Presentations

211: Meth Börse

Time: Tuesday, 19th Mar 2013: 8:00am - 8:30am · Location: KG I, HS 1199

211: 1

Flexible Methoden bei der Planung und Durchführung von klinischen Studien

Sandra Ligges

Institut für Biometrie und Klinische Forschung (IBKF), Germany

tba

221: Statistics in Clinical and Preclinical Research -- Sample Size Calculations

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG II, Audimax
Session Chair: Tina Müller

221: 1

Fallzahlplanung für prospektive Studien zum Vergleich von Kosteneffektivitäten

Frank Krummenauer, Jessica Hirsch

University of Witten/Herdecke, Germany

Einleitung: Die Evaluation komplexer Versorgungskonzepte erfolgt meist simultan aus klinischer und ökonomischer Perspektive über deren Kosteneffektivitäten; die Planung vergleichender Evaluationen basiert dann auf deren inkrementeller Kosteneffektivität. Nachteil der „confidence box“-Methode nach Briggs & Fenn ist jedoch deren konservative Schätzung der Fallzahl durch separate Betrachtung der ökonomischen und klinischen Versorgungs-Unterschiede.

Methoden: Über den inkrementellen Nettonutzen wird ein Ansatz zur Fallzahlplanung vorgestellt, der die Berücksichtigung der Korrelation zwischen finanzieller Investition in die jeweilige Versorgung und resultierendem Patienten-seitigen Ergebnis gestattet. Der Ansatz wird illustriert an der Planung einer Klinischen Studie zur Einführung eines Klinischen Behandlungspfades in die Knieendoprothetik: Aus Patientenperspektive soll vor und nach Einführung des Pfades das funktionelle Ergebnis durch die dreimonatige Änderung des WOMAC-Index [%] beschrieben werden, aus Sicht der Klinik die investierte (Prozess-) Kostensumme [€].

Ergebnisse: Der Ansatz bedingt gegenüber dem nach Briggs & Fenn durchweg eine Reduktion der notwendigen Patientenzahl, für realistische Korrelationen von 0.2 bis 0.5 zwischen Kosten und Nutzen um bis zu 20%. Für die Planung der Studie zur Einführung eines Klinischen Pfades zeigte sich konkret eine Korrelation zwischen Leistungserbringer-seitig investierten (Prozess-) Kosten und Patienten-seitig attestiertem funktionellen Ergebnis von 0.27, entsprechend einer Reduktion der kalkulierten Fallzahl um 9 % von 282 auf 256 Patienten.

Diskussion: Der Ansatz zur Planung von Kosteneffektivitätsvergleichen komplexer Versorgungskonzeptze bedingt gegenüber dem nach Briggs & Fenn zum Teil merkliche Reduktionen der notwendigen Patientenzahl, erfordert aber Informationen zur Korrelation zwischen Kosten und Nutzen.

221: 2

Fallzahlplanung für Crossover Designs basierend auf vorgegebener Präzision

Benjamin Lang, Frank Fleischer

Boehringer Ingelheim Pharma GmbH & Co KG, Germany

In frühen Phasen basiert die Fallzahl von Crossover oft nicht auf formalen Power-Berechnungen, sondern auf einer vorspezifizierten Präzision des resultierenden Konfidenzintervalls. Derartige Berechnungen der Stichprobengröße werden in vielen Fällen mittels Software realisiert, die lediglich in der Lage ist ein paralleles oder verbundenes Stichprobendesign zu berücksichtigen. Für den Fall, dass ein 2x2x2 oder höheres Crossover durchgeführt wird, kann die tatsächlich benötigte Fallzahl demnach nur durch die des verbundenen Stichprobendesigns approximiert werden. Alle Studien mit intraindividuellem Vergleich müssen sich somit auf die Fallzahl des verbundenden Stichprobendesigns stützen und das tatsächliche Studiendesign bei der Fallzahlplanung vernachlässigen. Wir haben die Auswahl von möglichen Studiendesigns erweitert (ähnlich der Umsetzung im R-Paket PowerTOST), was uns erlaubt verschiedene Crossover-Designs auszuwählen und in die Berechnung der tatsächlich benötigten Fallzahl einfließen zu lassen. Wir haben resultierende Fallzahlen von verschiedenen Crossover-Designs mit der des entsprechenden verbundenden Stichprobendesigns verglichen. Unsere Ergebnisse zeigen, dass obwohl die Berechnungen für das tatsächliche Crossover-Design präziser ist, die Wahl der Stichprobengröße des verbundenden Designs in der Regel eine konservative Wahl darstellt. Hierbei haben wir zusätzlich den Fehler quantifiziert, der bei der Wahl des verbundenen Designs anstelle des tatsächlichen zur Fallzahlberechnung entsteht.

221: 3

Sample size determination and re-estimation for matched pair designs with multiple binary endpoints

Jin Xu

East China Normal University, People's Republic of China

Motivated by a recent symptom management trial to simultaneously assess multiple binary endpoints for cancer chemotherapy, we extend the univariate McNemar's test to multivariate cases for doubly blinded clinical trials with matched pairs. We propose a general method to test non-inferiority or quivalence. The method employs the intersection-union principle on the marginal score statistics to obtain an asymptotic alpha-level test. Power formula and sample size calculation are provided by a simple numerical method that accounts for the correlation structure among the endpoints. We further consider sample size re-estimation through internal pilot study. To avoid the need of unblinding for doubly blinded trials, we also propose a blinded approach for nuisance parameter estimation. The effectiveness of the proposed methods is demonstrated by simulation studies. Application to the cancer chemotherapy trial is illustrated.

221: 4

An exact approach to sample-size planning of two-arm noninferiority trials with binary outcome and the raw difference of proportions as the parameter of interest

Stefan Wellek

Department of Biostatistics, CIMH Mannheim/University of Heidelberg, Germany

Notwithstanding the serious objections which can be raised from the viewpoint of statistical theory to measuring the dissimilarity of two binomial distributions in terms of the raw difference of proportions, this parametrization is still fairly popular both among regulators and clinical trialists. There is broad consensus that the best way of addressing the corresponding hypotheses testing problem is to apply a likelihood-based asymptotic procedure and enforcing maintenance of the significance level in finite samples by means of an iterative algorithm for determining the largest admissible nominal level. In the first part of the talk results are presented which suggest that among a variety of variants of this approach the one based on the score statistic commonly named after Farrington & Manning (Stat Med, 1990) is the most satisfactory. In the second part, an algorithm is described for performing exact power and sample-size calculations with the level-corrected score test. The exact sample sizes are compared with the sample sizes calculated by means of standard software being available for the planning of noninferiority trials. Although the values obtained through applying the approximation used there are fairly close to the exact sample sizes calculated by means of our approach, using them in connexion with the testing procedure for which they are provided would be grossly misleading. The facts behind this statement are presented in the concluding part of the talk.

