means to understand the complexity of the relationship between genotype and phenotype. Part of this complexity can be attributed to epistasis or biological gene-gene-interaction that occurs when the phenotype for a genotype at one locus is dependent on genotypes at one or more other loci. However, geneticists and statisticians understand the term "interaction" differently which leads to difficulties in detecting biological interaction using statistical models (Cordell, 2002; Foraita et al., 2008). Caliebe et al. (2005) prefer the concepts of "synergy" and "antagonism" to describe biological mechanisms instead of the concept of statistical interaction.

We introduce a new index to measure interaction that exploits the idea of synergy and antagonism. The so-called Conditional Synergy Index (CSI) is appropriate to detect interaction as well as genetic heterogeneity on the penetrance scale. The index is restricted to qualitative phenotypes and two susceptibility genes to be categorized into exposed and unexposed genotypes. We present the properties of $\widehat{CSI}$, its estimated variance and results on the power obtained from simulation studies using different biological models. The power analysis of the test procedure based on the CSI shows that the CSI behaves conservative for known penetrances and is too liberal for penetrances that are estimated from the data.

References:

CALIEBE, A., FREITAG, S., and KRAWCZAK, M (2005): Stochastische Modelle für Interaktion und Effektmodifikation. *Medizinische Genetik, 17, 14–19.*

CORDELL, H. (2002): Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics, 11, 2463–2468.*

FORAITA, R., BAMMANN, K., and PIGEOT, I. (2008): Modeling gene-gene interactions using graphical chain models. *Human Heredity, 65, 47–56.*

Session:

S14: Genetic Epidemiology (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Robust Filtering by Weighted Repeated Medians

Roland Fried[1], Jochen Einbeck[2], and Ursula Gather[1]

[1] Department of Statistics, University of Dortmund, 44221 Dortmund, Germany; fried@statistik.uni-dortmund.de
[2] Department of Mathematical Sciences, Durham University, Durham City, DH1 3LE, UK

Abstract: We construct weighted repeated median smoothers and filters for robust non-parametric regression in general and for robust online signal extraction from time series in particular. An advantage of regression-based local linear fits as compared to local constant fits is that they do not deteriorate during (local) linear trends. The new methods proposed here allow to remove outlying sequences and to preserve shifts in the underlying regression function (the signal) even in the presence of such trends. Suitable weighting of the observations according to their distances in the design space reduces the bias arising from non-linearities and improves the efficiency of the ordinary repeated median filters suggested by Davies et al. (2004) and Gather et al. (2006) using larger bandwidths, while still distinguishing long-term shifts from outlier sequences. Comparisons are given to other localized robust regression techniques like S-, M- and MM-estimators as well as weighted L_1-regression.

The methods are illustrated using monitoring data from intensive care, where we need to extract the relevant information from high-frequency data in real time.

References:

DAVIES, P. L., FRIED, R. and GATHER, U. (2004): Robust Signal Extraction for On-line Monitoring Data. *Journal of Statistical Planning and Inference, 122, 65–78.*

FRIED, R., EINBECK, J. and GATHER, U. (2007): Weighted Repeated Median Smoothing and Filtering. *Journal of the American Statistical Association, 480, 1300–1308.*

GATHER, U., SCHETTLINGER, K. and FRIED, R. (2006): Online Signal Extraction by Robust Linear Regression. *Computational Statistics, 21, 33–51.*

SIEGEL, A. F. (1982): Robust Regression using Repeated Medians. *Biometrika, 68, 242–244.*

Session:

S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Heterogeneity of Treatment Effect Estimates across Design Stages in Adaptive Trials

Tim Friede[1] and Robin Henderson[2]

[1] Warwick Medical School, The University of Warwick, Coventry, UK;
t.friede@warwick.ac.uk
[2] Mathematics & Statistics, Newcastle University, UK;
Robin.Henderson@ncl.ac.uk

Abstract: The recently adopted CHMP reflection paper on flexible designs (CHMP, 2007) highlights a controversial issue regarding the interpretation of adaptive trials. The guideline suggests that heterogeneity of treatment effects across design stages may be investigated using techniques as in meta-analysis. In meta-analysis it is common practice to carry out a test for heterogeneity and if treatment effect estimates differ significantly between studies then results from different studies might not be combined in a formal analysis. In this presentation we investigate error rates for such a procedure when applied to design stages of adaptive trials. Since effects of calendar time are not uncommon in clinical research (Altman and Royston, 1988) we also study the properties of this procedure in the presence of calendar time effects. Furthermore, we describe alternative models allowing for calendar time effects and discuss what can be learned from meta-analysis.

References:

ALTMAN, D.G. and ROYSTON, P. (1988): The hidden effect of time. *Statistics in Medicine, 7, 629–637.*

CHMP (2007): Reflection paper on methodological issues in confirmatory clinical trials with flexible design and analysis plan. London, 18 October 2007, Doc. Ref. CHMP/EWP/2459/02.

Session:

S15: Flexible designs : Tuesday, 11/03/2008, 11:00am - 12:40pm

Limits of agreement for dependent data

Katja Frieler[1], Tania Schink[1], Hans Hoffmann[2], and Marc Dewey[2]

[1] Department of Medical Statistics and Clinical Epidemiology, Universitätsmedizin-Berlin, Germany; katja.frieler@charite.de, tania.schink@charite.de
[2] Department of Radiology, Universitätsmedizin-Berlin, Germany; hans.hoffmann@charite.de, marc.dewey@charite.de

Abstract: Limits of agreement (Altman and Bland, 1983; Bland and Altman, 1986) are a very intuitive tool to assess agreement between different methods of measurement as well as inter- or intrarater agreement. The method is well established in the case where the differences between pairs of measurements can be considered as independent. But applying the usual approach to dependent data may lead to an overestimation of agreement. In this case, classical one way analysis of variance allows transferring the concept to the situation of multiple pairs of measurements per individual and constant limits of agreement (Bland and Altman, 2007). We want to demonstrate that available methods on linear mixed models provide a natural extension to this approach. Especially, the approach allows for very flexible adjustments when limits of agreement are not constant but depend for example on the mean of the paired measurements.

We will apply the method to assess intrarater agreement in volumetric measurements of coronary artery plaques whose number varies from patient to patient.

References:

ALTMAN, D. G., and BLAND, J. M. (1983): Measurement in Medicine: The Analysis of Method Comparison Studies. *Statistician 32:307-317.*

BLAND, J. M., and ALTMAN, D. G. (1986): Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet i:307-310.*

BLAND, J. M., and ALTMAN, D. G. (2007): Agreement between Methods of Measurement with Multiple Observations per Individual. *Journal of Biopharmaceutical Statistics, 17, 571-582.*

Session:
S37: Clinical studies (III) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Rank Transformation in Haseman-Elston Regression Using Scores for Location-Scale Alternatives

Daniel Gerhard

Institute of Biostatistics, Leibniz University Hanover, 30419 Hannover, Germany;
gerhard@biostat.uni-hannover.de

Abstract: The Haseman-Elston method is a simple regression tool for detecting genetic linkage to quantitative traits in sib-pair studies. Although this method and especially the extended new Haseman-Elston approach is quite robust, there might be some loss of power for non-normal distributed traits. Here, rank transformation methods, which either combine the information of a trend in locations and in scales or detect a trend for only a subset of the trait variables, are presented. Simulation results for settings with several assumptions on the residual error distribution indicate a gain in power in comparison to recently suggested nonparametric methods. Under normality, the maintenance of almost the same characteristics of the original Haseman-Elston approach can be achieved. The straightforward application is illustrated on an exemplary dataset.

References:

CONOVER, W.J. and SALSBURG, D.S. (1988): Locally most powerful tests for detecting treatment effects when only a subset of patients can be expected to 'respond' to treatment. *Biometrics, 44, 189–196.*

FORREST, W. (2001): Weighting improves the 'New Haseman-Elston' method. *Human Heredity, 52, 47–54.*

HASEMAN, J.K. and ELSTON, R.C. (1972): The investigation of linkage between a quantitative trait and a marker locus. *Behaviour Genetics, 2, 3–19.*

KIM, M.K., HONG, Y.J., SONG, H.H. (2006): Nonparametric trend statistic incorporating dispersion differences in sib pair linkage for quantitative traits. *Human Heredity, 62, 1–1.*

VAN ZWET, W.R. and OOSTERHOFF, J. (1967): On combination of independent test statistics. *Annals of Mathematical Statistics, 38, 659–680.*

Session:

S03: Genetic Epidemiology (I) : Monday, 10/03/2008, 9:10am - 10:30am

Feature Extraction in Signal Regression: a Boosting Technique for Functional Data Regression

Jan Gertheiss and Gerhard Tutz

Department of Statistics, LMU Munich, Germany;
jan.gertheiss@stat.uni-muenchen.de, tutz@stat.uni-muenchen.de

Abstract: Signal regression means predicting or explaining a scalar response from a functional predictor (a *signal*). Main objectives of feature extraction in signal regression are the improvement of accuracy of prediction on future data and identification of relevant parts of the signal. A feature extraction procedure is proposed that uses boosting techniques to select the relevant parts of the signal. The proposed blockwise boosting procedure simultaneously selects intervals in the signal's domain and estimates the effect on the response. The blocks that are defined explicitly use the underlying metric of the signal. It is demonstrated in simulation studies and for real-world data that the proposed approach competes well with procedures like PLS, P-spline signal regression and functional data regression. Though motivation, illustration and comparisons of methods are mainly based on Ramsey's benchmark weather data set and NIR spectroscopy data, the method can in principle be used for the analysis of mass spectrometry data in proteomics as well.

References:

FRANK, I. E. and FRIEDMAN, J. H. (1993): A Statistical View of Some Chemometrics Regression Tools, *Technometrics, 35, 109–135.*

MARX, B. D. and EILERS, P. H. C. (1999): Genaralized linear regression on sampled signals and curves: A P-spline approach, *Technometrics, 41, 1–13.*

RAMSAY, J. O. and SILVERMAN, B. W. (2005): *Functional Data Analysis*, 2nd edition, Springer, New York.

TUTZ, G. and GERTHEISS, J. (2007): Feature Extraction in Signal Regression: A Boosting Technique for Functional Data Regression, *University of Munich, Department of Statistics: Technical Report, No.11.*

Session:
S27: Structuring high-dimensional data : Wednesday, 12/03/2008, 9:10am - 10:30am

A 2 × 2 Factorial Clinical Trial in Oncology

Frank Gilberg and Sandrine Laguerre

F. Hoffmann La Roche, Basel, Switzerland; frank.gilberg@roche.com, sandrine.laguerre@roche.com

Abstract: The presentation will describe design and analysis of a major study in oncology using a factorial design. The study started as a phase III 2-arm open-label non-inferiority study comparing XELOX with FOLFOX4. In 2003 the addition of bevacizumab (Bev) to irinotecan/5-FU/LV was shown to improve progression-free survival (PFS) and overall survival. In consequence the trial was amended to a 2 partially blinded study to assess the addition of Bev. The original 2-arm study recruited 634 pts; after transition to 2, an additional 1400 patients were recruited.

The baseline characteristics were well balanced between the arms. The 2 co-primary objectives of the study were met. XELOX was non-inferior to FOLFOX-4 in terms of PFS with a HR of 1.05 (95% CI $0.94 - 1.18$) the upper limit of the 97.5% CI being below the non-inferiority margin of 1.23.

The addition of Bev to oxaliplatin-based chemotherapy demonstrated a significant benefit in terms of PFS in the primary analysis (HR 0.83; 97.5% CI $0.72 - 0.95$, $p = 0.0023$). Prespecified analysis of PFS on treatment (defined as progressive disease or death within 28 days from the last dose of study treatment) and PFS analysis based on tumor assessments by an independent review committee (IRC) were consistent with the benefit observed in the primary analysis.

Design, study results and statistical issues specifically related to the factorial design of this trial will be presented and discussed.

References:

ROTHMAN, M. (2003): Design and analysis of noninferiority mortality trials in oncology. *Statistics of Medicine, 22, 239–264.*

Session:

S18: Clinical studies (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Testing and model selection for prediction in large sets of variables

Alexandra Goll and Peter Bauer

Institute of Medical Statistics, Medical University Vienna, 1090 Vienna, Austria;
alexandra.goll@meduniwien.ac.at

Abstract: The task of selecting useful markers with rather moderate effects from a large number of candidates in small samples and estimating suitable scores to be used for the prediction in future patients is a formidable exercise. It is known that model selection by multiple testing of individual model parameters under fairly general conditions asymptotically is a consistent selection procedure. Asymptotic relationships between model selection procedures and multiple tests controlling the false discovery rate (FDR) have been shown. However these are general results which do not help how to choose the selection boundaries in specific samples. We evaluated selection based on a multiple test controlling the FDR and construction of a linear score by considering the resulting receiver operating characteristic (ROC) in independent future patients. We simulated different scenarios with varying number of markers, proportion of effective markers and sample size. The effect sizes were chosen such that the optimal prediction of future patients, if known, would lead to a ROC-curve crossing through a certain benchmark point with pre-defined values for sensitivity and specificity. The best FDR to be used in selection which provides the ROC-curve with the largest area under the curve (AUC) varies over the different parameter constellation not known in advance. Hence, we considered cross validation to determine the optimal selection criterion in a specific sample. For that purpose we used a modified jackknife procedure: All possible pairs of a single responder and non-responder respectively were left out and the linear prediction score was repeatedly calculated for a grid of selection boundaries from each of the resulting training samples. The optimal boundary chosen is the one which leads to the score with the best discrimination when averaged over all possible pairs. This procedure allows choosing an appropriate selection criterion for constructing a prediction score and at the same time provides an estimate for the extent of false positive decisions. Moreover, this procedure when leading to low jackknife AUCs may indicate that there are no effective markers at all in a particular data set.

Session:
S36: Young Statistician Papers (IBS-ROeS) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

A modified Bonferroni procedure with plug-in estimator for the number of true hypotheses

Helmut Finner and Veronika Gontscharuk

German Diabetes Center, Leibniz Center at the Heinrich-Heine-University Düsseldorf, Institute of Biometrics and Epidemiology, Germany;
finner@ddz.uni-duesseldorf.de,
veronika.gontscharuk@ddz.uni-duesseldorf.de

Abstract: Schweder and Spjøtvoll (cf. SCHWEDER (1982)) studied the problem of estimating the number of true hypotheses n_0 (say) in various multiple testing scenarios with n hypotheses. They also mentioned that such an estimate may be utilized to modify the classical Bonferroni test by replacing the threshold α/n by $\alpha/\hat{n}_0$. We denote such test procedures by BPI (Bonferroni plug-in). There seem to be no theoretical results yet whether BPIs control the familywise error rate (FWER) in the strong sense. In this talk we give a first answer. If the p-values under the corresponding null hypotheses are iid uniformly distributed on $[0, 1]$ and if we take a special version of Storey's (cf. STOREY (2002)) estimator for $\hat{n}_0$ depending on some tuning parameter $\lambda \in (\alpha, 1]$, it will be shown that the FWER is controlled at the prespecified level α. Moreover, it turns out that so-called Dirac-uniform configurations are least favorable for the FWER. This results in explicit formulas and upper bounds for the actual FWER. Finally, we discuss the asymptotic behaviour of BPIs with alternative plug-in estimators.

References:

SCHWEDER, T. and SPJØTVOLL, E. (1982): P-value plots to evaluate many tests simultaneously. *Biometrika, 69, 493-502.*

STOREY, J.D. (2002): A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B, 64, Part 3, pp. 479-498.*

Session:

S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

Issues in comparing propensity score stratification with linear regression to assess a treatment effect

Erika Graf

Institute for Medical Biometry and Medical Informatics, Freiburg, Germany;
egr@imbi.uni-freiburg.de

Abstract: Over the last decade, propensity score (PS) based approaches have become popular as a means to adjust for confounding in the estimation of a treatment effect from non-experimental studies. Estimation from regression models is a more traditional technique.

Meaningful comparisons can be made either from a marginal perspective in which properties of estimators are considered assuming that the observed covariates would vary across samples according to an underlying probability distribution. Here, situations exist in which PS stratification and linear regression both produce unbiased estimators, but none of the two is uniformly better than the other in terms of marginal variance. Alternatively, conditional inference examines properties induced when the observed covariate distribution is constant across samples. Insight can be gained by noting that, under specific circumstances, the estimator derived from PS stratification is identical to the effect estimated from a full linear model. This is true even if the PS estimator is built on coarser covariate strata than the linear model.

As a consequence, the coarsening property of the PS — adjustment for a one-dimensional confounder instead of a high-dimensional covariate — may be viewed as a way to implement a pre-specified but richly parameterised linear model. It follows that the estimator based on PS stratification may be prone to substantial overfitting. This has gone largely unnoticed by statisticians up to now, because the PS strata are derived without making use of outcome data.

References:

RUBIN D. R. (2007): The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine, 26, 20-36.*

SENN, S., GRAF, E. and CAPUTO, A. (2007): Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure. *Statistics in Medicine, 26, 5529-5544.*

Session:

S37: Clinical studies (III) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Inference in Additive and Linear Mixed Models

Sonja Greven[1], Ciprian Crainiceanu[2], Fabian Scheipl[1], Helmut Küchenhoff[1], and Annette Peters[3]

[1] Department of Statistics, LMU Munich; sonja.greven@stat.uni-muenchen.de
[2] Department of Biostatistics, Johns Hopkins University, Baltimore, USA; ccrainic@jhsph.edu
[3] Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; peters@gsf.de

Abstract: Flexible regression models with random effects are widely used in epidemiology and medical statistics. Our motivating application is the assessment of non-linearity for air pollution dose-response functions. Using the penalized spline approach, testing for linearity in a semiparametric mixed model can be seen as a special case of testing for a zero random effect variance. This test is non-standard due to the tested parameter on the boundary of the parameter space. Furthermore, observations are generally not independent in linear mixed models. We compare several tests in extensive simulations, including different approximations to the null distribution of the restricted likelihood ratio test (RLRT), and several F-type tests. We find that our fast finite sample approximation to the RLRT null distribution (Greven et al., 2008; Scheipl et al., 2008; Crainiceanu and Ruppert, 2004; R package `RLRsim`) outperforms the χ^2-mixture approximation and parametric bootstrap currently used, as well as all F-type tests with regard to power, adherence to the nominal level and/or speed. The class of models we consider includes models with moderate numbers of clusters, unbalanced designs, or nonparametric smoothing. We also compare our test procedures to model selection using the Akaike Information Criterion (AIC), which is adversely affected by the boundary problem.

References:

CRAINICEANU, C. and RUPPERT, D. (2004): Likelihood ratio tests in linear mixed models with one variance component. *JRSS-B, 66, 165–185.*

GREVEN, S., CRAINICEANU, C., KÜCHENHOFF, H. and PETERS, A. (2008): Restricted Likelihood Ratio Testing for Zero Variance Components in Linear Mixed Models. *Journal of Computational and Graphical Statistics, to appear.*

SCHEIPL, F., GREVEN, S. and KÜCHENHOFF, H. (2008): Size and Power of Tests for a Zero Random Effect Variance or Polynomial Regression in Additive and Linear Mixed Models. *Computational Statistics & Data Analysis, to appear.*

Session:

S31: Advances in Statistical Modelling (III) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Imputation of missing longitudinal data: different approaches

Ulrike Grittner[1], Matthias Wicki[2], Gerhard Gmel[2], and Kim Bloomfield[3]

[1] Institute for Biometrics and Clinical Epidemiology, Charité–University Medicine Berlin, Germany; ulrike.grittner@charite.de
[2] Swiss Institute for the Prevention of Alcohol and Drug Problems, Lausanne, Switzerland; mwicki@sfa-ispa.ch,ggmel@sfa-ispa.ch
[3] Unit for Health Promotion Research, University of Southern Denmark, Esbjerg, Denmark; kbl@health.sdu.dk

Abstract: In longitudinal studies missing information is unavoidable. Due to differences between the participants with complete information and those participants with missing information in some waves, solely analysing complete data could distort the results and conclusions.

This study demonstrates different approaches for imputation of missing data by using a survey on alcohol consumption in Denmark as an example. Data from the Nordic Study of Changes in Drinking was used. It was collected in four waves in 2003, 2004, 2005 and 2006. The 2003 dataset contains 1771 cases aged 16 to 69. Only 35.7% of these cases participated in all four waves. 15.9% participated in 2003 only. 17.8% participated in both 2003 and 2004; 13.4% in 2003, 2004 and 2005. Those who dropped out in the second or third wave were asked to participate in the following wave and 16.9% responded again in a later wave.

Missing value imputation for alcohol consumption was both done with multiple imputation Markov Chain Monte Carlo methods using WINBUGS and Heckman models using STATA. The results of the different imputation models were compared with each other and the model-based predicted values of people with complete data were compared with the true values. Similarities and differences of both methods will be discussed.

References:

HECKMAN, J. (1979): Sample selection bias as a specification error. *Econometrica, 47, 153–161.*

GILKS, W.R.; RICHARDSON, S. and SPIEGELHALTER, D.J. (2003): *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Session:

S12: Freie Themen (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Estimating threshold values for minimum provider volumes in total knee replacement data

Ulrich Grouven[1], Ralf Bender[1], and Helmut Küchenhoff[2]

[1] Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany; ulrich.grouven@iqwig.de, ralf.bender@iqwig.de
[2] Statistical Consulting Unit, Department of Statistics, Ludwig-Maximilians Universität, Munich, Germany; Kuechenhoff@stat.uni-muenchen.de

Abstract: During the past decades a large number of studies investigating a possible association between provider volume and outcome for various surgical interventions have been published. We analyzed German health care data of total knee replacement in order to derive suitable threshold models for the calculation of minimum provider volumes [SCHRÄDER (2007)]. Two binary variables describing outcome quality, i.e. "postoperative insufficient mobility" and "postoperative infection" were analyzed accounting for potentially confounding factors. For modelling the association between provider volume and outcome multiple logistic regression models and generalized estimating equations (GEE) to account for possible cluster effects were used. The method of fractional polynomials (FP) was employed to model non-linear relationships between continuous confounders and outcome variables [ROYSTON (1999)]. Threshold values were estimated by Bender's risk limit approach and by an extension of Ulm's breakpoint model [GROUVEN (2008)]. The results and features of these approaches are compared and discussed.

References:

SCHRÄDER, P., GROUVEN, U. and BENDER, R. (2007): Können Mindestmengen für Knieprothesen anhand von Routinedaten errechnet werden? Ergebnisse einer Schwellenwertanalyse mit Daten der externen stationären Qualitätssicherung. *Orthopäde, 36, 570–576.*

ROYSTON, P., AMBLER, G. and SAUERBREI, W. (1999): The use of fractional polynomials to model continuous risk variables in epidemiology. *International Journal of Epidemiology, 28, 964–974.*

GROUVEN, U., KÜCHENHOFF, H., SCHRÄDER, P. and BENDER, R. (2008): Flexible regression models are useful tools to calculate and assess threshold values in the context of minimum provider volumes. *Journal of Clinical Epidemiology (accepted for publication).*

Session:
S02: Diagnostic studies : Monday, 10/03/2008, 9:10am - 10:30am

Neural networks modeling gene-gene-interactions

Frauke Guenther, Karin Bammann, and Nina Wawro

Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Germany; guenther@bips.uni-bremen.de

Abstract: The investigation of complex diseases plays an important role in genetic epidemiology. Besides the main effects of single covariates, particulary the interplay of potential risk factors such as genetic causes and environmental factors is considered. However, the statistical modeling of such gene-gene- or gene-environment-interactions causes a lot of problems because there is no unique overall definition of interaction in disciplines like biology and statistics. This leads to a discrepancy between what statistical tools are able to capture and what biologists have in mind (Foraita et al., 2008).

To resolve these discrepancy we exploit the high flexibility of artificial neural networks (Bishop, 1995) for modeling biological interaction. We investigate whether artificial neural networks are able to capture the complexity and structure of different biological interaction models and compare the results with those of logistic regression models and multifactor dimensionality reduction (MDR) (Ritchie et al., 2001). First simulation results will be presented.

References:

BISHOP C.M. (1995): *Neural networks for pattern recognition.* Oxford University Press, New York.

FORAITA, R., BAMMANN, K., and PIGEOT, I. (2008): Modeling gene-gene interactions using graphical chain models. *Human Heredity, 65, 47–56.*

RITCHIE M.D., HAHN L.W., ROODI N., BAILEY L.R., DUPONT W.D., PARL F.F., and MOORE J.H. (2001): Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *American Journal of Human Genetics, 69(1), 138–147.*

Session:

S14: Genetic Epidemiology (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

An algorithm for the calculation of exact confidence intervals for adaptive group sequential trials

Niklas Hack and Werner Brannath

Section of Medical Statistics, Medical University of Vienna, Vienna, Austria;
niklas.hack@meduniwien.ac.at, werner.brannath@meduniwien.ac.at

Abstract: We describe a new algorithm for the calculation of stage-wise adjusted confidence intervals for adaptive group sequential trials proposed in Brannath, Mehta and Posch (2007). The key idea, based on the method of Müller and Schäfer (2001), is to preserve the overall type I error rate after a possible design adaptation, by preserving the null conditional rejection probability of the remainder of the trial at each time of an adaptive change. The confidence intervals provide exact coverage in the case of no adaptive changes. In the case of adaptive changes the calculated confidence intervals are conservative. The algorithm for the calculation of stage-wise adjusted confidence intervals is described and an example for the full implementation in C is given. We will show by simulations that the stage-wise adjusted confidence intervals provide exact coverage probability and the calculated point estimates for the treatment difference are median unbiased.

References:

TSIATIS, A.A. and ROSNER, G.L. and MEHTA, C.R. (1984): Exact confidence intervals following a group sequential test. *Biometrics, 40, 797–804.*

MÜLLER, H.H. and SCHÄFER, H. (2001): Adaptive group sequential design for clinical trials: Combining the advantages of adaptive and of classic group sequential approaches. *Biometrics, 57, 886–891.*

BRANNATH, W. and MEHTA, C.R. and POSCH, M. (2007): Exact confidence bounds following adaptive group sequential tests. *submitted.*

Session:
S05: Adaptive Group Sequential Designs : Monday, 10/03/2008, 2:10pm - 3:30pm

Bivariate random effects meta-regression of diagnostic tests for Summary ROC curves

Taye H. Hamza[1], Hans C. van Houwelingen[2], Majanka H. Heijenbrok-Kal[1], and Theo Stijnen[1,2]

[1] Erasmus University Medical Center, Rotterdam, The Netherlands; t.hussienhamza@erasmusmc.nl
[2] Leiden University Medical Center, Leiden, The Netherlands

Abstract: The bivariate random effects (BRE) meta-analysis of diagnostic tests has been discussed and applied by many authors (for example, Reitsma et al., 2005 and Arends et al., 2008). It accounts for both the within- and between-study variability in sensitivities as well as specificities and their correlation. The method is relatively easy to fit by standard likelihood methods in statistical packages such as SAS (Proc Mixed or Proc NLMIXED). The bivariate approach is flexible in that it can handle different outcome measures of diagnostic test accuracy (Reitsma et al., 2005) and also different types of summary ROC curves can be chosen once the model is fitted (Arends et al., 2008). Although the BRE is easily extended with covariates for specificity and for sensitivity, it is less straightforward to study the influence of covariates on summary ROC curves. In this paper, at the hand of a case-study, we show how the BRE can be used to investigate the effect of covariates on the performance of a diagnostic test as characterized by a summary ROC curve.

References:

REITSMA J.B., GLAS A.S., RUTJES A.W.S., SCHOLTEN R.J.P.M., BOSSUYT P.M. and ZWINDERMAN A.H. (2005): Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology 58: 982–990.*

ARENDS L.R., HAMZA T.H., VAN HOUWELINGEN J.C., HEIJENBROK-KAL M.H., HUNINK M.G.M. and STIJNEN T. (2008): Meta-analysis of ROC curves using bivariate normal distribution of sensitivities and specificities. *Medical Decision Making (To appear).*

Session:

S42: Meta-analysis and Meta-regression : Thursday, 13/03/2008, 11:00am - 12:40pm

Using CMAT for Data Mining of Large Data Sets

Wolfgang M. Hartmann

D-69117 Heidelberg, Germany; cmat.wolfgang@gmail.com,
www.cmat.pair.com/cmat

Abstract: In Data Mining we are confronted with either many observations (1), many variables (2), or many of both (3). In this presentation we discuss some newer strategies for tackling (1) and (2), for example,

1. the possible use of the transposed data set and the Sherman-Morrison-Woodbury formula;
2. the case of many observations (chunking, sampling),
3. the case of many variables: use of some variable selection and dimension reduction methods.

The strategies are discussed in connection with known statistical analysis methods, like principal components, variable clustering, SVM, SCAD, and MDS.

References:

FUNG, G. and MANGASARIAN, O.L. (2003): A Feature Selection Newton Method for Support Vector Machine Classification, *Computational Optimization and Aplications, 1-18.*

HASTIE, T., and TIBSHIRANI, R., (2004): Efficient quadratic regularization for expression arrays, *Biostatistics, 5, 329-340.*

MANGASARIAN, O.L. and THOMPSON, M.E. (2006): Massive data classification via unconstrained support vector machines, *Journal of Optimization Theory and Applications. Technical Report 06-07, Data Mining Institute, University of Wisconsin, Madison, Wisconsin.*

MANGASARIAN, O.L. and THOMPSON, M.E. (2006): Chunking for massive nonlinear kernel classification, *Technical Report 06-07, Data Mining Institute, University of Wisconsin, Madison, Wisconsin.*

Session:

S29: Advances in Statistical Modelling (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

A Multivariate Dunnett Procedure for Correlated Endpoints

Mario Hasler

Institute of Biostatistics, Leibniz Universität Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany; hasler@biostat.uni-hannover.de

Abstract: This talk considers a single step procedure for simultaneous comparisons of several treatments with a control by ratios of means when there is more than one primary endpoint of interest. An approximate multivariate t-distribution is used to obtain quantiles for test decisions and simultaneous confidence intervals. Simulation results show that this approach asymptotically controls the familywise error rate strongly.

References:

BLOCH, D.A., LAI, T.L., and TUBERT-BITTER, P. (2001): One-sided tests in clinical trials with multiple endpoints. *Biometrics, 57, 1039–1047.*

DILBA, G., BRETZ, F., and GUIARD, V. (2006): Simultaneous confidence sets and confidence intervals for multiple ratios. *Journal Of Statistical Planning And Inference, 136(8), 2640–2658.*

DILBA, G., BRETZ, F., GUIARD, V., and HOTHORN, L.A. (2004): Simultaneous confidence intervals for ratios with applications to the comparison of several treatments with a control. *Methods Of Information In Medicine, 43(5), 465–469.*

GENZ, A., BRETZ, F., and R support by HOTHORN, T. (2006): mvtnorm: Multivariate Normal and t Distribution, R package version 0.7-5.

HOTHORN, T., BRETZ, F., and GENZ, A. (2001): On multivariate t and Gauß probabilities in R. *R News, 1(2), 27–29.*

R DEVELOPMEMT CORE TEAM (2007): R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

TONG, Y.L. (1990): *The multivariate normal distribution.* Springer, New York.

Session:
S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

Choice of delta in non-inferiority studies: difference- versus ratio-based approach

Dieter Hauschke[1] and Ludwig A. Hothorn[2]

[1] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; dieter.hauschke@imbi.uni-freiburg.de
[2] Institute of Biostatistics, Leibniz University of Hannover, Germany; hothorn@biostat.uni-hannover.de

Abstract: In a recent paper Brown et al. (2006) presented some important features of the European CHMP guideline on the choice of the non-inferiority margin. With regard to the definition of the non-inferiority margin, the authors argued again that it is not appropriate to define the non-inferiority margin as a proportion between active comparator and placebo. Such ideas were formulated with the aim of ensuring that the test product was superior to (a putative) placebo; however they may not achieve this purpose. If the reference product has a large advantage over placebo this does not mean that large differences are unimportant, it just means that the reference product is very efficacious. These statements are too strong and demand further discussion (Hauschke and Hothorn, 2007, Brown and Day, 2007). The purpose of this lecture is to point out that (i) the ratio-based approach is superior to the difference-based approach with regard to sample size allocation, which is also of importance from an ethical viewpoint, and (ii) the ratio-based approach often reflects clinical practice and is applied by statisticians and clinicians to define what constitutes a clinically irrelevant difference.

References:

BROWN, D., VOLKERS, P. and DAY, S. (2006): An Introductory Note to CHMP Guidelines: Choice of the Non-inferiority Margin and Data Monitoring Committees. *Statistics in Medicine, 25, 1623–1627.*

HAUSCHKE, D. and HOTHORN, L. (2007): Letter to the Editor. *Statistics in Medicine, 26, 230–233.*

BROWN, D. and DAY, S. (2007): Author's Reply. *Statistics in Medicine, 26, 234–236.*

Session:

S09: Non-Inferiority Trials (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

A GIANT on its way: meta-analyses of genome-wide association studies

Iris M. Heid[1,2], Claudia Lamina[2], the GIANT consortium, Christian Gieger[1,2], and H.-Erich Wichmann[1,2]

[1] Helmholtz Center Munich, Institute of Epidemiology, Neuherberg, Germany; heid@gsf.de
[2] Institute of Medical Informatics, Biostatistics, and Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany

Abstract: Genome-wide association (GWA) analyses require particularly large sample sizes due to the high-dimensional testing. Meta-analyses of worldwide available genome-wide scans now enable to improve power, but pose several methodological issues on how to meta-analyse GWA studies.

We illustrate the challenges of GWA meta-analyses on the example of a pooling initiative of GWA data on body-mass-index (BMI), the worldwide GIANT (Genome-wide investigation of anthropometric parameters) initiative. Challenges include (i) population stratification varying per study, (ii) phenotype homogenization, (iii) only partially joint SNP panels from different genotyping platforms, (iv) statistical model standardization, (v) an appropriate pooling method, which is compromised by the fact that genome-wide testing of heterogeneity is not meaningful.

In the GIANT project, we ensured comparability of phenotype distribution and model adjustments by computing normalized Z-scores for BMI. Lambda factors for population stratification were computed per study. Linear regression on the Z-scores was performed and the weighted Z-score method to pool p-values was applied for genome-wide scan tests. For the strongest associations, testing for heterogeneity between study effects was performed and beta-estimates were computed by fixed or random effects models. We analyzed 6 studies involving 15,881 participants (and currently add 6 more studies) at 2,2 Million SNPs. While p-values in single studies did not fall below 10-5, pooled p-values were as low as 10-14 in the pooled analysis. The fact that known candidate genes for BMI were re-identified is a proof-of-principle. We conclude that with the pooled GWA studies using appropriate standardization, we have the power to infer new stable associations from GWA studies without tedious and expensive multiple replication steps.

Session:
S20: Genome-wide association studies (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Bayesian model selection by Markov Chain Monte Carlo

Leonhard Held

University of Zurich, Switzerland; leonhard.held@ifspm.uzh.ch

Abstract: In this talk, I will provide an introduction to Bayesian model selection based on Markov chain Monte Carlo simulation. I will illustrate the methodology using two illustrative examples. The first describes Bayesian model selection in fractional polynomial regression (Sabanes Bove and Held, 2007). The second outlines Bayesian variable selection for detecting adaptive genomic differences among populations (Riebler, Held and Stephan, 2007).

References:

RIEBLER, A., HELD, L. and STEPHAN, W. (2008): Bayesian variable selection for detecting adaptive genomic differences among populations. *Genetics, to appear.*

SABANES BOVE, D. and HELD, L. (2007): Bayesian fractional polynomials. *Technical Report, University of Zurich.*

Session:

S34: Bayesian models in biostatistics (II) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Two level frailty models for institutional quality assessment

Volkmar Henschel[1] and Jutta Engel[2]

[1] Institute for Medical Informatics, Biometry and Epidemiology (IBE), University of Munich, Germany; henschel@ibe.med.uni-muenchen.de
[2] Munich Cancer Registry at IBE; engel@ibe.med.uni-muenchen.de

Abstract: Register data contain limited information about individual cases. Tumour registries routinely document specific prognostic factors but no comorbidity, for example. Individual random effects can be used to compensate this missing information. In the analysis of survival data such a random effect is called frailty.

The institutional effect on the prognosis can be assessed by introducing institution as a covariable or by a random effect on the institutional level. Assessing institutional quality and adjusting for unobserved prognostic factors consequently establishes a two level random effects structure.

We analyze data on the survival of 9943 colon cancer patients whose data have been collected in the Munich Cancer Registry. The main interest of this analysis is the influence of clinics on patients' survival. 48 clinics with very different hospital volume are involved.

A model for such a hierarchical survival analysis can easily be constructed with a proportional hazards model by a Bayesian approach. The Cox proportional hazards model can be considered as a generalized linear model where the baseline hazard is known. Gamerman (1997) describes how to sample effectively the regression coefficients in a generalized linear model. The log baseline hazard can be modeled by constant or cubic B-splines. The coefficients can be sampled from a Gaussian Markov Random Field, see Rue (2001). The above model incorporates easily the two level random effects structure which is specific for the problem above. Log-normal as well as gamma frailty terms can be used.

References:

GAMERMAN, D. (1997): Sampling from the posterior distribution in generalized linear mixed models. *Statistics and Computing, 7, 57–68.*

RUE, H. (2001): Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society (B), 63, 325–338.*

Session:

S32: Hierarchical models : Wednesday, 12/03/2008, 11:00am - 12:40pm

Calculating the NNT in RCTs with time to event data: a systematic review

Mandy Hildebrandt[1], Elke Vervölgyi[2], and Ralf Bender[2]

[1] Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany, and Institute for Medical Biometry, Epidemiology, and Informatics (IMBEI), University of Mainz, Germany; mandy.hildebrandt@iqwig.de
[2] IQWiG, Cologne, Germany; elke.vervoelgyi@iqwig.de, ralf.bender@iqwig.de

Abstract: The NNT is the estimated average number of patients needed to be treated with a new treatment to prevent one adverse event. Adequat NNT estimation for time-to-event data can be performed either by means of risks estimated by the Kaplan-Meier survival curve [Altman (1999)] or by means of estimated hazards if important limitations are taken into account [Lubsen (2000), Mayne (2006)]. We searched PubMed for RCTs published in 2003-2005 in frequently cited journals. We included RCTs with parallel group design and individual randomization. We investigated the use of NNTs to present results, adequacy of the calculation method, and reporting of confidence intervals (CI). Of 734 eligible articles, 376 reported time-to-event outcomes and 63 NNTs, 38 with time-to event and 25 binary data. Of those 38 articles only 18 applied appropriate methods for NNT calculation. CIs were presented in 21 of 63 RCTs. For time-to-event data incorrect methods were applied frequently. CIs should be reported to indicate the precision of the estimate. In summary, there is much room for improvement in the application of the NNT to present results of RCTs, especially where the outcome is time to an event.

References:

ALTMAN, D. and ANDERSEN, P.K. (1999): Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ, 319, 1492-1495.*

LUBSEN, J., HOES, A., and GROBBEE, D. (2000): Implications of trial results: the potentially misleading notions of number needed to treat and average duration of live gained. *The Lancet, 356, 1757-1759.*

MAYNE, T.J., WHALEN, E., and VU, A. (2006): Annualized was found better than absolute risk reduction in the calculation of number needed to treat in chronic conditions. *Journal of Clinical Epidemiology, 59, 217-223.*

Session:

S28: Clinical studies (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Simulation and Inference for a Spatio-temporal Stochastic Epidemic Model

Michael Höhle[1,2]

[1] Department of Statistics, University of Munich, Germany;
michael.hoehle@stat.uni-muenchen.de
[2] Munich Center of Health Sciences, Germany

Abstract: This work contains a contribution to the cooperation of virologists, epidemiologists and statisticians on determining ecological drivers of infectious disease dynamics. A stochastic model based on conditional intensities is proposed, which extends the stochastic susceptible-infected-recovered (SIR) model to a stochastic spatio-temporal process. In this extended model the transmission rate is parametrized by a baseline hazard function and covariates, e.g. distance kernel and (possibly time-varying) covariates of the susceptible and infectious individuals.

Simulation from the model is based on Ogata's modified thinning algorithm, while inference can be based on either the full likelihood of the stochastic process or the partial likelihood approach proposed by Diggle (2006). Here, conditioning on the locations and time-points of the events a partial log-likelihood can be formulated, which eliminates the baseline hazard. Inference for covariate effects is then straightforward by using standard maximization algorithms such as Broyden-Fletcher-Goldfarb-Shanno. We compute maximum likelihood estimators, determine standard errors, compare with simulations and use Akaikes information criterion to decide on the inclusion of covariates and distance kernels.

As an illustration we analyse data provided by the Federal Research Centre for Virus Diseases of Animals, Wusterhausen, Germany, on the incidence of classical swine fever virus in Germany during 1993-2004. The presented methods will be made available in the new version of the R-package `RLadyBug`.

References:

DIGGLE, P. (2006): Spatial-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research, 15, 325–336.*

HÖHLE, M. and FELDMANN, U. (2007): RLadyBug – An R package for stochastic epidemic models. *Computational Statistics and Data Analysis, 52, 680-686.*

Session:
S24: Spatial Analysis of Surveillance Data : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Model choice in Cox-type additive hazard regression models with time varying-effects

Benjamin Hofner[1], Wolfgang Hartl[2], Helmut Küchenhoff[1], and Thomas Kneib[3]

[1] Statistical Consulting Unit, Department of Statistics, Ludwig-Maximilians-Universität München, Akademiestr. 1, 80539 Munich, Germany; benjamin.hofner@stat.uni-muenchen.de
[2] Department of Surgery, Campus Großhadern, LMU München
[3] Department of Statistics, LMU München

Abstract: In recent years, flexible hazard regression models based on penalised splines have been developed that allow to extend the classical Cox-model via the inclusion of time-varying and nonparametric effects. Despite their immediate appeal in terms of flexibility, these models introduce additional difficulties when performing model choice and variable selection. In our application, we present an analysis of retrospective data where the aim is to explore whether surgical patients with severe sepsis have a treatment benefit in terms of 90-day survival from an activity-guided antithrombin III (AT 3) therapy. Due to the retrospective nature of the study, a sensible confounder model has to be selected before testing for an AT 3 effect.

We propose a two-stage stepwise selection strategy that devises a model choice procedure with competing modelling possibilities within the choice set of possible confounder variables. The procedure is called two-stage since it proceeds both on the stage of the covariates (which covariates should be included in the model?) and the different modelling possibilities (in which way should a covariate enter the model). For categorical covariates, competing modelling approaches are fixed effects and time-varying effects, whereas nonparametric modelling provides a further alternative in case of continuous covariates. In each forward step, all covariates not already contained in the model are included with any possible type of effect and the best-fitting variant according to the conditional AIC is added. All remaining modelling possibilities for this covariate are removed from the choice set in sub-sequent steps. Each forward step is followed by a backward selection step where each covariate is tested for AIC improvement when being excluded from the model.

In our data analysis we identified a confounder model containing both smooth effects and time-varying effects. No treatment benefit (neither time-varying nor time-constant) for AT 3 could be shown when using the flexible confounder model.

Session:
S13: Event Data Analysis (I) : Tuesday, 11/03/2008, 9:10am - 10:30am

Adaptive change of hierarchy of endpoints in clinical trials after an interim analysis

Jakob Duncker[1] and Gerhard Hommel[1]

[1] Institute of Medical Biostatistics, Epidemiology and Informatics, D-55101 Mainz, Germany; hommel@imbei.uni-mainz.de

Abstract: When an interim analysis is performed within a clinical trial, it is possible to make adaptations of the design, e.g. changing the preplanned number of patients. Moreover, it is also possible to modify the structure of testing the hypotheses, e.g. to drop endpoints, to change an a priori order of endpoints or even to include new endpoints (Hommel, 2001). Nevertheless, the use of these techniques in practice can be very problematic because of the danger of loss of scientific integrity.

Kieser (2005) has shown that a complete change of an a priori order of endpoints does not only lead to integrity problems, but results also in a reduced power. We investigated whether a more "careful" change would lead to better results. Such a change could only be allowed when large power differences between the endpoints can be expected from the interim results; moreover, instead of a strict a priori order weights for the endpoints can be assigned.

For the case of two bivariate normally distributed endpoints we performed a simulation study comparing different strategies. It turned out that careful changes lead to more powerful results than a rigid change of a priori order. Nevertheless, the most robust method consists in assigning equal weights and no change after the interim analysis, as proposed by Kieser et al. (1999).

References:

HOMMEL, G. (2001): Adaptive Modification of Hypotheses after an Interim Analysis. *Biometrical Journal, 43, 581–589.*

KIESER, M. (2005): A Note on Adaptively Changing the Hierarchy of Hypotheses in Clinical Trials with Flexible Design. *Drug Information Journal, 39, 215–222.*

KIESER, M., BAUER, P. and LEHMACHER, W. (1999): Inference on Multiple Endpoints in Clinical Trials with Adaptive Interim Analyses. *Biometrical Journal, 41, 261–277.*

Session:

S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

Boosting for high-dimensional data: Applications and software

Torsten Hothorn

Institut für Statistik, LMU München, Munich, Germany;
torsten.hothorn@stat.uni-muenchen.de

Abstract: Classically, boosting or functional gradient descent algorithms for optimizing various empirical risk functions have been implemented using relatively complex base-learners such as regression trees. The resulting regression fit is typically a 'black box machine', i.e., one can predict the outcome based on the covariate status of a new observation but the fitted functional form is too complex to be readable by human beings.

Recently, boosting algorithms for fitting generalized linear or additive models have been suggested. The key innovation is the application of componentwise linear models or smoothing splines which allows us to reformulate the regression fit in terms of classical linear or additive models. In the former case, the regression coefficients can be interpreted in the usual way. Moreover, those boosting algorithms have been demonstrated to be useful for variable selection in high-dimensional situations.

The 'mboost' add-on package implements generalized linear ('glmboost'), generalized additive ('gamboost') and classical tree-based ('blackboost') models which are fitted via functional gradient descent. Similar to 'glm', the users are allowed to specifiy new loss functions via 'Family' objects whose empirical risk is to be optimized. Both classical and corrected AIC formulations are available for the Gaussian situation which can be used to determine the optimal number of boosting iterations.

We briefly sketch the underlying theory and demonstrate how to actually fit regression models with many covariates in the R system for statistical computing using 'mboost'.

Session:
S30: Boosting for biomedical data : Wednesday, 12/03/2008, 11:00am - 12:40pm

Extremes in biometrics

Jürg Hüsler

Department of Mathematical Statistics, University of Bern, Switzerland;
juerg.huesler@stat.unibe.ch

Abstract: In our talk we discuss several recent applications of extremes in biometrics. These applications focus on the wide possibilities of extreme value analysis and point on methodological trends of this topic.

Session:
Close: Closing session : Thursday, 13/03/2008, 12:50pm - 1:30pm

Association between a Prognostic Gene Signature and Functional Gene Sets

Manuela Hummel[1], Klaus H. Metzeler[2], Stefan K. Bohlander[2], Christian Buske[2], and Ulrich Mansmann[1,3]

[1] IBE, University of Munich, Germany; hummel@ibe.med.uni-muenchen.de
[2] Laboratory of Leukemia Diagnostics, Department of Internal Medicine III, University Hospital Großhadern, University of Munich, Germany
[3] Department of Statistics, University of Munich, Germany

Abstract: The development of expression based gene signatures for predicting survival time or a binary clinical endpoint is a very popular and competing task in microarray data analysis. Besides stringent validation of the propagated signatures, we think another important point is the exploration of their biological context.

In this talk we propose some ideas for the assessment of the relation between a prognostic signature and functional gene sets like pathways or Gene Ontology categories. We apply the methods to an AML data set and a corresponding score for the prediction of survival based on the expression of selected genes. The connection between the single signature genes and pathways is explored by 1) hierarchical variable selection and 2) gene association networks. The signature is used to develop a continuous risk score for patient survival. We explore 3) the effects of the risk score and its interaction with other known risk factors on the expression patterns within Gene Ontology categories.

Such kind of analysis helps to deepen the insights into the biological mechanisms of a signature.

References:

GOEMAN, J.J. and MANSMANN, U. (2008): Multiple testing on the directed acyclic graph of Gene Ontology. *Bioinformatics*.

HUMMEL, M. and MEISTER, R. and MANSMANN, U. (2008): GlobalANCOVA: exploration and assessment of gene group effects. *Bioinformatics, 24(1), 78–85.*

MEINSHAUSEN, N. (to appear): Hierarchical testing of variable importance. *Biometrika*.

METZELER, K.H. et al. (Work in progress): An 86-probeset gene expression signature can predict survival in AML with normal karyotype independently of FLT3 internal tandem duplications and NPM1 mutational status.

SCHAEFER, J. and STRIMMER, K. (2005): A shrinkage approach to large-scale covariance estimation and implications for functional genomics. *Statist Appl Genet Mol Biol, 4(32).*

Session:

S38: Biological Networks : Thursday, 13/03/2008, 9:10am - 10:30am

Challenges and Some Regulatory Experiences with Non-Inferiority Trial Design without Placebo Arm

H.M. James Hung

Division of Biometrics I, Office of Biostatistics, OTS, CDER, FDA;
hsienming.hung@fda.hhs.gov

Abstract: For a non-inferiority design without a placebo arm, the direct comparison between the test treatment and the selected positive control is in principle the only basis for statistical inference. Therefore, evaluating the test treatment relative to the non-existent placebo presents extreme challenges and requires some kind of bridging from the past to the present because of no current placebo data. For such inference based partly on an indirect bridging manipulation, fixed margin method and synthesis method are two widely discussed methods in recent literature. There are major differences in statistical inference paradigm between the two methods. The fixed margin method employs the historical data for the performances of the active control versus a placebo to guide the selection of the non-inferiority margin. Such guidance is not part of the ultimate statistical inference in the non-inferiority trial. In contrast, the synthesis method connects the historical data to the non-inferiority trial data for making broader inferences relating the test treatment to the non-existent current placebo. The issues and further thoughts surrounding these methods and choice of the non-inferiority margin will be discussed through some typical experiences from reviews of regulatory submissions.

Session:
S11: Non-Inferiority Trials (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Confidence rectangles in calibration using a parametric bootstrap

Bernd-Wolfgang Igl and Lutz Dümbgen

IMBS, Universität zu Lübeck, Germany; igl@imbs.uni-luebeck.de

Abstract: In certain multivariate regression scenarios, a vector of dependent variables y is linearly related to a vector of explanatory variables x. In multivariate calibration problems for example arising in analytical chemistry estimates of the regression parameters are available. Then, the main aim is to predict x from replicated measurements of y. Obviously, besides of point estimates of x one is interested in confidence regions. However, in many experimental situations it is useful to compute intervals for each component of the vector x. To this end, we propose a confidence rectangle for x based on a parametric bootstrap. In addition to inference about x, these confidence regions may be used to assess the quality of the training data and the difficulty to reconstruct x from y.

Session:
S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Combination of diagnostic markers with regard to either 100% sensitivity or 100% specificity

Klaus Jung and Edgar Brunner

Department of Medical Statistics, Georg-August-University Göttingen, D-37073 Göttingen, Germany; kjung1@uni-goettingen.de

Abstract: In medical diagnostics, where a patient is to be assigned to one of two classes (e.g. healthy or diseased), a combination of multiple markers is mostly designed in such a way that the area under the ROC curve is maximized (cf. Pepe *et al.*, 2006). Furthermore, a reasonable balance between sensitivity and specificity is usually desired. Oftentimes, the cut-off points for the markers are then chosen with regard to a maximal Youden-index (= sensitivity + specificity - 1).

In some other situations, however, a false negative classification is much worse than a false positive one. In the extreme case, 100% sensitivity is demanded, whereas specificity is only of secondary note. The cut-off point is then chosen from that position at the ROC curve where specificity is maximal under the condition that sensitivity is 1. (Of course, one can also consider situations where specificity is more important than sensitivity.)

In my talk, I will focus on these extreme cases when in addition more than one marker is available. I propose a combination of markers in a multi-stage procedure under the major aspect of achieving 100% sensitivity on the one hand, and reducing the number of false positives as minor aspect on the other hand. I demonstrate the usefulness of this procedure on an example of real clinical data. The classification errors are evaluated by cross-validation and bootstrapping (cf. Efron and Tibshirani, 1997).

References:

EFRON, B. and TIBSHIRANI, RW. (1997): Improvements on Cross-Validation: The .632+ Bootstrap Method. *JASA, 92, 548–560.*

PEPE, M.S., CAI, T. and LONGTON, G. (2006): Combining Predictors for Classification Using the Area under the Receiver Operating Characteristic Curve. *Biometrics, 62, 221–229.*

Session:

S18: Clinical studies (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Survival Models with Gene Groups as Covariates

Kai Kammers and Jörg Rahnenführer

Technische Universität Dortmund, Fakultät Statistik, 44221 Dortmund, Germany;
kammers@statistik.uni-dortmund.de

Abstract: Our goal is to obtain prediction models for survival times from high dimensional gene expression data, with high level of interpretability of the estimated models. The models should be complex enough to have high prediction accuracy but should also have a parsimony property to make proper interpretations.

Typical univariate approaches use single genes as covariates in survival time models, multivariate models perform dimension reduction through gene selection. Analysis of time-dependent ROC curves and the area under the curves (AUC) can be used to assess the predictive performance [see Gui and Li (2005)]. The interpretability can be improved by combining genes to gene groups (biological processes or molecular functions) and using these groups as covariates. The hierarchically ordered "GO groups" (Gene Ontology) are particularly suitable. Cox models are used for detecting covariates that are significantly correlated with survival times. Based on these models statistical procedures like the LARS-Cox-method (L_1 shrinking method) [see Gui and Li (2005)] are applied for variable selection.

Our aim is the combination of methods for survival prediction with biological *a priori* knowledge. First, we compare the prediction performance of models using single genes as covariates with models using gene groups as covariates. Then, we integrate GO graph structure in the models [see Alexa, Rahnenführer and Lengauer (2006)] in order to cope with the high correlations between neighboring GO groups.

References:

GUI, J. and LI, H. (2005): Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics 21(13), 3001–3008.*

ALEXA, A., RAHNENFÜHRER, J. and LENGAUER, T. (2006): Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics 22(13), 1600–1607.*

Session:

S01: Systems Biology and Bioinformatics : Monday, 10/03/2008, 9:10am - 10:30am

Application of canonical variate analysis for presenting ecological requirements of aquatic plant communities

Dariusz Kayzer[1] and Krzysztof Szoszkiewicz[2]

[1] Department of Mathematical and Statistical Methods, August Cieszkowski Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; dkayzer@au.poznan.pl
[2] Department of Ecology and Environmental Protection, August Cieszkowski Agricultural University, Pitkowska 94C, 60-649 Poznań, Poland; kszoszk@au.poznan.pl

Abstract: Analysis based of results of river survey carried out in UK and Northern Ireland in the period 1995-2001 where 288 releves were described. Field assessment was undertaken according to phytosociological method by Braun-Blanquet approach. Common types of plant communities were analysed, which were classified as 12 phytosociological associations. These represent a wide ecological range. The environmental requirements were classified basing on the 10 variables, which reflect different kinds of geographical conditions as geographical, landscape, geology, land use of the watershed and water quality. The aim of the study was presenting usefulness of canonical variate analysis for estimating environmental requirements of plant communities. The position of different vegetation types were shown in the space of canonical variates. Canonical variates appeared as useful for detecting ecological needs of vegetation in the multivariate matrix.

Session:
S43: Statistical Methods in Environmental Monitoring : Thursday, 13/03/2008, 11:00am - 12:40pm

Blinded Sample Size Reassessment in Non-inferiority Studies: Review and New Results

Meinhard Kieser[1] and Tim Friede[2]

[1] Institute of Medical Biometry and Informatics, University of Heidelberg, 69120 Heidelberg, Germany; meinhard.kieser@imbi.uni-heidelberg.de
[2] Warwick Medical School, University of Warwick, Coventry CV4 7AL, United Kingdom; T.Friede@warwick.ac.uk

Abstract: In the planning phase of a clinical trial there is generally some uncertainty with respect to parameters relevant for sample size calculation. Hence, performing a study in a fixed sample size design implies a considerable risk of resulting in an inadequate sample size. Several designs have been proposed that allow a mid-course sample size reassessment. The recent CHMP Reflection Paper on adaptive designs (CHMP, 2007) gives clear preference to methods for sample size recalculation that are based on blinded data only compared to unblinded procedures. The internal pilot study design is appropriate for this purpose and can be described as follows (Wittes and Brittain, 1990). After an initial sample size calculation in the planning stage, the sample size is reassessed based on nuisance parameter estimates (e.g., variance or overall event rate) that are obtained when a portion of the planned sample has completed the study. In our presentation, we show for various test problems how blinded sample size reassessment can be performed in non-inferiority trials with internal pilot study design (see for example Friede and Kieser, 2003, and Friede et al., 2007). We investigate the characteristics of the procedures with particular emphasis on: ability to assure the desired power under misspecifications in the planning stage, actual type I error rate and methods to control the nominal significance level, and differences to superiority trials.

References:

COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (2007): *Reflection Paper on Methodological Issues in Confirmatory Clinical Trials Planned with an Adaptive Design* (CPMP/EWP/2459/02). European Medicines Agency, London.

FRIEDE, T. and KIESER, M. (2003): Blinded sample size reassessment in non-inferiority and equivalence trials. *Statistics in Medicine, 22, 995–1007.*

FRIEDE, T., MITSCHELL, C., MÜLLER-VELTEN, G. (2007): Blinded sample size reestimation in non-inferiority trials with binary endpoints. *Biometrical Journal, to appear.*

WITTES, J. and BRITTAIN, E. (1990): The role of internal pilot studies in increasing the efficiency of clinical trials. *Statistics in Medicine, 9, 65–72.*

Session:

S09: Non-Inferiority Trials (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Maximum likelihood estimator in growth curve model with serial correlation structure

Daniel Klein and Ivan Žežula

Institute of Mathematics, Šafárik University, Jesenná 5, 04154 Košice, Slovak Republic; daniel.klein@upjs.sk, ivan.zezula@upjs.sk

Abstract: The standard growth curve model with serial correlation structure is considered. Maximum likelihood estimator of correlation coefficient is derived.

References:

POTTHOFF, R.F. and ROY, S.N. (1964): A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika, 51, 313–326.*

WU, Q.G. (2000): Some results on parameter estimation in extended growth curve models. *Journal of Statistical Planning and Inference, 88, 285–300.*

Session:

S21: Advances in Statistical Modelling (I) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Model choice and variable selection in geoadditive regression models

Thomas Kneib, Torsten Hothorn, and Gerhard Tutz

Department of Statistics, Ludwig-Maximilians-University Munich, Ludwigstraße 33, 80539 Munich, Germany; thomas.kneib@stat.uni-muenchen.de

Abstract: Model choice and variable selection are issues of major concern in practical regression analyses. This is not only true in generalised linear models but becomes even more obvious in geoadditive regression models comprising spatial effects, nonparametric effects of continuous covariates, interaction surfaces, random effects, and varying coefficient terms: Should a continuous covariate be included into the model at all and if so as a linear effect or as a nonparametric, flexible effect? Is a spatial effect required in the model, i.e., is spatial correlation present beyond the spatial variation accounted for by spatially varying covariates? Are some of the covariate effects spatially varying? To answer these questions, we propose a systematic, fully automated componentwise boosting procedure. The major modelling component are penalised splines and their bivariate tensor product extensions. All base-learning procedures can be cast into a general modelling framework leading to simple penalised least-squares fits. This allows to devise a generic componentwise boosting procedure for a comprehensive model class. One major difficulty is to obtain base-learners that are comparable in complexity to avoid biased selection towards more flexible effects. The equivalent degrees of freedom of a flexible effect will be used as a general measure of complexity for different types of base-learners. We demonstrate the versatility of our approach with two examples: a geoadditive Poisson regression model for species counts in habitat suitability analyses and a geoadditive logit model for the analysis of forest health.

References:

KNEIB, T., HOTHORN, T. and TUTZ, G. (2007): Model choice and variable selection in geoadditive regression models, Department of Statistics Technical Report No. 3. Available from http://www.statistik.lmu.de/~kneib

Session:

S30: Boosting for biomedical data : Wednesday, 12/03/2008, 11:00am - 12:40pm

EU regulatory guidance on phase II / III combination trials

Armin Koch

Bundesinstitut für Arzneimittel und Medizinprodukte, Germany; a.koch@bfarm.de

Abstract: European Guidance on adaptive designs in late stage drug development is now available. The respective reflection paper also comments on the acceptability of phase II / phase III combination trials planned with such an approach. This presentation summarizes the key considerations and the argumentation for the positions that are held. Regulatory experience with results from clinical trials planned with an adaptive design is still limited. However, various proposals have been presented for discussion and this will be summarized, in addition.

Session:
S23: Adaptive seamless design for combining phase II / III clinical studies : Tuesday, 11/03/2008, 4:00pm - 6:10pm

What are genome-wide association studies?

Inke R. König and Andreas Ziegler

Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, 23538 Lübeck, Germany; inke.koenig@imbs.uni-luebeck.de

Abstract: The past year has seen a surge in publications of genome-wide association studies (GWAs) investigating the genetic background of complex diseases. To name just a few, this study design was used to identify and validate genetic factors for myocardial infarction (Samani et al., 2007), colorectal cancer (Zanke et al., 2007), breast cancer (Hunter et al., 2007), and six further common diseases (The Wellcome Trust Case Control Consortium, 2007). Thus, GWAs appear to promise a significant step forward to successfully identify the genetic background of many common diseases.

This presentation will provide an overview of the recent developments of study designs for genetic association. Specifically, it will describe in what ways GWAs are different from other study designs (Ziegler et al., 2007), which will be outlined using examples from the past months. Finally, possible future paths will be discussed.

References:

HUNTER, D.J. ET AL. (2007): A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nature Genetics, 39, 870–874.*

SAMANI, N.J. ET AL. (2007): Genome-wide association analysis of coronary artery disease. *The New England Journal of Medicine, 357, 443–453.*

THE WELLCOME TRUST CASE CONTROL CONSORTIUM (2007): Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447, 661–678.*

ZANKE, B.W. ET AL. (2007): Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nature Genetics, 39, 989–994.*

ZIEGLER, A., KÖNIG, I.R., THOMPSON, J.R. (2007): Biostatistical Aspects of Genome-Wide Association Studies. *Biometrical Journal, in press.*

Session:

S07: Genome-wide association studies (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Treatment Selection in Adaptive Group-sequential Designs

Franz König[1], Nigel Stallard[2], and Tim Friede[2]

[1] Section of Medical Statistics, Medical University of Vienna, Vienna, Austria; franz.koenig@meduniwien.ac.at
[2] Warwick Medical School, The University of Warwick, Coventry, UK; n.stallard@warwick.ac.uk, t.friede@warwick.ac.uk

Abstract: When competing medical treatments exist it might be desirable to intially compare all of these to a common control and then reduce the set of treatments in a stepwise fashion as information regarding their efficacy and safety is accrued. The ultimate aim is to demonstrate that at least one of these treatments is efficacious. Whereas traditionally treatment selection and proof of efficacy would have been conducted in different trials, recently clinical trial designs have been proposed to achieve both treatment selection and proof of efficay within a single study (Bauer and Kieser, 1999; Stallard and Todd, 2003; Bretz et al, 2006; Koenig et al, 2007).

In this presentation we describe a flexible design based on group-sequential tests described by Kelly et al (2005) and Stallard and Friede (2007), and the closed test principle. Full flexibility is achieved through either application of the conditional error function approach or formulation of the group-sequential test in terms of combination functions. The different approaches are compared in a simulation study and other adaptions such as sample size reassessment are discussed.

References:

BAUER, P. and KIESER, M. (1999): Combining different phases in the development of medical treatments within a single trial. *Stat Med, 18, 1833–1848.*

KELLY, P.J., STALLARD, N. and TODD, S. (2005): An adaptive group sequential design for phase II/III clinical trials that select a single treatment from several. *Journal of Biopharmaceutical Statistics, 15, 641–658.*

KOENIG, F., BRANNATH, W., BRETZ, F. and POSCH, M. (2007): Adaptive Dunnett tests for treatment selection. *Statistics in Medicine, to appear.*

STALLARD, N. and FRIEDE, T. (2007): Flexible group-sequential designs for clinical trials with treatment selection. *Submitted.*

STALLARD, N. and TODD, S. (2003): Sequential designs for phase III clinical trials incorporating treatment selection. *Statistics in Medicine, 22, 689–703.*

Session:
S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

Nonparametric analysis of high dimensional repeated measures designs

Frank Konietschke and Edgar Brunner

Department of Medical Statistics Universitity of Göttingen, Germany;
fkoniet@gwdg.de, brunner@ams.med.uni-goettingen.de

Abstract: Time course experiments are very frequent in medical research. In such designs n subjects are observed repeatedly at d time points. Such data are called high dimensional if the number of dimension d is much larger than the number of independent subjects n. Also repeated measures designs with d conditions can be covered by similar statistical models.

So far, a few parametric procedures are known which can handle high dimensional data. In the case, where the data do not follow a normal distribution, in particular for ordinal data, there does not exist any procedure for the analysis of such data taking into account an arbitrary covariance structure.

In this talk we present some new ideas for the nonparametric analysis of such data. The so-called relative marginal effects are used to define nonparametric effects which can be consistently estimated by different types of rankings.

Within this framework it is also possible that the repeated measures may have a factorial structure. Statistics for testing general nonparametric hypotheses about the underlying factors within the repeated measures are derived and their asymptotic distributions are approximated by χ^2-distributions. It is demonstrated by simulations that the large sample approximations are quite satisfactory, even for sample sizes as small as 10. The application of the new methods is demonstrated by means of a real data example.

References:

AHMAD, R., WERNER, C., BRUNNER, E. Analysis of high dimensional repeated measures designs. The one sample case. *Computational Statistics and Data Analysis. Submitted.*

WERNER, C., BRUNNER, E. (2004). Dimensionsstabile Approximation für Verteilungen von zufälligen quadratischen Formen im Repeated-Measures Design. *Technical Report. Abteilung für Medizinische Statistik, Universität Göttingen.*

Session:

S35: Nonparametric and Parametric Multivariate Tests in High Dimensions : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Assessing agreement in imaging studies with continuous measurements - a case study

Annette Kopp-Schneider[1], Marc Weimer[1], and Michael Fabel[2]

[1] Dept. Biostatistics, DKFZ, D-69120 Heidelberg, Germany;
[2] Dept. Radiology, DKFZ, D-69120 Heidelberg, Germany; kopp@dkfz.de

Abstract: A common goal in radiological studies is the replacement of standard imaging techniques by modern methods which are faster or more convenient to evaluate and/or reduce the burden for the patient. It is essential to make sure that the new measurement is reliable and accurate before it is used in practise. Since the true value of the measurement is usually not known, the novel method is compared to the standard. In addition, reliability is assessed by comparing results from different readers.

Data were obtained from a study investigating the feasibility of semi-automated volumetric analysis of lymph node metastases compared to manual volumetric analysis and RECIST. Two readers evaluated about 200 lesions with three methods. An evaluation will be presented using Bland-Altman plots (Bland and Altman 1999) and Concordance Correlation Coefficient (Barnhart et al. 2007). The use of different methods to quantify agreement will be discussed and their interpretations will be compared. Conclusions will be shown regarding reliability and accuracy of the imaging techniques.

References:

BLAND, J.M. and ALTMAN, D.G. (1999): Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8, 135–160.*

BARNHART, H.X., HABER, M.J. and LIN L.L. (2007): An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics, 17, 529–569.*

Session:

S02: Diagnostic studies : Monday, 10/03/2008, 9:10am - 10:30am

What a difference a year makes: lessons learned from early genome-wide association scans

Peter Kraft

Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Bldg 2 Room 207, 665 Huntington Ave, Boston, MA 02115 USA;
pkraft@hsph.harvard.edu

Abstract: Long argued to be well-suited to detect common genetic variants with modest effects hypothesized to influence complex traits, genome-wide association studies (GWAS) only recently became practical thanks to advances in knowledge of patterns of genetic variation and decreasing genotyping costs. Last year was a watershed, with scores of GWAS for multiple traits appearing, showing not only that GWAS were technically feasible, but that indeed detectable loci existed. For some complex traits such as diabetes, prostate and breast cancers, and inflammatory bowel disease, an order of magnitude more susceptibility loci were discovered in 2007 than in two decades of painstaking linkage and candidate gene studies.

Using the Cancer Genetic Markers of Susceptibility (CGEMS) project as an example, I review some of the main statistical challenges inherent in GWAS, including adjustment for multiple testing, cost-efficient design and analysis, and interpretation. I then look ahead to one of the possible next steps after discovery of multiple susceptibility loci by GWAS and discuss the difficulties in building and using the prognostic genetic risk scores envisioned by proponents of "personalized medicine".

Session:
S07: Genome-wide association studies (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Multivariate Tests Based on Pairwise Distance Measures

Siegfried Kropf[1], Daniela Kose[1], and Ludwig A. Hothorn[2]

[1] Institute of Biometry and Medical Informatics, Otto-von-Guericke-Universität Magdeburg, D-39120 Magdeburg, Germany; Siegfried.Kropf@med.ovgu.de, Daniela.Kose@med.ovgu.de
[2] Institute of Biostatistics, Leibniz Universität Hannover, D-30419 Hannover, Germany; Hothorn@biostat.uni-hannover.de

Abstract: If in multivariate tests the dimension of the sample vectors exceeds the sample size, then usually dimension reducing techniques as summary scores or principal components are applied. An alternative way for constructing exact multivariate tests consists in the utilization of pairwise similarity or distance measures of the sample elements as known, for example, in cluster analysis.

The use of permutation tests is straight-forward. But also exact parametric tests have been considered in the context of comparing independent samples (Kropf et al., 2007). It has been shown that tests based on pairwise distance or similarity measures often have a good power in high-dimensional data.

In the talk, a parametric version of such tests for a general linear model is considered. It utilizes rotation tests (Langsrud, 2005; Läuter et al., 2005). This is a Monte-Carlo technique for sample matrices with spherically distributed data. As a special application, we consider the use of these global multivariate tests for testing equivalence in multivariate data.

References:

KROPF, S., LUX, A., ESZLINGER, M., HEUER, H., and SMALLA, K. (2007): Comparison of Independent Samples of High-Dimensional Data by Pairwise Distance Measures. *Biometrical Journal, 49, 230–241.*

LANGSRUD, Ø. (2005): Rotation Tests. *Statistics and Computing, 15, 53–60.*

LÄUTER, J., GLIMM, E., and ESZLINGER, M. (2005): Search for Relevant Sets of Variables in a High-Dimensional Setup Keeping the Familywise Error Rate. *Statistica Neerlandica, 59, 298–312.*

Session:

S09: Non-Inferiority Trials (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Exact vs. Common Logistic Regression. A Comparison based on the Criteria "Feasibility" and "Maintaining of alpha-Level" for Testing the Regression Coefficient

Bertram Krumm

Central Institute of Mental Health, Department of Biostatistics, D-68159 Mannheim, Krumm; Bertram.Krumm@zi-mannheim.de

Abstract: The study was performed for the model of logistic regression (LR) with a binary dependent variable Y and one independent variable x. The most important hypothesis is H_0: $b = 0$ (independence). Common logistic regression (M-LR): The test statistic for H_0: $b = 0$ is based on the ML-estimator $\hat{b}$ of b and the large-sample normal distribution of $\hat{b}$. It can be proven, that the ML-estimator does not exist, if there is complete separation or quasi-complete separation of data points. Furthermore it is known, that asymptotic tests for testing the hypothesis of independence in contingency tables do not maintain the alpha-level if the sample is small. Exact logistic regression (E-LR): Here, for testing the hypothesis H_0: $b = 0$ a conditional exact test is performed which is based on the exact distribution of the sufficient statistic of the model. However, there are configurations of sample points, where the p-value cannot be calculated and exact tests are considered to be conservative. We compared the two methods of LR by a simulation study using the criteria feasibility (specific definition for each method) and rejection rate. The number of observations, the distribution of x and $P(Y = 1)$ for $x = 0$ were variied. Surprisingly, there was no exceedance of alpha-level with M-LR and E-LR was more liberal than M-LR.

References:

ALBERT, A. and ANDERSON, J.A. (1984): On the Existence of Maximum Likelihood Estimates in Logistic Regression Models. *Biometrika 71, 1, 1 - 10.*

CYTEL SOFTWARE CORPORATION (2004): LogXact, Version 6 with Cytel Studio, User Manual, Cytel Software Corporation, Cambridge, Massachusetts.

HIRJI, K.F., MEHTA, C.R. and PATEL, N.R. (1987): Computing Distributions for Exact Logistic Regression, *Journal of the American Statistical Association 82, 110 - 117.*

Session:

S21: Advances in Statistical Modelling (I) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Runoff Simulation in River Catchment Using Spatial Weather Generator

Leszek Kuchar and Sławomir Iwański

Department of Mathematics, Wrocław University of Environmental and Life Sciences, Poland; Leszek.Kuchar@gmail.com, slawek@aqua.ar.wroc.pl

Abstract: For the few past decades a lot of research work has been dedicated to the problem of changes in natural environment and their potential consequences. In particular, different climate change scenarios are considered, expected water cycle in river catchment, and hydrological risk. In the paper application of spatial weather generator SWGEN to simulate runoff in river catchment for future climate conditions is presented. Weather data required for hydrological rainfall-runoff-model are generated according to selected climate scenarios. Three parameter Gamma probability distribution to fit daily runoff on Kaczawa river (southwest of Poland) at discharge point Piatnica for present conditions (2000) and conditions of 2030 according to GISS, CCCM and GFDL are used, and gives a full information on moment characteristics, confidence intervals and critical values being an important tool for decision support system.

References:

GATES, W.L. (1985): The use of general circulation models in the analysis of the ecosystem impacts of climatic change. *Clim. Change, 7, 267-284.*

KATZ, R.W. (1996): Use of conditional stochastic models to generate climate change scenarios. *Clim. Change, 35, 397-414.*

KUCHAR, L. (2004): Using WGENK to generate synthetic daily weather data for modelling of agricultural processes. *Math. Comput. Simul., 65, 69-75.*

NILSEN, S.A. and HANSEN, E. (1973): Numerical solution of the rainfall runoff process on a daily basis. *Nordic Hydrology, 4, 171-190.*

RICHARDSON, C.W. (1985): Weather simulation for crop management models. *Trans. ASAE., 28, 1602-1606.*

Session:

S43: Statistical Methods in Environmental Monitoring : Thursday, 13/03/2008, 11:00am - 12:40pm

Clinical Validation of New Alarm Algorithms in Intensive Care

Silvia Kuhls[1], Ursula Gather[1], Michael Imhoff[2], Sylvia Siebig[3], and Christian Wrede[3]

[1] Fakultät Statistik, Technische Universität Dortmund, Germany; kuhls@statistik.uni-dortmund.de, gather@statistik.uni-dortmund.de
[2] Abteilung für Medizinische Informatik, Biometrie und Epidemiologie, Ruhr-Universität Bochum, Germany; mike@imhoff.de
[3] Klinik und Poliklinik für Innere Medizin I, Universitätsklinikum Regensburg, Germany; sylvia.siebig@klinik.uni-regensburg.de, christian.wrede@klinik.uni-regensburg.de

Abstract: In intensive care units the vital signs of a patient are continuously monitored. Alarms occur when the measured values surpass or fall below predefined alarm limits for a certain time. A large number of these so called threshold alarms are caused by outliers or short fluctuations around the alarm limit and have no clinical consequence (Imhoff and Kuhls, 2006). We describe the design and execution of a clinical validation study for a new alarm algorithm which is based on the method of robust online signal extraction (Gather, Schettlinger and Fried, 2006). Alarms are produced by comparing not the original measurements but the filtered signal with the alarm limits. The validation is performed off-line on the basis of an annotated reference database. This database comprises all physiological measurements sampled once per second, all alarms of the currently used alarm system, all alarm settings, and clinical annotations for 68 different patients with a total of 982 hours of monitoring. One third of the reference data is used for determining the appropriate window size of the filter and for developing additional alarm rules. The fully specified new algorithm is then validated on the remaining data using performance criteria such as properly adapted sensitivity and specificity.

References:

IMHOFF, M. and KUHLS, S. (2006): Alarm Algorithms in Critical Care Monitoring. *Anesthesia and Analgesia, 102, 1525–1537.*

GATHER, U., SCHETTLINGER, K. and FRIED, R. (2006): Online Signal Extraction by Robust Linear Regression. *Computational Statistics, 21, 33–51.*

Session:
S12: Freie Themen (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Accounting for haplotype phase uncertainty in LD estimation

Bettina Kulle Andreassen[1,2], A. Frigessi[1], H. Edvardsen[3,4], V. Kristensen[3,4], and L. Wojnowski[5]

[1] Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, PO Box 1122 Blindern, 0317 Oslo, Norway; bkulle@medisin.uio.no
[2] Department of Mathematics, University of Oslo, Norway
[3] Department of Genetics, Institute for Cancer Research, RR Medical Centre, Oslo, Norway
[4] Medical Faculty, University of Oslo, Oslo, Norway
[5] Department of Pharmacology, Johannes Gutenberg University Mainz, Germany

Abstract: The characterization of linkage disequilibrium (LD) is applied in a variety of studies including the identification of molecular determinants of the local recombination rate, the migration and population history of populations, and the role of positive selection in adaptation. LD suffers from the phase uncertainty of the haplotypes used in its calculation, which reflects limitations of the algorithms used for haplotype estimation. We introduce a LD calculation method, which deals with phase uncertainty by weighting all possible haplotype pairs according to their estimated probabilities as evaluated by PHASE. In contrast to the EM algorithm as implemented in the HAPLOVIEW and GENETICS packages, our method considers haplotypes based on the entire genetic information available for the candidate region. We tested the method using simulated and real genotyping data. The results show that, for all practical purposes, the new method is advantageous in comparison with algorithms that calculate LD using only the most probable haplotype or bilocus haplotypes based on the EM algorithm. The new method deals especially well with low LD regions, which contribute strongly to phase uncertainty. Altogether, the method is an attractive alternative to standard LD calculation procedures, including those based on the EM algorithm. We implemented the method in the software suite R, together with an interface to the popular haplotype calculation package PHASE.

References:

KULLE, B. and FRIGESSI, A. and EDVARDSEN, H. and KRISTENSEN, V. and WOJNOWSKI, L. (2008): Accounting for haplotype phase uncertainty in linkage disequilibrium estimation. *Genet Epidemiol, 32(2), 168–178.*

Session:

S03: Genetic Epidemiology (I) : Monday, 10/03/2008, 9:10am - 10:30am

Do Randomized and Non-Randomized Trials Yield Different Answers in Similar Populations? - Evidence from a 'Meta-Propensity Score' Analysis in Cardiac Surgery

Oliver Kuss[1], Thomas Legler[1], and Jochen Börgermann[2]

[1] Institute of Medical Epidemiology, Biostatistics, and Informatics, University of Halle-Wittenberg, Halle (Saale), Germany; oliver.kuss@medizin.uni-halle.de
[2] Department of Cardiac and Thoracic Surgery, Friedrich Schiller University Jena, Germany

Abstract: Randomized controlled trials (RCT) have limited external validity. A simple consequence of this fact is the limited *internal* validity of all systematic comparisons of RCTs and Non-RCTs: If RCTs are conducted in highly selected populations, but Non-RCTs in general populations, potential differences between RCTs and Non-RCTs are not necessarily due to missing randomisation. They might also arise from the different populations being involved. Ideally, we would like to conduct a 'meta-randomized' trial to systematically compare RCTs and Non-RCTs. That is, investigators willing to conduct a study on a specific clinical question would be randomly selected to perform a RCT or a Non-RCT. Obviously, conducting such a 'meta-study' would be difficult and maybe even unethical.