222: Observational Studies - Study Designs and Methods for Confounder Control and Modelling in Observational Studies

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1199
Session Chair: Sigrid Behr

222: 1

Assessment of the case-chaos analysis as an alternative to the case-control design

Sam Doerken, Peggy Sekula, Martin Schumacher

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Recently, the case-chaos analysis design for risk estimation was proposed [1]. In contrast to the cohort or the case-control design, it is based on data obtained from cases only. Case-chaos controls are artificially generated by random permutation of the exposure information in terms of several binary factors for every single case. In analogy to matched controls of a case-control study, these case-chaos controls are thus available for risk estimation. Obviously, such an approach saves effort, time and costs in comparison to a case-control study. The aim of this investigation is to assess this method in more detail, both theoretically and practically; the latter being done in the context of the EuroSCAR-study on severe cutaneous adverse reactions (SCAR).

Using case information obtained in the EuroSCAR case-control study, case-chaos controls were repeatedly generated in different ratios (1:1,1:3,1:10). For risk estimation, conditional logistic regression was applied. Estimated odds ratios (OR) show a strong positive correlation to observed exposure prevalences in cases. Theoretical considerations confirm this observation and show that the case-chaos control approach reflects nothing else than this prevalence that would also be used in an informal analysis of the cases only.

Although the case-chaos design seems to be appealing at a first glance, it is applicable only under certain strong assumptions. In general, it does not provide additional information and is definitely not suitable in the context of SCAR.

References:

[1] Gillespie IA, et al. The “Case-Chaos Study” as an Adjunct or Alternative to Conventional Case-Control Study Methodology. Am J Epidemiol 2012; 176(6):497-505.

222: 2

Bewertung der Balancierung in Propensity-Score-gematchten Behandlungsgruppen

Lena Kemper, Dirk Enders, Walter Schill

BIPS - Institute for Epidemiology and Prevention Research, Germany

In Beobachtungsstudien erfolgt die Behandlungszuweisung nicht randomisiert. Eine Möglichkeit, trotzdem eine Balancierung der Kovariablen in den Behandlungsgruppen zu erzielen, ist das Matchen nach dem Propensity Score. Als Gütemaß wird oft die c-Statistik verwendet, die sich jedoch nicht zur Überprüfung der Kovariablenbalanciertheit eignet. Alternative Beurteilungskriterien werden benötigt.

Am Beispiel einer Studie, die das Herzinfarktrisiko von Typ-2-Diabetikern in zwei Insulinbehandlungsgruppen untersucht, wurden Methoden zur Beurteilung der Kovariablenbalanciertheit für vier verschiedene Propensity-Score-Modelle verglichen. In Anlehnung an Austin (2009) wurde für jede Kovariable die standardisierte Differenz berechnet und für die stetige Variable Alter wurden zusätzlich das Varianz-Ratio ermittelt sowie grafische Methoden (Darstellung der empirischen Verteilungsfunktionen, Quantil-Quantil-Plot, Boxplot) angewendet. Als zusammenfassendes Maß wurde der Durchschnitt der standardisierten Differenzen der einzelnen Kovariablen betrachtet. Dieses Maß wurde weiterentwickelt, indem die Stärke des Einflusses der einzelnen Kovariablen auf das Outcome in Form von Gewichten berücksichtigt wurde. Die Gewichte wurden aus der ungematchten Stichprobe ermittelt, wobei eine Kovariable ein umso höheres Gewicht erhielt, je mehr sie zur Erklärung des Outcomes beitrug. Als Maß dafür wurde die auf der Likelihoodfunktion [fn]L[/fn] basierende Anpassungsstatistik [fn]-2 logL[/fn] verwendet.

Für den Vergleich der Propensity-Score-Modelle ergaben sich konsistente Ergebnisse für alle Balancierungsmaße. Im Gegensatz zu Maßen für einzelne Kovariablen lässt sich mit zusammenfassenden Maßen eine Rangfolge der Balancierungsgüte zwischen verschiedenen Modellen aufstellen. Da die Reduktion auf eine Kenngröße jedoch mit einem Informationsverlust einhergeht, haben zusammenfassende grafische Darstellungen, z.B. der standardisierten Differenz, Vorteile gegenüber den zusammenfassenden Maßen.

References:

[1] PC Austin (2009). Statistics in Medicine 28, 3083-3107.

222: 3

Dose-response modelling for bivariate covariables with and without a ‘spike at zero’: Theoretical results and applications in epidemiology

Eva Lorenz[1], Carolin Jenkner[2], Willi Sauerbrei[2], Heiko Becher[1]

[1]Institute of Public Health, Medical Faculty, University of Heidelberg, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In observational studies a common goal is to estimate the effect of covariables on an outcome quantitatively by using appropriate regression models. Such covariables often have individuals with a value of zero while the distribution of exposed is continuous (variables with a spike at zero). For both cases with and without a spike at zero, we have investigated the theoretical shape of the dose-response curve under various distributional assumptions in the univariate case (Becher et al., 2012). A procedure to fit a dose-response relation for a given dataset where the distribution is unknown based on an extended fractional polynomial procedure has also been developed.

In practice, however several covariables have to be considered simultaneously for confounder adjustment. Therefore, further investigations for the multivariate case are required. While it is infeasible to investigate the multivariable case theoretically in a general way, we consider the bivariate normal and log normal case in the framework of a logistic regression model. We derive the theoretical bivariate dose-response curve. In the bivariate normal and log normal case without spike, it is straightforward to develop conditions for the occurrence of confounding and interaction based on the expectation and covariance matrix in diseased and non-diseased. For the spike at zero situation additional assumptions must be made on the covariance matrix and expectation of [fn]X = (X_{1j},X_{2j})[/fn], where [fn]X_{ij}[/fn] can take values zero with positive probability. Theoretically obtained regression coefficients will be presented and compared to empirical estimates from a laryngeal cancer study.

References:

Becher et al. (2012). Biom.Journal. 54(5)686-700

222: 4

Modeling continuous predictors with a 'spike' at zero: multivariable approaches

Carolin Jenkner[1], Eva Lorenz[2], Heiko Becher[2], Willi Sauerbrei[1]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Institute of Public Health, Medical Faculty, University of Heidelberg, Germany

In epidemiology and clinical research, predictors often consist of an amount of individuals with a value of zero while the distribution of the remaining ones is continuous (variables with a spike at zero). Examples in epidemiology are smoking or alcohol. For one spike variable, an extension of the fractional polynomial (FP) procedure was proposed to deal with such situations (Royston et. al. 2010) and recently slightly modified (Becher et al 2012). To indicate whether or not a value is zero, a binary variable is added to the model. In a two-stage procedure, it is assessed whether the binary variable and/or the continuous FP function for the positive part is required (FP-spike). This procedure can also be used if adjustment for other variables is required. Therefore it can easily be incorporated in the multivariable fractional polynomial (MFP) procedure.

If more than one spike variable is present, several approaches are possible. Methods of handling them are strongly dependent upon the bivariate or multivariable distribution of the zero and non-zero values. In the easiest case, one can assume independence of the spike variables and use the univariate approach . Otherwise, combinations of dummy variables are considered which vary the influence of observations that are positive in one variable and zero in the other one. Within this talk, different statistical and distributional issues are analyzed and illustrated through three distinct datasets.

References:

Royston et al. Stat. Med. 2010; 29: 1219-27.

Becher et al. Biom. Journal 2012;54: 4. 686-700.

223: Applied Econometrics

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1098
Session Chair: Axel Werwatz
Session Chair: Winfried Pohlmeier

223: 1

Occurrence dependence and zero-inflation in count data models

Rainer Winkelmann

Universität Zürich

Excess zeros are encountered in many empirical count data applications. We provide a new explanation of extra zeros, related to the underlying stochastic process that generates events. The process has two rates, a lower rate until the first event, and a higher one thereafter. We derive the corresponding distribution of the number of events during a fixed period and extend it to account for observed and unobserved heterogeneity. An analysis of the effect of a health care reform on the individual number of doctor visits in Germany illustrates the usefulness of the new approach.

223: 2

When a Random Sample is Not Random. Identification of the Effects of Migration on Remaining Household Members Using Survey Data

Andreas Steinmayr

University of St. Gallen, Switzerland

We investigate the effects of migration of an adult household member on school attendance of children who stay behind in rural Mexico. This paper is the first to address the double-selection problem using observational data. First, households select into migration. Second, in some households only one or a subset of household members migrate, while in other households all members migrate. In the latter case, the household will not be included in cross-sectional survey data at all. The resulting sample-selection problem has largely been ignored in the literature This paper uses principal stratification to model the behavior of household members and to identify bounds for the effects of migration on school attendance of children who stay behind under a transparent set of assumptions. The results suggest that the effect of adult migration on school attendance of boys is negative and between -0.2 and -0.02 while the direction of the effect is ambiguous for girls.

223: 3

Quantile regression for the estimation of technical efficiencies

Hauke Rennies[1], Elisabeth Waldmann[2], Thomas Kneib[2]

[1]Universität Oldenburg, Germany; [2]Universität Göttingen, Germany

Technical efficiencies calculated as ratios of achieved output and maximal possible output given a certain input are an important measure for companies to benchmark their performance in comparison with other competitors. While key perfomance indicators are more or less limited in their use for benchmarking, the estimation of technical efficiencies implies the estimation of production frontier functions of whole industries. This contribution describes corrected least squares fitting and stochastic frontier analysis as simple methods for the estimation. By applying quantile regression as a generalization of median regression, it is possible to completely determine the conditional distribution of the output given a fixed input for an industry, whereas the previous methods focus on for example adjusted conditional means. In comparison, the more elementary models tend to underestimate the efficiencies. Furthermore, quantile regression displays the influence of different cause variables in a more detailed way. The analyses are based on data from farm business surveys, including information from about 1200 farms per year in England and Wales from 2003 to 2007.

223: 4

Disclosure Risk From Interactions and Saturated Models in Remote Access

Gerd Ronning

Universitaet Tuebingen, Germany

Empirical research using micro data via remote access has been advocated in recent time by statistical offices since confidentiality is easier warranted for this approach. However, disclosure of single values and units cannot be completely avoided. Binary regressors (dummy variables) bear a high risk of disclosure, especially if their interactions are considered as it is done by definition in saturated models. This is well known for linear models. The paper considers saturated specifications of the most popular nonlinear microeconometric models (logit, probit, Poisson and negative binomial) and shows that in all cases the disclosure risk is high if some design points are represented by a (very) small number of observations. We also draw attention to the fact that interactions of binary regressors can be used to construct ”strategic dummy variables” which lead to hight disclosure risk as shown, for example, in Bleninger et al. (2010) for the linear model. In this paper we extend the analysis to the set of established nonlinear models, in particular logit, probit and count data models.

Keywords: logit model , probit model , poisson regression , negative binomial regression model , strategic dummy variable , tabular data.

224: Research Synthesis and Meta Analysis -- Diagnosis and application

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1010
Session Chair: Peter Schlattmann
Session Chair: Gerta Rücker

224: 1

Meta-analysis of paired-comparison studies of diagnostic test data: a Bayesian modelling approach

Pablo Emilio Verde

Heinrich Heine Universität Düsseldorf, Germany

Diagnostic paired-comparison studies arise when two diagnostic tests are applied to the same group of patients. The main problem in meta-analysis of this type of data is the lack of published information to make direct comparison between tests and to account for intra-study correlation of accuracy characteristics (e.g. sensitivities and specificities). The common statistical practice is to ignore these issues and proceed with simple meta-analytical techniques. In this work we proposed a new Bayesian hierarchical model to make meta-analysis of diagnostic paired-comparison studies. The observed diagnostic characteristics are modeled as the marginal results of unobserved data which allow direct comparison between tests. Variability between studies is modeled by extending a Bayesian hierarchical model for meta-analysis of diagnostic test (Verde, 2010). We propose new techniques of model diagnostics based on posterior predictive simulations and visualization techniques. Statistical methods are illustrated with two systematic reviews: The first one investigates the diagnostic accuracy of doctors aided with decision tools (e.g. neural networks) compared with unaided doctors in patients with acute abdominal pain (Liu et al. 2006). The second one compares the diagnostic accuracy of positron emission tomography with computer tomography in the detection of lung cancer (Birim et al. 2005).

Keywords: multivariate meta-analysis, diagnostic test, generalized evidence synthesis, MCMC, ecological inference.

224: 2

Meta-analysis for the comparison of two diagnostic tests to a common gold standard: A new idea using quadrivariate statistical models

Oliver Kuß[1], Annika Hoyer[2]

[1]Institut für Medizinische Epidemiologie, Biometrie und Informatik, Medizinische Fakultät, Martin-Luther-Universität Halle-Wittenberg; [2]Institut für Statistik, Ludwig-Maximilians-Universität München

Methods for the meta-analytic comparison of two diagnostic tests to a common gold standard have been called for and proposed. In these meta-analyses the parameters of interest are the differences of sensitivities and specificities (with their corresponding confidence intervals) between the two diagnostic tests while accounting for the various associations within single studies, between the tests and within probands. Current methods have deficiencies and we propose to use statistical models with a quadrivariate response (where sensitivity of test 1, specificity of test 1, sensitivity of test 2, and specificity of test 2 are the four responses) as a sensible approach to this task. Using a quadrivariate Generalized linear mixed model (GLMM) naturally generalizes the common standard model of meta-analysis for a single diagnostic test. Quadrivariate copula models are also possible. In the talk we report on first experiences with the respective models and estimation methods using an example data set to compare two drugs in pharmacological stress echocardiography for the diagnosis of coronary artery disease.