To approach a 'meta-randomized' trial as closely as possible we conducted a 'meta-Propensity Score' analysis. In a systematic review we collected all RCTs and Propensity Score (PS) analyses comparing off- an on-pump surgery in myocardial revascularization. RCTs and PS analyses were 'meta-matched' for relevant 'meta-confounders'. Estimated treatment effects from RCTs and PS analyses were compared for ten clinical outcomes in the 'meta-matched' sample as differences in odds ratios.

With 28 PS analyses and 51 RCTs initially retrieved, data from 10 PS analyses and 29 RCTs were available after 'meta-matching', including information for 186 clinical outcomes from more than 28.000 patients. For all clinical outcomes the observed odds ratio difference was smaller than 0.15 in absolute value, the overall odds ratio difference across all clinical outcomes was 0.03 (95%-CI: -0.07, 0.12).

In our example, treatment effects from RCTs and PS analyses were very similar in a 'meta-matched' population, indicating a small effect of randomisation itself. Of course, our study has limitations and needs independent replication in a different setting. Even if replicated we do not think that RCTs would be obsolete, but the current practice of excluding well conducted Non-RCTs from systematic reviews of treatment effects could at least be questioned.

References:

ROTHWELL, P.M. (2005): External validity of randomised controlled trials: 'to whom do the results of this trial apply?'. *Lancet, 365(9453), 82–93.*

Session:

S42: Meta-analysis and Meta-regression : Thursday, 13/03/2008, 11:00am - 12:40pm

Near-factorial experiments in a block design with nested rows and columns for plant protection research

Agnieszka Łacka and Maria Kozłowska

Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; aga@riders.pl, markoz@au.poznan.pl

Abstract: Block designs with nested rows and columns are widely used in experiments carried out on heterogeneous experimental material. Such situation occurs in plant protection experiments, where localisation plays an important role. When planning plant protection experiments, one has to consider irregular occurring and spreading of diseases, migration or raid of pests and also the possibility of reinfection. Other uncontrolled sources of variability that have to be considered are: dissemination of weed seeds and spores of fungal diseases or increased germinating of weed in more humid field areas. Having such heterogeneous experimental material, one has to use systems of blocks perpendicularly to the directions of occurring variabilities. Just block designs with nested rows and columns ensure eliminating of the heterogeneities described above.

An important aspect in plant protection experimentation is planning of research for two experimental factors, for example to compare the effectiveness of chosen plant protection products used in various concentrations and terms. To specify the effectiveness of used methods of plant protection, one has to analyse their activity when related to the untreated control. Due to economical reasons, in given examples of research problems, factorial experiments with one control treatment (near factorial experiments) are considered.

The aim of this paper is studying the efficiency factors estimation of contrasts of effects of two experimental factors, their combination and control treatment by the derived mixed model of observations in the block design with nested rows and columns. We carried out stratum analysis and obtained the efficiency factors estimation of treatment contrasts in the strata. We describe a block design with nested rows and columns having certain chosen properties.

Session:
S16: Agricultural science (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Statistische Entscheidungen trotz riesiger Lösungsvielfalt – biometrische Dateninterpretation und Modellbildung

Jürgen Läuter

Otto-von-Guericke-Universität Magdeburg und Interdisziplinäres Zentrum für Bioinformatik der Universität Leipzig, Germany; juergen.laeuter@med.ovgu.de

Abstract: Biometrie ist in ihren frühen Anfängen vor allem Volkszählung, die Zählung der Geburten, der Heiraten und der Sterbefälle, die Zählung der Pesttoten usw. Vom 17. Jahrhundert an wurden die Begriffe der Wahrscheinlichkeit und der mathematischen Erwartung entwickelt. Man lernte, zwischen Grundgesamtheit und Stichprobe zu unterscheiden, sodass im 20. Jahrhundert der große Fundus von einfachen und komplizierten statistischen Modellen entstehen konnte.

In der Anfangsphase hatte die Biometrie die Aufgabe, die große Anzahl auftretender Fälle, der geborenen und gestorbenen Menschen, der beobachteten Erkrankungen usw., sinnvoll zu beschreiben. Aber in der jüngeren Zeit treten komplexe Fragestellungen, mehr von qualitativem Charakter, besonders in den Vordergrund. Daher ist die Biometrie heute stärker an inhaltlichen Entscheidungen der Landwirtschaft, der Biologie, der Medizin usw. beteiligt.

In multivariaten Anwendungen spielt das Problem der Variablenauswahl eine große Rolle. Oftmals werden Verfahren benutzt, die sich stark an der Methode der kleinsten Quadrate orientieren und die sich auf multivariate Tests für alle Variablen oder bedingte Tests für einzelne Variablen, evtl. in Gestalt des Likelihood-Quotienten, stützen. Die Programmpakete bieten entsprechende Schrittverfahren der Vorwärts- und Rückwärtsselektion an.

Diese Verfahren haben jedoch viele Mängel: Durch die mehr oder weniger starke Bindung an die Methode der kleinsten Quadrate und durch die Restriktionen der Schrittprozeduren werden wichtige Einzelvariablen oder wichtige Teilmengen von Variablen nicht sicher erkannt. Es entsteht auch ein großer Selektionsbias, und insbesondere, wenn viele Variablenmengen gesucht werden, wird ein vorgegebener "familywise error" nicht eingehalten. Bei extrem hoher Dimension, wie z.B. in der Genexpressionsanalyse, ergeben sich außerdem Schwierigkeiten in Bezug auf die Rechenzeit.

Am Interdisziplinären Zentrum für Bioinformatik der Universität Leipzig wurden auf strenger mathematischer Grundlage neue Selektionsverfahren entwickelt, die die genannten Mängel vermeiden. Das Ziel besteht darin, innerhalb eines hochdimensionalen Datensatzes (Dimension $p > 20000$) Mengen wesentlicher Variablen zu suchen und ihre Relevanz exakt nachzuweisen. Jede der gegebenen p Variablen wird zunächst als Quellelement von möglichen Auswahlmengen betrachtet. Diese Auswahlmengen bestehen aus Variablen, die mit der Quellvariablen eine bestimmte Mindestkorrelation erreichen. Dann wird das bekannte auf Resampling beruhende Westfall-Young-Prinzip angewandt, sodass für ein Niveau α die signifikanten Mengen gefunden werden. Man kann zeigen, dass höchstens mit Wahrscheinlichkeit α irgendeine "falsch-signifikante"

Auswahlmenge auftreten kann, dass also der "familywise error α" nicht überschritten wird. Neben diesem nicht-parametrischen Verfahren liegt auch ein entsprechendes parametrisches Verfahren vor, das nach dem von Läuter, Glimm und Kropf entwickelten sphärischen Testprinzip arbeitet.

Diese Prozeduren sind allgemein anwendbar, ohne dass ein übermäßiger Rechenaufwand entsteht. Sie genügen den Anforderungen an ein stabiles multivariates Verfahren; hohe Korreliertheit der Variablen bzw. Redundanz trägt zum Erfolg des Verfahrens bei, aber verursacht weder "Overfitting" noch Instabilität. Mengen von Variablen sind gegenüber Einzelvariablen auch besser interpretierbar. Die Strategie der durch beliebige Quellvariablen generierten Auswahlmengen setzt keine speziellen Kovarianzstrukturen voraus.

Die Prozeduren werden vielfältig in der Forschung zur Genexpression eingesetzt. Die Ergebnisse lassen sich durch Anwendung von Projektionen, durch Biplots, durch "Heat Maps" und durch Kurvendiagramme veranschaulichen. Einige neue Ideen zur übersichtlichen Präsentation der Resultate werden dabei genutzt.

Session:
Open: Opening session and Welcome : Monday, 10/03/2008, 11:00am - 12:40pm

Estimating adjusted number needed to treat (NNT) measures in the Cox regression model

Rüdiger P. Laubender[1] and Ralf Bender[2]

[1] IBE, LMU München, Germany; Ruediger.Laubender@t-online.de
[2] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), 51105 Köln, Germany; Ralf.Bender@iqwig.de

Abstract: In medical research number needed to treat (NNT) measures for survival times have been proposed for use in randomized controlled trials without consideration of covariates (Altman and Andersen 1999). In this paper, we develop adjusted NNT measures for survival times within the framework of the Cox proportional hazards regression model for epidemiological studies taking the distribution of confounders into account. Estimation of adjusted NNT measures is performed by using the recently proposed average risk difference approach (Bender et al. 2007). Within this approach the effect of exposures is described by means of the NNT measures 'number needed to be exposed' (NNE) and 'exposure impact number' (EIN). To determine standard errors and confidence intervals for these estimators we use two approaches, the delta method and bootstrapping and compare each other. The performance of these estimators is scrutinized and assessed by performing Monte Carlo simulations (Laubender 2007). The use of these estimators is numerically and graphically illustrated by means of data of the Düsseldorf Obesity–Mortality–Study (Bender et al. 1998). As the estimation of the adjusted NNT measures is not implemented in commonly used statistical software, we present software solutions which are coded in 'R'.

References:

ALTMAN, D.G. and ANDERSEN, P.K. (1999): Calculating the number needed to treat where the outcome is time to an event. *British Medical Journal, 319, 1492–1495.*

BENDER, R., KUSS, O., HILDEBRANDT, M. and GEHRMANN, U. (2007): Estimating adjusted NNT measures in logistic regression analysis. *Statistics in Medicine, 26, 5586–5595.*

BENDER, R., TRAUTNER, C., SPRAUL, M. and BERGER, M. (1998): Assessment of excess mortality in obesity. *American Journal of Epidemiology, 147, 42–48.*

LAUBENDER, R. (2007): Berechnung des Maßes 'Number Needed to Treat" (NNT) aus Überlebenszeiten auf der Basis der Cox–Regression mit Berücksichtigung von Kovariablen. *Magisterarbeit, München.*

Session:
S17: Event Data Analyses (II) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Exact tests of non-inferiority based on higher order asymptotic pivotals

Chris J. Lloyd

Melbourne Business School, Melbourne, Australia; c.lloyd@mbs.edu

Abstract: Recent advances in likelihood asymptotics (Reid, 2003) lead to pivotal quantities that are closer to standard normal than standard pivotals and also respect some kind of conditionality. It is less clear the extent to which these methods work for discrete models. On the other hand, in the context of binomial trials conditional pivotals can lead to more efficient unconditional inferences, see Boschloo (1970) and Lloyd and Moldovan (2007). This suggests that second order pivotals that respect local conditionality might provide more powerful exact tests. For testing the rate ratio from independent binomial samples, we investigate 5 first order pivotals and the second order pivotal. Each of these is sued to generate an exact test by maximising with respect to the nuisance parameter. We also consider the effect of pre-estimating the nuisance parameter.

References:

BOSCHLOO, RD. (1970): Raised conditional level of significance for the 2×2 table when testing the equalities of probabilities. *Statistica Neerlandica 1–35.*

LLOYD, C.J. and MOLDOVAN, M. (2007): Unconditional efficient one-sided confidence limits for the odds ratio based on conditional likelihood. *Statistics in Medicine, 26.*

REID, N. (2003): Asymptotics and the theory of inference. *Ann. Statist. 31 1695–1731.*

Session:

S09: Non-Inferiority Trials (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Microarray meta-analysis identifies a taxane response gene signature

Marion Schütt[1,2], Barbara Nicke[1], Charles Swanton[1], Manfred Berres[2], and Maik Kschischo[2]

[1] Cancer Research UK, 44 Lincoln's Inn Fields, London, WC2A 3PX, UK; mschuett@rheinahrcampus.de, barbara.nicke@cancer.org.uk, charles.swanton@cancer.org.uk
[2] RheinAhrCampus Remagen, Südallee 2, 53425 Remagen; berres@rheinahrcampus.de, kschischo@rheinahrcampus.de

Abstract: Microtubule stabilisers (MTS) like taxanes, are established chemotherapeutic drugs in the clinic, used to treat breast, lung and ovarian cancer. The prediction of drug-sensitivity or -resistance to MTS could allow the selection of patients who might benefit from therapy. In this work a meta-analysis method, based on the binomial test and a gene-wise false discovery rate, was developed to derive a gene expression signature of MTS treatment for late time points using data from publicly available microarray experiments with several cancer types. To support the biological hypothesis that chromosomal instability might render cells resistant to MTS treatment, the MTS response signature was tested against a published signature of genes over-expressed in tumours with chromosomal instability. A significant intersection of genes repressed in the MTS response signature and genes over-expressed in cells with CIN was found. This work represents a valid basis to pursue work in the future trying to predict sensitivity or resistance to MTS in patients.

References:

CARTER, S.L., EKLUND, A.C., KOHANE, I.S., HARRIS, L.N. and SZALLASI, Z. (2006): A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nature Genetics, 38(9).*

SWANTON, C., MARANI, M., PARDO, O., WARNE, P.H., KELLY, G., SAHAI, E., ELUSTONDO, F., CHANG, J., TEMPLE, J., AHMED, A.A. et al. (2007): Regulators of Mitotic Arrest and Ceramide Metabolism Are Determinants of Sensitivity to Paclitaxel and Other Chemotherapeutic Drugs. *Cancer Cell, 11(6): 498-512.*

Session:

S06: Biostatistic education at universities of applied sciences : Monday, 10/03/2008, 2:10pm - 3:30pm

Medical biometry in academic medicine: perspectives and challenges

Jochen Mau

Institute of Statistics in Medicine, Heinrich Heine University Hospital,
P.O. Box 10 10 07, 40225 Düsseldorf, Germany; ismmau@uni-duesseldorf.de

Abstract: Biometry has certainly established itself firmly in pharmaceutical research and development towards providing effective drugs for clinical medicine, during the last decades. Its guiding role has also become acknowledged in clinical research in Germany since the 1960's as evidenced by the growing number of dedicated chairs of medical statistics until the end of the 1980's. However, apart from the ubiquitous accessibility of user–friendly software for statistical planning and analysis, the ongoing re-structuring of academic medicine towards cost-effective services and resource management affect all medical departments whether engaged in patient care, or not. Hence, it is timely to consider the perspectives for biometry within the medical faculties of the future. The historical development is reflected briefly, from the seminal impetus by the French doctor Pierre C. A. Louis (1787–1872) via Florence Nightingale and others to the present situation with FDA rulings, GCP standards, and EBM guidance, a plethora of biometrical technicalities for research in clinical medicine and its practice. However, engineering research is not the same as doing or driving it. Can today's biometricians adequately prepare doctors for their professional duties as doctors? Would today's biometricians in medical faculties not have the duty to recruit the young talents for their field from their own students, the doctors as suggested by Mau (2007)? Today's biometricians are handicapped by their self–imposed limitation to transfers of statistical methodology, and they should become able – and accepted – to set off themselves, then, to translate the transfer results towards the formation of original medical knowledge: Thus, today's biometricians would transform into tomorrow's medical scientists and define themselves in a new authentically medical discipline, the Quantitative Medicine. Present missions are illustrated in biotechnology, dental and human medicine, health, and the environment and some challenges that arise in the Global World are mentioned.

References:

BOLLET A. J. (1973): Pierre Louis: The numerical method and the foundation of quantitative medicine. *The American Journal of the Medical Sciences, 266, 92–101.*

MAU, J. (2007): Institut für Statistik in der Medizin. In: Th. Halling, J. Vögele (Eds.): *100 Jahre Hochschulmedizin in Düsseldorf 1907–2007.* Verlag düsseldorf university press, Düsseldorf, 309–311.

Session:

S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Combined phase II/III adaptive designs: why are they needed?

Michael Branson and Willi Maurer

Novartis Pharma AG, CH-4056 Basel, Switzerland;
michael.branson@novartis.com, willi.maurer@novartis.com

Abstract: In the early phases of drug development, where learning and hypothesis finding is the prime focus, the place of adaptive designs is largely uncontested. However, in late stage development there are also situations where there is a need for adaptive designs in order to efficiently answer the remaining questions. This is particularly the case where in the earlier phases it was not yet possible to settle with sufficient confidence on just one possible hypothesis to be confirmed in phase III trials. In these situations adaptive designs are a promising tool. Leveraging adaptive designs reliably selects from the remaining options, reducing the set of hypotheses and simultaneously providing adequate evidence and characterization of the treatment effect and the safety profile. The advantages of such an approach are usually reduced number of patients exposed to potentially ineffective doses or with inferior benefit risk ratios and shorter development timelines. Other advantages include selecting from pre-defined subgroups, reducing the number of patients belonging to a subgroup having unfavorable predictive factors for benefiting from the treatment under investigation. In our short introduction to the session we will discuss in further details such situations and the main advantages and additional efforts necessary to implement confirmatory clinical trials with an adaptive design as part of a drug development program.

Session:
S23: Adaptive seamless design for combining phase II / III clinical studies : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Modeling Longitudinal Data with Application to the Multicenter Aids Cohort Study

Cyntha A. Struthers and Don L. McLeish

University of Waterloo, Waterloo, Ontario, Canada N2L 3G1;
castruth@uwaterloo.ca, dlmcleis@uwaterloo.ca

Abstract: Motivated by an example in the book (DIGGLE 2005) and a number of papers analyzing earlier versions of the data, we analyze data obtained from the MACS (Multicenter Aids Cohort Study) by fitting a stochastic process to the longitudinal data obtained from each study participant. The response is the CD4 count, fitted using a stochastic process with time-varying parameters which are functions of time-dependent covariates. There are extensive missing data in this data which is partially accommodated in our analysis. The parameters are estimated using both Bayesian and frequentist methods. We obtain the asymptotic covariance matrix of the estimators and Bayesian credible regions for the parameters, the latter by Gibbs sampling.

References:

DIGGLE, P.J., HEAGERTY, P., LIANG, K-Y, and ZEGER, S.L. (2005): *Analysis of Longitudinal Data.* Second Ed., Oxford, London.

Session:
S34: Bayesian models in biostatistics (II) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Experimental Design and Analysis approaches for Large and Complex Time-course Gene Expression Microarray Studies

Andrew Mead

Warwick HRI, University of Warwick, Wellesbourne, Warwick, CV35 9EF, UK;
andrew.mead@warwick.ac.uk

Abstract: The current interest in developing systems approaches to understanding the genetic pathways and processes associated with a whole range of biological activities has resulted in a dramatic increase in the number and size of microarray gene expression studies being performed. Within Warwick HRI, much of the focus has been on studies using the model plant, Arabidopsis thaliana, in a diverse range of applications areas including senescence, flowering, pathogen interactions, seed dormancy and nutrient response, but other research groups are interested in gene expression studies in mushrooms and bees!

Within this paper I will discuss some of the issues concerned with the design and analysis of a range of experiments within the A. thaliana research programme, using a customized 2-colour microarray technology, CATMA. The research aims within this programme often suggested the need to consider a relatively large number of unique biological samples (approaching 100 in some cases), always involving a time-course, often with a number of qualitative treatments, sometimes with a rather complex treatment structure, and usually wanting to include samples for a number of biological replicates for each treatment/time combination. Even with relatively little technical replication, a large number of microarrays would be necessary, so an efficient design was of great importance. The challenge, therefore, was to devise appropriate experimental designs to allow the extraction of the maximum possible information from these valuable datasets. Various approaches based on extensions and modifications of the commonly used "loop" design will be presented and discussed.

An apparently often neglected element of the analysis of microarray data is the data quality checking that we encourage in almost all other areas of experimentation. Based on some procedures in the R Library MAANOVA, I will discuss some of the issues that we have encountered, particularly associated with checking the quality of such large datasets, and the software tools that we have developed. I will also briefly discuss some of the issues associated with the formal analysis of these gene expression data sets using the MAANOVA approach.

Session:
S22: Agricultural science (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Experiences in Teaching Applied Biometry

Reinhard Meister

TFH Berlin, FB II Mathematik, Physik, Chemie, 13353 Berlin, Germany;
Reinhard.Meister@tfh-berlin.de

Abstract: The Fachhochschulen, nowadays called Universities of Applied Sciences prepare their students for immediate practical work when leaving University. This paradigm of education abides for the new concept of bachelor and master degrees.

We will present some experiences from more than 20 years of teaching at the Technische Fachhochschule Berlin, discussing how to achieve this goal. We will present examples from teaching for different disciplines: mathematics, medical informatics, bioinformatics.

The key to success in teaching applied biometry is application. Therefore, contact to industry, other universities and research institutions is highly important. Cooperation is a really interactive process. How interaction can work will be demonstrated.

Session:
S06: Biostatistic education at universities of applied sciences : Monday, 10/03/2008, 2:10pm - 3:30pm

Repeated Youden square with split units

Stanisław Mejza[1] and Danuta Kachlicka[2]

Agricultural University, Wojska Polskiego 28, PL-60-637 Poznan, Poland;
smejza@au.poznan.pl, kachol@au.poznan.pl

Abstract: In field experiments quite often we use a row-column design in order to eliminate real or potential orthogonally disposed heterogeneity of experimental material. Then the Latin square is the appropriate design. However, this design uses many experimental units. We can reduce the number of experimental units by using the Youden square (see e.g.. Cox, 1958).

Let us assume that the experimental material is divided into k_0 superblocks. Each superblock constitutes a row-column design with k_1 rows and k_2 columns. On each unit of the row-column design that is treated as a whole plot, the whole plot treatments (A) are arranged. Additionally, each whole plot is divided into k_3 small plots (called subplots) and on each of them subplot treatments (B) are arranged.

In the paper we examine statistical properties of the design in which each superblock has the Youden square structure. Moreover, it is assumed that a subdesign of the Youden square with respect to columns is a balanced incomplete block design or a group divisible partially balanced incomplete block design with two association classes, (cf. Clatworthy, 1973).

In the paper the statistical properties (general balance, efficiency balance) of above design are examined under the mixed linear model of observations. The dispersion structure of linear model results from the scheme of randomization applied.

References:

CLATWORTHY, W.H. (1973): Tables of two associate classes partially balanced designs. *NBS Applied Math. Ser. 63. Washington, D.C, USA.*

COX, D.R. (1958): *Planning of Experiments.* Wiley, New York.

KACHLICKA, D. and MEJZA, S. (1996): Repeated row-column designs with split units. *Comp. Statist. and Data Analysis 21, 293-305.*

Session:

S16: Agricultural science (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Discussion of a model for the Carcinogenic Potency of Fibers

Oliver Melsheimer[1], Joachim Kunert[1], Markus Roller[2], and Peter Wardenbach[3]

[1] TU Dortmund, Fakultät Statistik, Germany;
melsheimer@statistik.uni-dortmund.de, kunert@statistik.uni-dortmund.de
[2] BMR; markus.roller@bmr-online.de
[3] Bundesanstalt für Arbeitsschutz und Arbeitsmedizin;
Wardenbach.Peter@baua.bund.de

Abstract: Inhalation of asbestos fibers has caused numerous cases of lung cancer, mesotheliomas and asbestosis in humans. As a consequence, other fibrous materials (e.g. man-made vitreous fibers) have been tested for possible adverse health effects. This was usually done in experiments with rats.

Berry (1969) presented several models for describing the temporal relationship between exposure to asbestos fibers and the occurrence of mesotheliomas. These models implicitly assumed that inhaled fibers deposited in the lungs would contribute continuously to carcinogenesis. Because man-made vitreous fibers persist in the lung for a shorter time than asbestos fibers, Berry (1999) concluded that his model might be inappropriate for these fibers. Fibers which are eliminated in just a few months or even weeks do not seem liable to be a permanent cause of danger.

Therefore, Berry (1999) proposed an extended model containing the time dependant elimination of fibers as an additional term. He applied the modified model to data from a cohort study of the Wittenoom crocidolite asbestos mine in Western Australia. Although the available data on mesothelioma incidence and time since exposure at Wittenoom did not allow discrimination between his early models and the elimination model, Berry (1999) claimed that there is sufficient biological rationale to suggest that the elimination model is plausible.

However, the extended Berry model has some consequences for the attempt to extrapolate results of experiments with rats to humans. Under the extended model virtually no fiber with a sufficiently low biopersistence should ever have an effect on humans, even if it shows severe consequences for rats [see Rödelsperger (2004)].This is due to the much longer life span of humans

Since it is not possible to perform experiments with humans, there is no final answer on the validity of the extrapolation. However, it is possible to examine whether Berry's (1999) extended model describes data from experiments with rats better than the original model.

We therefore applied his models to data from animal intraperitoneal injection experiments with fibers of varying lung biopersistence [for details see Roller et al. (1996)]. We found no indication that the modified model would be an improvement. In fact, for fibers with very short half-times we observed a better fit for the original model than for the modified model.

The presentation will discuss a possible explanation for the poor performance of the extended model.

References:

BERRY, G. and WAGNER, J.C. (1969): The Application of a Mathematical Model Describing the Times of Occurrence of Mesotheliomas in Rats Following Inoculation with Asbestos. *British Journal of Cancer, 23, 582–586.*

BERRY, G. (1999): Models for Mesothelioma Incidence Following Exposure to Fibers in Terms of Timing and Duration of Exposure and the Biopersistence of the Fibers. *Inhalation Toxicology, 11, 111–130.*

ROLLER, M., POTT, F., KAMINO, K., ALTHOFF, G.H. and BELLMANN, B. (1996): Results of current intraperitoneal carcinogenicity studies with mineral and vitreous fibres. *Exp Toxic Pathol, 48, 3–12.*

RÖDELSPERGER, K. (2004): Extrapolation of the Carcinogenic Potency of Fibers from Rats to Humans. *Inhalation Toxicology, 16, 801–807.*

Session:

S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

On Using Extended Maximum Likelihood Principle to Hierarchical General Linear Models

Andrzej Michalski

Department of Mathematics, Wrocław University of Environmental and Life Sciences, Poland; max@ozi.ar.wroc.pl

Abstract: In the paper some aspects of statistical inference by using maximum likelihood principle to estimation of parameters in mixed linear models are presented. The linear models with structure including unobserved random variables require to effective estimation using a particular definition of general likelihood. Many statisticians attempted to find a theoretical basis for choosing a particular form of general likelihood (e.g. Bayarri *et al.* (1987), Bjørnstad (1996)). It was proved that one key property of likelihood inference is an invariance with respect to some group of transformations such as in the problem of estimation of variance components in mixed linear normal models (Gnot *et al.* (2004)). In the paper extended likelihood and h-likelihood for hierarchical generalized linear model by selected examples are presented (see Lee and Nelder (1996); Lee *et al.* (2006)).

References:

BJØRNSTAD, J.F. (1996): On the generalization of the likelihood function and likelihood principle. *J. of the American Statistical Association, 91, 791- 806.*

GNOT, S., MICHALSKI, A. and URBAŃSKA-MOTYKA, A. (2004): On some properties of ML and REML estimators in mixed normal models with two variance components. *Discussiones Mathematicae - Probability and Statistics, 24, 109-126.*

LEE, Y., NELDER, J.A. (1996): Hierarchical generalized linear model (with discussion). *Journal of the Royal Statistical Society B, 58, 619- 656.*

LEE, Y., NELDER, J.A. and PAWITAN,Y. (2006): *Generalized Linear Models with Random Effects. Unified Analysis via H-likelihood.* Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton London New York.

Session:

S32: Hierarchical models : Wednesday, 12/03/2008, 11:00am - 12:40pm

The assessment of non-inferiority in a gold standard design with censored, exponentially distributed endpoints

Matthias Mielke

Institute for Mathematical Stochastics, Georg-August-University of Göttingen, Maschmühlenweg 8-10, D-37073 Göttingen, Germany;
mmielke@math.uni-goettingen.de

Abstract: The objective of this paper is to develop statistical methodology for non-inferiority hypotheses to censored, exponentially distributed time to event endpoints. Motivated by a recent clinical trial in depression we consider a gold standard design where a test group is compared to an active reference and to a placebo group.

The test problem is formulated in terms of a retention of effect hypothesis. Thus, the proposed test procedure assures that the effect of the test group is better than a pre-specified proportion Δ of the treatment effect of the reference group compared to the placebo group. A sample size allocation ratio to the three groups to achieve optimal power is presented, which only depends on the pre-specified Δ.

In addition, a pretest is presented for either the reference or the test group to ensure assay sensitivity in the complete test procedure. The actual type I error and the sample size formula of the proposed tests is explored asymptotically and by means of a simulation study showing good small sample characteristics. To illustrate the procedure a randomized, double blind clinical trial in depression is evaluated.

This paper is a joint work with A. Munk (University of Göttingen) and A. Schacht (Lilly Deutschland GmbH).

Session:
S11: Non-Inferiority Trials (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Blinded Assessment of Treatment Effects Utilizing Information about the Randomisation Procedure

Frank Miller[1], Meinhard Kieser[2], and Tim Friede[3]

[1] AstraZeneca, Södertälje, Sweden; frank.miller@astrazeneca.com
[2] Institute of Medical Biometry and Informatics, University of Heidelberg, Germany; meinhard.kieser@imbi.uni-heidelberg.de
[3] Warwick Medical School, The University of Warwick, Coventry, UK; t.friede@warwick.ac.uk

Abstract: It is generally believed that inference about the treatment effect is impossible in a randomised, double-blind trial as long as data is blinded, i.e. as long as it is unknown to which treatment the patients were assigned to. Surprisingly, if information about the randomisation procedure is utilized, it is possible to construct treatment effect estimators and tests for the null-hypothesis of no treatment effect against the two-sided alternative based on blinded data. If application of these methods would result in reliable information, this would lead to severe problems since the integrity of the trial would be in danger if persons involved in the study could gain treatment effect knowledge during its course.

We consider block-randomised trials with two treatment arms. Van der Meulen (2005) described estimators for the treatment effect such trials with block sizes 2 and 4. He investigated their characteristics using a single simulated dataset. We investigate these estimators thoroughly using repeated simulations. Moreover, we define a blinded significance test and compare its power with the power of the unblinded t-test. We see that it is hard to get a reasonable power for the blinded test if the actual treatment effect is not very much higher than assumed for sample size calculation. However, if the block size is small and the study is overpowered, there might be a chance to obtain information from blinded data. Therefore, we suggest avoiding small block sizes especially in situations where the assumption about the true treatment effect and standard deviation are vague and the sample size was calculated in a conservative way. If these precautions are undertaken, information about the treatment effect gained from blinded data is of little value.

References:

VAN DER MEULEN, E.A. (2005): Are we really that blind? *Journal of Biopharmaceutical Statistics, 15, 479–489.*

Session:
S18: Clinical studies (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Measuring Heterogeneity in Meta-Analyses

Martina Mittlböck and Harald Heinzl

Core Unit for Medical Statistics and Informatics, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria; martina.mittlboeck@meduniwien.ac.at, harald.heinzl@meduniwien.ac.at

Abstract: A meta-analysis is the combination of clinical trial results in order to arrive at summary conclusions about a medical research question. A major problem is to decide whether the trials under review are homogeneous or heterogeneous. This important step determines the statistical methods to be used and the interpretation of the results. Tests of heterogeneity are not always reliable because of low power for sparse data and the detection of irrelevantly small amount of heterogeneity when many studies are involved. Thus measures of heterogeneity are better suited to determine the amount of between-study variance and its impact on meta-analysis results.

We review measures for heterogeneity with different scaling. Simulation study results are used to assess various properties of the heterogeneity test and heterogeneity measures.

Heterogeneity test and heterogeneity measures are not directly related to the absolute amount of between-study variance but to the relative increase of variance due to heterogeneity. A heterogeneity measure scaled to a fixed interval needs reference values for proper interpretation. A measure defined by the relation of between- to within-study variance has a more natural interpretation but no upper limit. Investigated heterogeneity measures depend on the variance of the individual effects and thus on the number of patients in the studies.

References:

HIGGINS, J. and THOMPSON, S. (2002): Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21, 1539–1558.*

MITTLBÖCK, M. and HEINZL, H. (2006): A simulation study comparing properties of the heterogeneity measures in meta-analyses. *Statistics in Medicine, 25, 4321–4333.*

Session:

S39: Meta-analysis : Thursday, 13/03/2008, 9:10am - 10:30am

Additivity tests for the mixed model in the two way ANOVA with single sub-class numbers. A new approach to evaluate interaction in block design

Karl Moder

University of Natural Resources and Applied Life Sciences, Institute for Applied Statistics and Computing, Gregor-Mendel-Strasse 33, A1080 Vienna, Austria;
karl.moder@boku.ac.at

Abstract: In a block design a test on interaction is difficult because of no replications within subclasses. In this situation it is not possible to separate interaction from the error term. [Tukey (1949)] introduced some restrictions to the structure of interaction, in particular that the interaction effect $\tau_{ij} = \lambda\alpha_i\beta_j$ (α_i ... effect of the factor, $i = 1, \ldots, a$; β_j ... block effect, $j = 1, \ldots, b$). This means, that interaction is the product of factor and block effects. Firstly if there is no influence of one of these effects no interaction exists. Secondly interaction depends on main effects in a non additive way. Similar restrictions can be found with [Mandel et al. (1961)] and [Johnson et al. (1972)]. In the proposed method no constraint to the structure of interaction are made, but there are some restriction to the design of the experiment. A well known design which allows estimation of these interactions is latin square. But in addition to the ordinary assumption that $\sum \gamma_i = 0$ (γ_i,... column effect, $i = 1, \ldots, a = b$), we also assume that all interaction effects within a column sum up to zero. This is very intuitive, as this is the common assumption for interactions within blocks. As there is nor real distinction between blocks and columns this assumption seems to be reasonable. The method as such however is not restricted to latin squares but can be generalized to a very broad range of experimental designs.

References:

JOHNSON, D.E. and GRAYBILL, F.A. (1972): An analysis of a two-way model with interaction and no replication. *Journal of the American Statistical Association 67, 862-868.*

MANDEL, J. (1961): Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association 56, 878-888.*

TUKEY, J.W. (1949): One degree of freedom for nonadditivity. *Biometrics 5, 232-242.*

Session:

S40: Additivity tests for mixed models : Thursday, 13/03/2008, 9:10am - 10:30am

Comparison of one-stage and two-stage analysis in series of experiments

Jens Möhring and Hans-Peter Piepho

Institute for Crop Production and Grassland Research, Universität Hohenheim, 70599 Stuttgart, Germany; moehring@uni-hohenheim.de, piepho@uni-hohenheim.de

Abstract: Series of plantbreeding trials have the aim to evaluate a great number of genotypes as early as possible and as fast as possible. Therefore plantbreeding trials have some specifics. Each year a lot of genotypes are tested, mostly with a low number of replications. Some of them are selected and tested again in the following year with a higher number of replications. This testing is parallel with the testing of genotypes of the next year, thats why there is a high unbalances concerning genotypes, years and locations. It is common to use mixed models to evaluate unbalanced data coming from several years. If the model should account for covariances coming from ancestry, the genetic effects has been taken random. Genotypic effects are then estimated by using BLUP (best linear unbiased prediction). Often a one-stage analysis is not feasible cause of time or computational problems. To safe a quick selection the analysis as been done in the short period between harvest and sowing. That is the reason for analysing data in two steps. In step one adjusted means from the genotypes coming from one experimental unit j are estimated. These estimates have a variance-covariance matrix R_j. In step two, the adjusted means are further analyses through the experimental units. A open question is, which weighting procedure for the adjusted means should be used in step two. Mostly diagonal matrices are used to approximate either R_j or R_{j-1}. If the real variance-covariance matrix is nearly diagonal, estimated hardly differ between one-stage and two-stage analysis. In the presentation we used five published weighting procedures (Piepho et al., 2000; Smith et al. 2001, 2005) and three new weighting procedures to evaluate one-stage and two-stage analysis for several series of plantbreeding trials. The correlation between this analyses was used as a criteria of evaluating. The new weighting methods based on the approximation of the variance of a difference of genotype means. We found a high correlation with the one-stage analysis for models with fixed genotype effect with a simple genetic structure. For complex genetic structure and using BLUP the correlation gets weaker, so results coming from one-stage and two-stage analysis differ. In both cases we observed marginal differences between different weighting procedures.

References:

PIEPHO, H.P. and MICHEL, V. (2000): Überlegungen zur regionalen Auswertung von Landessortenversuchen. *Informatik, Biometrie und Epidemiologie in Medizin und Biologie 31, 123–136.*

SMITH, A., CULLIS, B. and GILMOUR, A. (2001): The analysis of crop variety evaluation data in Australia. *Australian and New Zealand Journal of Statistics 43, 129–145.*

SMITH, A.B., CULLIS, B.R. and THOMPSON, R. (2005): The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *Journal of Agricultural Science 143, 449–462.*

Session:

S16: Agricultural science (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Confidence Sets Following a Modified Group Sequential Test

Hans-Helge Müller and Roman Pahl

Institute of Medical Biometry and Epidemiology, Philipps-University, D-35037 Marburg, Germany; muellerh@med.uni-marburg.de, pahlr@med.uni-marburg.de

Abstract: Consider monitoring a clinical trial where the analysis is based on a carefully planned group sequential design for a Brownian motion with the drift parameter reflecting the treatment difference. Suppose that during the course of the trial a change of the design is advisable and that the effect size measure is unchanged.

In order to control the type I error rate, trials are re-designed using the Conditional Rejection Probability (CRP) principle proposed by Müller and Schäfer (2004). An important issue is estimating the effect size parameter. While the naive confidence intervals (CIs) are inadequate, adequate approaches reflecting early stopping for both significance and futility have been presented. Confidence sets after modification have been suggested by Müller and Schäfer (2001), and the repeated CIs of Jennison and Turnbull were investigated by Mehta et al. (2007).

In this contribution, we treat CIs after modification using Tsiatis et al. (1984) and the flexible CRP approach. The application is illustrated for a survival study.

References:

MEHTA, C.R., BAUER, P., POSCH, M., and BRANNATH, W. (2007): Repeated confidence intervals for adaptive group sequential trials. *Statistics in Medicine, 26, 5422–5433.*

MÜLLER, H.-H. and SCHÄFER, H. (2001): Adaptive group sequential designs for clinical trials: combining the advantages of adaptive and of classical group sequential approaches. *Biometrics, 57, 886–891.*

MÜLLER, H.-H. and SCHÄFER, H. (2004): A general statistical principle for changing a design any time during the course of a trial. *Statistics in Medicine, 23, 2497–2508.*

TSIATIS, A.A., ROSNER, G.L., and MEHTA C.R. (1984): Exact confidence intervals following a group sequential test. *Biometrics, 40, 797–803.*

Session:

S05: Adaptive Group Sequential Designs : Monday, 10/03/2008, 2:10pm - 3:30pm

Quality control measures in genome-wide association studies

Bertram Müller-Myhsok

MPI Psychiatry Munich, 80804 München, Germany; bmm@mpipsykl.mpg.de

Abstract: Genome-wide association studies (GWAS) are characterized by a large number of polymorphic markers being genotyped and analyzed for association with one or several traits of interests. Even though genotyping technology has made significant advances recently and error rates have been dramatically reduced, the sheer number of markers makes a large absolute number of errors likely to be present in any given data set. Quality control is thus vital in the analysis of GWAS and the choice of procedure and sets of thresholds is important for the later success or failure of a study. I will present own empirical data on the influence of thresholds on the conclusions drawn from GWAS.

Session:
S07: Genome-wide association studies (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Robustness of estimates in the cross-over model against variance heterogeneity

Ralf Sigmund and Gerhard Nehmiz

Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach, Germany;
ralf.sigmund@boehringer-ingelheim.com,
gerhard.nehmiz@boehringer-ingelheim.com

Abstract: We investigate the random-effects complete cross-over model without carry-over, sequence group effects and repetition:

$$Y_{jkm} = \mu + \tau_k + \pi_j + s_m + \epsilon_{jkm},$$

whereby μ is the overall mean, the τ_k are the treatment effects ($k = 1, \ldots, K$), the π_j are the period effects ($j = 1, \ldots, J$, here: $J = K$), the s_m are the random subject effects ($m = 1, \ldots, M$) and ϵ_{jkm} is the residual error. Note that the indices j and k are dependent from each other, k is already sufficient. The between-subject variability is

$$s_m \sim N(0, \sigma^2) \text{ i.i.d., and all } s_m \text{ are independent from all } \epsilon_{km}.$$

The following models for the residual variability are investigated:

$$\begin{array}{ll} (I) & \epsilon_{km} \sim N(0, \sigma^2) \text{ i.i.d. for all } k \\ (II) & \epsilon_{km} \sim N(0, \sigma_1^2) \text{ i.i.d. } (k = 1, ..., K-1) \text{ and} \\ & \epsilon_{km} \sim N(0, \sigma_2^2) \text{ for } k = K, \text{ indep. between } k\text{'s.} \end{array}$$

The question is how far an inequality of σ_2^2 and σ_1^2 has an influence on the estimate of σ^2 in model (I), i.e. how robust model (I) is against an undetected variance heterogeneity. This has then consequences for the CIs of any treatment contrasts, even if none of the 2 treatments in the contrast is treatment K, e.g. in bioavailability investigation of pharmaceutical formulations. For large up- or downward deviations from equality, σ^2 of (I) is seriously over- and underestimated, respectively, and the CIs of treatment contrasts are relevantly too wide and too narrow, respectively. We recommend to leave out the data of the outlying treatment, but to analyse the remaining data in the periods in which they were actually given. The principle that the data were observed "in the same trial" is then still followed. An additional example from lung function testing shows however that large deviations from equality do practically not need to be expected in this area.