224: 3

Modelling SROC curves in meta-analysis of diagnostic studies using a proportional hazards model

Dankmar Böhning[1], Heinz Holling[2]

[1]University of Southampton, United Kingdom; [2]University of Münster, Germany

Meta analysis of diagnostic studies has become an important area of assessing diagnostic accuracy of a given diagnostic test of interest. The summary receiver operating characteristic has become a standard device in this arena since it can naturally cope with the cut-off value problem. It is assumed that for each study one pair of sensitivity and false-positive rate is observed. We further assume that this pair arises as realization of a parametric ROC curve from this study. This parametric ROC curve is characterized by a one-parameter proportional hazards model which is identifiable within the study. Since the within-study variance is determined by the mean parameters, there is an option for a variance component decomposition into within-study variance and variance between studies associated with the 'study' random effect. If study characteristics like diagnostic test or gold standard variation are available these can also be taken into account as fixed effects. This leads to a mixed model characterized by a specific covariance structure. It is shown that this model can easily fitted with SAS 'proc mixed'. A concrete meta-analysis on the Mini-Mental State Examination as a diagnostic test for dementia or mild cognitive impairment illustrates the method.

224: 4

Economic Implications of the Dynamic Relationship between Antibiotic Use and Hospital-Acquired Infections

Klaus Kaier

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Objectives: The emergence and spread of antimicrobial resistance (AMR) is still an unresolved problem worldwide. Recent evidence shows correlations between the volume of broad-spectrum antibiotics used in the hospital setting and the incidence of multidrug-resistant bacteria. According to this dynamic relationship, loss of antibiotic activity can be modeled as a negative externality of antibiotic consumption.

Methods: The present study proposes to present an economic model describing the probability of antibiotic treatment failure as a function of antimicrobial use and alcohol-based hand rub use. Furthermore, the results of recently conducted time-series analyses and cost-of-illness (COI) studies are applied to the model in order to determine the externalities of antibiotic consumption and alcohol-based hand rub use with respect to the costs of hospital-acquired infections.

Results: According to our calculations, consumption of third-generation cephalosporins and fluoroquinolones is associated with the highest negative externalities (€143 and €101, respectively) because their use has been shown to be associated with most types of hospital-acquired infections. In contrast, use of alcohol-based hand rub solution for hand disinfection is associated with a positive externality of 41 cent per single disinfection of the hands.

Conclusions: The externalities presented in this work represent a possible application of COI data in order to quantify the impact of antibiotic use on antimicrobial resistance. In addition, the results indicate that most economic research on the topic is biased in assuming the overall use of antibiotics to be responsible for the spread of antimicrobial resistance.

225: Time Series Analysis

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1015
Session Chair: Carsten Jentsch

225: 1

Identifiability of regular and singular multivariate autoregressive models from mixed frequency data

Manfred Deistler[1], Brian D O Anderson[2], Elisabeth Felsenstein[1], Bernd Funovits[1], Mohsen Zamani[2]

[1]TU Wien, Austria; [2]Australian National University, Australia

This paper is concerned with identifiability of an underlying high frequency multivariate AR system from mixed frequency observations. Such problems arise for instance in economics when some variables are observed monthly whereas others are observed quarterly. If we have identifiability, the system and noise parameters and thus all second moments of the output process can be estimated consistently from mixed frequency data. Then linear least squares methods for forecasting, nowcasting and interpolation of non-observed output variables can be applied. The main results show that on a generic, i.e., open and dense, subset of the parameter space identifiability holds. We deal with AR systems with nonsingular as well as with singular innovation variance matrices, the latter being important in the context of generalized dynamic factor models. For the case of singular innovation covariance matrices, exact (i.e. noise free) interpolation is discussed.

225: 2

Multivariate Linear Processes with Observations taken at Different Frequencies in each Component

Tobias Niebuhr, Marco Meyer, Jens-Peter Kreiß

TU Braunschweig, Germany

The class of linear processes is one of the most frequently used classes of processes in time series analysis. For many statistical quantities, among them sample autocovariances and autocorrelations, central limit theorems are available in the literature. For classical univariate linear processes it was recently shown that only one particular time scale possesses a linear structure. Thus, observing a linear process at a lower frequency leads to substantially different asymptotic results for standard statistical quantities. We review central limit results for the univariate setting and present a multivariate extension of this non-standard observation approach. Corresponding to the univariate case, the linear structure is destroyed as well once we observe the multivariate process at a lower frequency, and again we obtain different asymptotic results. Furthermore, the multivariate results are generalized in two reasonable ways. On the one hand the multivariate linear process is assumed to be observed on a different frequency in each component. Thus, we can only observe certain components at each specific time point but not necessarily the entire random vector. On the other hand the components of the multivariate process are each assumed to be univariate linear processes but not necessarily with the same underlying linear time scale for each component. Thus, there is no clear linear structure for the whole multivariate process although there is one for each component. It is shown that the generalized model contains the well-known standard setting as a special case.

225: 3

Nonlinear IV Panel Unit Root Testing under Structural Breaks in the Error Variance

Matei Demetrescu[1], Christoph Hanck[2]

[1]Universität Bonn, Germany; [2]Universität Duisburg-Essen, Germany

The paper examines the behavior of a generalized version of the nonlinear IV unit root test proposed by Chang (J Econometrics 2002) when the series' errors exhibit nonstationary volatility. The leading case of such nonstationary volatility concerns structural breaks in the error variance. We show that the generalized test is not robust to variance changes in general, and illustrate the extent of the resulting size distortions in finite samples. More importantly, we show that pivotality is recovered when using Eicker-White heteroskedasticity-consistent standard errors. This contrasts with the case of Dickey-Fuller unit root tests, for which Eicker-White standard errors do not produce robustness and thus require computationally costly corrections such as the (wild) bootstrap or estimation of the so-called variance profile. The pivotal versions of the generalized IV tests -- with or without the correct standard errors -- do however have no power in 1/T-neighbourhoods of the null. We also study the validity of panel versions of the tests considered here.

225: 4

Extensions of the Autoregressive Sieve Bootstrap to Multivariate Time Series and Random Fields

Marco Meyer, Jens-Peter Kreiss

TU Braunschweig, Germany

Extending the results of Kreiss, Paparoditis and Politis (2011), we explore the limits of the autoregressive (AR) sieve bootstrap. This procedure fits a finite-order AR model to the given data and uses this fit to generate bootstrap replicates of the time series. Based on these bootstrap samples the distribution of many statistics like the sample mean or sample autocorrelations can be approximated. The order of the fitted AR model depends on the number of given observations and increases as the sample size tends to infinity. For certain statistics and

data stemming from a large class of processes, the validity of the procedure can be shown. This is due to the fact that, under relatively mild conditions, many processes inherit a certain autoregressive structure which is mimicked correctly by the AR sieve bootstrap. With this work, we extend these results to the case of multivariate time series and random fields. The goal is to provide a general check criterion for the validity of the bootstrap which enables us to decide whether the procedure does or does not work for a specific statistic. In the latter case it should also point out the exact reason which causes the bootstrap to fail. This can be used as a starting point to implement specific modifications to the procedure in order to compensate its shortcomings.

227: Non- and Semi-Parametric Statistics

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3042
Session Chair: Arne Bathke

227: 1

Causal pitfalls in the decomposition of wage gaps

Martin Huber

University of St. Gallen, Switzerland

The decomposition of gender or ethnic wage gaps into explained and unexplained components (often with the aim to assess labor market discrimination) has been a major research agenda in empirical labor economics. This papers demonstrates that conventional decompositions, no matter whether linear or non-parametric, are equivalent to assuming a (probably too) simplistic model of mediation (aimed at assessing causal mechanisms) and may therefore lack causal interpretability. The reason is that decompositions typically control for post-birth variables that lie on the causal pathway from gender/ethnicity (which are determined at or even before birth) to wage but neglect potential endogeneity that may arise from this approach. Based on the newer literature on mediation analysis, we therefore provide more realistic identifying assumptions and demonstrate non-parametric identification based on reweighting.

227: 2

Misspecification testing in a class of conditional distributional models

Christoph Rothe[1], Dominik Wied[2]

[1]Columbia University, United States of America; [2]TU Dortmund, Germany

We propose a specification test for a wide range of parametric models for the conditional distribution function of an outcome variable given a vector of covariates. The test is based on the Cramer-von Mises distance between an unrestricted estimate of the joint distribution function of the data, and a restricted estimate that imposes the structure implied by the model. The procedure is straightforward to implement, is consistent against fixed alternatives, has non-trivial power against local deviations of order [fn]n^{−1/2}[/fn] from the null hypothesis, and does not require the choice of smoothing parameters. In an empirical application, we use our test to study the validity of various models for the conditional distribution of wages in the US.

227: 3

Pricing of houses and their characteristics using multiple nonparametric regression

Harry Haupt[1], Joachim Schnurbus[1], Rolf Tschernig[2]

[1]University of Passau, Germany; [2]University of Regensburg, Germany

"What the hedonic approach attempted was to provide a tool for

[1] estimating missing prices ... of particular bundles not observed, ...

[2] detection of the relevant characteristics of a commodity and ...

[3] estimation of their marginal market valuation." (Griliches, 1990).

This paper pursues those three issues by reflecting the recent emphasis on the necessity of nonlinear methods for modeling hedonic price functions. The resulting statistical problem of model selection and validation of such methods may pose a problem as respective statistical criteria may not coincide in a tentative parsimonious model which also admits an economically sound interpretation. We apply multiple nonparametric regression and cross-validation for estimation, model validation and selection to analyze an urban hedonic house price function. Our analysis of three different housing data sets is complemented by an extensive Monte Carlo-study.

227: 4

Horizontal Changes in Conditional Distributions at a Quantile from the Marginal Distribution and a Rank Preservation Test

Anthony Strittmatter

University Freiburg, Germany

When estimating horizontal changes in potential outcome distributions after a policy intervention, one has to specify a quantile at which these changes are estimated. Researchers typically choose a quantile from the same potential outcome distributions for which the horizontal changes in distributions are estimated. In this study, we suggest to estimate horizontal changes in conditional potential outcome distributions at a quantile from the marginal potential outcome distribution. First, this is a new policy parameter enabling answers to different policy questions than existing parameters. Secondly, comparing the influence of different characteristics on distributional policy parameters could

be difficult, when the policy parameters are estimated at different values of the outcome. This difficulty might be caused by quantiles that refer to different distributions. It disappears when horizontal changes in potential outcome distributions are measured at a quantile from the same marginal potential outcome distribution. Finally, this approach allows a rank preservation test to be implemented.

228: Statistik in der Öffentlichkeit und in der Politik

Time: Tuesday, 19th Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3043
Session Chair: Walter Krämer

228: 1

Die Unstatistik des Monats - was machen Medien und Politik im Umgang mit Statistik falsch?

Thomas Bauer

RWI Essen, Germany

tba

228: 2

Die Macht der Daten

Uwe Saint-Mont

FH Nordhausen, Germany

„Die Macht der Daten, wie Information unser Leben bestimmt“ ist der Titel eines Buches, das in diesem Frühjahr im Springer-Verlag erscheint. Es handelt von Statistik, Informatik, Wissenschaft und Philosophie, also Feldern, die zumeist nicht in einem Atemzug genannt werden.

Im Vortrag wird der Autor die Idee einer durchgängig datenbasierten Informationskultur erläutern. Alle Institutionen der Gesellschaft sind genauso wie das sich schnell wandelnde soziale Leben und die meisten wirtschaftliche Vorgänge auf einen funktionierenden Daten- und Informationsfluss angewiesen. Den Informationswissenschaften Informatik und Statistik fällt deshalb die zentrale Aufgabe zu, Daten systematisch zu erheben, zu organisieren und zu analysieren. Insbesondere wird auf dieser Infrastruktur aufbauend eine konsequent sachorientierte Politik möglich.

Tatsächlich sind aussagekräftige Daten sogar der entscheidende Rohstoff aller erfolgreichen empirischen Unternehmungen. Ohne Daten keine zuverlässigen Erklärungen, präzisen Prognosen und wirkungsvollen Maßnahmen, und nichts bewahrt uns so gründlich vor Illusionen wie ein Blick in den Spiegel der Realität, der sich aus unzähligen Datenpunkten zusammensetzt.

Nicht zuletzt stützen sich auch alle empirischen Wissenschaften bis hin zu deren Theorie (Philosophie) maßgeblich auf valide und umfangreiche Datenbestände. Erst sie ermöglichen umfassende abstrakte Konzepte, die gleichwohl präzise genug sind, um wieder unmittelbar angewandt werden zu können, insbesondere in Form nichttrivialer Technik. Insgesamt bilden Statistik, Informatik und die empirischen Wissenschaften, in Verbindung mit dem modernen, zugleich empirisch-konzeptionellen und analytisch-quantitativen Denken eine natürliche Einheit, die mehr und mehr das gesellschaftliche Leben prägt.

228: 3

Zur Novellierung des Bundeswahlgesetzes

Friedrich Pukelsheim

Universität Augsburg, Germany

tba

230: Opening session (Invited Speaker: Fahrmeir)

Time: Tuesday, 19th Mar 2013: 10:40am - 12:00pm · Location: KG II, Audimax
Session Chair: Martin Schumacher

230: 1

Progression

Ludwig Fahrmeir

Ludwig-Maximilians-University Munich, Germany

In this talk I take a subjective view at progress in regression. Actually, I am going to make a journey through the land of regression, motivated and illustrated by applications in various fields, and powered by likelihood-based and Bayesian machinery. The journey starts with a trip back to generalized linear models and their extensions. Then it leads us to regularization and selection of predictors based on penalized likelihood and related Bayesian concepts. The journey ends with an outlook at regression problems where responses are more complex objects, such as high-dimensional responses, curves, images, or networks.