References:

SENN, S. and EZZET, F. (1999): Clinical cross-over trials in phase I. *Statistical Methods in Medical Research, 8, 263–268.*

SHEINER, L.B. (1992): Bioequivalence revisited. *Statistics in Medicine, 11, 1777–1788.*

TEN HAVE, T.R. and LOCALIO, A.R. (1999): Empirical Bayes Estimation of Random Effects Parameters in Mixed Effects Logistic Regression Models. *Biometrics, 55, 1022–1029.*

Session:
S32: Hierarchical models : Wednesday, 12/03/2008, 11:00am - 12:40pm

It can be important to know the rate of imprinting

Nadine Neugebauer, Volker Guiard†, and Norbert Reinsch

Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere (FBN), Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany;
neugebauer@fbn-dummerstorf.de

Abstract: Can statistical genetics provide answers regarding the relative contribution of genomic imprinting to total genetic variation?

Imprinted genes are involved in many aspects of development in humans, other mammals, plants and perhaps birds. Consequently, alteration of normal imprinting patterns may give rise to some genetic human diseases. Since regulation of growth is also affected by imprinting, it may play a role in food production. More knowledge about imprinting can affect the profit of this production.

Genomic imprinting inhibits gene expression. The imprinted alleles are not transcribed. The mechanism of imprinting is complex and not completely understood, but it results in expression of either the maternal or paternal allele. A special type of imprinting, known as partial imprinting, influences allele expression in one or several of the following ways: 1. an allele is understated in its expression, 2. the allele expression is cell-type-specific, 3. the expression changes in different phases of life.

Existing models for estimating imprinting effects do not account for partial imprinting. A new adapted model accommodates all these types of imprinting. This 'compact model' separates imprinting variance from the other variances. Use of this model requires a deep family tree. At the moment this model is applied to analyse imprinting effects on livestock phenotypes. Slaughter data from more than 20,000 pigs were analysed. The attendant pedigree has up to 20 generations. It was shown that imprinting may be responsible for up to 15% of the additive genetic variance. Such results are useful for implementing more efficient breeding programmes for farm animals. Furthermore, the methods used in this study can easily be applied to other (non-agricultural) data sets to estimate imprinting variance provided a data set with suitable structure is available. The separation of imprinting variance, maternal effects and other sources of irritation is difficult and will be discussed.

Additionally, it may be useful to gain information about the relative contributions of imprinted loci to the total genetic variation prior to more elaborate positional studies, e.g. mapping of imprinted QTLs. The method is assumed to be of special interest for particular human data sets, farm animals and probably other organisms.

Session:
S04: Mapping Approaches in Plants and Animals : Monday, 10/03/2008, 2:10pm - 3:30pm

A Review on Joint Models in Biometrical Research

Anneke Neuhaus[1], Thomas Augustin[2], Christian Heumann[2], and Martin Daumer[1]

[1] Sylvia Lawry Centre, Hohenlindenerstr. 1, 81677 Munich, Germany; neuhaus@slcmsr.org, daumer@slcmsr.org
[2] Department of Statistics, University of Munich, Ludwigstr. 33, 80539 Munich, Germany; thomas.augustin@stat.uni-muenchen.de, christian.heumann@stat.uni-muenchen.de

Abstract: In biometrical research joint modelling of longitudinal measures and event time data has become a very popular tool to account for informative dropouts, to incorporate longitudinal trends and to link disease processes that occur simultaneously. However, in some disease areas the power of this methodology has not yet been explored despite deficiencies of standard methodology. An example is the neurodegenerative disease multiple sclerosis which is the most common neurological disease in young adults. Against this background we review recent fruitful research in the area of joint models by classifying approaches in three categories: approaches with focus on serial trends, approaches with focus on event time data and approaches with equal focus on both outcomes. Typically longitudinal measures and event time data are modeled jointly by introducing shared random effects or by considering conditional distributions together with marginal distributions. We present the approaches in an uniform nomenclature, comment on sub-models applied to longitudinal measures and event time data outcomes individually and exemplify applications in bio- metrical research. The increasing variety of joint model approaches is a promising framework to shed light on the clinically relevant and controversial discussion about interactions between exacerbations and permanent disability in multiple sclerosis. We will discuss extensions of current approaches to address that issue.

Session:
S29: Advances in Statistical Modelling (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Regionalisation of Forest Stand Variables in Baden-Württemberg

Arne Nothdurft[1] and Joachim Saborowksi[2]

[1] Forest Research Institute Baden-Württemberg, Department of Biometry and Informatics, 79007 Freiburg, Germany; arne.nothdurft@forst.bwl.de
[2] University of Göttingen, Chair of Ecoinformatics, Biometrics and Forest Growth, Büsgenweg 4, 37073 Göttingen, Germany; jsaboro@gwdg.de

Abstract: This study aims at the development of a model to predict forest stand variables in management units (stands) from sample plot inventory data. For this purpose we apply a non-parametric nearest-neighbour (NN) approach. The study area is the municipal forest of Waldkirch, 13 km north-east of Freiburg, Germany, which comprises 328 forest stands and 834 sample plots. Low-resolution laser scanning data, classification variables as well rough estimations from the forest management planning serve as auxiliary variables. In order to avoid common problems of k-NN-approaches caused by asymmetry at the boundaries of the regression spaces and distorted distributions, forest stands are tessellated into subunits with an area approximately equivalent to an inventory sample plot. For each subunit only the one nearest neighbour is consulted. Predictions for target variables in stands are obtained by averaging the predictions for all subunits. Additionally, global multipliers for bias correction are derived from the unbiased Horvitz-Thompson estimates. After formulating a random parameter model with variance components, we calibrate the prior predictions by means of sample plot data within the forest stands via BLUPs (best linear unbiased predictors). The averaged and calibrated results are shown to be approximately unbiased. Based on bootstrap simulations, prediction errors for most management units finally prove to be smaller than the design-based sampling error of the mean.

Session:
S43: Statistical Methods in Environmental Monitoring : Thursday, 13/03/2008, 11:00am - 12:40pm

Quality Management and Analysis of Genomic High-throughput Data

Michael Nothnagel, Amke Caliebe, Olaf Junge, and Michael Krawczak

Institute of Medical Informatics and Statistics, University of Kiel, Brunswiker Str. 10, 24105 Kiel, Germany; nothnagel@medinfo.uni-kiel.de

Abstract: Genetic high-throughput data have recently been used for a number of purposes, but mainly for screening whole genomes for potential associations with phenotypes and diseases. The validity of model assumptions in the statistical analysis of datasets is crucial for reliable results. We have developed quality control protocols to avoid a number of violations of those assumptions, including the independence of observations, the homogeneity of the study population, and the detection of systematic genotyping errors. We note that genetic association testing is only one among a number of objectives for the analysis of high-throughput datasets and that criteria for quality control depend on those objectives.

We also note that association testing is usually only the initial step which is followed by more focused and more advanced analyses of subsets of variables and samples. We have therefore implemented data storage, quality control, and statistical analysis in an integrated system by combining the R statistical environment with a database, providing direct and fast access to this database. This solution allows the structured and consistent storage of raw as well as derived variables, such as quality control features and previous analysis results, and the access to the whole range of statistical analyses implemented in R.

The amount of information obtained from genomic high-throughput scans is prohibitive for any kind of comprehensive manual inspection of the raw data. However, we show that while such an inspection is not feasible prior to the analysis, it is mandatory as an additional validation of the results from the statistical analysis.

Session:
S07: Genome-wide association studies (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Determining the parameter settings of different randomization methods for specific study designs

Petra Ofner-Kopeinig, Maximilian Errath, and Andrea Berghold

Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria; petra.ofner@meduni-graz.at

Abstract: Randomization is an important feature of clinical trials. Different methods are used for random allocation of patients to treatment groups. Frequently used methods are permuted block randomization, urn design, minimization, biased coin randomization or big stick randomization. The randomization methods differ according to balance behaviour and selection bias. The choice of the randomization procedure and the setting of the parameters depend on the design features of the study. It is important to have a tool to study how to choose the parameters of the methods under certain design features. The Randomizer, a web-based patient randomization service for multi-center clinical trials, developed at the Institute for Medical Informatics, Statistics and Documentation at Medical University of Graz, can be used for this purpose. The application provides unrestricted randomization as well as 5 different restricted randomization methods and has an integrated simulation tool. This tool can be used to explore the behaviour of the different methods for various settings. For the different randomization method some scenarios are simulated. Simulation results will be discussed and compared to theoretical results.

Session:
S28: Clinical studies (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Multivariate Modelling of Infectious Disease Surveillance Data

Michaela Paul and Leonhard Held

Biostatistics Unit, ISPM, University of Zurich, Switzerland;
michaela.paul@ifspm.uzh.ch, leonhard.held@ifspm.uzh.ch

Abstract: A major challenge in infectious disease epidemiology remains the statistical analysis of spatially and temporally stratified counts of notifiable diseases collected by national surveillance systems. Such data often show a regular pattern over time such as long-term trends or seasonality, but may also contain occasional outbreaks due to the infectiousness of the disease. Interdependencies between cases caused by different pathogens, or pathogen transmission between different population subgroups (e.g. age groups, geographical areas) might particularly be of interest to further understand the dynamics of the diseases. In this talk, we will describe and extend the framework for the analysis of multivariate time series of infectious disease counts proposed by Held *et al.* (2005). The model is based on a Poisson or negative binomial observation model with two components: one component describing endemic seasonal patterns and the other one describing epidemics. We will illustrate the approach analyzing weekly counts on influenza and meningococcal disease in Germany, obtained from the Robert Koch Institute, and data on influenza mortality in the USA, obtained from the Centers for Disease Control and Prevention (CDC). The predictive performance of the models is assessed by means of one-step-ahead predictions and proper scoring rules (Gneiting and Raftery, 2007).

References:

HELD, L., HÖHLE, M. and HOFMANN, M. (2005): A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling, 5, 187–199.*

GNEITING, T. and RAFTERY, A. E. (2007): Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association, 102, 359-378.*

Session:
S33: Bayesian models in biostatistics (I) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Shrinkage Regression with Polytopes

Sebastian Petry

Department of Statistics of the Ludwig-Maximilians-University Munich, Ludwigstr. 33, 80539 Munich, Germany; Sebastian.Petry@stat.uni-muenchen.de

Abstract: Hoerl and Kennard (HOERL (1970)) gave a geometrical interpretation of the results of James and Stein (JAMES (1961)) as follows: $|\boldsymbol{\beta}_{OLS}|$ tends to be longer than $|\boldsymbol{\beta}_{true}|$. This effect can be corrected by restricting the parameter domain to regions that are point symmetric around the origin. Existence and uniqueness of the solution suggest to postulate compactness and convexity of the region. Geometrical objects with these properties can be found in the class of polytopes.

There are two ways to describe polytopes: As convex hulls of a finite set of points (V-polytopes) or as intersections of closed halfspaces (H-polytopes). Established constraint regression problems use only H-polytopes (cf. LASSO in TIBSHIRANI (1996) and OSCAR in BONDELL (2007)). Here these procedures are analyzed from the polytope theoretic perspective. It is demonstrated that there exists a class of polytopic penalty regions that allows for variable selection and clustering of predictors. New polytopic penalty regions of this class are presented by using both V- and H-representation. Chances, limitations, and problems are shown. Algorithmic aspects are discussed.

References:

BONDELL, H. D., REICH, B. J. (2007): Simultaneous regression shrinkage, variable selection and supervised clustering of predictors with OSCAR. *Biometrics (To appear).*

HOERL, A. E., KENNARD, R. W. (1970): Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics Vol .12, No. 1. 55–67.*

JAMES, W., STEIN, C. (1961): Estimation with quadratic loss. *Proc. 4th Berkeley Symp. Math. Statist. Prob., Vol. 1, University of California Press, Berkeley, pp. 361–380.*

TIBSHIRANI, R. (1996): Regression Shrinkage and Selection via the LASSO. *Journal of Royal Statistical Society. Series B (Methodologial), Vol. 58, No. 1, pp. 267–288.*

Session:

S31: Advances in Statistical Modelling (III) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Issues in the analysis of plant breeding and variety trials

Hans-Peter Piepho

Institute for Crop Production and Grassland Research, Universität Hohenheim, Fruwirthstrasse 23, 70599 Stuttgart, Germany; piepho@uni-hohenheim.de

Abstract: This talk reviews methods for nearest neighbour analysis that adjust for local trend in one direction. Such methods are commonly used in plant breeding and variety testing. The focus is on simple differencing methods, including first differences and the Papadakis method. We discuss mixed model representations of these methods on the scale of the observed data. Modelling observed data has a number of practical advantages compared to differencing, for example the facility to conveniently compute adjusted cultivar means. Most models considered involve a linear variance-covariance structure and can be represented as state-space models. The reviewed methods and models are exemplified using two datasets. We will also briefly discuss extension of linear variance and state-space models in two dimensions using separable processes.

Session:
S16: Agricultural science (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Statistical Analysis of Adaptive Designs: Solutions and Open Problems

Martin Posch

Section of Medical Statistics, Core Unit of Medical Statistics and Informatics, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria;
Martin.Posch@meduniwien.ac.at

Abstract: A minimal requirement for the implementation of adaptive designs in confirmatory clinical trials is the use of analysis methods that control the type I error rate and give conservative point estimates and confidence intervals (EMEA, 2007). Statistical methodology that meets these requirements is available for a wide range of scenarios. In this presentation we give an overview of the present portfolio of statistical procedures for the analysis of adaptive clinical trials and discuss some of the limitations and open problems. Especially, we address adaptive designs for survival trials, point estimation, the construction of simultaneous confidence bounds, and the assessment of heterogeneity between stages.

References:

EMEA (2007): Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design.

Session:
S23: Adaptive seamless design for combining phase II / III clinical studies : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Flexible Methods in Clinical Trials

Michael Proschan

NIAID, 6700B Rockledge Drive, MSC 7609 Bethesda, MD 20892, USA;
proscham@niaid.nih.gov

Abstract: For years statisticians involved in clinical trials followed rigid rules about not deviating from the protocol-specified sample size, primary endpoint, etc. Then a new frontier of adaptive designs emerged that allowed mid-course changes in trial design. Excitement became tempered with the realization that there was a tradeoff: the flexibility afforded by adaptive designs also carries with it some loss of efficiency. This talk gives a history of flexible designs, focusing mostly on two-stage and multi-stage methods for sample size modification, but also including more drastic modifications such as changing the endpoint, dropping arms, and combining different phases of trials. The talk aims to present a balanced view of the advantages and disadvantages of adaptations, as well as the most promising future directions for research in adaptive methods.

Session:
S15: Flexible designs : Tuesday, 11/03/2008, 11:00am - 12:40pm

Testing for Association of GO Pathways in Genome Wide Association Studies (GWAS)

Franz Quehenberger

Institute of Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria; franz.quehenberger@meduni-graz.at

Abstract: GWAS scan the whole human genome for association of a phenotype with single bi-allelic markers due to linkage-disequilibrium. They are typically designed as case-control studies. Huge sample sizes are needed because of small effect sizes and correction for multiple testing.

There are now several GWAS that did not provide convincing evidence for the association of single gene with a disease phenotype. Lesnick et al. (2007) report that, although no single SNP had shown significant association, they were able to build a logistic regression model that predicted Parkinson's disease from SNPs located in genes from the axon guidance pathway.

Reanalysis of the publicly available data by various tests and model building techniques ($L1$-regularization, boosting and random forests) did not provide convincing evidence for association, although the methods were powerful on simulated data. As the statistical methods in Lesnick et al. are not reported in sufficient detail, it is not clear, why it was not possible to confirm their result.

Despite the negative result in the example, reanalysis of GWAS with the proposed methods could potentially find evidence for the association of GO pathways with disease phenotypes due to small associations in many genes of a pathway.

References:

LESNICK, T. et al. (2007): A Genomic Pathway Approach to a Complex Disease: Axon Guidance and Parkinson Disease. *PLoS Genetics, 3, e98, doi:10.1371/journal.pgen.0030098* .

Session:

S20: Genome-wide association studies (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Additivity tests for the mixed model in the two-way ANOVA with single sub-class numbers. Introduction to the topic and motivation by practical problems

Dieter Rasch

University of Natural Resources and Applied Life Sciences, Institute for Applied Statistics and Computing, Gregor-Mendel-Strasse 33, A1080 Vienna, Austria;
dieter.rasch@boku.ac.at

Abstract: The above topic was selected for the work of an international research group with scientists from Prague, Vienna (Boku and Uni Wien) and Rostock because it plays an important role in many applications of biometry. Two of them will be discussed in detail:

The first problem is that of variety testing. Usually $a = v$ varieties are tested at different (b) locations, each variety at just one location (farm). For this complete block design we must assume that no interaction exists. The situation is modelled by a two-way ANOVA. To test for missing interaction for special structures of the latter tests (Tukey, Mandel, Boik) are known for a model with fixed effects of both factors. But locations must be considered to be random and in this situation no test is known.

The second problem occurs in psychological testing. Usually items (questions which can be either answered correctly or not) are put to b persons. Also in this case a mixed two-way ANOVA model must be used, the items are fixed, the persons random. Here a further difficulty stems from the fact that we have binary responses and use the Rasch model (a generalized linear model with a logistic link function). For the application of the Rasch model again the absence of interactions is needed.

References:

BOIK, R.J. (1993): Testing additivity in two-way classifications with no replications: the locally best invariant test. *Journal of Applied Statistics, 20, 41–55.*

CORSTEN, L.C.A. and VAN EIJNSBERGEN, C.A. (1972, 1974): Multiplicative effects in two-way analysis of variance. *Statistica Neerlandica 26, 61-68, addendum 27, 51.*

GHOSH, M.N. and SHARMA, D. (1963): Power of Tukey´s Test for Non-Additivity. *Jour. Roy. Stat. Soc. , Ser,. B, 25, 213-219.*

HARTER, H.L. and LUM, M.D. (1962): An interpretation and extension of Tukeyś one-degree of freedom for nonadditivity. *Aeronautical Research Laboratory Technical Report, ARL, pp. 62-313.*

JOHNSON, D.E. and GRAYBILL, F.A. (1972): An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc., 67, 862–868.*

MANDEL, J. (1961): Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association 56, 878-888.*

RASCH, D. (1995): Mathematische Statistik. *Johann Ambrosius Barth, Heidelberg - Leipzig und Wiley, Berlin (851 p.).*

SCHABENBERGER, O., GREGOIRE, T. G., and KONG, F. (2000). One degree of freedom for non-additivity. *Biometrics, 5, 232-242.*

TUKEY, J.W. (1949): One degree of freedom for nonadditivity, *Biometrics 5, 232-242.*

TUSELL, J.W. (1990): Testing for interaction in two-way ANOVA tables with no replication. *Computational Statistics and Data Analysis, 10, 29–45.*

YATES, F. (1972): A Monte-Carlo trial on the behaviour of the non-additivity test with non-normal data. *Biometrika 59, 253-261.*

Session:

S40: Additivity tests for mixed models : Thursday, 13/03/2008, 9:10am - 10:30am

Analysing and interpreting gene expression in heterogeneous tissues

Dirk Repsilber[1], Sabine Kern[2], Joachim Selbig[2,3], and Marc Jacobsen[4]

[1] FBN Dummerstorf; repsilber@fbn-dummerstorf.de
[2] University of Potsdam, Germany; skern@rz.uni-potsdam.de
[3] MPIMP, Potsdam-Golm, Germany; selbig@mpimp-golm.mpg.de
[4] MPIIB, Berlin, Germany; jacobsen@mpiib-berlin.mpg.de

Abstract: Gene expression experiments in multicellular organisms, e.g. humans, plants etc., involves isolating transcripts from tissues. Tissues are a mixture of different cell types. Hence, analysing gene expression data from tissues comes with problems concerning interpretation and confounding with cell type proportions. For some tissues and not too large sample sizes microdissection is a possible solution, such that only homogeneous parts of the tissue in question are analysed. However, this comes with additional costs and is impossible at all for most tissues. Under certain circumstances cell type specific gene expression patterns and cell type proportions can be estimated from a decomposition of the original data matrix. Approaches along this line already exist (Venet et al., 2001), but do not consider

- realistic values for noise and differential gene expression,
- validation with experimental data (on both tissue and isolated constituting cells),
- sample size plans for studies of differential gene expression.

We have improved existing approaches with respect to these issues and also concerning important algorithmic details. Both simulation study and application to an experimental validation dataset on gene expression in blood PBMC of different groups of tuberculosis patients and their contacts will be presented. We conclude that, given typical sample sizes in gene expression experiments as part of clinical studies or surveys, interpretation and analysis of gene expression data from heterogeneous tissues can be enhanced by using a decomposition approach as proposed.

References:

VENET, D., PECASSE, F., MAENHAUT, C. and BERSINI, H. (2001): Separation of samples into their constituents using gene expression data. *Bioinformatics, 17, S1, S279–S287.*

Session:
S01: Systems Biology and Bioinformatics : Monday, 10/03/2008, 9:10am - 10:30am

Bayesian Age-Period-Cohort Models for Detecting Temporal Patterns in Multivariate Mortality Data

Andrea Riebler and Leonhard Held

Biostatistics Unit, ISPM, University of Zurich, Switzerland;
andrea.riebler@ifspm.uzh.ch, leonhard.held@ifspm.uzh.ch

Abstract: In epidemiology risk factors responsible for a disease are often unknown and if there exists knowledge about such factors the available data are frequently limited. In most developed countries, however, area-level incidence and mortality rates stratified by gender and age are available. These data can be analysed by age-period-cohort (APC) models using three different time scales: age, period (calendar period during which the incidence or mortality rates were observed) and cohort (time of birth). A joint analysis of mortality rates from different regions may borrow strength from shared age effects while possibly identify different period or cohort effects. In this talk, we will describe a multivariate Bayesian APC model for inference based on Markov chain Monte Carlo (MCMC) using auxiliary variables (Frühwirth-Schnatter and Wagner 2006; Frühwirth-Schnatter *et al.* 2007). The method will be applied to mortality rates from chronic obstructive pulmonary disease (COPD) in England and Wales, split into three areas: Greater London, metropolitan counties excluding Greater London and areas outside conurbations. We were able to identify different period effects in the three areas, possibly related to different levels of air pollution. A comparison with a likelihood analysis is given.

References:

FRÜHWIRTH-SCHNATTER, S. and WAGNER, H. (2006): Auxiliary mixture sampling for parameter-driven models of time series of counts with applications to state space modelling. *Biometrika, 93, 827–841.*

FRÜHWIRTH-SCHNATTER, S., FRÜHWIRTH, R., HELD, L. and RUE, H. (2007): Improved Auxiliary Mixture Sampling for Hierarchical Models of Non-Gaussian Data. *IFAS Research Report 2007-25, http://www.ifas.jku.at, Johannes Kepler University Linz.*

Session:
S33: Bayesian models in biostatistics (I) : Wednesday, 12/03/2008, 11:00am - 12:40pm

On attempts to gain regulatory and ethics committees' acceptance for the 2-arm active control non-inferiority trial

Joachim Röhmel

Im Holze 36a, 28355 Bremen, Germany; joachim.roehmel@gmx.de

Abstract: Clinical trials testing non-classical shifted null-hypotheses are in use for many purposes. Except for cases which are aiming at the demonstration of relevant superiority a commonly used phrase for this type of trial is "non-inferiority". Although the phrase "non-inferiority" (NI) induces a sense of the opposite of inferiority this is often only loosely reflected in the design of such studies. Frequently, for example, the methodology is used for the (indirect) demonstration of superiority over placebo. Also, the frequently quoted "50% retention" rule offers little confidence regarding true claims of being non-inferior. The origin for this informal use is the lack of guidance from regulators, statisticians and clinicians. Sponsors are routinely referred to ICH 10 [ICH E 10]. Meanwhile, however, it has got around that the grapes are hanging high, and that ICH E10 resembles more of an oracle than of an expert document of practical help. A recent document [Guidance for Industry Antibacterial Drug Products] expresses very well that the wisdom regarding recommendations on how to select a proper NI margin can still grow: "For an NI study, having an adequately justified NI margin is essential to having an informative study".

Although for the proper design of NI studies input from clinicians has always been demanded and required, the research front is actually digging at a different place. Terms like "Point estimate method", "Two 95% CI rule", "synthesis method", "discounted synthesis method", "putative placebo estimate", "δ-margin", "λ% preservation-margin", "X-trial comparison" are buzzing through the journals and bear witness that at least the statisticians are avidly doing their homework. Finally, however, all such statistical work remains unfinished if there is no timely echo from the clinical site. Clinicians are struggling with the use of placebo and advocate the use of NI trials as the more ethical alternative [MICHELS (2003)]. For diseases which may result in serious complications such as e.g. depression or osteoporosis the willingness to accept placebo controlled trials decreases towards zero in ethics committees while regulatory experts reject NI trials particularly in regard of their perception of NI trials bringing along an increased risk for licensing ineffective drugs. This contrast is a great challenge for the design of clinical studies in the future.

References:

ICH E 10 guidance for industry E10 Choice of Control Group and Related Issues in Clinical Trials.

Guidance for Industry Antibacterial Drug Products: Use of Noninferiority Studies to Support Approval (draft, Oct 2007).

MICHELS, K. and ROTHMAN K. (2003): Update on unethical use of placebos in randomised clinical trials. *Bioethics 17*.

Session:
S11: Non-Inferiority Trials (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Handling Covariates in the Design of Clinical Trials

William F. Rosenberger[1] and Oleksandr Sverdlov[2]

[1] Department of Statistics, George Mason University, Fairfax, Virginia, USA; wrosenbe@gmu.edu
[2] Bristol-Myer Squibb, New York, New York, USA

Abstract: There has been a split in the statistics community about the need for taking covariates into account in the design phase of a clinical trial. There are many advocates of using stratification and covariate-adaptive randomization to promote balance on certain known covariates. However, balance does not always promote efficiency or ensure more patients are assigned to the better treatment. We describe these procedures, including model-based procedures, for incorporating covariates into the design of a clinical trials, and give examples where balance, efficiency, and ethical considerations may be in conflict. We advocate a new class of procedures, covariate-adjusted response-adaptive (CARA) randomization procedures that attempt to optimize both efficiency and ethical considerations, while maintaining randomization. We review all these procedures, present a few new simulation studies, and conclude with our philosophy.

Session:
S41: Cost efficient designs for biostatisticians : Thursday, 13/03/2008, 11:00am - 12:40pm

No events in both treatment arms - methods for meta-analysis including zero studies

Gerta Rücker[1], Guido Schwarzer[1], and James Carpenter[1,2]

[1] Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; ruecker@imbi.uni-freiburg.de
[2] Medical Statistics Unit, London School of Hygiene & Tropical Medicine, London, UK

Abstract: Zero trials (trials in which no event was observed in both treatment arms) are frequently encountered in meta-analyses of clinical trials. For trials with no events, neither risk ratio nor odds ratio are defined, wherefrom these trials are often excluded from meta-analysis. However, these may be several trials and also large trials, so that potentially valuable information is disregarded.

We investigate different methods of meta-analysis, taking as example the recent large rosiglitazone meta-analysis. We use two-stage (meta-analytic) and one-stage (regression-based) methods. Meta-analytic methods include various effect measures (risk ratio, odds ratio, risk difference, and arcsine difference), several methods of continuity correction, and different weighting methods. For the arcsine difference, a conservative variance estimate is used. Regression methods comprise logistic regression, a non-central hypergeometric model, Poisson regression with and without scale parameter, and negative binomial regression.

The systematic and comprehensive comparison of methods results in a sensitivity analysis of the given data.

References:

DIAMOND, G.A., BAX, L. and KAUL, S. (2007): Uncertain effects of rosiglitazone on the risk for myocardial infarction and cardiovascular death. *Annals of Internal Medicine, 147(8),578–581.*

KUSS, O. and GROMANN, C. (2007): An exact test for meta-analysis with binary endpoints. *Methods of Information in Medicine, 46,662–668.*

NISSEN, S.E. and WOLSKI, K. (2007): Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular diseases. *NEJM, 356,2457–2471.*

Session:

S39: Meta-analysis : Thursday, 13/03/2008, 9:10am - 10:30am

Constrained estimation of regression models with ordinal explanatory variables

Kaspar Rufibach and Leonhard Held

Biostatistics Unit, Institute of Social and Preventive Medicine, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland; kaspar.rufibach@ifspm.uzh.ch, leonhard.held@ifspm.uzh.ch

Abstract: Ordinal explanatory variables are commonplace in applied regression problems. Unconstrained estimation treats such variables in effect as unordered, whereas modelling as a continuous variable requires quantitative levels of the variable, which may not be available.

We propose a new constrained regression estimator when several ordinal predictors are present. Our approach ensures that the estimated parameter corresponding to the dummy variable of a "higher" group of an ordinal predictor is at least as big as those of "lower" groups. We show that the constrained estimator $\widehat{\boldsymbol{\beta}}$ of the regression coefficients can be computed in finitely many steps and very efficiently via an active set algorithm. A characterization of $\widehat{\boldsymbol{\beta}}$ is given and some simple cases where direct computation of $\widehat{\boldsymbol{\beta}}$ is possible are indicated. The fast algorithms allow bootstrap inference in these models.

The new method will be compared to a Bayesian regression approach, where the order constraints are imposed through the prior distribution. Inference is then based on the output from a Markov chain Monte Carlo algorithm, where point and interval estimates are readily available based on samples from the posterior distribution.

The two approaches will be applied to a logistic regression example analyzed in Taussky et al. (2005).

References:

TAUSSKY, D., RUFIBACH, K., HUGUENIN, P., ALLAL, A. (2005): Risk factors for developing a second aerodigestive cancer after radiotherapy with or without chemotherapy for head and neck cancer: An exploratory outcome analysis. *Int. J. Radiat. Oncol. Biol. Phys., 62, 684–689.*

Session:

S31: Advances in Statistical Modelling (III) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Additivity tests for the mixed model in the two-way ANOVA with single sub-class numbers - Robustness of additivity tests with binary observations

Thomas Rusch[1], Dieter Rasch[2] and Klaus D. Kubinger[3]

[1] Institute of Developmental Psychology and Psychological Assessment, University of Vienna, Austria; thomas.rusch@univie.ac.at
[2] Institute of Applied Statistics and Computing, University of Natural Resources and Applied Life Sciences, Vienna, Austria; dieter.rasch@boku.ac.at
[3] Institute of Developmental Psychology and Psychological Assessment, University of Vienna, Austria; klaus.kubinger@univie.ac.at

Abstract: In the case of two-way layouts, testing the interaction effect is difficult if only a single observation is present in each cell. Regarding model type I - that is both factors are fixed - special tests have been developed of which the Tukey test (1949) is best known. However, empirical data very often refer to the mixed model. Furthermore, the tests which have been developed for model type I all use the standard assumptions of a normal distribution of errors and constant variance. Hence, even if they perform satisfactorily in a mixed model layout with normal errors, this will not necessarily be the case for non-normal data such as binary observations. In this paper we evaluate model type I tests in case of the mixed model and dichotomous data to find a valid and appropriate test for checking if interactions are present. For this a simulation study is applied. The question is which of the tests is robust in the sense that the type-I-error, given an $\alpha = .05$, actually differs not more than 20 % from α, i.e. $.04 \leq \widehat{\alpha}_{act} \leq .06$.

References:

TUKEY, J.W. (1949): One degree of freedom for nonadditivity. *Biometrics, 5, 232–242.*

YATES, F. (1972): A Monte-Carlo trial on the behaviour of the non-additivity test with non-normal data. *Biometrika, 59, 253–261.*

Session:

S40: Additivity tests for mixed models : Thursday, 13/03/2008, 9:10am - 10:30am

Approaches to a large scale forest inventory by subsampling of current two-phase designs for districts

Joachim Saborowski[1], Felix Mader[1], and Thomas Böckmann[2]

[1] Abteilung Ökoinformatik, Biometrie und Waldwachstum, Georg-August-Universität Göttingen, Germany; {jsaboro, fmader}@gwdg.de
[2] Niedersächsisches Forstplanungsamt, Wolfenbüttel, Germany; Thomas.Boeckmann@nfp.niedersachsen.de

Abstract: Forest district inventories in Niedersachsen are currently performed according to a two-phase sampling design. In the first phase, potential sample plots are assigned to one of 8 strata (4 age classes of 2 dominating species groups) by aerial photo interpretation. In the second phase, a subsample of concentric circular plots is drawn from the first phase sample points of each stratum and tree characteristics (e.g. species, height and diameter at breast height) are measured. All districts are sequentially inventoried in a period of about 10 years. As a control of country-wide sustainable management and efficient assessment of calamities, it might be appropriate to carry out a rapid extensive inventory for the entire country, which is based on existing second-phase terrestrial sample plot locations. To reduce travelling costs, we seek a new design that additionally provides spatially concentrated samples, as far as possible. Therefore, we study different two-stage-two-phase and three-phase designs and compare the respective sampling errors for basal area per ha of different target populations (species and diameter classes), based on existing inventory data of 22 forest districts.

Session:
S43: Statistical Methods in Environmental Monitoring : Thursday, 13/03/2008, 11:00am - 12:40pm

Investigation of interactions between treatment and continuous covariates by fractional polynomial methodology

Willi Sauerbrei[1] and Patrick Royston[2]

[1] Institute of Medical Biometry and Informatics, University Medical Center Freiburg, Germany; wfs@imbi.uni-freiburg.de
[2] MRC Clinical Trials Unit, London, UK.

Abstract: With larger clinical trials there is considerable interest in investigating whether a treatment effect is similar in all patients or whether a subgroup of patients profits more from a treatment than the remainder. For a continuous covariate Z the usual approach to analysis is to categorise Z into groups according to cutpoint(s) and to analyse the interaction in a model with main effects and multiplicative terms. The cutpoint approach raises several well-known and difficult issues for the analyst. Recently Royston & Sauerbrei (2004) extended the multivariable fractional polynomial (MFP) approach to investigate treatment-covariate interactions (MFPI). Cutpoints are avoided. Estimating a treatment effect in overlapping subgroups defined with respect to a continuous covariate, Bonetti & Gelber (2000) introduced the 'subpopulation treatment effect pattern plot' (STEPP) method. In an example we will show that STEPP may be used as a simple and effective check of postulated interactions from MFPI modelling.

References:

BONETTI, M. and GELBER, R.D. (2000): A graphical method to assess treatment-covariate interactions using the Cox model on subsets of the data. *Stat Med, 19, 2595-2609.*

ROYSTON, P. and SAUERBREI, W. (2004): A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Stat Med, 23, 2509-2525.*

SAUERBREI, W. and ROYSTON, P. and ZAPIEN, K. (2007): Detecting an interaction between treatment and a continuous covariate: a comparison of two approaches. *Computational Statistics and Data Analysis, 51, 4054-4063.*

Session:

S31: Advances in Statistical Modelling (III) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Using Neighborhood Graphs for the Investigation of Gene Clusters

Theresa Scharl[1,2] and Friedrich Leisch[3]

[1] Department of Statistics and Probability Theory; Vienna University of Technology; Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria; theresa.scharl@ci.tuwien.ac.at
[2] Department of Biotechnology; University of Natural Resources and Applied Life Sciences; Muthgasse 18, A-1190 Vienna, Austria
[3] Department of Statistics; University of Munich Ludwigstraße 33, D-80539 München, Germany

Abstract: Gene expression microarray experiments yield large and complex multivariate datasets that consist of several thousands of genes under multiple conditions. Cluster methods have been frequently applied to gene expression data in order to find homogeneous subgroups of co-expressed genes that may imply co-regulation or relation in functional pathways. The display of cluster solutions particularly for a large number of clusters is very important in exploratory data analysis. Visualization methods give practitioners an understanding of the underlying cluster structure and makes it easier to interpret the cluster results. In this work the utility of neighborhood graphs for the graphical representation of centroid-based partitioning cluster objects is shown. Neighborhood graphs allow for visual assessment of relationships between adjacent clusters. Beside the graphical representation of the cluster structure neighborhood graphs are used to visualize further information about specific clusters. Different colors or shapes are used for the node representation in order to highlight potentially interesting clusters. Additional information include cluster properties like cluster size or cluster tightness as well as external knowledge from differential expression analysis or functional grouping. The new visualization method is demonstrated on microarray data from E. coli.
Acknowledgement: The project was funded by the Austrian Kind/Knet Center of Biopharmaceutical Technology (ACBT).

Session:
S36: Young Statistician Papers (IBS-ROeS) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Nonparametric Tests for Sparse Marked Spatial Point Processes, with an application to FISH Images

Fabian Scheipl

Institut für Statistik, Ludwig-Maximilians-Universität München, D-80539 München, Germany; Fabian.Scheipl@stat.uni-muenchen.de

Abstract: Recent progress in cell and nuclear biology shows an inherent relationship between the spatial organization and the function of cell nuclei. The spatial arrangement of chromatin loci in the nucleus and the interaction between nuclear bodies and chromatin loci is seen as a critically important factor in transcriptional regulation.

Conventional marked point process methodology such as cross-type K-function estimators and related quantities is frequently not applicable to the analysis of multicoloured fluorescence in situ hybridization (FISH) labeled sequences due to the limited number of observed marks (e.g. 2 alleles per gene per nucleus).

As a robust alternative, we present some permutation tests for interaction between the localizations of marks in replicated point patterns in the context of clustering between heterologous pairs or triples of genomic loci and clustering between genomic loci and nuclear bodies. We also introduce tests for the analysis of attraction of FISH signals to the nuclear center or border. The performance of the tests has been assessed in simulation studies. An exemplary application to mouse neuron data is discussed.

References:

DIGGLE, P.J., LANGE, N., BENES, F.M. (1991): Analysis of Variance for Replicated Spatial Point Patterns in Clinical Neuroanatomy. *JASA, 86, 618–625*

RONNEBERGER, O., BADDELEY, D., SCHEIPL, F. et al. (2008): Spatial quantitative analysis of fluorescently labeled nuclear structures: problems, methods, pitfalls. *submitted*

Session:

S24: Spatial Analysis of Surveillance Data : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Genome wide Association Studies of Complex Traits and Strategies for Marker Validation

André Scherag[1], Johannes Hebebrand[2], and K.H. Jöckel[1]

[1] Institute of Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Essen, Germany; andre.scherag@uk-essen.de; k-h.joeckel@uk-essen.de
[2] Department of Child and Adolescent Psychiatry and Psychotherapy, University of Duisburg-Essen, Essen, Germany; johannes.hebebrand@uni-due.de

Abstract: Advances in high-throughput genotyping technology lead to the realization of genome wide association (GWA) studies which helped identify new susceptibility loci of complex traits. Such studies usually start with genotyping fixed arrays of single nucleotide polymorphisms (SNPs) in an initial sample. Out of these lists of markers are compiled which will be further genotyped in independent samples. Due to the very low a priori probability of a true positive association, the fast majority of marker signals will turn out to be false positive.

As shown by Frayling (2007) for type II diabetes mellitus, some of the true associations will not be among those with the smallest p-values. Here we report similar data from a GWA scan for the phenotype "early onset extreme obesity" (Hinney et al., in press). No marker within classical supported candidate genes (Rankinen et al., 2006) was found among the top 1,000. Thus, we evaluate and discuss alternative ways of sorting single SNP data (e.g., Wakefield, 2007) and demonstrate their degree of overlap using the real data example.

References:

FRAYLING, TM. (2007): Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet, 8(9),657–662.*

HINNEY, A., et al. (in press): Genome-wide Scan for Early Onset Extreme Obesity Supports the Role of Fat Mass and Obesity Associated Gene (FTO) Variants. *PLoS One.*

WAKEFIELD, J. (2007): A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet, 81(2), 208–227.*

RANKINEN, T., et al. (2006): The human obesity gene map: the 2005 update. *Obesity, 14, 529–644.*

Session:

S20: Genome-wide association studies (II) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Estimating relative risk of pressure ulcers

Ramona Scheufele and Ekkehart Dietz

Department of Medical Statistics and Clinical Epidemiology,
Charite-Universitätsmedizin Berlin, 10098 Berlin, Germany;
ramona.scheufele@charite.de

Abstract: Pressure ulcers are one of the major challenges in nursing care. A large number of individuals develop pressure ulcers which cause pain and restrictions to the life of the individuals concerned, increase costs for health care systems, and are connected to higher mortality rates. Risk assessment is recommended as the first step in pressure ulcer prevention. Doing this by logistic regression analysis of available prevalence studies leads to estimates of prevalence odds ratios. It is well known, however, that the interpretation of prevalence odds ratios as relative risks can be heavily misleading. Therefore, at analysis of the *GPUP2005* (german pressure ulcer 2005 prevalence) study, we applied a computational simple method to adjust for a potential PI (prevalence-incidence) bias, which had been suggested recently (Dietz et al. 2007). The method combines the results of the usual logistic regression analysis with results of an survival analysis of disease duration. It is also possible to use observations of the duration to date of prevalent cases. Thus, the method could be applied to respective data available for the GPUP2005-study. These data are interval censored observations, however, so that an EM-algorithm had to be implemented. The estimated relative risks obtained by the method were quite different from the respective prevalence odds ratios but there was no effect reversion.

References:

MERTENS, E.I., HALFENS, R.J.G., DIETZ, E., SCHEUFELE, R., DASSEN, Th. (2008): Pressure Ulcer Risk Screening with the Care Dependency Scale (CDS). *Journal of Evaluation in Clinical Practice, to appear.*

DIETZ, E., SCHEUFELE, R., and BÖHNING, D. (2008): Estimating relative risk from prevalence studies: prevalence-incidence bias correction. *Statstics in Medicine, submitted.*

Session:
S17: Event Data Analyses (II) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Comparison of four estimators of the heterogeneity variance for meta-analysis when the distribution of random effects is not normal

Peter Schlattmann

Dept. of Biostatistics and Clinical Epidemiology, Charité Universitätsmedizin, 10117 Berlin, Germany; peter.schlattmann@charite.de

Abstract: The analysis of heterogeneity is a crucial part of each meta-analysis. In order to analyze heterogeneity often a random effects model which incorporates variation between studies is considered. It is assumed that each study has its own (true) exposure or therapy effect and that there is a random distribution of these true exposure effects around a central effect. The variability between studies is quantified by the heterogeneity variance.

In order to compare the performance of four estimators of the heterogeneity variance a simulation study was performed. This study compared the Dersimonian-Laird (1986) estimator with the maximum-likelihood estimator based on the normal distribution for the random effects. Further comparators were the simple heterogeneity (SH) variance estimator proposed by Sidek and Jonkman (2005).

All of the afore mentioned methods assume a normal distribution for the random effects. This assumption may be true or not. Thus an alternative estimator of the heterogeneity variance is based on a finite mixture model (Böhning, Dietz, and Schlattmann, 1998).

This simulation study investigates these four estimators, when sampling from discrete distributions, i.e. the major assumption of a normal distribution for the random effects is not fulfilled. In this setting the simulation study investigates bias, standard deviation and mean square error (MSE) of all four estimators.

Additionally, coverage probabilities of 95% confidence intervals for the heterogeneity variance based on the various methods are considered.

References:

BÖHNING, D., DIETZ, E. and SCHLATTMANN, P. (1998): Recent developments in computer assisted mixture analysis. *Biometrics, 54, 283-303*

DERSIMONIAN, R. and LAIRD, N. (1986): Meta-analysis in clinical trials. *Controlled Clinical Trials, 7, 177-188*

SIDEK, K. and JONKMAN, J. (2005): Simple heterogeneity variance for meta-analysis. *JRSS Series C, 54, 367-384*

Session:

S39: Meta-analysis : Thursday, 13/03/2008, 9:10am - 10:30am

Boosting Additive Models using Component-wise P-Splines

Matthias Schmid[1] and Torsten Hothorn[2]

[1] Institut für Medizininformatik, Biometrie und Epidemiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Waldstr. 6, D-91054 Erlangen, Germany; matthias.schmid@imbe.imed.uni-erlangen.de
[2] Institut für Statistik, Ludwig-Maximilians-Universität München, Ludwigstr. 33, D-80539 München, Germany; Torsten.Hothorn@stat.uni-muenchen.de

Abstract: In recent years boosting has developed into one of the most important techniques for regularization and model fitting in high-dimensional data settings. The technique can be considered as a gradient-descent algorithm in function space and can therefore be used for a large variety of loss functions and statistical problems. An important example is boosting with the least squares (L2) loss function in combination with smoothing spline baselearners, which has been established by Buhlmann and Yu (2003) as a technique for fitting additive regression models with continuous outcomes.

From a computational point of view, Bühlmann and Yu's L2-Boosting algorithm can be further improved, since the integrated squared second-order derivative penalty of smoothing spline baselearners is relatively time-consuming to evaluate. We therefore consider an efficient approximation of L2-Boosting with component-wise smoothing splines. Smoothing spline baselearners are replaced by P-spline baselearners, which yield similar prediction errors but are computationally more efficient. By means of theoretical results and a simulation study, we give a detailed analysis on the effect of various P-spline hyper-parameters on the boosting fit. Our results are illustrated by a recent application example from clinical research.

References:

BÜHLMANN, P. and YU, B. (2003): Boosting with the L2 loss: regression and classification. *Journal of the American Statistical Association 98, 324–339.*

SCHMID, M. and HOTHORN, T. (2007): Boosting additive models using component-wise P-splines. Technical Report 002, Department of Statistics, University of Munich.

Session:

S30: Boosting for biomedical data : Wednesday, 12/03/2008, 11:00am - 12:40pm

Parallelized preprocessing algorithms for high-density oligonucleotide array data

Markus Schmidberger and Ulrich Mansmann

Chair of Biometrics and Bioinformatics, IBE, University of Munich, Germany;
schmidb@ibe.med.uni-muenchen.de

Abstract: Studies of gene expression using high-density oligonucleotide microarrays have become standard in a variety of biological contexts. The data recorded using the microarray technique are characterized by high levels of noise and bias. These failures have to be removed, therefore preprocessing of raw-data has been a research topic of high priority over the past few years.

Actual research and computations are limited by the available computer hardware. For many researchers the available main memory limits the number of arrays that may be processed. Furthermore most of the existing preprocessing methods are very time consuming and therefore not useful for first and fast checks in laboratories. To solve these problems, the potential of parallel computing should be used. In microarray technologies and statistical computing parallel computing does not appear to have been used extensively. For parallelization on multicomputers, message passing (MPI) methods and the R language will be used.

This presentation proposes the new R language package `affyPara` for parallelized preprocessing of high-density oligonucleotide microarray data. Partition of data could be done on arrays and therefore parallelization of algorithms gets intuitive possible. In view of machine accuracy, the same results as serialized methods will be achieved. The partition of data and distribution to several nodes solves the main memory problems and accelerates the methods by up to the factor ten.

References:

GENTLEMAN, R., et al. (2005): Bioinformatics and Computational Biology Solutions Using R and Bioconductor. *Springer, Statistics for Biology and Health.*

ROSSINI, A. (2003): Simple Parallel Statistical Computing in R. *UW Biostatistics Working Paper Series, 193.*

SEVCIKOVA, H. (2003): Statistical Simulations on Parallel Computers. *Journal of Computational and Graphical Statistics, 13, 886-906.*

IRIZARRY, R. A., et al (2003): Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics 4 (2003), Apr, Nr. 2, 249-264.*

Session:

S29: Advances in Statistical Modelling (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Comparing linear mixed models in assessment of Parkinson dysphagia

Irene Schmidtmann[1], Tobias Pirot[2], and Peter Paul Urban[3]

[1] IMBEI, Klinikum der Johannes Gutenberg-Universität Mainz, D-55101 Mainz, Germany; schmidtm@imbei.uni-mainz.de
[2] Abteilung für stereotaktische Neurochirurgie, Neurozentrum, Uniklinik Freiburg, 79106 Freiburg, Germany; Tobias.Piroth@uniklinik-freiburg.de
[3] Asklepios Klinik Barmbek, 22291 Hamburg, Germany; p.Urban@asklepios.com

Abstract: 29 patients suffering from idiopathic Parkinson's disease were examined by means of mouth base EMG during deglutition after administration of four different boluses: 5 ml mush, 5 ml, 10 ml and 15 ml water. Each patient swallowed each type of bolus eight times. Various parameters characterising the process of deglutition were recorded. Previously 60 controls had been examined in exactly the same way. We analysed data from 28 idiopathic Parkinson patients and 41 controls after excluding 10 controls aged under 30 due to lack of young Parkinson patients in our study and 1 patient and 9 controls due to poor quality of the EMG. Here we only deal with the duration of apnoe.

We investigated the dependence of duration of apnoe on group (Parkinson patient vs controls), bolus, dysphagia and age. We analysed log(duration of apnoe) using SAS PROC MIXED. The models we considered always included effects of group and bolus (fixed), some additionally contained dysphagia and/or age effects (fixed). Patient effects were considered random. We applied various covariance structures and compared the various models using the Akaike information criterion (AIC).

We found that – among the models considered – the best model included fixed group and bolus effects plus interaction, random patient effects and patient-bolus-interaction. Different variance estimates for controls, Parkinson patients with dysphagia and Parkinson patients without were obtained. However, only when testing for bolus, we observed $p < 0.05$. Among the controls there is the least variation whereas Parkinson patients with manifest dysphagia show most variation in duration of apnoe. Also the interaction of bolus and patient is most pronounced in Parkinson patients with manifest dysphagia.

References:

VERBEKE, G. and MOHLENBERGHS, G. (2001): *Linear Mixed Models for Longitudinal Data*. Springer, Berlin.

Session:

S32: Hierarchical models : Wednesday, 12/03/2008, 11:00am - 12:40pm

Adaptive Designs with Optimal Sample Size by Step Functions

Astrid Schneider, Gerhard Hommel, and Andreas Faldum

Institute of Medical Biostatistics, Epidemiology and Informatics, Langenbeckstr.1, D-55101 Mainz, Germany; schneider@imbei.uni-mainz.de

Abstract: When planning a trial with an adaptive trial there are many possibilities for choosing an appropriate conditional error function. One of the most popular designs is the Inverse Normal Method which starts with a group sequential design. With adaptive designs it is possible to minimize the average and the maximal sample size while controlling a prespecified power simultaneously. In order to approximate optimal conditional error functions, step functions can be used. The step functions are compared with the best possible Inverse Normal Method. The assessment of the conditional error functions depends on the quality of the a priori estimation of the treatment effect, on the choice of the "stop-for-futility" bound and on the conditional power. Therefore the impact of those parameters on the choice of the best step function is analysed. The procedure can help choosing an appropriate adaptive design with a continuous or a step-wise continuous conditional error function.

References:

BAUER, P. and KÖHNE, K. (1994): Evaluation of experiments with adaptive interim analyses. *Biometrics, 50, 1029–1041.*

LEHMACHER, W. and WASSMER, G. (1999): Adaptive sample size calculations in group sequential trials. *Biometrics, 55, 1286–1290.*

Session:

S05: Adaptive Group Sequential Designs : Monday, 10/03/2008, 2:10pm - 3:30pm

Sample Size Estimation for Cluster-Randomized Clinical Trials

Carsten Schwenke

SCO:SSiS, Zeltinger Str. 58g, D-13465 Berlin, Germany;
carsten.schwenke@scossis.de

Abstract: Cluster-randomized studies constitute a very popular study design in several areas like, e.g., clinical trials where diagnostic studies evaluate new imaging contrast agents in several vessels in the same patient or dermatologic studies evaluate several skin sites on the body. The statistical challenge in cluster-randomized trials is to take into account the correlation of multiple observational units per cluster. Before analyzing a trial, the trial has to be planned adequately also with regard to the required sample size to power the trial appropriately. Several suggestions for corrections for the effect of clustered data were published. As an easy approach, one may use the intracluster correlation coefficient to account for the cluster effect assuming a fixed average number of units within each cluster as introduced by Donald and Donner (1997). Another example is given by Kang et al. (2003) who also proposed sample size formulae for clustered data and an estimate with equal weights across observational units to estimate the required number of clusters for dichotomous outcomes. In case of varying cluster sizes, they also propose to use the average cluster size for the sample size estimation. However, the correlation coefficient and the average cluster size have to be evaluated carefully as they have huge impact on the sample size. These effects will be demonstrated by examples. Pitfalls with regard to the sample size will be described and solutions provided how to take into account the effects of clustered data and how to power a trial appropriately.

References:

DONALD, A. and DONNER, A. (1997): Adjustment to the Mantel-Haenszel Chi-Square Statistic and Odds Ratio Variance Estimator when the Data are Clustered. *Statistics in Medicine, 6, 491–499.*

KANG, S.-H., AHN, C. and JUNG, S.-H. (2003): Sample Size Calculation for Dichotomous Outcomes in Cluster Randomization Trials with Varying Cluster Size. *Drug Information Journal, 37,109–114.*

Session:
S28: Clinical studies (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Analysis of case series - assumptions for risk estimation using case-control like approaches

Peggy Sekula and Martin Schumacher

University Medical Center Freiburg, Institute of Medical Biometry and Medical Informatics, Germany; ps@imbi.uni-freiburg.de

Abstract: In certain circumstances, the analysis of case series provides a valuable approach for risk assessment of factors causing transient change in risk. Since each case serves as its own control by comparing the exposure at different times (so-called case and referent periods), it is obvious that their choice is crucial. Different strategies for selecting referent periods have been proposed on which various methods for estimation of the risk are based. In principle, the cohort-like approach, where the time of the event is considered as random given the observed exposure, and the case-control-like approach, where the exposure is regarded as random given the time of the event, can be distinguished. Here, we will concentrate on unidirectional referent selection strategies and related estimators (Mantel-Haenszel-estimator, conditional logistic regression, maximum likelihood estimator) based on case-control-like approaches and will derive necessary assumptions for valid estimation of risk. The methodology will be illustrated using data collected within the EuroSCAR study to estimate the risk of certain drugs to cause a severe cutaneous adverse reaction.

References:

MACLURE, M. (1991): The case-crossover design: a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology, 133, 144–153.*

GREENLAND, S. (1999): A unified approach to the analysis of case-distribution (case-only) studies. *Statistics in Medicine, 18, 1–15.*

VINES, S.K. and FARRINGTON, C.P. (2001): Within-subject exposure dependency in case-crossover studies. *Statistics in Medicine, 20, 3039–3049.*

MOCKENHAUPT, M. et al. (2007): Stevens-Johnson syndrome and toxic epidermal necrolysis: Assessement of medication risks with emphasis on recently marketed drugs. The EuroSCAR-study. *Journal of Investigative Dermatology, online available.*

Session:

S17: Event Data Analyses (II) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Flexible Matching Coefficients for Clustering Genetic Data

Silvia Selinski[1] and Katja Ickstadt[2]

[1] Fakultät Statistik, Technische Universität Dortmund, 44221 Dortmund, Germany; selinski@statistik.uni-dortmund.de
[2] Fakultät Statistik, Technische Universität Dortmund, 44221 Dortmund, Germany; ickstadt@statistik.uni-dortmund.de

Abstract: The search for patterns in categorical data sets, for instance genetic data from association studies of complex diseases, can be performed by a variety of clustering and classification approaches. We employ hierarchical cluster analysis which requires the specification of a suitable similarity measure. For SNP data, our main application, the term 'similarity' is still vague. The general structure of these data presents a major problem to account for: the minor allele frequency is often rather small compared to the major allele frequency which leads to a masking effect due to the common occurrence of the homozygous reference sequences.

Therefore, we developed Flexible Matching Coefficients that generalise the conventional Matching Coefficients, e.g. Simple Matching and the Jaccard coefficient, often used in the cluster analysis of categorical data. We compare the performance of the Flexible Matching Coefficients, conventional Matching Coefficients and coefficients based on the χ^2 statistic using 40 simulated SNP data sets of four different scenarios and apply the Flexible Matching Coefficients and Pearson's Corrected Coefficient of Contingency to a SNP data set from the GENICA case-control study of sporadic breast cancer.

For the detection of the causative SNPs Flexible Matching Coefficients perform best whereas Pearson's Corrected Coefficient of Contingency is more suitable to detect the general block structure of related SNPs.

References:

NOTHNAGEL, M. (2002): Simulation of LD block-structured SNP haplotype data and its use for the analysis of case-control data by supervised learning methods. *Genetics, 71 (suppl.), A2363.*

Session:

S27: Structuring high-dimensional data : Wednesday, 12/03/2008, 9:10am - 10:30am

Non-inferiority Trials – Towards a New Paradigm

Stephen Senn

Department of Statistics, University of Glasgow, Scotland;
stephen@senns.demon.co.uk

Abstract: Conventional discussions of statistical analysis of clinical trials have assumed that the Neyman-Pearson framework of hypothesis testing was appropriate. This had led to the sometimes absurd pantomime of decision-making in which we have been invited to consider as rational, behaviour in which a group of scientists are mandated to make a decision once and for all on behalf of all scientific posterity.

In the case of non-inferiority trials this framework is now under such extreme strain that it will shortly crack and splinter. According to current regulatory norms, we are asked to believe that a sponsor has the chief responsibility for choosing a margin of non-inferiority and that the most important feature of this margin is not that it is narrow enough but that it is pre-specified.

An alternative paradigm would see clinical trials as delivering information which different individuals (patients, prescribers, reimbursers) may judge has different relevance. There is then no question of a given trial proving that a drug is significantly non-inferior to a comparator. The trial delivers information for others to use.

Furthermore, the increasing integration of information from different sources, means that the age of judging trial results or even programme results in isolation is over. Current information initiatives from the Cochrane Collaboration to the Clinical Data Interchange Standards Consortium (CDISC) and, of course, the world wide web mean that all trials will in future be studied in terms of an 'ineluctable extended global context'. We might as well retire the idea of a sponsor chosen non-inferiority margin today and accept that the immediate deliverable of a non-inferiority trial is not decisions but information.

Session:
S19: Non-Inferiority Trials (III) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Additivity tests for the mixed model in the two-way ANOVA with single sub-class numbers – On a modification of Tukey's test

Petr Šimeček and Marie Šimečková

Institute of Animal Science, Přátelství 815, Prague, Czech Republic;
simecek.petr@vuzv.cz; simeckova.marie@vuzv.cz

Abstract: In this contribution we discuss a problem of testing for an interaction in the two–way array with just one observation per cell. The known results are reviewed and a modification of Tukey's additivity test is derived.

In a simulation study we show that when the interaction is a product of the main effects, the power of the modified test appears to be similar to the power of Tukey's or Mandel's test and outperforms the Johnson–Graybill, locally best invariant (LBI) and Tussel's test. When the interaction scheme is more general the power of the modified test is not as good as for the Johnson–Graybill, LBI and Tussel's test but is still much better than Tukey's and Mandel's tests.

All tests have been implemented in R package AdditivityTests.

References:

ALIN, A. and KURT, S. (2001): Testing non-additivity (interaction) in two-way ANOVA tables with no replication. *Statistical Methods in Medical Research, 15, 63–85.*

BOIK, R.J. (1993): Testing additivity in two-way classifications with no replications: the locally best invariant test. *Journal of Applied Statistics, 20, 41–55.*

JOHNSON, D.E. and GRAYBILL, F.A. (1972): An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc., 67, 862–868.*

MANDEL, J. (1961): Non-additivity in two-way analysis of variance. *J. Amer. Statist. Assoc., 56, 878–888.*

TUKEY, J.W. (1949): One degree of freedom for non-additivity. *Biometrics, 5, 232–242.*

TUSELL, J.W. (1990): Testing for interaction in two-way ANOVA tables with no replication. *Computational Statistics and Data Analysis, 10, 29–45.*

Session:

S40: Additivity tests for mixed models : Thursday, 13/03/2008, 9:10am - 10:30am

Use of the heterogeneity interval to distinguish between the results of fixed and random effects models in meta-analytic results

Guido Skipka

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), Dillenburger Str. 27, 51105 Köln, Germany; guido.skipka@iqwig.de

Abstract: Meta-analyses are widely used to combine the results of clinical studies by calculating statistics for overall treatment effects, for example, quantified as odds ratios. Basically, two different models are applied in meta-analyses. The fixed effect model (FEM) assumes that in each study the same treatment effect θ is measured. Different estimates for θ are expected to arise solely from sampling error. In contrast, the random effects model (REM) incorporates the between-study variation τ^2, taking heterogeneous true effects into account [Higgins et al. (2002)]. Although the two approaches - FEM and REM - estimate different parameters (true effect versus expectation of the distribution of true effects), in practice the results are presented in the same way. The point and interval estimate of θ is commonly presented graphically in a forest plot as a diamond, irrespective of the model chosen. The estimate of the between-study variation τ is therefore not considered in the presentation of the results of REMs. Following the REM, we propose a new interval, called the *heterogeneity interval*. Assuming that the pooled estimate $\widehat{\theta}$ is approximately normally distributed, the interval $[\widehat{\theta} - 1.96 \times \widehat{\tau};\ \widehat{\theta} + 1.96 \times \widehat{\tau}]$ provides a region in which about 95% of the true study effects are to be expected. This is in contrast to the commonly presented confidence interval for the average effect θ which only quantifies the precision of the estimate for θ, and the width of the interval becomes smaller the more studies are included in the meta-analysis. The heterogeneity interval, on the other hand, is almost independent of the number of studies. We propose an extension of the forest plot which provides a graphical illustration of the extent of heterogeneity.

References:

HIGGINS, J.P.T. and THOMPSON, S.G. (2002): Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539–1558.

Session:

S39: Meta-analysis : Thursday, 13/03/2008, 9:10am - 10:30am

Space-time cluster detection using scan statistics with examples in disease surveillance and drug development

Christian Sonesson

AstraZeneca R&D, Mölndal, Sweden; Christian.Sonesson@astrazeneca.com

Abstract: Different ways of constructing space-time scan statistics based on surveillance theory are presented. We bridge the ideas from space-time disease surveillance, public health surveillance and industrial quality control and show that previously suggested space-time scan statistics methods can be fitted into a general CUSUM framework. Crucial differences between the methods studied are due to different assumptions about the spatial process. An example is the specification of the spatial regions of interest for a possible cluster, another is the increased rate to be detected within a cluster. The methods are applied to the detection of an increased incidence of Tularemia in Sweden. Different situations where this type of analyses could be useful in drug development are also discussed.

Session:
S24: Spatial Analysis of Surveillance Data : Tuesday, 11/03/2008, 4:00pm - 6:10pm

A simulation study on conditional and marginal inference of propensity score based approaches in comparison to linear modeling

Susanne Stampf[1], Erika Graf[1,2] and Claudia Schmoor[1,2]

[1] Institute of Medical Biometry and Medical Informatics, Freiburg, Germany; susta@fdm.uni-freiburg.de
[2] Center of Clinical Trials, University Medical Center, Freiburg, Germany

Abstract: Propensity score methods are being applied more and more frequently in the analysis of medical research data. Stratification by propensity score is one of the popular approaches to deal with confounding when estimating treatment effects as also done by linear regression techniques. But both approaches differ in their fundamental framework and in the handling of covariates. In our simulation study we distinguish between conditional and marginal inference as discussed in the recent work of Senn et al. They derive theoretically the properties of the propensity score effect estimator compared to an estimator from linear modeling in a conditional setting characterized by fixed covariates. The question of a correct variance estimation remains open and shall be investigated in the framework of our simulation study by using several proposals. Regarding marginal inference, in general no unique best linear unbiased estimator (BLUE) for the effect of treatment exists. This is illustrated in the simulation study: situations where the usual least square estimator from linear modeling outperforms the propensity score and vice versa are contrasted.

References:

SENN, S. and GRAF, E. and CAPUTO, A. (2007): Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure. *Statistics in Medicine, 26(30):5529–5544.*

HAHN, J. (1998): On the role of propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica, 66(2):315–331.*

POPPER SHAFFER, J. (1991): The Gauss-Markov theorem and random regressors. *The American Statistician, 45:269–272.*

Session:
S37: Clinical studies (III) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Construction of optimal designs for estimating the treatment×cell line interaction effect in two-colour cDNA microarray experiments

Sven Stanzel and Ralf-Dieter Hilgers

Institute for Medical Statistics, RWTH Aachen, Germany; sstanzel@ukaachen.de, rhilgers@ukaachen.de

Abstract: Two-colour cDNA microarrays are a powerful tool for gene expression analyses. Landgrebe et al. proposed a gene-specific fixed effects linear model for statistical evaluation of the gene expression data generated in two-colour cDNA microarray experiments (Landgrebe et al., 2006).

In this talk we will consider a specific three-factorial layout of the experiments with N microarrays, two colours, $K \geq 2$ cell lines and $L \geq 2$ treatments. We will present ϕ_p-optimal block designs for estimating the linear contrast related to the treatment×cell line interaction term of the Landgrebe model. We adopt modified versions of the generalized equivalence theorems suggested by Pukelsheim (Pukelsheim, 1972) to proof optimality. We demonstrate the interesting result that the solution of the optimization problem depends on the relation between the numbers K and L. The independency of this solution from the optimality criterion chosen can be interpreted as robustness characteristic.

For practical purposes, the optimality results obtained can be transferred into direct recommendations for the choice of an efficient design for a concrete experiment with given numbers of microarrays, treatments and cell lines.

References:

LANDGREBE, J., BRETZ, F. and BRUNNER, E. (2006): Efficient design and analysis of two-colour factorial microarray experiments. *Computational Statistics and Data Analysis, 50, 499–517.*

PUKELSHEIM, F. (1972): *Optimal Design of Experiments.* Wiley, New York.

Session:

S41: Cost efficient designs for biostatisticians : Thursday, 13/03/2008, 11:00am - 12:40pm

Mixed-model association mapping approaches in wheat

Benjamin Stich[1], Jens Möhring[2], Hans-Peter Piepho[2], Martin Heckenberger[1], Edward S. Buckler[4,5], and Albrecht E. Melchinger[1]

[1] Institute for Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany; stich@uni-hohenheim.de
[2] Institute for Crop Production and Grassland Research, University of Hohenheim, 70593 Stuttgart, Germany.
[3] Institute for Genomic Diversity, Cornell University, Ithaca, New York 14853, USA.
[4] Department of Plant Breeding and Genetics, Cornell University, Ithaca, New York 14853, USA.
[5] United States Department of Agriculture-Agricultural Research Service.

Abstract: Association mapping methods promise to overcome the limitations of linkage mapping methods. Based on the phenotypic and genotypic data of 303 soft winter wheat (Triticum aestivum L.) inbreds, (i) various association mapping methods were evaluated, (ii) a marker-based kinship matrix using a restricted maximum likelihood (REML) estimate of the probability of two alleles at the same locus being identical in state but not identical by descent was determined, and (iii) the results of association mapping approaches based on adjusted entry means (two-step approaches) with the results of approaches in which the phenotypic data analysis and the association analysis were performed in one step (one-step approaches) were compared. Spearman rank correlation between P values calculated based on one- and two-stage association mapping methods ranged from 0.63 and 0.93. The mixed-model association mapping approaches using a kinship matrix estimated by REML are more appropriate for association mapping than the recently proposed QK method with respect to (i) the adherence to the nominal alpha level and (ii) the adjusted power for detection of quantitative trait loci. Furthermore, we showed that our data set could be analyzed by using two-step approaches of the proposed association mapping method without substantially increasing the empirical type I error rate in comparison to the corresponding one-step approaches.

Session:
S04: Mapping Approaches in Plants and Animals : Monday, 10/03/2008, 2:10pm - 3:30pm

Setting cause-specific hazard ratios into relation: derivation of a mathematical function to be applied in decision-analytic screening models

Björn Stollenwerk and Uwe Siebert

Dept. of Public Health, Medical Decision Making and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i. T., Austria; bjoern.stollenwerk@umit.at

Abstract: Due to joint risk factors, positive screening results are often not only associated with mortality specific for the target disease of screening, but also with an increased mortality due to other causes. Most decision-analytic models do not consider this association leading to potentially biased results. Our goal was to develop a tool which helps to adjust for such joint mortality. We developed a function $HR_{other}(HR_{target})$ characterizing the relationship between two mortality hazard ratios (HRs). HR_{target} compares the mortality rate due to the target disease among those at high risk versus low risk. HR_{other} compares the respective mortality rates not directly related to the target disease. We postulated several properties that must be fulfilled by this function: (1) if the screening test result does not explain the mortality of the target disease (i. e. $HR_{target} = 1$) it also does not explain mortality due to other causes (i. e. $HR_{other}(1) = 1$); (2) the function $HR_{other}(HR_{target})$ is strictly increasing in HR_{target}; (3) the relationship of HRs below 1 is derived by taking the reciprocal values of the corresponding HRs above 1; (4) the function has an upper bound; (5) in the range of HRs above 1, HR_{target} exceeds HR_{other}; (6) the function $HR_{other}(HR_{target})$ is continuously differentiable. We created a function sufficing all postulated properties and applied it in the context of predictive coronary artery disease (CAD) screening. We fitted the function based of published HRs. Finally, we applied this function to hypothetic screening methods, for which only the HRs due to CAD were assumed to be known. The generated HR_{other} can be used by decision-analysts for incorporation into their CAD decision-analytic models. In conclusion we created a useful function that can be used for the adjustment of differential non-target-disease related mortality among risk groups in decision-analytic screening models. This should result in more valid modeling results.

Session:
S17: Event Data Analyses (II) : Tuesday, 11/03/2008, 11:00am - 12:40pm

A 46-Item Checklist for the Statistical Evaluation of Medical Research Manuscripts

Alexander M. Strasak[1], Karl-Peter Pfeiffer[1], Gerhard Marinell[2], Georg Goebel[1], and Hanno Ulmer[1]

[1] Department of Medical Statistics, Informatics and Health Economics, Innsbruck Medical University, Austria; alexander.strasak@i-med.ac.at, karl-peter.pfeiffer@i-med.ac.at, georg.goebel@i-med.ac.at, hanno.ulmer@i-med.ac.at

[2] Institute of Statistics, Faculty of Economics and Statistics, University of Innsbruck, Austria; gerhard.marinell@uibk.ac.at

Abstract: There is widespread evidence of the extensive use of statistical methods in modern medical research. Just the same, application standards are known to be generally low and a growing body of literature points to persistent statistical errors in most medical journals (Strasak et al. 2007a). The "misuse" of statistics in medical research has therefore been widely discussed, and it has been pointed out that it is both unethical and can have serious clinical consequences (Altman 1981, 2000; Strasak et al. 2007b). In order to further enhance statistical quality in medical research, we present a standardized, comprehensive 46-item checklist for the statistical evaluation of medical research manuscripts, developed on the basis of manifold literature related to the topic. The checklist may either be used just for the broad and in-depth statistical evaluation of submitted/published journal contributions (Strasak et al. 2007a) or as a useful guideline when planning, conducting and presenting medical research (Strasak et al. 2007b). Although our assessment tool clearly cannot cover aspects of all statistical methods incorporated in modern medical research, it is to our knowledge one of the most comprehensive lists presented yet. While predominately focusing on the issue of statistical significance testing, the 46-item checklist includes multifaceted statistical aspects of study design, statistical analysis, documentation of applied statistical methods, as well as presentation and interpretation of study findings. Consideration of issues included in our checklist, when planning, conducting and preparing medical research manuscripts, should help to further enhance statistical quality in medical journals. In addition, however, statisticians should be involved early in study design, as mistakes at this point can have major repercussions, negatively affecting all subsequent stages of medical research.

References:

ALTMAN D.G. (1981): Statistics and Ethics in Medical Research. Improving the Quality of Statistics in Medical Journals. *British Medical Journal, 282, 44-7.*

ALTMAN D.G. (2000): Statistics in Medical Journals: Some Recent Trends. *Statistics in Medicine, 19, 3275-89.*

STRASAK A., ZAMAN Q., MARINELL G., PFEIFFER K.P., ULMER H. (2007*a*): The Use of Statistics in Medical Research: A Comparison of The New

England Journal of Medicine and Nature Medicine. *American Statistician, 61, 47-55.*

STRASAK A., ZAMAN Q., PFEIFFER K.P., GOEBEL G., ULMER H. (2007*b*): Statistical Errors in Medical Research - A Review of common Pitfalls. *Swiss Medical Weakly, 137, 44-9.*

Session:

S28: Clinical studies (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

Ordered multiple comparisons with the best and their applications in dose-response studies

Klaus Strassburger[1], Frank Bretz[2], and Helmut Finner[1]

[1] German Diabetes Center, Leibniz Center at Heinrich-Heine-University Düsseldorf, Institute of Biometrics and Epidemiology, Düsseldorf, Germany;
strass@ddz.uni-duesseldorf.de, finner@ddz.uni-duesseldorf.de
[2] Biostatistics, Novartis Pharma AG, CH-4002 Basel, Switzerland;
frank.bretz@novartis.com

Abstract: In this contribution we consider the problem of comparing several treatments (dose levels, interventions, etc.) with the best, where the best treatment is unknown and the treatments are ordered in some sense. Order relations among treatments often occur quite naturally in practice. They may be ordered according to increasing risks, such as tolerability or safety problems with increasing dose levels in a dose–response study, for example. We tackle the problem of constructing a lower confidence bound for the smallest index of all treatments being at most marginally less effective than the (best) treatment having the largest effect. Such a bound ensures at confidence level $1 - \alpha$ that all treatments with lower indices are relevantly less effective than the best competitor. We derive a multiple testing strategy that results in sharp confidence bounds. The proposed lower confidence bound is compared with those derived from other testing strategies. We further derive closed-form expressions for power and sample size calculations. Finally, we investigate real data sets to illustrate various applications of our methods.

References:

STRASSBURGER, K., BRETZ, F. and FINNER, H. (2007): Ordered multiple comparisons with the best and their applications in dose-response studies. *Biometrics 63, 1143–1151 (doi:10.1111/j.1541-0420.2007.00813.x).* Supplementary Web Material: http://www.biometrics.tibs.org/datasets/060350.pdf.

Session:
S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

A simple semiparametric approach to the estimation of local and tail area-based false discovery rates

Korbinian Strimmer

Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Germany; strimmer@uni-leipzig.de

Abstract: The false discovery rate (FDR) criterion plays a prominent role in many high-dimensional testing and model selection procedures. Consequently, FDR methodologies are ubiquitous in the analysis of high-throughput data, such as in differential expression, SNP selection, peak detection in proteomic mass spectrometry data, or network edge detection.

In my talk I present a simple semiparametric approach for estimating FDR based on a variant of the Grenander density estimator. This approach not only allows for the simultaneous determination of both local FDR and tail area-based FDR but also can be applied to a diverse variety of test statistics. Furthermore, where possible, the empirical null is estimated (including any free parameters) by censored maximum likelihood where the truncation point is chosen with the false *non*discovery rate as guiding criterion.

This FDR estimation procedure is implemented in the R package "fdrtool" available from CRAN. In the talk I investigate the statistical performance of this approach and compare it against competing packages such as "qvalue" or "locfdr".

Session:
S01: Systems Biology and Bioinformatics : Monday, 10/03/2008, 9:10am - 10:30am

Testing Variable Importance in Random Forests – Strong Assumptions and Weak Hypotheses

Carolin Strobl[1] and Achim Zeileis[2]

[1] Department of Statistics, LMU Munich, Germany;
Carolin.Strobl@stat.uni-muenchen.de
[2] Department of Statistics and Mathematics, WU Wien, Austria;
Achim.Zeileis@wu-wien.ac.at

Abstract: Besides their high prediction accuracy, random forests (Breiman, 2001) have become increasingly popular in applied research due to the variable importance measures they provide: Especially in high dimensional problems, for example in genetics and the neurosciences, an advantage of the random forest permutation importance is that it reflects the impact of an explanatory variable in complex interactions. While sensitive applications, such as Lunetta et al. (2004), rely only on descriptive rankings of the variables, Breiman and Cutler (2007) suggest a statistical test for the variable importance scores, that has been widely applied in the meantime. In the talk we point out that (i) the assumptions and the construction of the suggested test are not justified and thus results can be highly misleading and (ii) even the null hypothesis to be tested is not at all clear. Different permutation schemes and the corresponding null hypotheses are explored and first steps towards a statistically more sound approach are outlined.

References:

BREIMAN, L. (2001): Random Forests. *Machine Learning, 45(1), 5–32.*

BREIMAN, L. and CUTLER, A. (2007): Random Forests - Classification Manual, *http://www.math.usu.edu/∼adele/forests.*

LUNETTA, K. L. and HAYWARD, L. B. and SEGAL, J. and VAN EERDEWEGH, P. (2004): Screening Large-Scale Association Study Data: Exploiting Interactions Using Random Forests, *BMC Genetics*, 5:32.

Session:

S21: Advances in Statistical Modelling (I) : Tuesday, 11/03/2008, 2:10pm - 3:30pm

Bayesian models for variables given on disparate scales: an epidemiological example

Sibylle Sturtz[1,2] and Katja Ickstadt[1]

[1] Fakultät Statistik, Technische Universität Dortmund, Germany;
sibylle.sturtz@iqwig.de, ickstadt@statistik.uni-dortmund.de
[2] Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany

Abstract: We analyse the relation of London's traffic-related benzene emissions given on a grid of 1km^2 squares and incidence of childhood leukaemia given on ward level. The usual modelling approach is to aggregate data and covariates to a common spatial scale leading to the problem of ecological fallacy.

Hierarchical Poisson/gamma models originally introduced by Wolpert and Ickstadt (1998) and generalised by Best et al. (2000) for an application in epidemiology allow data and covariates to be analysed on their original spatial scale. Additionally, risk factors can be modelled multiplicatively or additively leading to different interpretations. Based on the real leukaemia example from London, we perform a simulation study designed to analyze the behaviour of Poisson/gamma hierarchical models. Moreover, we compare the results to those of the CAR-model and the clustering approach of Knorr-Held and Rasser (2000).

Compared to other spatial models, the simulation study identifies Poisson/gamma random field models to be more flexible and easier to interpret for modelling different spatial patterns with and without latent risk sources. For the observed data we find a positive association between a multiplicative influence of benzene and leukaemia.

References:

BEST, N.G., ICKSTADT, K., and WOLPERT, R. (2000): Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association 95, 1076–1088.*

KNORR-HELD, L. and RASSER, G. (2000): Bayesian detection of clusters and discontinuities in disease maps, *Biometrics 56, 13–21.*

WOLPERT, R. and ICKSTADT, K. (1998): Poisson/gamma random field models for spatial statistics. *Biometrica 85, 251–267.*

Session:

S33: Bayesian models in biostatistics (I) : Wednesday, 12/03/2008, 11:00am - 12:40pm

The analysis of properties of the likelihood ratio test for testing association between a quantitative trait and SNP polymorphisms under different parametrisations using linear mixed models

Tomasz Suchocki[1] and Joanna Szyda[1,2]

[1] Wrocław University of Life Sciences, Institute of Animal Genetics, ul. Kozuchowska 7, 51-631 Wrocław, Poland; suchocki@gen.ar.wroc.pl
[2] Institute of Natural Sciences, pl. Grunwaldzki 24, 50-365 Wrocław, Poland; szyda@ar.wroc.pl

Abstract: Using a standard linear mixed model

$$Y = X\beta + Zu + \epsilon$$

applied to a dependent quantitative variable representing individual's phenotype, using the maximum likelihood, we estimate fixed effects comprising additive, dominance or epistatic (i.e. interaction) effects of SNP–type polymorphisms and a random effect representing an additive polygenic effect of individuals with assumption that the both variance components (additive polygenic and residual) are known. The main aim of this study is to assess type I error rates and power of the likelihood ratio test under different parametrisations of fixed effect's incidence matrix X as well as the analysis of the quality of parameter estimates. Additionally, the applicability of the mBIC (modification of standard BIC in situation where model contains epistatic e effects) to select the true model is considered.