241: Statistics in Clinical and Preclinical Research -- Early Development and Non-inferiority Trials

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG II, Audimax
Session Chair: Martin Wolfsegger

241: 1

Establishing bioequivalence in AB/BA cross-over trials using the ratio of AUCs estimated from sparse sampling designs

Philip Pallmann[1], Thomas Jaki[2], Martin J. Wolfsegger[3]

[1]Leibniz University Hannover, Germany; [2]Lancaster University, United Kingdom; [3]Baxter Innovations GmbH, Vienna, Austria

Showing bioequivalence of two treatments by comparing their AUCs (area under the concentration-versus-time curves) is a standard approach in pharmaceutical research. This could be done using a parallel group design, but major regulatory guidelines encourage running cross-over trials as statistical inference is based on the within-subject variability allowing frequently to use a smaller sample size than with a parallel group design. Estimating treatment effects and associated standard errors from such designs is, however, not trivial. We provide three methods for obtaining estimators with standard errors from an AB/BA cross-over trial. Posing an additional challenge, incomplete data settings are a common occurrence, i.e. not all study participants are sampled at all pre-specified time points, which complicates estimation of AUCs and their standard errors. Assuming at least asymptotic normality, Fieller-type confidence intervals can be constructed for the ratio of AUCs estimated using a non-compartmental approach in a sparse sampling setting from a two-treatment, two-period, two-sequence cross-over trial. In particular we focus on a flexible batch design, which includes traditional serial sampling and complete data designs as special cases. Simulation results indicate that the proposed intervals have nominal coverage and keep the type I error even for small sample sizes. Software for estimating AUCs and their standard errors in sparse sampling designs is available in the R package 'PK'.

241: 2

Simultaneous confidence intervals for benefit assessment in early diagnostic trials

Antonia Zapf

University Medical Center Göttingen, Germany

In many diagnostic trials in early phases the aim is to select the most promising markers. The selection criterion in such studies is often the area under the Receiver Operating Characteristic (ROC) curve, named AUC [1]. Then the confidence intervals for the AUC's are much more interesting than the p-values. But to control the global type one error, simultaneous confidence intervals have to be used [2]. Furthermore sometimes there is more than one observation per patient (called clustered data). For example, if a marker is measured in different brain regions. The dependency of the observations belonging to the same patient has to be taken into account [3].

In this talk I will present nonparametric methods for simultaneous confidence intervals in general and also for the case of clustered data. Moreover I will present the results both of simulation studies and of an example.

References:

[1] EMA, Committee for Medicinal Products for Human Use (2010), Guideline on Clinical Evaluation of DIagnostic Agents. Doc.Ref.CPMP/EWP/1119/98/Rev.1.

[2] Konietschke, Bathke, Hothorn, Brunner (2010), Testing and Estimation of nonparametric relative effects in repeated measures designs. Computational Statistics & Data Analysis 54, 1895-1905.

[3] Obuchowski (1997), Nonparametric Analysis of Clustered ROC Curve Data. Biometrics 53, 567-578.

241: 3

Network meta-analysis and meta-regression in design and analysis of non-inferiority trials

Heinz Schmidli

Statistical Methodology, Novartis Pharma AG, Switzerland

In non-inferiority clinical trials, a test treatment is compared to an active-control rather than to placebo. Such designs are considered when placebo is unethical or not feasible. The critical question is whether the test treatment would have been superior to placebo, had placebo been used in the non-inferiority trial. This question can only be addressed indirectly, based on information from relevant historical trials with data on active-control and placebo. The network meta-analytic-predictive (NetMAP) approach to non-inferiority trials is based on a network meta-analysis of the data from the historical trials and the non-inferiority trial, and the prediction of the putative test vs. placebo effect in the non-inferiority trial. The approach extends previous work by incorporating between-trial variability for all relevant parameters, and focusing on the parameters in the non-inferiority trial rather than on population means.

References:

Schmidli H, Wandel S, Neuenschwander B (2012) The network meta-analytic-predictive approach to non-inferiority trials. Statistical Methods in Medical Research, DOI:10.1177/0962280211432512.

Witte S, Schmidli H, O'Hagan A, Racine-Poon A (2011) Designing a non-inferiority study in kidney transplantation: a case study. Pharmaceutical Statistics; 10, 427-432.

241: 4

Eine flexible multiple Teststrategie für dreiarmige Nichtunterlegenheitsstudien

Kathrin Stucke, Meinhard Kieser

Universität Heidelberg, Germany

In dreiarmigen Nichtunterlegenheitsstudien erhalten die Patienten entweder das Prüfmedikament, ein Standardmedikament oder ein Placebo. Dieses Studiendesign gilt als "Goldstandard" für Nichtunterlegenheitsstudien und wird von aktuellen Guidelines (siehe z.B. [1]) empfohlen.

Wird neben dem Nachweis der Nichtunterlegenheit des Prüfmedikaments gegenüber der Standardtherapie auch der Nachweis der Überlegenheit des Prüfmedikaments gegenüber Placebo und der Überlegenheit des Standardmedikaments gegenüber Placebo (sogenannte „assay sensitivity“) in die primäre konfirmatorische Analyse aufgenommen, so führt dies zu einem multiplen Testproblem. Ein hierarchisches Vorgehen ermöglicht ein Testen der drei Einzelhypothesen ohne Adjustierung des Fehlerniveaus. Beginnend mit dem Vergleich des neuen Medikaments gegen Placebo kann nach dem Ablehnen der zugehörigen Hypothese und der damit gezeigten Überlegenheit gegenüber Placebo, der Nachweis der Nichtunterlegenheit und der assay sensitivity durchgeführt werden [2]. Dieses Vorgehen hat allerdings den Nachteil, dass bei einem Nichtablehnen der ersten Null-Hypothese die Nichtunterlegenheit und auch die assay sensitivity nicht getestet werden können.

Im Vortrag wird basierend auf einem Ansatz von Alosh und Huque [3] für Subgruppenanalysen das hierarchische Vorgehen dahingehend modifiziert, dass auch bei einem Nichtablehnen der ersten Null-Hypothese die weiteren Tests unter Einhaltung der multiplen Irrtumswahrscheinlichkeit durchgeführt werden können. Die Power der vorgeschlagenen multiplen Testprozedur hängt von mehreren vor Studienbeginn zu spezifizierenden Größen ab. Es werden Vorschläge zur Wahl dieser Größen unterbreitet. Die Eigenschaften des Ansatzes werden mittels analytischer Berechnungen untersucht und mit denen des hierarchischen Ansatzes verglichen.

References:

[1] EMA Reflection Paper. EMA/759784/2010.

[2] Röhmel J, Pigeot I. Biopharmaceutical Statistics 2010; 20: 911-926.

[3] Alosh M, Huque MF. Statistics in Medicine 2009; 28:3-23.

242: Survival and Event History Analysis -- Planning and sample size calculation

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG III, HS 3044
Session Chair: Andreas Wienke

242: 1

Chances and challenges of effect measures based on prioritized outcomes

Geraldine Rauch

Institut für Medizinische Biometrie und Informatik, Heidelberg, Germany

We discuss different approaches to combine multiple endpoints in a univariate outcome measure. In case of binary or time-to-event variables, composite endpoints which combine several event types within a single event or time-to-first-event analysis are most often used to assess the overall treatment effect. A main drawback of this approach is that the interpretation of the composite effect can be difficult as a negative effect in one component can be masked by a positive effect in another. Recently, more general approaches bases on a priority ranking of outcomes were proposed which moreover allow to combine outcome variables of different scale levels. These new combined effect measures assign a higher impact to more important endpoints which is meant to simplify the interpretation of results. Whereas statistical tests and models for binary and time-to-event variables are well understood, the latter methods have not been investigated in detail so far. We will derivate the statistical properties of prioritized combined outcome measures. A systematical comparison to standard composite measures, such as the all-cause hazard ratio in case of multiple time-to-event variables or the absolute rate difference in case of multiple binary variables, will be performed to derive recommendations for the use of either effect measure in different clinical trial scenarios.