References:

COCKERHAM, C. C. (1954): An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics, 39, 859–882.*

ZENG, Z-B., WANG, T., ZOU, W. (2004): Modeling Quantitative Trait Loci and Interpretation of Models. *Genetics, 169, 1711–1725.*

ALVAREZ-CASTRO, J.M. and CARLBORG, O. (2007): A Unified Model for Functional and Statistical Epistasis and Its Application in Quantitative Trait Loci Analysis. *Genetics, 176, 1151–1167.*

Session:

S14: Genetic Epidemiology (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Variable and Split Selection in Classification Trees based on the Gini Index – What if Values are Missing Not at Random?

Viola Svejdar, Thomas Augustin, and Carolin Strobl

Department of Statistics, LMU Munich, Germany; Viola.Svejdar@gmx.de, Thomas.Augustin@stat.uni-muenchen.de, Carolin.Strobl@stat.uni-muenchen.de

Abstract: The popular Gini Index split selection criterion in classification trees suffers from one fundamental problem: In the comparison of several continuous predictor variables the variable selection with the Gini Index is biased in favor of those variables with many missing values. This unwanted bias, that artificially privileges variables with many missing values, can be corrected by means of a new criterion. This new criterion combines the well known Gini Index with the statistical approach of maximally selected statistics and has been applied successfully when values are missing completely at random (MCAR) (Strobl et al., 2007). However, motivated by an application to veterinary data on uterine infection in dairy cows with values apparently missing not at random (MNAR), a new simulation study was conducted to address systematically the question whether the mechanism of missingness – missing completely at random (MCAR), missing at random (MAR) or missing not at random (MNAR) – has an effect on variable selection bias. Another aim was to give an overview which imputation techniques seem suitable to replace the missing values under each condition. The talk shows simulation results comparing the results for the Gini Index with those for the new criterion under the null hypothesis and for different degrees of information content of the predictor variables as well as an application to the veterinary data.

References:

STROBL, C., BOULESTEIX, A.-L., AUGUSTIN, T. (2007): Unbiased Split Selection for Classification Trees Based on the Gini Index. *Computational Statistics & Data Analysis, 52(1), 483–501.*

Session:

S29: Advances in Statistical Modelling (II) : Wednesday, 12/03/2008, 9:10am - 10:30am

How to survive in a non-normal world: are mutual information or adaptive methods sensible alternatives to parametric and nonparametric ANOVA?

Silke Szymczak, Andreas Ziegler, and Bernd-Wolfgang Igl

Institute of Medical Biometry and Statistics, University at Lübeck, Lübeck, Germany; silke.szymczak@imbs.uni-luebeck.de, ziegler@imbs.uni-luebeck.de, igl@imbs.uni-luebeck.de

Abstract: A challenging aspect in genetic association studies is to compare quantitative phenotypes observed in different genotype groups. Often, data distributions are highly skewed, heavy tailed or contaminated with outliers and hence, common statistical model assumptions like homoscedastic normal errors are not fulfilled. On these grounds, we focus on robust nonparametric procedures and adaptive methods as alternatives to common analysis of variance mechanisms. To this end, we analyze an entropy based "quantitative mutual information score" (QMIS) proposed by Tsalenko et al. (2006). In addition, we consider certain adaptive methods tracing back to Hogg (1974) and use also robust variants analogue to Keselman et al. (2007). Therein, in a first step we estimate the underlying phenotype distribution conditional on the genotype. Based on these results, we specify a statistical model to determine location shifts in a second step.

The main aim of our work is to compare QMIS and adaptive methods with parametric analysis of variance and its nonparametric counterpart, the Kruskal-Wallis test using simulated gene expression and genotype data. More precisely, we determine the number of significant p-values of these methods considering various types of distributions. We focus on different location, variance and skewness parameters of low and also heavy tailed distributions. In addition, we vary genotype frequencies.

References:

HOGG, R.V. (1974): Adaptive Robust Procedures: A Partial Review and Some Suggestions for Future Applications and Theory. *Journal of the American Statistical Association, 69, 909–923.*

KESELMAN, H.J., WILCOX, R.R., LIX, L.M., ALGINA, J. and FRADETTE, K. (2007): Adaptive Robust Estimation and Testing. *British Journal of Mathematical and Statistical Psychology, 60, 267–293.*

TSALENKO, A. et al. (2006): Analysis of SNP-Expression Association Matrices. *Journal of Bioinformatics and Computational Biology, 4, 259–274.*

Session:

S03: Genetic Epidemiology (I) : Monday, 10/03/2008, 9:10am - 10:30am

Cluster analysis: Data preparation? Which algorithm? How many clusters?

Matthias Templ[1,2], Peter Filzmoser[3], and Clemens Reimann[4]

[1] Department of Statistics and Probability Theory, Vienna University of Technology, 1040 Vienna, Austria; templ@tuwien.ac.at
[2] Statistics Austria, 1110 Vienna, Austria, matthias.templ@statistik.gv.at
[3] Department of Statistics and Probability Theory, Vienna University of Technology, 1040 Vienna, Austria; p.filzmoser@tuwien.ac.at
[4] Geological Survey of Norway, N-7491 Trondheim, Norway; clemens.reimann@ngu.no

Abstract: Cluster analysis is a method for finding groups in multivariate data without providing any information about group membership (unsupervised classification). Many different clustering algorithms have been proposed in the literature, and many methods are implemented in R. Unfortunately, for real data sets without obvious grouping structure, different cluster algorithms will in general give slightly different results, sometimes even completely different results. For the user it would thus be important to know which clustering methods are ideally suitable for analyzing the data at hand. The results of cluster analysis applied on complex data depends strongly on different parameters and assumptions. Firstly, one have to consider an appropriate transformation, especially when the data are of compositional nature, for example. For most of real data it is also useful to standardise it in respect to the different scale of variables. Furthermore, the user has to consider an appropriate distance measure, since most of the existing clustering algorithms will use the distance matrix to start. In addition to that, the selected number of clusters plays an important role in finding good clusters. When using bootstrap samples of the data, one can also consider the stability of results arising from different clustering algorithms and different distance measures with the RAND index. Most of the additionally available cluster validity measures are quite useful regarding to (artificial) data with a strong clustering structure but they provide more or less poor results when applied on real complex data. We want to give deeper insights into these problems of cluster analysis by using real complex data sets from geochemistry. A package (clustTool) for clustering data is developed and downloadable on CRAN. Especially for the clustering of geochemical data, there is a graphical user interface implemented in package clustTool. Since cluster analysis and diagnostic tools for cluster results on complex geochemical data were in hand of specialists we want to provide methods and some diagnostic tools to the users with this GUI.

References:

TEMPL, M. and FILZMOSER, P. and REIMANN, C. (2006): Cluster analysis applied to regional geochemical data: problems and possibilities. *Research Re-*

ports of Group Computational Statistics, Vienna University of Technologie, Research Report CS-2006-5.

LEISCH, F. (2006): A Toolbox for K-Centroids Cluster Analysis. *Journal of Computational Statistics and Data Analysis, 51 (2), 526–544.*

Session:

S36: Young Statistician Papers (IBS-ROeS) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Meta-analyses of observational epidemiological studies based on individual participant data: combining within and between study information

Simon G. Thompson

MRC Biostatistics Unit, Cambridge, UK

Abstract: Collations of individual participant data (IPD) from epidemiological studies are increasingly common. Such extensive data can yield precise estimates of risk associations, help to resolve controversy, and enable exploration of heterogeneity. Here I discuss the statistical issues arising in undertaking meta-analyses of the relationship between fibrinogen and coronary heart disease risk, based on 150,000 people in 31 prospective studies. Heterogeneity across studies – in the distribution of fibrinogen, in the risk relationship, and in confounder effects – favoured a two-stage analysis, based on a sex-stratified proportional hazards model. The assumption of proportional hazards was assessed, and different methods of adjusting for age and other confounders were explored.

One principal benefit of IPD meta-analysis is the ability to explore interactions; for example, does the fibrinogen-risk relationship differ between men and women, or depend on the level of other risk factors? The nature of the relevant variable determines whether the information on the interaction derives from between-study information, within-study information, or a combination of both. In each case, a different statistical analysis is required.

Session:
S42: Meta-analysis and Meta-regression : Thursday, 13/03/2008, 11:00am - 12:40pm

A Hierarchical Poisson Model applied to Solexa Gene Expression Data

Helene H. Thygesen[1], Peter-Bram 't Hoen[2], and Johan den Dunnen[2]

[1] Dpt. of Appl. Stat., Lancaster University, UK; h.thygesen@lancaster.ac.uk
[2] Dpt. of Human Genetics, Leiden University Medical Center, The Netherlands

Abstract: The Solexa platform (Illumina, Inc.) is a promising new technology for sequencing of genomic DNA and cDNA libraries. The data produced are sequencing counts, which means that a natural model for the data is a hierarchical Poisson model, in which the measure of expression of tag (gene) g in library (tissue sample) i is Poisson(λ_{gi}) distributed, and the inference is with regard to the across-genes and across-tissue-samples distribution of the latent variable λ. The Bayesian model for SAGE developed by Vencio et al. works well for Solexa data but since it analyses each tag in isolation, there is scope for improvement by exploiting the across-tag distribution of the parameters.

We propose a multivariate lognormal model for λ_{gi} in which the per-tag between-library covariance matrix Σ_g takes the form $\sigma_g^2\Sigma$ where σ_g^2 follows an inverse-gamma distribution across tags. We applied the model to hippocampus cDNA libraries from 4 transgenic mice to 4 wildtype mice. We demonstrate that the model provides a good fit to the data and that the introduction of the across-tag distribution of σ_g^2 improves the accuracy of the effect size estimates.

References:

THYGESEN, H.H. and ZWINDERMAN, A.H. (2006): Modeling Sage data with a truncated gamma-Poisson model. *BMC Bioinformatics 2006, 7:157.*

VENCIO, R.Z.N. and BRENTANI, A. and PATRAO, A.F.C. and PEREIA, C.A.B. (2004): Bayesian model accounting for within-class biological variability in serial analysis of gene expressions (sage). *BMC Bioinformatics 2004, 5:119.*

Session:

S32: Hierarchical models : Wednesday, 12/03/2008, 11:00am - 12:40pm

Lineare Modelle und Interpolation auf unregelmäßigen Dreiecksnetzen für eine effektive Nutzung der geographischen Datenbasis bei ökologischen Aufgaben

Jürgen Tiedge

Hochschule Magdeburg-Stendal, FB Wasser- und Kreislaufwirtschaft, Postfach 3680, 39011 Magdeburg, Germany; juergen.tiedge@hs-magdeburg.de

Abstract: Biometrische und chemometrische Daten im Zusammenhang mit ökologischen Aufgabenstellungen haben in der Regel räumlichen Bezug. Mit digitalen Geländemodellen (DGM) wird das Ziel verfolgt, an jedem Punkt (x, y) die Höhe $z := z(x, y)$ und gegebenenfalls den Gradienten (zumindest als Schätzung) verfügbar zu machen. Datenbasis hierfür war zunächst das DGM10, das die z-Werte in einem $10m \times 10m$-Gitter bereitstellt. Mit der rasanten Entwicklung z.B. des Laser-Scannens stehen Rohdaten als DGM1 oder in noch feineren Gittern zur Verfügung. Diese umfangreichen Rohdaten sind mit zufälligen Fehlern behaftet.

Die statistische Bearbeitung ökologischer Aufgabenstellungen (aber auch schon die Visualisierung) verlangt nach einem effektiven Umgang mit den anfallenden geographischen Daten.

Eine Möglichkeit dazu bietet die sinnvolle Kopplung klassischer Ergebnisse zu unregelmäßigen Dreiecksnetzen, Interpolationsalgorithmen und Regression.

Ausgangspunkt bildet eine bezüglich der Datendichte relativ grobe Thiessen-Vermaschung. Zur Interpolation werden Algorihmen herangezogen, die einerseits eine stückweise lineare Interpolationsfunktion mit C_0- Eigenschaft, andererseits eine spezielle stückweise kubische Interpolationsfunktion mit C_1-Eigenschaft benutzen. In beiden Fällen liegt Linearität in den Parametern (z-Werte bzw. z-Werte und Gradienten in den Knoten der Vermaschung) vor. Die Schätzung der Parameter erfolgt mit Regressionsmethoden, die ausgehend von den klassischen Voraussetzungen auf einige besondere Anforderungen aus der Geländemodellierung zugeschnitten werden.

Auch auf die Implementierung innerhalb eines Experimental-GIS und erste Erfahrungen wird eingegangen.

Session:
S06: Biostatistic education at universities of applied sciences : Monday, 10/03/2008, 2:10pm - 3:30pm

Multiple Comparisons in Model–Based Testing with Examples from Larynx Cancer Research

Zdeněk Valenta[1], Aleš Slavíček[2], Libor Černý[3], Radka Lohynská[4], and Abdulrahman Bahannan[2]

[1] Dept. of Medical Informatics, Institute of Computer Science AS CR, v.v.i., Prague, Czech Republic; valenta@euromise.cz
[2] Dept. of Otorhinolaryngology, Head and Neck Surgery, University Hospital Motol, Charles University, Prague, Czech Republic; aslav@seznam.cz
[3] Dept. of Phoniatrics, General University Hospital, Prague, CR
[4] Dept. of Radiotherapy & Oncology, University Hospital Motol, Prague, CR

Abstract: Multiple testing problems arise in many areas of clinical research. Typically they appear in the context of One-way Analysis of Variance upon evaluating pairwise treatment differences between individual groups. Often more complex models, such as Analysis of Covariance, are required to adjust for the differences characterising the subjects at their entry point. The model–based multiple testing framework was described by Searle in 1971, which allows for simultaneous testing of several model–based contrasts, while keeping the experiment–wise Type I error rate at the pre–selected level (often set to $\alpha = 0.05$). In situations where treatment may interact with other covariates, one particularly needs to adjust for multiple testing when assessing significance of multiple linear contrasts simultaneously. We discuss the multiple testing framework for model–based inference in the context of larynx cancer research on selected examples. Analyses were performed using statistical system R.
This work was partly supported by the Institutional Research Plan *Computer Science for the Information Society: Models, Algorithms, Applications, AV0Z10300504*, and by the grant No. IGA 8430-3 of the Czech Ministry of Health.

References:

SEARLE, S.R. (1971): *Linear Models*. John Wiley & Sons, New York.

DUNNETT, C.W. (1955): A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association, 50(272), 1096–1121.*

R DEVELOPMENT CORE TEAM (2007): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Session:

S18: Clinical studies (I) : Tuesday, 11/03/2008, 11:00am - 12:40pm

Curves of Constant Power for "One-Sided" Bivariate Location Tests

Michael Vock

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland; michael.vock@stat.unibe.ch

Abstract: There have been several proposals of tests for "one-sided" (or restricted) multivariate location alternatives. Therefore, it is an interesting task to determine whether a test is suitable for a specific problem or not and to compare different tests. This task may be much more demanding than for univariate problems, particularly if a composite null hypothesis is necessary, which is the case in many applications.

In the bivariate location problem, the adequacy of a test for different types of one-sided null and alternative hypotheses can be graphically assessed by the consideration of curves of constant power. Those are the curves consisting of all points in the parameter space that lead to a specified power value of the test. The curves of constant power depend on the shape of the distribution of the data, and therefore, this distribution has to be specified up to the location parameter.

The most interesting choice of the power value is usually the nominal significance level α of the test examined; the curve of power α indicates the region in the parameter space where the nominal significance level is respected. The curves of power α of several tests can be easily compared. Further, the curves of power α of the same test but based on different distribution families can be used to assess the robustness of a test with respect to distributional assumptions.

The approach based on curves of constant power has two major advantages over other graphical representation techniques proposed earlier: The tests are visualized directly in the parameter space, which facilitates the interpretation, and the method is suitable for the comparison of tests based on entirely different statistics.

References:

VOCK, M. (2006): Graphical comparison of multivariate nonparametric location tests for restricted alternatives. *Statistics & Probability Letters, 76, 1529–1535.*

Session:
S36: Young Statistician Papers (IBS-ROeS) : Wednesday, 12/03/2008, 2:10pm - 3:30pm

Estimation of a four state model when not all patients are in one state initially

Reinhard Vonthein

Institut für Medizinische Biometrie, Universitätsklinikum Tübingen, 72070 Tübingen, Germany; reinhard.vonthein@uni-tuebingen.de

Abstract: Patients in intensive care often need thiopental against their high intracranial pressure. Whether that barbiturate then facilitates the development of pneumonia was the subject of a prospective observational study. The rate of incident pneumonia was estimated using different techniques: proportional hazards regression with a time-varying covariable, maximum-likelihood estimation of a four state model and application of the R-package changeLOS to the cases starting in the first state. The four states were: ventilated, ventilated and receiving thiopental, pneumonia, and discharge or death for other reasons than pneumonia. The first two were transient and the latter two absorbing. Some patients were on thiopental when ventilation commenced. Transition rates into discharge were assumed to be zero on the first day and into pneumonia on the first two days for the maximum-likelihood estimation of the four state model. The three approaches yielded results that were very similar for all practical purposes, namely that pneumonia developed in patients receiving thiopental with two times the rate observed in other patients. Even though this may be confounded with the initial ailment and with high intracranial pressure, physicians will now either look for other means against high intracrational pressure or take better precautions against pneumonia in these patients.

References:

WANGLER, M. and BEYERSMANN, J. (2007): *changeLOS: Change in LOS*. Version 2.0.6
http://cran.r-project.org/src/contrib/Descriptions/changeLOS.html

Session:

S12: Freie Themen (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Auxiliary Mixture Sampling for Dynamic Survival Models

Helga Wagner

Johannes Kepler Universität Linz, Austria; helga.wagner@jku.at

Abstract: Dynamic survival models, introduced by Gamerman (1991) are a useful extension of the popular Cox model as the effects of explanatory variables are allowed to change over time. The baseline log-hazard as well as covariate effects are assumed to be piecewise constant with a correlated prior process. Models of this type can be estimated in a Bayesian framework by Markov chain Monte Carlo methods, requiring however a Metropolis-Hastings step. In this presentation a new auxiliary mixture sampler is proposed, which requires only draws from simple distributions and need no tuning. This sampler is an extension of the auxiliary mixture sampling scheme for count data developed in Frühwirth-Schnatter and Wagner (2006a, 2006b). By introducing two sequences of latent data a representation of the dynamic survival model as a conditionally Gaussian model is achieved, where the multidimensional latent state vector consisting of baseline log-hazard and time varying covariate effects can be sampled in one move. The proposed auxiliary mixture sampler can deal with more general models, e.g. models including frailty or spatial effects and for data observed under censoring schemes different from right censoring. Application of the sampler will be illustrated on well-known data sets from the literature.

References:

FRÜHWIRTH-SCHNATTER, S. and WAGNER, H. (2006a): Auxiliary mixture sampling for parameter-driven models of time series of small counts with applications to state space modelling. *Biometrika*.

FRÜHWIRTH-SCHNATTER, S. and WAGNER, H. (2006b): Data augmentation and Gibbs sampling for regression models of time series of small counts. *Student*, 5:221–234.

GAMERMAN, D. (1991): Dynamic Bayesian models for survival data. *Applied Statistics*, 40:63–79.

Session:

S33: Bayesian models in biostatistics (I) : Wednesday, 12/03/2008, 11:00am - 12:40pm

When are confirmatory adaptive designs appropriate?

Sue-Jane Wang

Office of Biostatistics Office of Translational Sciences Center for Drug Evaluation and Research, U.S. Food and Drug Administration; suejane.wang@fda.hhs.gov

Abstract: In this presentation, I will lay out the principles of exploratory adaptive design trial versus that of confirmatory adaptive design trial. Using patient subpopulation adaptation as an example, some practical strategies in alternative design considerations and analytical approaches that allow wide flexibility in learning versus limited learning for confirming will be discussed and illustrated. The distinct utility in the application to patient subpopulation selection versus that of patient subpopulation enrichment will be highlighted using typical case examples.

Session:
S23: Adaptive seamless design for combining phase II / III clinical studies : Tuesday, 11/03/2008, 4:00pm - 6:10pm

On Multiarmed Adaptive Designs

Gernot Wassmer

IMSIE University of Cologne, Germany; gernot.wassmer@uni-koeln.de

Abstract: In multiarmed clinical trials, treatment selection is a straightforward application when using an adaptive test design. Testing and estimation procedures for designs with adaptive treatment selection when considering normally distributed data were recently proposed (e.g., Posch et al., 2005). When considering more than two treatment arms, Dunnett's test is often applied, but other testing procedures can also be used, e.g., Bonferroni tests, Simes tests. The procedure is generally based on the closed testing procedure and the inverse normal method for combining the stage results of the trial. This procedure can also be used for survival data where the test utilizes the independent increment property of the log-rank test statistic. In the talk, the procedures and designing options relevant for assessing the statistical performance of such designs are given.

References:

POSCH, M., KÖNIG, F., BRANSON, M., BRANNATH, W., DUNGER-BALDAUF, C. and BAUER, P. (1999): Testing and Estimation in Flexible Group Sequential Designs with Adaptive Treatment Selection. *Statistics in Medicine, 24, 3697–3714.*

Session:
S15: Flexible designs : Tuesday, 11/03/2008, 11:00am - 12:40pm

Meta-Regression of Response Ratios

Jürgen Wellmann

Institute of Epidemiology and Social Medicine, University of Münster, Germany;
wellmann@uni-muenster.de

Abstract: In meta-analyses of continuous data sometimes the "response ratio" (the ratio of mean outcome in the experimental group to that in the control group) might be a sensible summary statistic (Hedges et al., 1999). Since it is a dimensionless quantity, it is especially attractive when the individual studies measure the outcome of interest on different scales. The variability of the response ratio can readily be quantified from published standard errors by means of Fieller's theorem or the delta-method. The approach carries forward to meta-regression of log-transformed mean outcomes.

The meta-analysis of response ratios is illustrated using experimental data on the efficacy of granulocyte-colony stimulating factor in animal models for stroke (Minnerup et al.). The outcome of interest is the volume of affected area of the brain and was obtained in rats as well as in mice.

References:

HEDGES, L. V., GUREVITCH, J. and CURTIS, P. S. (1999): The meta-analysis of response ratios in experimental ecology. *Ecology, 80(4), 1150–1156.*

MINNERUP, J., HEIDRICH, J., WELLMANN, J., ROGALEWSKI, A., SCHNEIDER, A. and SCHÄBITZ, W.-R. Meta-analysis of the efficacy of granulocyte-colony stimulating factor in animal models of focal cerebral ischemia. Accepted for publication in *Stroke*.

Session:

S42: Meta-analysis and Meta-regression : Thursday, 13/03/2008, 11:00am - 12:40pm

Univariate and bivariate quadratic hazard models

Andreas Wienke

Institute for Medical Epidemiology, Biostatistics and Informatics, University Halle, Germany; andreas.wienke@medizin.uni-halle.de

Abstract: To combine the flexibility of the normal distribution with advantages of frailty models which provide an analytic survival function Yashin and Iachine (1996) suggested the quadratic hazard frailty model. Let frailty $Z = W^2$ with $W \sim N(\mu, \sigma^2)$ and let the conditional hazard in the univariate survival model be

$$\lambda(t, Z) = Z\lambda_0(t) = W^2\lambda_0(t),$$

where $\lambda_0(t)$ denotes the conditional baseline hazard and covariates are dropped from the model for ease of presentation. The unconditional survival function is

$$S(t) = \frac{1}{\sqrt{1 + 2\sigma^2\Lambda_0(t)}} e^{-\frac{\mu^2\Lambda_0(t)}{1+2\sigma^2\Lambda_0(t)}}$$

with $\Lambda_0(t) = \int_0^t \lambda_0(s)\,ds$. The identifiability constraint $\mathbf{E}Z = \mathbf{E}W^2 = 1$ in frailty models implies the relation $\sigma^2 + \mu^2 = 1$, which restricts possible values of μ and σ^2. One problem to circumvent this problem would be another parameter constraint, for example $\mu = 0$. In this simplified case a correlated bivariate frailty model can be derived, characterized by the marginal bivariate survival function

$$S(t_1, t_2) = (S_1^{-2}(t_1)S_2^{-2}(t_2) - \rho^2(S_1^{-2}(t_1) - 1)(S_2^{-2}(t_2) - 1))^{-1/2}.$$

Applications to real data illustrate advantages and limitations of the model.

References:

YASHIN, A.I. and IACHINE, I.A. (1996): Random effect models of bivariate survival: quadratic hazard as a new alternative. In: G. Kristensen (Ed.): *Transactions of Symposium i Anvendt Statistik, 22-24 Januar 1996.* Economic Institute, Odense University, Odense, 87–101.

Session:

S13: Event Data Analysis (I) : Tuesday, 11/03/2008, 9:10am - 10:30am

Introduction of Highly Mobile Pathogens into Germany and their Early Detection - the National Surveillance Program for Influenzaviruses in Wild Birds

Hendrik Wilking[1], A. Globig[2], F. Unger[1], S. Kowalcyk[1], S. Richter[1], M. Ziller[1], T. Harder[2], F.J. Conraths[1], and C. Staubach[1]

[1] Friedrich-Loeffler-Institut, Institute of Epidemiology, Wusterhausen, Germany; hendrik.wilking@fli.bund.de
[2] Friedrich-Loeffler-Institut, National Reference Laboratory for Avian Influenza, Insel Riems, Germany

Abstract: Domestic birds and poultry are intensively monitored for different infectious diseases. The monitoring of wild birds, and especially migratory birds, for pathogens is much more difficult. Highly pathogenic avian influenza virus of the subtype H5N1 Asia has become endemic in poultry and wild birds in some countries in Southeast Asia. A similar situation may also become possible in Europe. Migratory birds are suspected to be a main carrier of this virus and proved their ability to cause outbreaks associated with wetlands. Nationwide surveillance data on the Influenza-A status of wild birds are collected in a web based database along with geographical information on municipality level, the taxonomic identification of 479 bird species and groups of birds. The sampling of sick and dead animals as well as living and hunted birds allows for both a passive and an active part of the monitoring. The system describes the present status regarding avian influenza in wild birds and will allow the early detection of a new intensive introduction of highly pathogenic influenza viruses into Germany. When cases of avian influenza in wild birds are detected, the assessment of the true number of affected, positive animals is often difficult if not impossible. This results in limitations in calculating prevalence estimates. In addition, surveillance data in wild animals imply problems of autocorrelation in time and space, which affects the calculation of confidence limits. When the large number of different bird species is taken into account, it becomes obvious that a profound analysis of these data is a challenge for biometricians and veterinary epidemiologists. For the evaluation of the surveillance system, we propose a model for estimating confidence intervals for (i) prevalence calculations in outbreak situations or (ii) the absence of disease in certain time intervals for certain regional units in Germany. This surveillance database provides information on diagnostic results over a long time period and a broad geographical range for a highly mobile pathogen in large populations of wild migratory birds. This study may prove useful for the detection of time intervals, geographical units and population subgroups where monitoring is insufficient to detect a pathogen with low prevalence. In the future, it may be possible to develop a risk-based approach and to target specific animals with a better chance of detecting

the virus. Altogether this supports the risk analysis for introduction of Influenza-A viruses into poultry and the food-chain in Germany.

Session:
S24: Spatial Analysis of Surveillance Data : Tuesday, 11/03/2008, 4:00pm - 6:10pm

Diplotype μ-scores for Multipoint Screening

Knut M. Wittkowski[1] and Tingting Song

The Rockefeller University, Center for Clinical and Translational Science, 1230 York Ave Box 322, New York, NY 10065, USA; kmw@rockefeller.edu

Abstract: As the density of SNPs available for analysis increases, more than two adjacent SNPs can be in linkage disequilibrium with an unknown disease locus. Moreover, having several mutations in a single gene, but between different SNPs may increase disease severity. As the degree of linkage disequilibrium between disease loci and SNPs is unknown, the order between the SNP profiles and the risk they convey is only partial. U-scores for multivariate data are well suited for partial orderings, in general. When scoring diplotypes, knowledge about the SNPs' sequence on the chromosome provides additional information. Utilizing this information by computing μ-scores in a hierarchical fashion, starting with computing the partial ordering of each SNP interval, then combining these partial orderings, and, finally, computing the μ-scores based on this combination of SNP interval orderings, increases information content. Genetic factors in several genes are typically related to the common complex diseases. Adding another level in this hierarchy extends the method to screening for epistasis. Packages for R and S-Plus, as well as a Web server for distributed processing of whole genome association studies are available from http://muStat.rockefeller.edu.

References:

WITTKOWSKI, K. M., LEE, E., et al. (2004): Combining several ordinal measures in clinical studies. *Statistics in Medicine, 23, 1579–1592.*

WITTKOWSKI, K., HAIDER, A., et al. (2006): Bioinformatics Tools Enabling U-Statistics for Microarrays. *Conf Proc IEEE Eng Med Biol Soc, 1, 3464–3469.*

Session:

S14: Genetic Epidemiology (II) : Tuesday, 11/03/2008, 9:10am - 10:30am

Comparison of different methods for adjusting the area under the curve (AUC) for covariates

Antonia Zapf and Edgar Brunner

Department of Medical Statistics, University Göttingen, 37073 Göttingen, Germany; azapf@uni-goettingen.de

Abstract: In clinical studies, comparing diagnostic or therapy methods, the area under the receiver operating characteristic curve (AUC) is a widely accepted measure to assess the overall accuracy of a diagnostic test or the measure of separation of a therapy method.

Baseline values, however, can have a strong effect on a diagnostic result or the success of a therapy. In this case adjusting the statistics for baseline observations is necessary to account for a potential bias. The method of simply subtracting the baseline value has often been criticized in literature (Senn, p. 95–111) and it has been suggested to use analysis of covariance methods to adjust for baseline observations. Moreover, this simple method can only be used for metric data and is not applicable to adjust for other covariates.

It is the aim of this talk to generalize the covariance procedures in such a way, that the response variable as well as covariates may be either metric or ordinal data.

The different methods of adjusting for covariates will be compared in a simulation study. The application of the proposed nonparametric method will be demonstrated by a real data example from a clinical trial.

References:

BRUNNER, E. et al. (2002): The multivariate nonparametric Behrens-Fisher problem. *Journal of Statistical Planning and Inference, 108, 37–53.*

KAUFMANN, J. et al. (2005): Nonparametric methods for analyzing the accuracy of diagnostic tests with multiple readers. *Statistical Methods in Medical Research, 14, 129–146.*

SENN, S. (1997): *Statistical Issues in Drug Development.* John Wiley & Sons, Chichester.

Session:

S02: Diagnostic studies : Monday, 10/03/2008, 9:10am - 10:30am

Kidney exchange programs: a simulation study

Katarína Cechlárová[1], Vladimír Lacko[1], Ivan Žežula[1], and Ľuboslav Beňa[2]

[1] Faculty of Science, Šafárik University, Košice; katarina.cechlarova@upjs.sk, vladimir.lacko@upjs.sk, ivan.zezula@upjs.sk
[2] Faculty of Medicine, Šafárik University, Košice; luboslav.bena@upjs.sk

Abstract: To overcome the shortage of cadaveric kidneys for transplantation and not to lose willing living donors, who are incompatible with their intended recipients, several countries, transplantcenters and transplant organizations started various programs for kidney exchanges. Such programs require cooperation among several entities, which all have partially different objectives. Moreover, because many algorithms could be used, a careful evaluation of their efficiency is necessary. We will present three different approaches that incorporate individual as well as overall optimality criteria (Edmonds', Irving's and TTC algorithm) and compare their strengths and weaknesses. We simulated kidney exchanges between randomly generated donor-recipient pairs in attempt to find dual, triple, and longer cycles of exchanges. The results obtained from all algorithms were evaluated and compared using standard statistical methods.

Session:
S10: Freie Themen (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Online

Biostatistical Aspects of Quality Control in Genome-Wide Association Studies Using the Affymetrix Chip Technology

Andreas Ziegler, Daniel F. Schwarz, John R. Thompson, and Inke R. König

Universität zu Lübeck, Med. Biometrie u. Statistik, 23538 Lübeck, Germany;
ziegler@imbs.uni-luebeck.de

Abstract: To search the entire human genome for association is a novel and promising approach to unravelling the genetic basis of complex genetic diseases. In these genome-wide association studies (GWAs), several hundreds of thousands of single nucleotide polymorphisms (SNPs) are analyzed at the same time, posing substantial biostatistical and computational challenges. The validity and completeness of the genotyping procedures is a relevant criterion for the success of a GWA. In this talk, we discuss a number of biostatistical aspects in the quality control of GWAs in detail, and we restrict our attention to the Affymetrix platform. We show that specific quality control procedures heavily depend on the algorithm used for genotypes calling from signal intensities. We specifically show that signal intensity plots are a sine qua non condition in today's GWAs, and we discuss the possibility of automated reading of signal intensity plots. Data from several GWAs performed on the Affymetrix 500K SNP array as well as the Affymetrix Genome-Wide Human SNP Arrays 5.0 and 6.0 will be used for illustrating the different quality control approaches.

References:

SAMANI et al. (2007): *N Engl J Med, 357, 443-53*
ZIEGLER et al. (2008): *Biom J, invited*

Session:
S07: Genome-wide association studies (I) : Monday, 10/03/2008, 4:00pm - 6:10pm

Resampling-based Procedure to Control for the Generalized Familywise Error Rate

Astrid Zierer and Iris Pigeot

Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Germany; zierer@bips.uni-bremen.de, pigeot@bips.uni-bremen.de

Abstract: There are many applications where a large number of variables is analyzed simultaneously, especially in the field of genetics. For example in association studies or gene expression analysis the number of genes or gene-gene interactions considered is usually high, up to several thousands. Each decision, for instance about correlation between phenotype and a specific combination of SNPs, is based on an appropriate statistical test. Consequently a large multiplicity problem occurs. To adjust for multiplicity the Familywise Error Rate (FWER) seems too restrictive and alternative methods have been proposed to control for the amount of rejected true null hypotheses.

Our focus is on procedures controlling the generalized FWER (gFWER). The gFWER has the appealing property to control for a tolerable number of false positives k which can be specified by the user. We present a single-step procedure based on a bootstrap approach. The aim is to combine a good practicability with respect to the computer capacity while at the same time taking into account the dependency structure of the variables, being present in many genetic applications. We compare our proposal with other methods, regarding error rate and power, like the generalization of the Bonferroni and Bonferroni-Holm Procedure (Hommel and Hoffmann, 1987) and the Augmentation Procedure (van der Laan et al., 2004).

References:

HOMMEL, G. and HOFFMANN, T. (1987): Controlled Uncertainty. In: P. Bauer, G. Hommel and E. Sonnemann (Eds.): *Multiple Hypotheses Testing*. Springer, Berlin, 154–161.

VAN DER LAAN, M.; DUDOIT, S. and POLLARD, K. (2004): Augmentation Procedures for Control of the Generalized Familiy-Wise Error Rate and Tail Probabilities for the Proportion of False Positives. *Stat Appl Genet Mol Biol, 3, Article 15.*

Session:

S08: Multiple Testing : Monday, 10/03/2008, 4:00pm - 6:10pm

MCMC methods for gene expression profiling via Bayesian variable selection

Manuela Zucknick[1,2] and Sylvia Richardson[2]

[1] DKFZ, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany;
m.zucknick@dkfz.de
[2] Biostatistics Centre, Imperial College, Norfolk Place, London W2 1PG, UK

Abstract: Gene expression microarrays and other high-throughput technologies produce measurements for several thousand genes but the sample size is typically much smaller than that. A common application of microarray data is the construction of gene expression profiles for class prediction based on expression measurements of a small number of selected genes.

Bayesian variable selection methods are well suited to the problem, since priors can be imposed so that full modelling is possible even if the number of variables is much larger than the sample size. In addition, the uncertainty related to the role of each candidate gene can be assessed through posterior probabilities. However, the model space is very large and standard MCMC algorithms are computationally unfeasible. Several strategies can be used and ultimately combined to improve feasibility. Here, we focus on the 'block update' component of an MCMC strategy. We propose to employ the dependence structure in the data to decide which variables should always be updated together and which are nearly conditionally independent and do not need to be considered together.

For binary classification, logistic regression is traditionally preferred in medical and biological applications because of the good interpretability. We follow the implementation of the Bayesian logistic regression model by Holmes and Held (2006).

We investigate several MCMC samplers using the dependence structure in different ways. The mixing and convergence performances of the resulting Markov chains are evaluated and compared to standard samplers using simulated data and in an application to a gene expression data set related to ovarian cancer. In the latter, we also explore the additional benefit of combining the block update with a parallel tempering strategy.

References:

HOLMES, C.C. and HELD, L. (2006): Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis, 1, 55–67.*

Session:

S33: Bayesian models in biostatistics (I) : Wednesday, 12/03/2008, 11:00am - 12:40pm

Education in Biostatistics with the Support of e-learning Tools

Jana Zvárová, Patrícia Martinková, and Karel Zvára

EuroMISE Centre of Charles University and Academy of Sciences CR, Institute of Computer Science AS CR, 18207 Prague, Czech Republic; zvarova@euromise.cz

Abstract: The long-term effect of education in the field of biostatistics on efficiency and quality of biomedical research is discussed. Selected educational methods using e-learning tools are presented and their applications in graduate, doctoral and lifelong education are shown. Since 1994 the EuroMISE Center has been developing educational materials and e-learning tools for teaching biostatistics. The edition Biomedical Statistics, covering till now three books in the Czech language, has been published by the Printing House of Charles University. Interactive electronic versions of the books are serving to students and teachers in the EuroMISE courses and pdf versions are available for registered users on the web pages http://www.euromise.cz/. Since 1998 the ExaMe system for evaluation of a targeted knowledge has been developed. New features of the ExaMe evaluation system and statistical issues of evaluation will be described. Knowledge bases of the ExaMe (sets of questions/items) so far cover courses in biostatistics, genetics, telemedicine and biomedical informatics. The properties of items in ExaME system are investigated by means of item analysis. The reliability of fixed tests is assessed by means of Cronbach's alpha and by means of its counterpart logistic alpha introduced in 2007.

Acknowledgment: The work was partially supported by projects 1ET200300413 AS CR and by AV0Z10300504 ICS AS CR

References:

ZVAROVA J., ENGELBRECHT R., VAN BEMMEL JH. (1997): Education in Medical Informatics, Statistics and Epidemiology. *International Journal of Medical Informatics 45, 1/2, 1997, 3–8.*

ZVAROVA J., ZVARA K. (2000): Evaluation of Knowledge Using ExaMe Program on the Internet. In: Iakovidis I, Maglaveras N, Trakatellis A (Eds.). *User Acceptance of Health Telematics Applications*, Amsterdam, IOS Press; 145–151.

MARTINKOVA P., ZVARA K. jr., ZVAROVA J., ZVARA K.(2006): The New Features of the ExaMe Evaluation System and Reliability of its Fixed Tests. *Methods of Information in Medicine 45, 3, 310–315.*

MARTINKOVA P., ZVARA K. (2007): Reliability in the Rasch Model. *Kybernetika 43, 3, 315–326.*

Session:

S06: Biostatistic education at universities of applied sciences : Monday, 10/03/2008, 2:10pm - 3:30pm

Online

Part II

Abstracts for posters

Online

Generalized linear mixed models for counting processes

Christine Adrion[1], Simon Rückinger[2], and Ulrich Mansmann[1]

[1] Institute of Medical Informatics, Biometry and Epidemiology (IBE); adrion@ibe.med.uni-muenchen.de
[2] Division of Epidemiology, Institute of Social Pediatrics and Adolescent Medicine, Ludwig-Maximilians University, Munich, Germany

Abstract: Generalized linear mixed effects models (GLME) have become an increasing popular choice for the modeling of longitudinal, non-Gaussian responses in health related studies. Clinicians commonly use grouped or clustered data to estimate temporal trends in counts or certain repeated measurements in a regression setting, for example the number of vertigo attacks within a certain period. Are GLMEs appropriate to analyze these data?

Analytical methods for model diagnosis and goodness-of-fit criteria for GLMEs are limited so far (in contrast to methods available in LMEs for normal data). Although simulation studies have shown that in linear mixed models inference on fixed effects are comparatively robust to non-normality of random effects, this no longer holds for non-linear or generalized linear mixed models: there are findings of inconsistencies in fixed as well as random-effects estimation under misspecification of the random-effects distribution. This requires due diligence by the statistician.

A fully Bayesian approach can be applied to analyze hierarchical nonnormal data based on diffuse or even weakly informative priors for the parameter vector and involves the use of Markov Chain Monte Carlo (MCMC) methods. There are several methods for model validation and comparison in the Bayesian framework. To assess the fit of different hierarchical models we apply proper scoring rules, DIC or posterior predictive model checks to get information which model performs best. We will compare these Bayesian tools with frequentist approaches and apply them to longitudinal data of vertigo attacks.

References:

MOLENBERGHS G., VERBEKE G. (2005): *Models for Discrete Longitudinal Data*. Springer, New York.

GNEITING T., RAFTERY A.E. (2007): Strictly proper scoring rules, prediction, and estimation. *J. Am. Statist. Assoc.*, 102(477): 359–378.

BROWN H., PRESCOTT R. (2006): *Applied Mixed Models in Medicine*. 2nd edition. Wiley, Chichester.

An empirical efficiency comparison of three factor designs with nested and crossed treatment structures

Katarzyna Ambroży[1] and Iwona Mejza[2]

[1] Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; ambrozy@au.poznan.pl
[2] Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; imejza@au.poznan.pl

Abstract: The purpose of the paper is to examine the effectiveness of three factor designs with respect to the accuracy of the estimation of treatment parameters using empirical relative efficiency. There are considered a split–split–plot design, a split–block-plot design (otherwise a strip–split–plot design) and a split–plot x split–block design. Situations when the designs are equally efficient and when one of them is more efficient than the others are discussed. To illustrate the theory presented in the paper, we consider an experiment designed to test the effects of nitrogen fertilization and a chemical preparation – a growth regulator on the grain yields of wheat varieties.

References:

AMBROŻY, K. and MEJZA, I. (2002): Some incomplete split–block–plot designs. *Scientific Papers of Agriculture University of Poznań, Agriculture, Vol. 3, 3–12.*

AMBROŻY, K. and MEJZA, I. (2003): Some split–plot x split–block designs. *Colloquium Biometryczne, 33, 83–96.*

GOMEZ, K.A. and GOMEZ, A.A. (1984): *Statistical procedures for agricultural research.* Wiley, New York.

LeCLERG, E.L., LEONARD, W.H. and CLARK, A.G. (1962): *Field plot technique.* Burgess, Minneapolis.

Online

Small sample comparisons of binomial probabilities: Liebermeister's test, Fisher's exact test, exact odds ratio confidence intervals

Karl-Ernst E. Biebler, Bernd P. Jäger, and Michael Wodny

Institut für Biometrie und Medizinische Informatik, D-17475 Greifswald, Germany;
biebler@biometrie.uni-greifswald.de

Abstract: Binomial probabilities shall be compared in a two sample design. It is generally agreed that FISHER's exact test is conservative. Therefore LANCASTER's mid-P test was recommended as an adjustment for FISHER's test (BERRY and ARMITAGE 1995). In addition, the LIEBERMEISTER test (1877) has to be taken into account. SENETA/PHIPPS (2001) conclude that LANCASTER's and LIEBERMEISTER's tests give better approximations to the exact unconditional p-values than FISHER's exact test. For small samples the LIEBERMEISTER test is recommended. The mid-P test should be preferred in excessively unbalanced cases.

We compared in simulation studies FISHER's test, LIEBERMEISTER's test and the test being based on the exact confidence interval of the odds ratio. The latter is calculated with respect to incomplete beta functions. These simulations also offer the possibility of the empirical calculation of the power of the respective tests. The deviations from the null hypothesis are parameterized by the odds ratio. For every parameter situation the simulation size was 10 000. The calculations were fast and without problems executed on a PC in a SAS environment.

This method makes the choice of the best test as well as sample size calculations possible for the respective application situation.

References:

BERRY, G. and ARMITAGE, P. (1995): Mid-P confidence intervals: a brief review. The Statistician, 44, 417-423.

LACHIN, J.M. (2000): Biostatistical methods. The assessment of relative risks. John Wiley & Sons, New York.

LIEBERMEISTER, C. (1877): Über Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik. Sammlung klinischer Vorträge Innere Medizin 110, No. 31-64, 935-962.

PHIPPS, M.C. (2003): Liebermeister's quasi-exact test for two binomials. Proc. of the Second Annual Hawaii International Conference on Statistics and Related Fields. Honolulu. Hawaii.

SENETA, E. and PHIPPS, M.S. (2001): On the comparison of two observed frequencies. Biometrical Journal, 43, 1, 23-43.

SENETA, E., SEIF, F. J., LIEBERMEISTER, H. and DIETZ, K. (2004): Carl Liebermeister(1833-1901): A pioneer of the investigation and treatment of fever and the developer of a statistical test. Journal of Medical Biography 12, 215-221

Assessment of Water Supply Systems in Turkey Using Multivariate Statistical Methods

Hülya Boyacioglu[1] and Hayal Boyacioglu[2]

[1] Dokuz Eylul University, Faculty of Engineering, Department of Environmental Engineering, Tinaztepe Campus Buca 35160 Izmir Turkey; hulya.boyacioglu@deu.edu.tr

[2] Ege University, Faculty of Science, Department of Statistics, Bornova 35100, Izmir Turkey; hayal.boyacioglu@ege.edu.tr

Abstract: This paper examines water supply systems of Turkey using statistical methods. In this context, results of questionnaire surveys conducted annually by Turkish Statistical Institute (TURKSTAT) between 2001-2004 to investigate 'amount of water abstracted to drinking water networks by type of resources' and 'status of drinking water treatment plants' were evaluated. In the questionnaire, water sources were grouped under five categories as spring, (artificial) lake, river, reservoir and well. Additionally physical, conventional and advanced treatment plants were examined with respect to number of plants, capacity and amount of water treated in 81 provinces. In the study, the dimensionality reduction technique 'Factor Analysis' was applied to find latent factors among provinces. The correlation matrix of variables was generated and factors extracted by the Centroid method, rotated by Varimax rotation using 'Statistical Package for the Social Sciences Software-SPSS 10.0 for Windows'. Based on the results of factor analysis and distribution of water among the sources in the provinces correlated with first factor-F1 (which explained highest percentage of the variance in data set) it was concluded that F1 represented 'groundwater dependent provinces'. Therefore water supply systems in the country were governed by groundwater sources (spring and/or well). In addition mainly conventional treatment plants were used to provide good quality water to the settlements. This study showed that statistical methods can be used successfully in 'data-information' transformation process in environmental studies.

References:

SPSS- 10. Statistical Package for the Social Sciences. SPSS Inc, Chicago, USA.

TURKISH STATISTICAL INSTITUTE- [TURKSTAT] (2007b). Amount of water abstracted to drinking water networks by type of resources, 2004 Turkey's Statistical Yearbook, Environmental Statistics Compendium of Turkey II.

Effect of Box-Cox transformation in model of growth of fungus *Trichoderma harzianum*

Maria Kozłowska[1], Anna Budka[1], and Romuald Górski[2]

[1] Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; markoz@au.poznan.pl, abudka@au.poznan.pl

[2] Department of Plant Protection Methods, Agricultural University, Zgorzelecka 4, 60-198 Poznań, Poland; rgorski@au.poznan.pl

Abstract: The Box-Cox family of transformations has become a widely used tool to make data behave according to a linear regression model. The Box-Cox transformation is often used when the true functional form of the regression equation is unknown. The parameter lambda defining the transformation gives possibility of choice between, for example, normal or log-normal equation. The dependent variable, transformed according to the Box-Cox procedure, is usually assumed to be normally distributed with constant variance. The regression coefficients are generally estimated by least squares method. This transformation was applied to detect relationship between time and growth of area of colonies of fungus *Trichoderma harzianum*. The model of growth of the fungus was founded with or without the use of certain fungicides. Unfortunately, near normality and homoscedasticity were not simultaneously attained with this single transformation in some considered cases.

Evaluation of the needs and route for long-term rehabilitation in survivors of acute pancreatitis

Izabela Chmiel, Antoni Czupryna, Maciej Górkiewicz, and Tomasz Brzostek

Jagiellonian University Medical College; izabela_chmiel@wp.pl, msmichal@cyf-kr.edu.pl, gorkiewicz@poczta.fm, mmbrzost@cyfronet.krakow.pl

Abstract: The importance of the quality-of-life examination, e.g. with SF-36 questionnaire, in choosing strategies of rehabilitation, has steadily increasing in recent years. Nevertheless, the result of ordinary statistical procedures often should be treated with caution there, because of violation the basic condition to use them, like non-normality and/or variances' heterogeneity and/or mean-variance correlation. Moreover, since there don't exist any a-priori hypothesis about the pattern of SF-36 items, the multiple comparisons task arises there. With aim to overcome above difficulties in this study the bootstrap resampling method was used with Authors own Excel Macros tailored to course of computing. The medical example was based on SF-36 data obtained in $N = 142$ survivors of acute pancreatitis, divided into three groups with respect to their individual medical background of the disease. The raw SF-36 data were standardised as usually to range $0 - 100\%$ ability for 9 scales separately and then used to generate $L = 1.000$ random patterns of scale's means. For each a pattern the ranks of scale's means were computed. Besides, the probabilities $P(mean_j < mean_k)$ were estimated for each pair of scales. The homogeneity of the patient's population was confirmed by the comparison between the three groups, and then by rank Spearman correlation between the bootstrap ranks pattern and the measured patient's individual profiles.

References:

CARPENTER J. and BITHELL J. (2000): Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine 19 (9): 1141–1164.*

BORKOWF C.B. (2000): A new nonparametric method for variance estimation and confidence interval construction for Spearman's rank correlation. *Computational Statistics & Data Analysis 34: 219–241.*

CHALONEN K., PETILA V., LEPPANIEMI A., KEMPPAINEN E., PUOLOKKAINEN P., HAAPIAINEN R. (2003): Long-term health-related quality of life in survivors acute pancreatitis *Intensive Care Med. 29:782–786.*

'CMA' - Steps in developing a comprehensive R-toolbox for classification with microarray data and other high-dimensional problems

Martin Slawski, Anne-Laure Boulesteix, and Martin Daumer

Sylvia Lawry Centre for MS Research, Hohenlindenerstr. 1, D-81677 München;
Martin.Slawski@campus.lmu.de, boulesteix@slcmsr.org, daumer@slcmsr.org

Abstract: Microarray studies have stimulated the development of new approaches and motivated the adaptation of known traditional methods for class prediction with high-dimensional data. There already exist numerous sofware packages implementing single methods for microarray-based classification and in addition two synthesis packages: `MLInterfaces` by V. Carey and R. Gentleman (2007) and `MCRestimate` by Ruschhaupt et al (*Stat Appl Genet Mol Biol 2004, 3:37*), available from the `www.bioconductor.org` platform. Conceptually, the `R` package `CMA` is more related to the second one, focussing on comparative model evaluation according to accepted 'good pratice' standards/guidelines (Dupuy and Simon, *J Natl Cancer Inst 2007, 99:147-157*), an aspect neglected by `MLInterfaces`, though still widely used. In a nutshell, `CMA` provides a uniform interface to a total of more than 20 supervised classification methods, comprising classical approaches such as discriminant analysis or penalized multinomial logistic regression, dimension reduction by Partial Least Squares, and more sophisticated methods, e.g. Support Vector Machines, Neural Networks or boosting techniques.

The evaluation of the constructed classifiers is based on repeated splittings into learning and test sets or related approaches (e.g. bootstrap). For each learning set separately, variable selection can be performed optionally, either by a collection of simple tests or by advanced techniques such as the lasso, elastic net or componentwise boosting. In the last step, hyperparameter optimization and model evaluation are carried out via a 'nested' cross-validation procedure. The outer loop is used for classifier evaluation while appropriate values for the hyperparameters are determined in the inner loop.

`CMA` is implemented entirely in `S4` classes (J. Chambers, *Programming with data*, 1998). Its modular construction makes the incorporation of new methods easy. Furthermore, it is intended to be user-friendly by providing a multitude of pre-defined methods for summarizing and visualizing classifier evaluation and comparison.

A preliminary version of `CMA` is planned to be available in the next `Bioconductor` release in April 2008.

Population statistics and analysis of STR polymorphisms data using R and package forensic

Václav Faltus and Jana Zvárová

Centre of Biomedical Informatics, Institute of Computer Science AS CR, v.v.i, Pod Vodárenskou věží 2, Prague 8, Czech Republic; faltus@euromise.cz

Abstract: Individual genetic identification of biological samples plays an important role in cases of identification of victims of a crime and also in cases of mass disasters like floods, tsunami, aircraft and train accidents. To identify an individual we usually need a database of genetic profiles. If a victims profile is not in such database, we can determine to which ethnic group or subpopulation it belongs to. The STR polymorphisms data are convenient for several reasons. The STR loci are very polymorphous and they usually lie in non-coding regions of the DNA. Therefore it is believed that selection pressures do not influence the inheritance patterns.

There are several population characteristics in forensic genetics which are used when analysing STR polymorphisms data. In our work we present functions to compute these characteristics using **R**, which can be easily extended by packages which add new functions. We use and add new functions to the package *forensic*. The functions that we present are related to identification and two of them are specific to paternity testing. The characteristics are the homo- and heterozygosity, average match probability, average discrimination power, polymorphic information, average exclusion probability and typical paternity index.

The work was supported by the grant 1M06014 of the Ministry of Education of the Czech Republic.

References:

MARUŠIAKOVÁ, M. (2007): *forensic: Statistical Methods in Forensic Genetics. R package version 0.2.* Center of Biomedical Informatics, Institute of Computer Science and Academy of Sciences of the Czech Republic.

R DEVELOPMENT CORE TEAM (2006): *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

A Probabilistic Fertility Model for First Conception

Mohammed Ali Z. Farooqui

M.D.College, Mumbai, India; nuzza@mail.com

Abstract: This paper considers a new finite range continuous model to study the waiting time for first conception as against the earlier model with infinite range.

References:

GINI, C. (1924): Premiers Recherches sur la Fecondabilite de la femme. *889-892.*

LAWLESS, T.F. (1982): *Statistical Models and Methods for Life Time Data.* John Wiley, New York.

MUKHERJEE and ISLAM (1983): A finite range distribution of failure times. *Nav. Res. Log. Quart. 30, 487–491.*

SINGH, S.N. (1964a): On the time of First Birth, *Sankhya, Vol. 26(B), 95–102.*

VINCENT, P. (1961): *Recherches sur la Fecondite Biologique.* Institute National D'Eludes demographic Paris France, Press Universities De France.

Co-Morbidities and Biomarkers in Heart Failure

Götz Gelbrich[1], Stefan Störk[2], Christiane Prettin[1], and Christiane Angermann[2]

[1] Coordination Centre for Clinical Trials Leipzig (KKSL), Härtelstr. 16-18, 04107 Leipzig; goetz.gelbrich@kksl.uni-leipzig.de
[2] Medical Clinic I of the University of Würzburg, Cardiology, Klinikstr. 6-8, 97070 Würzburg; Angermann_C@medizin.uni-wuerzburg.de

Abstract: Chronic heart failure is among the most common clinical syndromes and is associated with an annual mortality of 10 to 40 percent. A variety of co-morbid conditions is related with worsening of heart failure and increased hazard of death. Little is known about the mechanisms of the interplay between heart failure and co-morbidities. It is, therefore, crucial to understand whether co-morbidities play a causal role in heart failure or, whether they just represent markers of progression of the systemic disease. Further, it is presently unclear, whether the interrelationship between heart failure and co-morbidities is mediated by shared pathophysiological mechanisms.

Biometric key challenges arising in this context are (a) the adequate reflection of co-morbidities by variables derived from physical findings, symptoms, biomarkers and medication, (b) efficacious discrimination between quantitative and qualitative information, (c) incorporation of methodological knowledge regarding specific measurements, (d) understanding the need for multistage modelling. These issues will be illustrated focussing on inflammation, anaemia, nutritional parameters and depressive mood disorder. We will demonstrate the complex interplay between statistical methods and clinical a priori knowledge in the analysis of these issues.

Further, we will discuss the development of prospective intervention trials targeting the issue of co-morbidities based on results from observational studies. Knowledge gained from such trials will show whether causal or symptomatic treatment of a co-morbid condition is warranted. The currently ongoing MOOD-HF trial examining the effect of antidepressive drug treatment on cardiac mortality and morbidity will be presented as an example.

This work was supported by the Competence Network of Heart Failure, funded by the Federal Ministry of Education and Research (BMBF), grant FKZ 01GI0205.

Analyzing relapse free survival: right censored methods on interval censored data?

Ursula Gerhardinger[1], Gertraud E. Markl[1,2], and Ulrich Mansmann[1]

[1] IBE, Institute of Medical Informatics and Biometry, University of Munich;
Ursula.Gerhardinger@ibe.med.uni-muenchen.de
[2] CSC^{LMU}, Clinical Study Center, Hospital of the University of Munich;
Gertraud.Markl@med.uni-muenchen.de

Abstract: Oncological studies often look at interval censored events (ICE), e.g. relapse free survival. Nevertheless these data are in general analyzed with inappropriate statistical methods, namely Kaplan-Meier-Estimator (KM) and Cox-Regression. Correct univariate alternatives are the Turnbull- or Cutler-Ederer-Estimator [Cutler (1958), Turnbull (1974)]. Multivariate methods for ICE are also available (e.g. the R-package *intcox* which implements ideas from Pan [Pan (1999)]).

Our objective is to discuss the consequences of the current misuse of statistical methods. First of all, event times may be overestimated. This will have a misleading effect when designing future studies. Second, estimators for relative hazard will be biased and lead to misinterpretation. Third, the information content of ICE data is lower than that of right censored event time data. Thus, planning the study by using tools for right censored data and analyzing ICE by inappropriate methods, the power of a study is reduced and the detection of a true effect is prevented.

We quantify the expected effects in a simulation study of relevant scenarios. The insight gained by the simulation study will be used to discuss results on event free survival of soft tissue sarcoma, taking into account the events of local, distal recurrence, and death, whichever occurred first.

References:

CUTLER, S.J. and EDERER, F. (1958): Maximum utilization of the life table method in analysing survival. *Journal of Chronic Diseases, 8, 699–712.*

WEI PAN (1999): Extending the iterative convex minorant algorithm to the cox model for interval-censored data. *Journal of Computational and Graphical Statistics, 78, 109–120.*

BRUCE W. TURNBULL (1974): Nonparametric Estimation of a Survivorship Function with Doubly Censored Data. *Journal of the American Statistical Association, 69, 169–173.*

Use of baseline ECGs in the evaluation of thorough-QT studies with crossover design

Patricia Glomb[1] and Arne Ring[2]

[1] University of Heidelberg, Germany; patricia.glomb@boehringer-ingelheim.com
[2] Boehringer Ingelheim Pharma, Germany

Abstract: Since the guideline ICH-E14 was released, the use of baseline ECGs for evaluation of the "largest time-matched mean difference" between the drug and placebo had been controversially discussed. This discussion was driven by theoretical considerations, which took account for the rather large variability of QT(c) endpoints. On the other hand, ECG specific effects like the circadian rhythm of the endpoints and their potential correlation between subsequent days had often been neglected.

Most crossover-TQT studies follow one of the following designs (based on 4-periods, with active and passive control and two doses of test drug):

A) No baseline values, only the absolute difference between placebo and the drugs is evaluated;
B) One pre-dose value before the first dosing of each period;
C) One baseline day as a separate crossover period;
D) One baseline day before each crossover period.

The basis for the investigations were 4 clinical studies with design D), which enables to simulate Designs A)-C) to derive the standard error of the QT(c)-endpoints for each design as a metric of the statistical efficiency.

Individual circadian rhythms of QTc parameters can be observed, which are not stable but which only change slowly over time. Hence, there some auto-correlation of QTc parameters over time can be found, which decreases if the time distance between baseline and on-treatment values increases. This leads to increasing variability of the endpoint 'QTc change from baseline-' - by about 40% at steady state compared to day 1. The extent of the circadian rhythm is lower than the variability of the QTc endpoints, even if triple ECGs are recorded.

Due to the change of the circadian rhythm over time, the baseline QT(c)-intervals cannot assumed to be equal, so that baseline ECGs provide additional information. In single dose TQT studies, the number of ECGs to be recorded are lowest in design Design B), while the sample size is lowest in Design D). For multiple dose studies, the advantage of strategy B) over A) is less pronounced, while strategy D) cannot be recommended in this design. For a final decision on the choice of the baseline strategy, the variability of the endpoint is the most important, but not the only issue for consideration (e.g. need for heart rate correction).

References:

HEIMANN G. (2005): *The Statistical Analysis of Thorough QTc Studies.* Heidelberg, Workshop.

RIEK M. (2007): *Baseline Adjustment in Central Tendency Analyses of QTc: Inefficient for Cross-over Designs?* Washington, DIA Cardiac Safety Conference.

A prospectively planned pooled meta-analysis of the association between chromosome 9p21.3 and coronary artery disease

Anika Götz[1], Inke R. König[2], Jeanette Erdmann[3], Ludwig A. Hothorn[4], H.-Erich Wichmann[5], Nilesh Samani[6], Heribert Schunkert[7], Andreas Ziegler[8] on behalf of the Cardiogenics and the WTCCC group

[1] Institute of Medical Biometry and Statistics and Medical Clinic II, University at Lübeck, Lübeck; anika.goetz@imbs.uni-luebeck.de
[2] Institute of Medical Biometry and Statistics, University at Lübeck, Lübeck; inke.koenig@imbs.uni-luebeck.de
[3] Medical Clinic II, University at Lübeck, Lübeck; j.erdmann@cardiogenics.eu
[4] Institute of Biostatistics, Leibniz University, Hannover; hothorn@biostat.uni-hannover.de
[5] Institute of Epidemiology, GSF National Research Center for Environment and Health, Neuherberg, Germany; wichmann@gsf.de
[6] Department of Cardiovascular Sciences, University of Leicester, Leicester; njs@leicester.ac.uk
[7] Medical Clinic II, University at Lübeck, Lübeck; heribert.schunkert@innere2.uni-luebeck.de
[8] Institute of Medical Biometry and Statistics, University at Lübeck, Lübeck; andreas.ziegler@imbs.uni-luebeck.de

Abstract: Recently, genome-wide association studies identified chromosome 9p21.3 to affect the risk of coronary artery disease (CAD). We investigated the association of this locus with CAD in seven case-control studies and undertook a prospectively planned meta-analysis.

A single nucleotide polymorphism, rs1333049, representing the 9p21.3 locus was genotyped in seven case-control studies involving a total of 4645 patients with myocardial infarction (MI) or CAD and 5177 controls.

This study was planned as a type IV meta-analysis according to Blettner et al. (1999). Three different approaches were used for the analysis. First, we investigated the additive effect of rs1333049 on CAD/MI by random effect logistic regression models with adjustments for study. The logistic regression framework was also used to identify the underlying genetic model. Second, we followed Minelli et al. (2005) and estimated the ratio λ of the log odds ratio for the heterozygous individuals compared to homozygous individuals. Here, λ equals $0, 1/2, 1$, and > 1 if the genetic model is recessive, additive, dominant, and positive heterosis, respectively. Third, Hothorn and Hothorn (unpublished) recently proposed a modification of the MAX test approach of Freidlin et al. (2002). Finally a meta-analysis of the present data and all previously published samples was conducted.

The risk allele of the lead SNP rs1333049 was uniformly associated with CAD in each study ($p < 0.05$). In a prospectively planned pooled analysis the odds ratio per copy of the risk allele was 1.29 ($95\%CI : [1.22, 1.37],\ p = 0.0001$). In the different approaches, an autosomal additive mode of inheritance best explained the underlying association. The meta-analysis in 12004 cases and 28949 controls increased the overall level of evidence for association with CAD to $p = 6.04 \cdot 10^{-10}(OR\ 1.24\ [1.20, 1.29])$.

References:

BLETTNER M, SAUERBREI W, SCHLEHOFER B, SCHEUCHENPFLUG T, FRIEDENREICH C (1999): Traditional Reviews, Meta-Analyses and Pooled Analyses in Epidemiology. Int J Epidemiol., 28, 1-9.

MINELLI C, THOMPSON JR, ABRAMS KR, THAKKINSTIAN A, ATTIA J (2005): The Choice of a Genetic Model in the Meta-Analysis of Molecular Association Studies. Int J Epidemiol., 34, 1319-28.

FREIDLIN B, ZHENG G, LI ZH, GIEGER C (2002): Trend Tests for Case-Control Studies of Genetic Markers: Power, Sample Size and Robustness. Human Heredity., 53, 146-52.

Some notes about taking trials for estimation height curve

Katarzyna Kaźmierczak[1] and Małgorzata Graczyk[2]

[1] Department of Forest Management, Agriculture University of Poznań, Wojska Polskiego 71c, 60-625 Poznań, Poland
[2] Department of Mathematical and Statistical Methods, Agriculture University of Poznań, Wojska Polskiego 28, 60-637 Poznań, Poland; magra@au.poznan.pl

Abstract: In research connected with the estimation of productivity of the forest the most important is height-diameter curve. For the estimation of the height curve parameters, we measure the diameter of breast height for all trees and height of 10% of trees in the sample. In the practice the 10% of the sample is systematically collected. In the paper we propose random choice of the tress for sample and we compare these two methods.

References:

BRUCHWALD, A., DUDZIŃSKA, M. and WIROWSKI, M. (1979): Uniform heihgt curves for oak stands. *Ann. Warsaw Agricult. Univ. - SGGW, 45, 3-5.*

On Certain Regular D-optimal Spring Balance Weighing Designs

Krystyna Katulska and Katarzyna Przybył

Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Umultowska 87, 61-614 Poznań, Poland; krakat@amu.edu.pl, pkasik@amu.edu.pl

Abstract: The estimation problem of individual weights of objects in spring balance weighing design using the criterion of D-optimality is discussed. It is assumed that variances of errors are not equal and errors are not correlated. The upper bound of the determinant of the information matrix of estimators is obtained and the conditions for this upper bound to be attained are proved. Some methods of constructions regular D-optimal spring balance weighing designs are demonstrated.

References:

HARVILLE, D. A. (1997): *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag, New York, Inc.

NEUBAUER, M. G., WATKINS, W. and ZEITLIN, J. (1997): Maximal j-simplices in the real d-dimensional unit cube. *Journal of Combinatorial Theory, Ser. A 80, 1-12*.

NEUBAUER, M. G., WATKINS, W. and ZEITLIN, J. (1998): Notes on D-optimal designs. *Linear Algebra and its Applications, 280, 109-127*.

Some Aspects on Bivariate Meta–Analysis

Guido Knapp

Fakultät Statistik, Technische Universität Dortmund, 44221 Dortmund, Germany;
guido.knapp@uni-dortmund.de

Abstract: In this paper we consider the meta–analysis of controlled clinical trials with two outcomes of interest. Usually two separate univariate meta–analyses are applied to synthesize the evidence for each outcome independently. Recently, Riley et al. (2007) evaluate a general bivariate random effects meta–analysis model and illustrate the benefits and limitations of this approach. Considering this general bivariate model we will address the following two topics:

1. Which model should we use, a fixed effects or a random effects model? How do we model the between–study variance–covariance matrix in a random effects model?
2. Assume that some studies only provide results for one outcome, but not for the other one. How can we use this information in our bivariate model?

Using a real data set we will illustrate the methods. Assessments of the performance of the statistical methods will be done using simulated data.

References:

RILEY, R.D., ABRAMS, K.R., LAMBERT, P.C., SUTTON, A.J., and THOMPSON, J.R. (2007): An Evaluation of Bivariate Random–Effects Meta–Analysis for the Joint Synthesis of Two Correlated Outcomes. *Statistics in Medicine, 26, 78–97.*

Application of Chow test to comparison of results of image analysis and conventional measurement in plant protection

Maria Kozłowska[1], Radosław J. Kozłowski[2], and Romuald Górski[3]

[1] Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; markoz@au.poznan.pl
[2] Institute of Agricultural Engineering, Agricultural University, Wojska Polskiego 50, 60-637 Poznań, Poland; rkozlowski@au.poznan.pl
[3] Department of Plant Protection Methods, Agricultural University, Zgorzelecka 4, 60-198 Poznań, Poland; rgorski@au.poznan.pl

Abstract: This article introduces a new view on use of Chow test. The test was used to the compare of two models of growth of *Trichoderma harzianum*. These models were constructed on basis of results of two different measurement methods.

Present digital technologies offer a trouble-free acquisition and processing of images. By means of computer-based image analysis it was possible to automate the measurement of areas with colonies of the fungus. The results of image analysis were obtained using *FotoDetekt* program. Sizes of these colonies were also measured in conventional way. On basis of two different results the models of evaluation growth of the fungus were introduced. These models were analyzed and compared.

Tailoring the residential home facilities to structure of the elderly patients needs

Ilona Kuzmicz[1], Maciej Górkiewicz[1], and Tomasz Brzostek[1]

Jagiellonian University Medical College, Cracow, Poland;
ilonakuzmicz@interia.pl, gorkiewicz@poczta.fm,
mmbrzost@cyfronet.krakow.pl

Abstract: For ethical and economical reasons the elderly people's needs should be met at nursing home on an individual basis, taking into account the abilities of each person, e.g. 10 daily abilities measured with Barthel Scale of Activities of Daily Living. The particular abilities have their different numbers of rating grades of 5 points each, e.g. a satisfying self-reliance with transfer from bed to wheelchair gives 15 rating points, but full ability to wash himself gives 5 points only. Nevertheless, any ordering of the individual abilities was not presupposed there. The data for medical example compromised 160 patient's profiles obtained at two nursing homes. In this paper the relations between abilities were modelled with chain of pairs of depended abilities with the transition matrices estimated on given sample of profiles. The better arrangement among all possible chains was chosen with parametric bootstrap under criterion of maximum of sum of chi-squares for all pairs of abilities adjacent at the chain divided by total sum of degree of freedoms. Then the parametric bootstrap used this model to generate a set of $L = 1000$ artificial patient's profiles needed to distinguish clusters of profiles with the aim to identify patient subgroups. The validity of the chosen chain model was confirmed with standard tests for Markov chains (Eggar, 2002) and then by comparing relations between scores in the original sample and in a bootstrap sample at all pairs of abilities (not only for these neighbouring at the chain).

References:

WAGENMAKERS E.J., RATCLIFF R., GOMEZ P., IVERSONC G.J. (2004): Assessing model mimicry using the parametric bootstrap. *Journal of Mathematical Psychology 48, 28-50.*

EGGAR M.H. (2002): Validity of fitting a first-order Markov chain model to data. *Journal of the Royal Statistical Society, Series D, 51 (2), 259-265.*

HANCOCK G.A., WOODS B., CHALLIS D., ORRELL M. (2006): The needs of older people with dementia in residual care. *International Journal of Geriatric Psychiatry 21: 43-49.*

Online

Do we gain from imputing ungenotyped SNPs in genomewide association studies?

Claudia Lamina[1], Goncalo Abecasis[3], H.-Erich Wichmann[1,2], Florian Kronenberg[4], and Iris M. Heid[1,2]

[1] Helmholtz Center Munich, Institute of Epidemiology, Neuherberg, Germany; claudia.lamina@gsf.de
[2] IBE, LMU, Munich, Germany
[3] Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA
[4] Division of Genetic Epidemiology; Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria

Abstract: Meta-analysis of genomewide association (GWA) studies pose the problem of combining SNP information from different genotyping platforms. This implies that different sets of SNPs for the involved studies are available. It was our objective to provide a comprehensive SNP panel for the KORA study using an imputation method to enable pooling of GWA data with other studies. It was a further objective to evaluate the impact of uncertainty from imputation on association results.

The GWA on plasma levels of HDL-Cholesterol (HDLC) in the KORA F3 study (n=1644) included the 500K SNP panel genotypes of the Affymetrix platform ($\sim$ 0.5 Mio SNPs). We applied the program MACH using Hidden Markov Models to impute missing genotypes for all available HapMap-SNPs ($\sim$ 2.5 Mio SNPs) by utilizing HapMap information on minor allele frequency and LD structure. Association analyses were performed with and without accounting for the uncertainty of the imputed genotypes. On the example of HDLC candidate genes, we illustrate the gain in information, especially for fine-mapping endeavours.

68.7% of the SNPs had an imputation quality score of 80% or higher, 14.5% moderate quality between 50-80%, and 16.8% inferior quality below 50%. Accounting for this uncertainty resulted in an accentuation of estimators of 1.6% in average to the naive estimator for high quality SNPs and up to 15.8% for low quality SNPs, indicating increasing bias for increasing uncertainty.

Applying imputation methods results in a SNP panel that can be compared to any other HapMap based SNP set. It can be be utilized for cheap fine-mapping with the pay-off of severe or moderate uncertainty in about 1/3 of the SNP data.

Turnbull's nonparametric estimator and its application to the analysis of failure and suboptimal response in chronic myeloid leukaemia (CML)

Michael Lauseker and Markus Pfirrmann

IBE, LMU, Munich, Germany; lauseker@ibe.med.uni-muenchen.de, pfi@ibe.med.uni-muenchen.de

Abstract: For the prognosis of the outcome of chronic myeloid leukaemia (CML), it is important to define time points at which it can be decided if the current treatment is satisfactory. If the patient has not reached a specified response within these periods or he has lost an already existing response, his success is classified as "failure" or "suboptimal response". One problem in analysing "failure" and "suboptimal response" is that the patient's status is only known at the doctor's visits which take part every three to six months, another one is that the exact date of an event is unknown and observation delayed. Although often applied to interval censored data, the well-known Kaplan-Meier estimator is only defined for exactly observed data. Thus, there had to be found another method for dealing with interval censored data. One suggestion came from Turnbull [2] and was implemented in R by Giolo [1]. Based on the data of 408 patients receiving Imatinib of the German CML IV study, this nonparametric estimator was compared to the Kaplan-Meier estimator and the lifetable estimator of Cutler and Ederer. It was observed that the Turnbull estimator was most suitable to this problem, because it was the only one that didn't overestimate survival on the one hand and represented the specific structure of the data on the other hand.

References:

GIOLO, S.R. (2004), Turnbull's Nonparametric Estimator for Interval-Censored Data. Available: http://www.est.ufpr.br/rt

TURNBULL B. (1974), Nonparametric Estimation of a Survivorship Function with Doubly Censored Data.*Journal of the American Statistical Association, 69, 169-173.*

Nonparametric Trend Tests for the Analysis of Right Censored Survival Times

Sandra Leissen[1] and Markus Neuhäuser[2]

[1] Fakultät Statistik, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany; leissen@statistik.tu-dortmund.de
[2] Fachbereich Mathematik und Technik, RheinAhr-Campus Remagen, Südallee 2, 53424 Remagen, Germany; neuhaeuser@rheinahrcampus.de

Abstract: In pharmaceutical trials, e.g. in preclinical carcinogenesis or dose finding studies, survival times may arise in groups associated with ordered doses. Here the emphasis of analysis is placed on trend dependent differences in the survival times. So a *trend test* for *ordered alternatives* may be a suitable application. Various trend tests for survival data have already been proposed including the logrank trend test, the one by Gehan (1965) and Mantel (1967) and the modified logrank test (Liu et al., 1998), where the latter is a special case of the logrank trend test (Leissen, 2007). Each of these tests can be viewed as a *single contrast test*, and therefore all tests have in common that they are more sensitive for certain trends than for others. According to the construction principle of *multiple contrast tests*, new trend tests were developed on the basis of the tests mentioned above. In order to compare the new with the conventional tests, simulation studies based on different scenarios of, e.g. number of groups, group sizes, level of censoring and ties, distribution of the survival times and inherent trend among the groups were carried out.

References:

GEHAN, E. A. (1965): A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples. *Biometrika, 52 (1 and 2), 203–223.*

LEISSEN, S. (2007): *Nichtparametrische Trendtests zur Analyse von rechtszensierten Überlebenszeiten.* Diplomarbeit, FB Statistik, Universität Dortmund.

LIU, P.-Y., TSAI, W. Y. and WOLF, M. (1998): Design and Analysis for Survival Data under Order Restrictions with a Modified Logrank Test. *Statistics in Medicine, 17, 1469–1479.*

MANTEL, N. (1967): Ranking Procedures for Abitrarily Restricted Observation. *Biometrics, 23 (1), 65–78.*

Online

Comparing short-term efficacy of the health-promotion actions addressed to the same target

Iwona Malinowska-Lipien, Ewa Kawalec, Maciej Górkiewicz, and Tomasz Brzostek

Jagiellonian University Medical College, CraCow, Poland; imalin@poczta.onet.pl, e.kawalec@interia.pl, gorkiewicz@poczta.fm, mmbrzost@cyfronet.krakow.pl

Abstract: Over the past decade it has become evident that even the best prevention exercises for healthy people or the post-hospital rehabilitation programme for convalescents has been ineffective without proper patient's attitude. On that occasion a huge of innovative health promoting techniques and tricks has been increasingly offered every year. Nevertheless, the comparative analyses met there some serious difficulties, in these because of a fact, that the typical distributions of the exercises intensity even on an initial period, say at the first month after promoting action, have usually two distinctly separated modes. First mode arises at near to zero patient's activity. The second mode arises near to recommended intensity of the exercises. Thus, the purpose of the comparative statistical analyses was there twofold: the first aspect was to verify the real effect of a considered promotion manner on the percentage of the exercising patients. The second purpose was to examine distributions of the exercise's intensity, with regard to random proportion of the active patients in a random sample. The resampling approach creates opportunity to carry out both above analyses simultaneously. In this pilot study data on health promotion among n = 96 pupils and among n = 32 cardiac convalescents were used. Confidence intervals for probabilities and mean values were computed with Wood's (2005) bootstrap calculator, available on-line. The bias reduction procedure (Davidson & MacKinnon, 2007) did not change the observed relations between above estimates.

References:

CAMPBELL M.J, BROWNE D., WATERS W.E. (1985): Can general practitioners influence exercise? *British Medical Journal. 290, 1044–1046.*

WOOD M. (2005): The Role of Simulation Approaches in Statistics. *J. of Statistics Education, 13(3) www.amstat.org/publications/jse/v13n3/wood.html.*

DAVIDSON R., MacKINNON J.G. (2007): Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis 51, 3259 – 3281.*

Constructions of incomplete split-plot × split-block designs based on Kronecker type products

Iwona Mejza and Katarzyna Ambroży

Department of Mathematical and Statistical Methods, Agricultural University, Wojska Polskiego 28, 60-637 Poznań, Poland; imejza@au.poznan.pl, ambrozy@au.poznan.pl

Abstract: The Kronecker product and a semi-Kronecker (Khatri–Rao) product have been adapted to construct incomplete split-plot × split-block designs for three factor experiments. With both methods some generated designs have been obtained when levels of each factor have been allocated to a balanced square lattice design. A comparison of the methods of constructing the designs with respect to stratum efficiency factors for some contrasts of treatment parameters and with respect to a size of an experiment has been presented.

References:

AMBROŻY, K. and MEJZA, I. (2004): Split-plot × split-block type three factors designs. In: A. Biggeri, E. Dreassi, C.Lagazio and M. Marchi (Eds): *Proceedings of the 19th International Workshop on Statistical Modelling*, Florence, Italy, 291-295.