242: 2

Blinded and Unblinded Re-estimation of the Sample Size for Studies with Recurrent Event Data as the Primary Outcome

Katharina Ingel, Antje Jahn-Eimermacher

Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Mainz, Germany

The repeated occurrence of the same type of event, e.g. an episode of hospitalization, is of primary interest in many clinical trials. The Andersen-Gill model (Andersen and Gill, Ann. Stat., 1982) has been proposed to analyse recurrent event data. When calculating the test statistic, heterogeneity in individual baseline hazards, that might be caused by unobserved or unobservable covariates, has to be accounted for to assure the type I error rate. This can be done by the use of a robust variance estimate. However, existing sample size formulas for the Andersen-Gill model do not incorporate unexplained heterogeneity, resulting in a lower power than anticipated. Therefore, we propose an adjusted sample size formula to reach the targeted power even in the presence of unexplained heterogeneity. The adjustment is based on an inflation factor that considers the degree of heterogeneity and is derived from characteristics of the robust variance estimate (Al-Khalidi et. al., Biometrics, 2011).

In the planning phase of a trial there will usually be some uncertainty about the size of the inflation factor. We propose an internal pilot study design to re-estimate the inflation factor during the course of the trial and adjust the sample size accordingly. This re-estimation procedure requires unblinding. For blinded studies we propose an alternative re-estimation procedure which is derived from sample size methods for count data (Friede and Schmidli, Methods Inf. Med., 2010).

In this talk we will introduce the adjusted sample size formula, and evaluate the re-estimation procedures through simulations.

242: 3

Resampling validation of a sample size formula for clustered time to event data

Antje Jahn-Eimermacher[1], Valerie Horvath[2], Katharina Ingel[1]

[1]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Mainz, Germany; [2]Center for Thrombosis and Hemostasis, University Medical Center Mainz, Germany

In cluster randomized trials groups of individuals (clusters) are randomized to the particular treatments or interventions. Intra-class correlation reduces the efficient sample size of a trial and thus has to be considered in sample size planning. We derived a sample size formula for clustered time to event data based on a shared frailty model. As the distribution of trial data may differ from the statistical model assumed in the planning phase of a trial, it is of interest to evaluate the performance of the sample size formula in applications. We propose a resampling validation using data from 60.000 patients of the NIATx 200 trial as an example. The NIATx 200 trial is a cluster randomized trial to evaluate interventions randomly allocated to 201 addiction treatment centers in the U.S.. One of the primary research questions is which intervention has the greatest improvement in the time the patients have to wait for their first treatment. For different cluster sizes, we draw resamples from the trial data and calculate test statistics from each of them. We compare the resulting empirical power to the power resulting from the model-based sample size formula. The empirical power only marginally increases with larger cluster sizes, which is in accordance to the model-based formula. In conclusion, this method requires large datasets, which will not always be available. But if they are, resampling validation shows to be a good method to evaluate the performance of a model-based sample size formula in applied settings.

242: 4

Sample size calculation: cohort vs. nested case-control study

Kristin Ohneberg[1,2], Martin Schumacher[1], Jan Beyersmann[1,2], Martina Rauscher[1]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Freiburg Center for Data Analysis and Modeling, Germany

Schoenfeld's formula (1983) for the total number of events required to provide a desired level of power for a single binary covariate in a Cox proportional hazards model is well known within the framework of sample size calculation in a cohort study with a time to event endpoint. For analysis in a multiply matched case-control study design, Lachin (2008) proposed a formula using a conditional logistic regression model that also applies to nested case-control studies. This study design has become an increasingly used tool, as it allows for statistically efficient analysis within a subsample of an existing cohort to derive estimates that are similar to estimates obtained by the full cohort, with notable savings in resources. It turns out that the number of matched sets needed for the nested case-control study can be calculated by multiplying Schoenfeld's formula for the Cox model in the full cohort with a factor that only depends on the number of controls matched to each case. We present the respective sample size formulas required for analyzing the full cohort or using incidence density sampling in the nested case-control approach. For illustration, we use data from an epidemiological study conducted to investigate risk factors for nosocomial infections in German intensive care units.

References:

Lachin, J.M. (2008). Sample size evaluation for a multiply matched case-control study using the score test from a conditional logistic (discrete Cox PH) regression model. Statistics in Medicine 27(14), 2509-2523.

243: Observational Studies -- Observational Studies and Beyond

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1199
Session Chair: Iris Pigeot

243: 1

Advantages of using both observational and experimental data: some examples from biomedical research

Stephen Walter

McMaster University, Canada

Some tension exists between clinicians and epidemiologists in how they assess the value of observational or experimental evidence. Guidelines for good clinical practice often emphasise the superiority of randomised trials, to avoid bias in the evaluation of new treatments. On the other hand, some epidemiologists argue that observational data is often highly relevant to such issues, and should not be ignored. I will describe two biomedical research areas for which I conclude that both experimental and observational evidence are valuable. First, we will consider the use of non-randomised treatment assignments in the context of clinical trial designs: this approach permits the estimation of so-called preference and selection effects on outcomes, arising from patients or their clinicians. These effects may be more important than the usual direct effect of treatment as it is estimated in conventional trials where all patients are randomly assigned to treatments. We may also briefly examine the expertise-based design and preference-based analyses for randomised trials, which shed further light on individual determinants of treatment outcomes. Second, we will examine randomised and observational data on the benefit of breast cancer screening. Despite fundamental differences between the results of these two types of evidence, further investigation shows that the estimated benefit is actually consistent between the two types of design when an appropriate analysis is applied. These examples suggest that potentially important determinants of treatment outcomes may be unidentifiable from randomised studies alone. In contrast, study designs that are partly or completely observational can often enhance our understanding of the data beyond what is possible in randomised studies alone. Depending on the specific circumstances, there is clearly a case for considering both types of evidence.

243: 2

Bias of relative risk estimates in cohort studies as induced by missing information due to death

Nadine Binder, Martin Schumacher

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In most clinical and epidemiological studies, information on the outcome of interest (e.g. disease status) is usually collected at regular follow-up times. Often, this information can only be retrieved in individuals who are alive at follow-up, but will be missing for those who died before. This is of particular relevance in long-term studies or when studying elderly populations. Frequently, individuals with missing information because of death are excluded and analysis is restricted to the surviving ones. Such naive analyses can lead to serious bias in incidence estimates, translating into bias in estimates of hazard ratios corresponding to potential risk or prognostic factors. We investigate this bias in hazard ratio estimates by simulating data from an illness-death multi-state model with different transition hazards, considering the influence of risk factors following a proportional hazards model. We extend an approximate formula for the bias of the incidence estimate by Joly et al. (2002) for a binary risk factor. Analytical and simulation results closely agree, revealing that the bias can be substantial and in either direction, depending on differential mortality. The problem will be illustrated considering examples from biomedicine, economics and social sciences. Specifically, in an application to a Danish long-term study on nephropathy for diabetics, where complete status information is available, we artificially induce missing intermediate disease status. The naive risk factor analyses differ significantly from those obtained in the original analysis and even change sign. This supports the analytical and simulation results and underlines that missing intermediate outcome status cannot be ignored.