GUPTA, S. and MUKERJEE, R. (1989): A calculus for factorial arrangements. Lecture Notes in Statistics 59, Springer-Verlag.

KHATRI, C.G. and RAO, C. R. (1968): Solutions to some functional equations and their applications to characterization of probability distributions. *Sankya, A, 30,167-180.*

LeCLERG, E.L., LEONARD, W.H. and CLARK, A.G. (1962): *Field plot technique.* Burgess, Minneapolis.

RAO, C.R. and MITRA, S.K. (1971): *Generalized inverse of matrices and its applications.* Wiley, New York.

A Model Selection Procedure with FWER-control for the Binomial Change Point Problem

Xuefei Mi

Institute of Biostatistics University of Hannover, Germany;
mi@biostat.uni-hannover.de

Abstract: Extra information from order restriction, such as simple-order, change point and tree-order, is used to improve the power and control the error rate. In this poster we focus on building a test based model selection methodology which uses confidence intervals to control the Family-Wise Error Rate. Several approaches are available for these problems, such as max-t statistics according to Hirotsu and Srivastava (2000) which can be formulated as maximum contrast approach belonging to the broader class of multiple contrast tests (MCT). The disadvantage of MCT is that it can only reject the global null hypotheses. Recently, Robertson (1988), Vuong (1989) and Akritas (2007) developed a test-based log-likelihood method for model selection approaches under certain types of order restriction. These methods achieve better finding rate of the true pattern, but do not control the FWER. In this poster we compare these two methods under change point order restriction. Also we present a modification which can control the FWER. An extension to other complicated order restrictions and higher dimensional contingency table is also possible.

References:

HIROTSU C. AND SRIVASTAVA M. (2000): Simultaneous confidence intervals based on one-sided max t test. Statistics and Probability letters 49, 24-37.

LIAO S., AKRITAS M. (2007): Test-based classification: A linkage between classification and statistical testing. Statistics and Probablity Letters 77, 1269-1281.

ROBERTSON T., WRIGHT F.T. AND DYKSTRA R.L. (1988): Order restricted statistical inference. Wiley, New York.

VUONG Q. (1989): Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses Econometrica, 57, 307-333.

Frequent Itemsets and Association Rules as Supportive Analysis Tools in Genetic Association Studies

Tina Müller and Katja Ickstadt

Technische Universität Dortmund, Fakultät Statistik, 44221 Dortmund, Sonderforschungsbereich 475; tmueller@statistik.uni-dortmund.de

Abstract: Our analyses are based on genetic association studies investigating the relationship between complex diseases and *single nucleotide polymorphisms* (SNPs) as well as epidemiological variables (e.g., drug treatment or BMI). SNP studies have become popular recently as both costs and time of genotyping have decreased dramatically. However, the tools of extracting relevant, useful and hopefully also reproducible results from such studies are yet to be established as the studies usually contain more variables than observations and only small signals.

To face this challenge, we use adaptations of frequent item sets and association rules in several ways to detect risk groups with certain genetic profiles and/or specific combinations of epidemiological parameters and classify the observations as cases or controls. By using frequent item sets we find subgroups of patients who share similar genetic profiles. In each group a separate classification rule can be learned. Another idea is to use the frequent item sets for feature construction. A subsequent classification is then based on binary interaction variables.

Additionally, we mine association rules under the constraint that the consequent of a rule can only consist of the disease status. These rules are employed in a classification framework either by ranking them and using only the best rule for classification or by assessing a vote from all association rules. All approaches are tested on simulated data as well as on real data and compared to CART by their misclassification rates. The presented methods contain great potential for the analysis of SNP data.

References:

AGRAWAL, R., IMIELINSKI, T. and SWAMI, A. (1993): Mining association rules between sets of items in large databases. In: P. Buneman and S. Jajodia (Eds.), *Proceedings of the 1993 ACM SIGMOD*, USA, Columbia, 207-216.

LIU, B., HSU, W. and MA, Y. (1998): Integrating classification and association rule mining. In: *KDD 98*, AAAI Press.

Screening for chlymydia trachomatis in asymptomatic women in hungary. An epidemiological and cost-effectiveness analysis

Tibor Nyári

Dept. of Medical Informatics, University of Szeged, Hungary;
nyari@dmi.u-szeged.hu

Abstract: A multicentre survey was carried out in order to carry out a cost-effectiveness analysis of screening for chlamydial infection in women with asymptomatic genital infections. Independent predictors of chlamydial infection were assessed by using multiple logistic regression analysis.

An incremental cost-effectiveness analysis was performed to compare the strategies of screening with the ELISA method (the sensitivity and specificity of this test are 70% and 99%, respectively) for the detection of C. trachomatis (strategy B), screening with use of the amplified Gen-Probe method (the sensitivity and specificity of this test are 92% and 99%, respectively) for the detection of C. trachomatis (strategy C), and no application of screening methods (strategy A). Costs were based on local charges. Decision analysis was used to assess the potential outcome of chlamydial infection. Sensitivity analysis was performed for outcomes in the model in order to determine how changes in estimated values affected the results and to identify 'break-even' prevalence points.

According to the test, the prevalence of Chlamydia trachomatis among 1300 pregnant women was 4.5%. The group aged under 20 years displayed a very high rate of infection: 12.6%.

Cost-effectiveness analysis, with associated sensitivity analysis was carried out for women aged below 20 years. The Gen-Probe method was best provided the infection prevalence exceeded 16.7%, the PID rate exceeded 24% and the probability of tubal infertility in untreated women exceeded 25%.

According to the infection rate of 12.6%, the most cost-effective strategy was strategy B (screening for C. trachomatis by using the ELISA method). Strategy C (screening for C. trachomatis by using the amplified Gen-Probe method) was slightly less cost-effective than neither testing, nor treating (strategy A).

References:

NYÁRI T., NYÁRI C.S., WOODWARD M., et al. (2001): Screening for Chlamydia trachomatis in Asymptomatic Women in Hungary. An Epidemiological and Cost-Effectiveness Analysis. *Acta Obstet Gynecol Scand, 80 ,300-306.*

Online

Training in Genetic Epidemiology - Implementation of a technology based training course

Friedrich Pahlke[1], Inke R. König[1], Michael Bischoff[2], and Andreas Ziegler[1]

[1] Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck,Germany; ziegler@imbs.uni-luebeck.de
[2] oncampus Fachhochschule, Lübeck, Germany; bischoff@fh-luebeck.de

Abstract: Even though the importance of genetic epidemiology as a scientific field has been widely recognized over the past decades, only very few technology assisted training opportunities have been offered in the last years. Specifically, no comprehensive technology assisted training course with sound didactical justification is available for our field. The goal of our project is to construct a self-learning online training course with a content covering about a five days course, based on the textbook *A Statistical Approach to Genetic Epidemiology from* A. Ziegler and I. R. König. Until now, the content of about a quarter has been implemented as a highly interactive e-learning module. In this presentation, we describe the process of building the raw concept and the storyboard. Also, the implementation of multimedia elements is illustrated, e.g. interactive Flash-based pedigree diagrams and interactive problems with algorithm-based free-text correction. Acknowledging that the e-learning projects at universities are produced under different conditions than in industry, we used a specific procedure tailored for academic use, which allows a high degree of flexibility and emphasizes the didactical concept - instead of the technical implementation. With our course, students and scientists of very different fields of research will get a flexi-time and flexi-location training opportunity in genetic epidemiological methodology and design.

References:

ZIEGLER, A. and KÖNIG, I.R. (2006): A Statistical Approach to Genetic Epidemiology. Concepts and Applications. Wiley-VCH.

Sample Size Determination for Case-Control Studies

Tiberiu Postelnicu

Commission for Biometrics of the Academy, Bucharest, Romania;
tposteln15@yahoo.com

Abstract: One of the most important and challenging issues in epidemiology is the identification of "risk factors", i.e. etiologic agents that increase the risk of a certain disease. Case-control studies represent a typical epidemiologic approach to these problems. Case-control studies consider subject affected by a disease (cases) and units not affected by the same disease (controls). The interest focuses on establishing whether a risk factor increases the chance of being affected by the disease or not.

The association between the etiologic agent and the disease is standardly measured by either the relative risk, or the odds-ratio. Frequentist procedures for choosing the sample size require initial guesses on the true values of the parameters. Hence the resulting criteria are only locally optimal and can be quite sensitive to the values chosen.

In this paper we consider some design problems that arise in the Bayesian inference on the odds-ratio ψ in the analysis of case-control studies. Namely, we consider sample size determination and optimal allocation of units among cases and controls. Unless it is differently specified, the parameter of interest will be $\varphi = \log \psi$. We first consider sample size determination for interval estimation of φ. For this problem we will use the Length Probability Criterion (LPC) proposed by De Santis and Perone Pacifico (2001) as an alternative to the Average Length Criterion (ALC) previously introduced by Joseph, Wolfson and du Berger (1995). For ALC, the idea is to select the smallest sample size n such that the expected length of the posterior interval estimates for the parameter of interest is less than or equal to a fixed threshold. We turn to sample size choice for testing problems on φ. In this case the idea for sample size choice is to consider the smallest number of observations such that the evidence provided by the data in favor of either the null or the alternative is substantial. As a standard measure of evidence we use the posterior probabilities.

References:

ADCOCK, J. (1997). Sample size determination: a review, The Statistician, 46, 261-283.

DE SANTIS, F., PERONE PACIFICO, M. and SAMBUCINI, V. (2004). Optimal predictive sample size for case-control, studies, Journal of the Royal Statistical Society, series C, 53, 3, 427-441.

JOSEPH, L., WOLFSON, D. and DU BERGER, R. (1995). Sample size calculation for binomial proportions via highest posterior density intervals, The Statistician, 44, 143-154.

Environmentally specific assessment of genotypes in different cropping systems

Marcin Przystalski[1], Hans-Peter Piepho[2], and P. Krajewski[1]

[1] Institute of Plant Genetics PAS, Strzeszyńska 34, 60-479, Poznań, Poland; mprz@igr.poznan.pl, pkra@igr.poznan.pl
[2] Department of Bioinformatics, University of Hohenheim, Stuttgart, Germany; piepho@uni-hohenheim.de

Abstract: Mixed models and index selection theory (Falconer and Mackay 1996) are frequently used tools for the analysis of plant experiments. Recently, they have been applied to analyse a large collection of data from trials performed in different environments (sites or years) under organic and non-organic systems in several European countries (Przystalski et al., 2007). The objective of the analysis was to see if the rankings of genotypes differ in the two systems and if organic testing is providing significant information. For this aim, the estimated variance components were interpreted in terms of correlation and selection parameters. The aim of the study reported here was to see how to extend the model to the situation where similar questions are asked, but the answer should take into account the three-way genotype by environment by system interaction. The need for such an extension is dictated by different characteristics of the organic and non-organic trials. We show how the extended covariance structure of the model can be interpreted in terms of the genetic correlation of the genotype by environment interactions between the two systems. In consequence, the selection decisions and ranking comparisons may be applied not to the general, but to the environmentally specific, genotypic performances.

References:

FALCONER D. S. and MACKAY T. F. C. (1996): *Quantitative genetics*. Forth Edition, Prentice Hall, London.

PRZYSTALSKI M., OSMAN A., THIEMT E. M., ROLLAND B., ERICSON L., OSTERGARD H., LEVY L., WOLFE M., BÜCHSE A., PIEPHO H.-P., and KRAJEWSKI P. (2007): Do cereal varieties rank differently in organic and non-organic cropping systems? Submitted.

Bayesian analysis of a dynamical model for the spread of the USUTU virus

Jenő Reiczigel[1,2], Katharina Brugger[2] and Franz Rubel[2]

[1] Szent István University, Faculty of Veterinary Science, Budapest, Hungary; reiczigel.jeno@aotk.szie.hu
[2] Department for Natural Sciences, University for Veterinary Medicine, Wien, Austria

Abstract: The USUTU virus is an arbovirus transmitted by mosquitos and causing disease in blackbirds. The virus was first detected in Austria in 2001, and a major outbreak occurred in 2003. Rubel et al. (2008) developed a 9-compartment SEIR model to explain the spread of disease. The model has several parameters, among which some (those controlling the reproduction and behaviour of mosquitos) are temperature-dependent. Rubel et al. (2008) set the parameters to values taken from the literature, made some model tuning, and checked model fit qualitatively.

In the present paper, we develop a hierarchical Bayes model to enable estimation of parameters from the data. The core of the model is the above mentioned SEIR model. We analyse just the 8 temperature-independent scalar parameters, because the temperature-dependent ones are vector parameters, and there is not enough degrees of freedom in data to analyse them too. Uniform priors are considered for all parameters. Random factors are introduced in the daily temperature data, in the reproduction data of birds, and in the number of birds examined and found to be infected. The analysis is made by MCMC, using a random walk Metropolis scheme with uniform jump distribution with a width of 10% of the range for the uniform prior distribution. We calculate posterior means, medians, and credible intervals for the parameters of interest. The model is implemented in R, combined with a Fortran subroutine computing the original deterministic model.

This research was partly supported by the grant OTKA T049157.

References:

RUBEL, F., BRUGGER, K., HANTEL, M., CHVALA-MANNSBERGER, S., BAKONYI, T., WEISSENBÖCK, H. and NOWOTNY, N. (2008): Explaining USUTU virus dynamics in Austria: Model development and calibration. Accepted for publication in *Preventive Veterinary Medicine*.

Kombinierte Analyse experimenteller und klinischer Genexpressionsdaten - eine Anwendung hochdimensionaler Statistik

Maciej Rosołowski

Institut für Medizinische Informatik, Statistik und Epidemiologie Leipzig;
maciej.rosolowski@imise.uni-leipzig.de

Abstract: In der Genexpressionsanalyse großer Tumordatensätze wird eine biologisch sinnvolle Charakterisierung und Klassifizierung von Patienten angestrebt. In dem Vortrag wird ein Verfahren behandelt, bei dem zunächst anhand klinischer Expressionsdaten geeignete Genmengen gebildet und diese dann in biologisch kontrollierten Versuchen an Zellkulturen auf statistische Relevanz geprüft werden. Das multivariate Variablenselektionsverfahren ist bei sehr hoher Dimension anwendbar und genügt den Anforderungen des multiplen Testens (familywise error rate). Die vorliegenden Ergebnisse zeigen, dass biologische Grundlagenforschung für die Systematisierung unübersichtlicher klinischer Genexpressionsdaten nutzbar gemacht werden kann.

References:

BILD, A. et al. (2006): Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature, 439, 353–357.*

LÄUTER, J. (2007): Hochdimensionale Statistik Anwendung in der Genexpressionsanalyse. Leipzig Bioinformatics Working Paper, No. 15, http://www.izbi.uni-leipzig.de/izbi/working_papers.php

LÄUTER, J., GLIMM, E. und ESZLINGER, M. (2005): Search for relevant sets of variables in a highdimensional setup keeping the familywise error rate. *Statistica Neerlandica, 59, 298–312.*

Zeitveränderliche Variablen im Cox Modell

Mathias Schaller[1] and Andrej Wöhrmann[2]

[1] Institut für Gesundheitsökonomie und klinische Epidemiologie;
mathias.schaller@uk-koeln.de
[2] Kuratorium für Dialyse und Nierentransplantation e.V.

Abstract: Das QiN Programm (Qualität in der Nephrologie) betreut seit 1999 Dialysezentren bei der Qualitätssicherung. Durch die Dauer der Beobachtung ist es dabei auch möglich das Überleben der Patienten zu betrachten. Wie in der Literatur üblich werden die Modelle auf zu Beginn des Beobachtungszeitraumes vorliegenden Einflussparameter aufgebaut. In QiN liegen zusätzlich Beobachtungen der Einflussparameter aus dem Beobachtungszeitraum vor. Es wird untersucht, wie diese Informationen ein besseres Modell ermöglichen.

Der Einfluss verschiedener Parameter auf die Überlebensdauer unter Dialyse mit Hilfe eines Cox-Modelles erfasst. Dabei kann gezeigt werden dass die Annahme des über die Zeit proportionalen Risikos nicht immer erfüllt wird. Mit Hilfe der im Beobachtungszeitraum erfassten Einflussparameter kann das Cox-Modell jedoch erweitert werden, so dass zum Einen die Modellannahmen erfüllt und zum anderen der Güte der Anpassung verbessert wird.

Mit Hilfe des Prentice, William und Peterson Ansatzes wird sowohl der zeitliche Einfluss der Parameterschätzer variiert als auch die Werte der Einflussparameter aus dem Beobachtungszeitraum berücksichtigt. Dabei kann gezeigt werden, dass die Berücksichtigung der Einflussparameter aus dem Beobachtungszeitraum zu einer Einhaltung der Modellannahmen und einer signifikanten Verringerung der Loglikelihood führt.

References:

COX, D.R. (1972): Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society B, 34, 187-220.*

PRENTICE, R.L. and WILLIAMS, B.J. and PETERSON, A.V. (1981): On the regression analysis of multivariate failure time data. *Biometrika, 68, 373-379.*

Assessment and Comparison of Two Questionnaires of Quality of Life in Crohn's Disease Patients

Michaela Šedová[1], Patricia Martinková[1], and Libor Gabalec[2]

[1] Department of Medical Informatics, Institute of Computer Science AS CR, v.v.i., Pod Vodárenskou věží 2, Prague, Czech Republic; sedova@euromise.cz
[2] Department of Internal Medicine, District Hospital Ústí nad Orlicí, ČSA 1076, Ústí nad Orlicí, Czech Republic; libor.gabalec@uo.hospital.cz

Abstract: The use of quality of life (QoL) questionnaires must be based on reliable and valid instruments. In population of 103 Crohn's disease patients we assessed and compared the Czech versions of generic World Health Organization QoL (BREF) and disease-specific Inflammatory Bowel Disease (IBDQ) Questionnaires. The study was focused on IBDQ, since its Czech version has not been validated yet.

In order to assess the validity of the questionnaires, the QoL dimension scores were related to the disease activity measured by Crohn's Disease Activity Index. The internal consistency of the IBDQ was measured by Cronbach's alpha and split-half method. Factor analysis was used to examine the definition of the dimensions. The discriminant ability was assessed by comparing patients with active disease with those in remission. Sensitivity to change and test-retest reliability was evaluated in a subgroup of patients for which the second measurement was available.

Both questionnaires appeared to reflect sufficiently the disease activity as well as to contain other information about some clinical or demographic factors. IBDQ provided more information about the disease activity than BREF. Although the factor analysis revealed slightly different structure of factors, the internal consistency and other requirements on the IBDQ described above were satisfactorily met.

The work was partially supported by the research institutional plan AV0Z10300504 of the ICS AS CR.

References:

FAYERS, P.M. and MACHIN, D. (2000): *Quality of Life - Assessment, Analysis and Interpretation.* Wiley, New York.

Online

A multivariate non-parametric approach to select a biomarker signature: discriminating between two phagocytic syndromes

Martin P. Seybold[1], Knut M. Wittkowski[2], and Marion E. Schneider[1]

[1] Universitätsklinikum Ulm, Sektion Experimentelle Anästhesiologie, Steinhövelstr. 9, D-89075 Ulm, Germany; martin.seybold@gmx.de, marion.schneider@uni-ulm.de

[2] The Rockefeller University, Center for Clinical and Translational Science, 1230 York Ave Box 322, New York, NY 10065, USA; kmw@rockefeller.edu

Abstract: Biomarkers as diagnostic and disease staging tools (http://ospp.od.nih.gov/biomarkers/ClinicalParmacology.pdf).

The fundamental process underlying complex diseases and the functional relation between biomarkers is typically not known. We demonstrate a novel non-parametric approach to differentiate the two life threatening phagocytic syndromes, MAS and HLH. Patients with hemophagocytosis were provisionally designated as MAS if autoimmune characteristics were present and as HLH when Perforin, Munc13-4, Syntaxin-11 or Rab27a were mutated or when signs of autoimmune disease were lacking. Biomarkers used for diagnostics were: Culture and function of phagocytes; Leukocyte surface marker analysis by flow cytometry; quantification of plasma cytokines, and soluble receptors. Biomarker expression profiles were partially ordered by u-scores for multivariate data (http://muStat.rockefeller.edu). The scores were compared using Mann-Whitney type u-tests (http://cran.r-project.org/doc/packages/muStat.pdf). From a large panel of biomarkers, a signature consisting of plasma sCD25, IL-1AY (inverse polarity), cellborne CD2+/CD86+ adaptive Treg, and CD25+ lymphocytes (inverse polarity) discriminates best and appears to substitute for laborious diagnostics. These markers can ascertain diagnosis and follow successful treatment response by rapid chemiluminescence based ELISA and flow cytometry within 3 hours, which can be life-saving in this disease. Outliers may indicate novel disease subgroups.

References:

WITTKOWSKI, K. M., LEE, E., et al. (2004): Combining several ordinal measures in clinical studies. *Statistics in Medicine, 23, 1579–1592.*

Analysis of student responses to BDI-II through CFA and IRT

Barbara Oliveiros[1], Cristina Macedo[2], Emanuel Ponciano[1], and Alexandre Gomes da Silva[1,3]

[1] IBILI, Faculdade de Medicina, Universidade de Coimbra
[2] Instituto Piaget de Mirandela
[3] ISCAC, Instituto Politécnico de Coimbra, Quinta Agricola, Bencanta, 3040-316 Coimbra; asilva@iscac.pt

Abstract: BDI-II studies like other studies involving Likert-type surveys are commonly analysed compiling data and reporting means and standard deviations. More recently confirmatory factor analysis (CFA) is also used to explore this kind of data. This study presents the analysis of BDI –II type data using CFA and item response theory. The data consists of 852 Students of Piaget Institute of Mirandela which were evaluated with BDI-II scale of 21 (Likert-type) items. Individual performances and items consistency were analysed and an evaluation of procedures is made.

References:

BAKER, F.B. (2001). *The Basics of Item Response Theory.* 2ond edition. ERIC Clearinghouse on Assessment and Evaluation.

BECK, A.T., STEER, R.A. & BROWN, G. K. (1996). *Beck Depression Inventory – Second Edition* – Manual. San Antonio: Psychological Corp.

BECK, A.T., WARD, C.H., MENDELSON, M., MOCK, J., & ERBAUGH, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4, 53–63.*

BRADLEY, K. & BRADLEY, J. (2004). *Challenging the Validity of Higher Education Course Evaluations.* On-line publication.

BYRNE, B.M. (2001). *Structural equation modelling with AMOS, basic concepts, applications and programming.* Lawrence Erlbaum Associates, Publishers, London.

HAYS, R.D., MORALES, L.S. & STEVE, S.P. (2000). Item Response Theory and Health Outcomes Measurement in the 21st Century. *Med Care. September; 38(9 Suppl): II28–II42.*

RAMSAY, J.O. (2000). *TestGraf A Program for the Graphical Analysis of Multiple Choice Test and Questionnaire Data.* McGill University.

Additivity tests for the mixed model in the two way ANOVA with single sub-class numbers – Type-I-risk of several tests

Marie Šimečková[1] and Dieter Rasch[2]

[1] Institute of Animal Science Prague - Uhříněves; simeckova.marie@vuzv.cz
[2] University of Natural Resources and Applied Life Sciences, Vienna; dieter.rasch@boku.ac.at

Abstract: Testing for the interaction in the two-way ANOVA with one fixed and one random factor without replication is needed in many applied studies. A number of tests were developed for testing the hypothesis of no interaction in case of ANOVA models with both effects fixed. The aim of this contribution is to verify using these tests for mixed models.

Five tests are concerned on the level 5 %: Tukey's, Mandel's, Johnson - Graybill's, locally best invariant (LBI) and Tusell's test. Simulation was performed to examine whether the type-I-risk remains on 5 % level even for mixed ANOVA models. The number of levels of the fixed factor was chosen between 3 and 10, of the random factor between 4 and 50, the variance of the random factor equals 2, 5 or 10, the variance of the random error equals 1.

It was concluded that for 5 % type-I-risk all these tests hold the level of type-I-risk and therefore the tests developed for fixed models can be used for the mixed models as well.

References:

BOIK, R.J. (1993): Testing additivity in two-way classifications with no replications: the locally best invariant test. *Journal of Applied Statistics, 20, 41–55.*

JOHNSON, D.E. and GRAYBILL, F.A. (1972): An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc., 67, 862–868.*

MANDEL, J. (1961): Non-additivity in two-way analysis of variance. *J. Amer. Statist. Assoc., 56, 878–888.*

TUSELL, J.W. (1990): Testing for interaction in two-way ANOVA tables with no replication. *Computational Statistics and Data Analysis, 10, 29–45.*

Some simulation studies under a threshold animal model with single gene effect

Ewa Skotarczak[1], Tomasz Szwaczkowski[2], Krzysztof Moliński[1], and Anita Dobek[1]

[1] Department of Mathematical and Statistical Methods, August Cieszkowski Agricultural University of Poznań, Wojska Polskiego 28, 60-637 Poznań, Poland; efalsa@au.poznan.pl

[2] Department of Genetics and Animal Breeding, August Cieszkowski Agricultural University of Poznań, Wołyńska 33, 60-637 Poland

Abstract: In recent years the threshold animal model has been successfully used to estimate genetic parameters for many reproductive traits. The presented studies were aimed at checking the ability of threshold animal model to detect effects of single genes on the basis of phenotypic, categorical data only. Such analysis could be made as the first stage of studies, before the molecular research, as it does not demand much costs. Three data sets were simulated and analyzed via the threshold animal model with one single gene effect. Apart from the single gene effect the model included: vector of fixed effects, vector of random additive polygenic effects, vector of random permanent environmental effects and vector of random errors. In the first set of simulated data the binary observations were generated, one observation (0 or 1) for every recorded individual (for instance, data for fertility of mammal species are collected in this way). In the second data set ten binary observations were simulated for each recorded animal (hatchability for poultry can be here an example). In the third case the simulated observations were distributed into three categories (as calving difficulty for instance). Bayesian methods with Gibbs sampling algorithm were used in statistical analysis of the simulated data sets under the proposed threshold model. The significance of the single gene effect was checked on the basis of the highest posterior density regions. The mixing properties of Gibbs sampling and the accuracy of obtained estimators were compared. The main conclusion is that the implemented method is rather conservative, i.e. the major gene effect is usually underestimated.

Acknowledgment: Paper supported by Ministry of Education and Science grant no. 2 P06D 049 28

Comparing of different regression models regarding the inflammatory marker interleukin 8 in bitumen-exposed workers

Anne Spickenheuer[1], Benjamin Kendzia[1], Monika Raulf-Heimsoth[1], Katja Ickstadt[2], Thomas Bruening[1], and Beate Pesch[1]

[1] BGFA - Research Institute of Occupational Medicine of the German Social Accident Insurance, Buerkle-de-la-Camp-Platz 1, 44789 Bochum; spickenheuer@bgfa.de, kendzia@bgfa.de, Raulf-Heimsoth@bgfa.de, Bruening@bgfa.de, Pesch@bgfa.de

[2] Technische Universitaet Dortmund, Fakultaet Statistik, 44221 Dortmund; ickstadt@statistik.uni-dortmund.de

Abstract: We used data of a cross-shift study in 280 bitumen-exposed workers and 74 construction workers as referents to apply various regression models for the assessment of potential irritative effects of fumes of bitumen. Exposure to fumes of bitumen during shift was measured by personal air monitoring. To assess inflammatory processes in the lower airways we analyzed induced sputum pre- and post-shift for cellular and humoral composition. Interleukin (IL)-8, a cytokine with a high chemotactic potency was selected for modelling of irritative effects. A questionnaire was used to record potential confounders like age, smoking status, and nationality. Specific IgE against aeroallergens using the sx1-screening tool (Phadia, Sweden) was determined for atopy status. Different regression models were applied, comprising Least Squares Linear Regression, robust regression models like Least Median of Squares, Least Trimmed Squares, Reweighted Least Squares, and a Mixed Linear Model. Potential predictors of IL-8 levels were fumes of bitumen during shift and exposure group. All models were adjusted for age, smoking status, nationality, and atopy status. Median shift concentration of fumes of bitumen was 3.4 mg/m^3 in exposed workers. Median pre-shift level of IL-8 was 4839 pg/ml (post-shift 3512 pg/ml) in exposed workers compared to 2294 pg/ml (post-shift 1207 pg/ml) in referents. In all models exposure group and smoking had a significant influence on IL-8, whereas the concentration of fumes of bitumen had no significant effect. Current smokers and exposed workers had higher IL-8 than non-smokers and referents. The analyses revealed a sub-chronic irritative effect. The different regression models showed similar results, which supports the findings. IL-8 was taken by way of example for other parameters in this study to proof the performance of different regression models.

Online

A SAS Macro for the Calculation of Relative Risks from Non-Parametric Logistic Regression Models using B-Splines

Martin Gregory[1], Karl-Peter Pfeiffer[2], Hanno Ulmer[2], and Alexander M. Strasak[2]

[1] SAS Institute Inc., Heidelberg, Germany; martin.gregory@yahoo.co.uk
[2] Department of Medical Statistics, Informatics and Health Economics, Innsbruck Medical University, Austria; karl-peter.pfeiffer@i-med.ac.at, hanno.ulmer@i-med.ac.at, alexander.strasak@i-med.ac.at

Abstract: Due to a lack of alternative procedures, available in statistical standard software packages, current epidemiologic research, estimating risk ratios for the effects of independent exposure variables in studies of disease aetiology, is largely based upon the practice of categorizing continuous risk factors, in order to handle possible non-linearity in dose-response. However, it was long reported that this approach is associated with several problems, most notably, jumps in risks at category cut points not being biologically plausible and violating the assumption of actual risks varying smoothly with data. Moreover, it has been shown that these approaches are typically associated with a considerable loss of power and possible cut-point bias, introduced by categorization. We provide a SAS-based macro, performing non-parametric logistic regression analyses using a b-spline expansion of an independent variable under consideration and calculating relative risks with respective confidence intervals for each predicted value with respect to a supplied reference value. The macro allows specification of a number of parameters influencing how the spline expansion is carried out, producing listings and graphs in addition to the generated output datasets. The calculations are based on an algorithm described in Cao et al. (2006) and implemented by them as an S-Plus function. The relative risk for the predictor x with respect to a reference value x_{ref} is calculated as

$$rr(x, x_{ref}) = exp\left(\sum_{i=1}^{n} \beta_i [s_i(x) - s_i(x_{ref})]\right)$$

where n is the number of degrees of freedom of the spline expansion, β_i is the coefficient of the ith spline basis function estimated by the logistic regression and $s_i(x)$ is the value of the ith spline basis function at x. Beside estimating relative risks with respective confidence intervals, our macro allows to graphically determine the shape of the association between a given, continuous exposure and the risk of a binary outcome, retaining the independent variable under consideration in its initial, continuous form, concurrently adjusting for multiple confounding factors. As the SAS macro is based upon non-parametric spline regression, it avoids common problems associated with non-linearity, typically present in conventional parametric regression approaches.

References:

CAO J. and VALOIS M.F. and GOLDBERG M.S. (2006): An S-Plus function to calculate relative risks and adjusted means for regression models using natural splines. *Comput Methods Programs Biomed, 84, 58-62.*

Application of nonparametric smoothing technique to the assessment of acute health effects of air pollution in Kraków, Poland

Krystyna Szafraniec[1] and Bogdan Wojtyniak[2]

[1] Jagiellonian University Medical College, Institute of Public Health, Kraków, Poland; mygomola@cyf-kr.edu.pl
[2] National Institute of Hygiene, Department of Medical Statistics, Warsaw, Poland

Abstract: We apply a nonparametric smoothing technique to model data from the field of environmental epidemiology. Because environmental exposure is involuntary in nature, it affects all studied individuals in similar manner. So, while exposure is common, the risk tends to be low and consequently estimates of risk obtained by statistical procedures have the potentials to be biased easily.

Acute health effects of air pollution are studied by means of time-series approach. Daily counts of health outcome are regressed against daily concentration of air pollutant in the presence of strong confounding factors. In time-series studies particular attention is given to factors that vary on similar timescale as exposure and health outcome. They are weather variables, which influence both health and exposure, and some unmeasured variables like seasonality and long-term trends which are included in health outcome and air pollution data series. One of the statistical methods to account this problem is to employ generaliced additive models (GAM) of the form

$$Y_t \sim \text{Poisson}(\mu_t), \quad \ln \mu_t = \ln E(Y_t) = \sum\nolimits_{j=1}^{q} f_j(X_{tj}, \lambda_j) + \sum\nolimits_{i=1}^{p} \beta_i X_i,$$

where Y_t denotes the daily number of health outcome, β is the log relative rate of health outcome associated with increase in exposure, f_j is a nonparametric smooth function of a time-varying variable X_{tj} with λ_j df. The goal is to estimate β, the association between air pollution X and health outcome Y, in the presence of time-varying confounding factors. The difficulty here is to identify proper nonparametric smoothing function f_j and then to determine the degree of smoothness of f_j that maximally reduces the confounding bias in the estimates of β.

We applied this methodology to 9-year series of mortality from cardiovascular diseases and PM10 exposure in Kraków population. As nonparametric smoothing functions we choose smoothing splines. The appropriate number of degrees of freedom we evaluated on the basis of those that minimised Akaike's information criterion (AIC). Once the best-fitted model had been selected we inspected for residual autocorrelation using partial autocorrelation function PACF. Details of the model choice and the estimates of the percent change in daily number of deaths from cardiovascular diseases related to 10 μg/m^3 increase in PM10 concentration will be presented.

The study has been supported by the KBN grant 2P05D 020 28.

A Generic HMM Implementation for the Inference of Latent Chromosomal States

Helene H. Thygesen[1], Renee X. de Menezes[2], and Judith M. Boer[2]

[1] Dpt. of Appl. Stat., Lancaster University, UK; h.thygesen@lancaster.ac.uk
[2] Dpt. of Human Genetics, Leiden University Medical Center, The Netherlands

Abstract: Genomics researchers use DNA microarray technology to estimate chromosomal states. Some examples:

Chromosomal state	Type of microarray data
Copy number, Loss Of Heterozygosity	SNP arrays or array-CGH
Chromatin state	ChIP-on-chip, methylation arrays
Haplotype	SNP-array data from relatives

It is often a natural assumption that those (latent) states are discrete, and that they are constant within chromosomal regions. That assumption leads to a Hidden Markov Model (HMM). HMM is supported by software packages for specific types of states using data from specific platforms. What we present is a generic HMM implementation that allows users to infer latent states from data from any platform by plugging in their own platform-specific model for the measurement process. The implementation is available as an R Script.

Identifying Molecular Predictive Marker in Translational Clinical Trials

Wiebke Werft and Axel Benner

German Cancer Research Center, INF 280, 69120 Heidelberg; w.werft@dkz.de

Abstract: Recently, the analysis of high dimensional molecular data gained major attention due to the broad application of microarray experiments in biomedical research. In translational clinical trials one is interested in the usage of microarrays to identify genes that are differentially expressed in cells taken from patients' cancer tissue and one aims to analyse their ability to predict patients' response to a specific treatment. The statistical analysis of translational trials may use gene-wise logistic regression for the identification of predictive factors. This allows the examination of a dichotomous clinical endpoint e.g. response to treatment together with 'gene expression values at baseline' as well as the interaction between treatment and expression. Since one has to assume that the genes analysed are highly dependent we consider a resampling-based multiple testing procedure as introduced by Pollard et al. (2005) which incorporates the joint dependence structure between the test statistics and in consequence also between the genes. We extended this resampling-based multiple testing procedure to the case of logistic regression of a dichotomous clinical endpoint and test for interaction between gene expression and treatment. In a simulation study we analyse the choice of parameters to generate the test statistics null distribution of this respampling-approach and its impact on the control of the false discovery rate.

References:

POLLARD, K.S. et al. (2005): Multiple Testing Procedures: R multtest Package and Applications to Genomics. In: Gentleman, R. et al. (Eds.): *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, Berlin, 251–272.

Statistical modelling of within-litter variance in pigs

Dörte Wittenburg, Volker Guiard, Friedrich Teuscher, and Norbert Reinsch

Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere (FBN), Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany;
wittenburg@fbn-dummerstorf.de

Abstract: The study of sow reproduction traits is important in livestock science and production. One aim of present research is to reduce the variability of birth weight within litter and to increase piglet survival under the aspect of keeping litter size high. This paper studies genetic effects on the variability of birth weight within litter in Landrace pigs. In several studies variation in birth weight was described by the sample standard deviation of birth weights within one litter. In the present work we additionally considered the sex effect on piglet birth weight and on its variability. Thus, the sample variance of birth weights per litter and gender was assigned as the trait of sow. Main focus was set on model choice. Different transformations of the sample variance were fitted by linear mixed models. The sample variance itself was described by a generalized linear mixed model. Appropriate weights were suggested which included individual litter sizes. The presented models were compared via residual diagnostics. Furthermore, we analysed the influence of stillborn piglets on the estimates of genetic parameters. The heritability of variability within litter has been investigated based on data received from the experimental pig unit (EAS) in Dummerstorf consisting of 2211 litters of German Landrace. A second emphasis of this study was set on testing random effects. The distribution of the likelihood ratio test statistic is not exactly given under the null hypothesis in the context of animal models due to the relationship structure between individuals. Thus, we obtained an approximation via parametric bootstrapping. On the other hand and to reduce the number of bootstrap simulations the choice of an adequate distribution function led to a suitable distributional approximation of the test statistic and its threshold value for the null hypothesis testing problem.

Analyse von Strukturblöcken mit Nachbarschaftseffekten

Joanna Zyprych and Idzi Siatkowski

Der Lehrstuhl der Mathematischen und Statistischen Methoden, August Cieszkowski Landwirdschaftliche Universität in Poznań, Wojska Polskiego 28, 60-637 Poznań, Poland; zjoanna@au.poznan.pl

Abstract: Diese Arbeit ist eine Einleitung zur Analyse der optimalen Blockstrukturen mit Nachbarschaftseffekten. Kunert und Martin (2000) haben die Methode für die Bildung der optimalen Blockstrukturen in dem Modell mit linken und rechten Nachbarschaftseffekten beschrieben. Am Anfang stellen wir die Hauptform des Mitwirkungsmodels mit Nachbarschaftseffekten vor und wir zeigen ein Beispiel zu dieser Theorie. Aus diesem Vorbild lässt sich eine Informationsmatrix ableiten. Aus der Veröffentlichung von Kunert und Martin (2000), schätzen wir im Anschluss den universal optimalen Plan ab. Diese Schätzung führt zur Bestimmung der oberen Schranke der Spur der Informationsmatrix. Die Theorie der universal optimalen Plänen basiert auf der Kiefer Theorie.

Am Ende stellen wir einige Beispiele unter den in dieser Arbeit gegebenen Voraussetzungen vor, um zu zeigen, wie man das bestimmte Experiment für gegebene Einheiten-, Objekten- und Blöckenzahl planen kann. Als Beilage zu dieser Arbeit wurde die Matlabrechnungen und das Programm in Pascal eingefügt. Diese berechnen die Matrizen und ihre Umgestaltungen, um damit am Ende für entsprechende Sequenzen bestimmte Strukturmatrizen zu schaffen.

References:

KIEFER, J. (1975): Construction and optimality of generalized Youden designs. In: J. N. Srivastava (Ed.): *A Survey of Statistical Design and Linear Models.* North-Holland Elsevier, New York, 333–353.

KUNERT, J. (1983): Optimal design and refinement of the linear model with applications to repeated measurements designs. *Ann. Statist. 11, 247–257.*

KUNERT, J. and MARTIN, R.J. (2000): On the determination of optimal designs for an interference model. *Ann. Statist. 28, 1728–1742.*

List of Contributors

Online

Dear Friends, dear Colleagues,

[...] The organizers chose “Statistics and Life Sciences” as a motto for this first conference of the Central European Network, since molecular life sciences, medicine and epidemiology are the framing, trendsetting subjects, which cannot be appropriately compiled or scientifically interpreted without statistical methodology.

For today‘s statisticians new contexts immerge from cooperation with the life scientists and stimulated by this dialog, changes in thinking in statistical methods and development takes place. [...]