243: 3

Conceptual difficulties of modeling mediation for observational data

Andreas Klein, Holger Brandt

Goethe-Universität Frankfurt, Germany

In many research contexts in the social or health sciences, researchers are interested in identifying intermediate variables that stand in the pathway or may serve as early indicators of a developing disease or critical behavior. Causal effects that are indirect (mediated) and go through intermediate variables that act as components in a causal chain are also known as mediator effects. In the biostatistical literature, however, serious concerns about the conventional methodology such as path analysis used to explore those effects have been raised. Following this critique, a causal interpretation of a path-analytic mediator effect is hardly ever justified, and it can be shown that path analysis may often identify supposedly strong mediator effects that in reality are nothing but methodological artefacts. In this paper, we explain why the standard mediator model is conceptually flawed, and why it requires additional, untestable assumptions to become interpretable as a model that can represent a mediated causal process. Directions for future model development and consequences for statistical applications are discussed.

244: Applied Econometrics

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1098
Session Chair: Winfried Pohlmeier

244: 1

Semi-Nonparametric Estimation Of A Nonseparable Demand Function Under Shape Restrictions

Matthias Parey[1], Richard Blundell[2], Joel L. Horowitz[3]

[1]University of Essex, United Kingdom; [2]University College London, United Kingdom; [3]Northwestern University, United States of America

Economic theory rarely provides a parametric specification for a model, but it often provides shape restrictions. We consider nonparametric estimation of the demand for gasoline in the U.S. subject to the Slutsky restriction of economic theory. The demand function is specified as [fn]Q=D(p,Y,U)[/fn], where [fn]Q[/fn] is the quantity demanded by an individual, [fn]p[/fn] is the price, [fn]Y[/fn] is the individual's income, [fn]D[/fn] is an unknown function, and [fn]U[/fn] is an unobserved random variable that is independent of [fn]p[/fn] and [fn]Y[/fn]. [fn]D[/fn] can be estimated consistently by nonparametric quantile regression but, owing to random sampling errors, the resulting estimate is wiggly and non-monotonic in [fn]p[/fn]. We deal with this problem by imposing the Slutsky restriction on the nonparametric quantile estimator. The resulting estimate of [fn]D[/fn] is a smooth and monotone decreasing function in [fn]p[/fn]. The estimated function reveals considerable variation in price responsiveness across the income distribution. The results illustrate the improvements in the finite-sample performance of a nonparametric estimator that can be achieved by imposing shape restrictions based on economic theory.

244: 2

The Heterogeneous Effects of Training Incidence and Duration on Labor Market Transitions

Bernd Fitzenberger[1], Aderonke Osikominu[1], Marie Paul[2]

[1]University of Freiburg, Germany; [2]University of Duisburg-Essen, Germany

This paper estimates the impact of training incidence and duration on employment transitions. We develop a dynamic evaluation approach in discrete time that accounts for the endogeneity of program participation and duration. We specify a very flexible bivariate random effects probit model for employment and training participation and use Bayesian Markov Chain Monte Carlo (MCMC) techniques for estimation. We propose a simulation approach that uses the estimated coefficients and individual specific effects from the MCMC iterations to calculate the posterior distributions of different treatment effects of interest. Our estimation results imply positive effects of training on the employment probability of the treated, lying between 12 and 21 percentage points ten quarters after program start. The effects are higher for women than for men and higher in West Germany than in East Germany. Further, we find that the effect of training versus waiting underestimates the effect of training versus no training in the medium and long run by a third. Finally, our results show that longer planned enrolment lengths of three and four quarters as opposed to just two quarters lead to an increase in employment rates in the medium and long run by four to eleven percentage points.

244: 3

What is the Effect of being Awarded with a Training Voucher on Labor Market Outcomes?

Annabelle Dörr[1], Bernd Fitzenberger[1], Anthony Strittmatter[1], Marie Paul[2], Thomas Kruppe[3]

[1]University of Freiburg, Germany; [2]University of Duisburg, Germany; [3]Institute of Employment Research, Germany

This paper estimates the labor market effects of being awarded with a training voucher using an instrumental variable approach. In Germany all public sponsored further training programs are allocated through vouchers and this system, which we study here, thus represents a major case of the use of vouchers in the context of active labor market policies. We use process generated data in which we observe all training vouchers awarded in 2003 and 2004 as well as realized training participation. Results suggest that on average voucher recipients suffer from strong lock-in effects and only experience small positive employment effects and no earning gains four years after the voucher is awarded. Subgroups of treated individuals, like individuals not holding a vocational degree and those participating in a degree program, benefit more.

245: Research Synthesis and Meta Analysis – Study heterogeneity and small samples

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1010
Session Chair: Katrin Jensen
Session Chair: Oliver Kuß

245: 1

Applying Bayesian evidence synthesis in comparative effectiveness research

David Ohlssen

Novartis, United States of America

Motivated by the use of evidence based medicine to evaluate health technology, there has been an enormous increase in the use of quantitative techniques that allow data to be combined from a variety of sources. In a drug development setting, there have been a number of recent key works: The recommendations on the use and application of network meta-analysis were recently presented by the ISPOR task force; From a regulatory perspective, the work of the Canadian Agency (Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis) and the UK NICE Evidence synthesis series have recently been published; Further, the FDA also started a number of recent projects on comparative effectiveness research as part of a plan to enhance regulatory science. By drawing on examples from a drug development setting, this talk aims to examine these recent advances. In particular, emphasis will be placed on the application of Bayesian evidence synthesis methods when applied to drug safety evaluation and comparative effectiveness.

245: 2

Meta-Analyse mit 2 Studien in einem Modell mit festen und zufälligen Effekten sowie einem Modell mit skalenabhängiger Heterogenität

Andrea Gonnermann, Theodor Framke, Armin Koch

Institut für Biometrie, Medizinische Hochschule Hannover, Germany

In aller Regel sind für die Zulassung eines neuen Arzneimittels zwei zielführende klinische Studien vorzulegen. Im Rahmen der Kosten-Nutzen-Bewertung von Arzneimitteln verbleiben unter Berücksichtigung klinischer Homogenität häufig nur zwei Studien für die Bewertung des Zusatznutzens [1,2,3]. Obwohl der Anspruch besteht, die Evidenz aus ggf. zwei Studien gemeinschaftlich zu bewerten, ist über die Durchführung von Meta-Analysen mit nur zwei Studien wenig bekannt. Ob in diesem Fall die Zusammenfassung in einem Modell mit festen (FE-Modell), in einem Modell mit zufälligen Effekten (RE-Modell) oder der Ansatz von Hartung und Knapp [4] zuverlässiger ist, ist nicht untersucht. Wir haben Simulationen zur Überprüfung des Fehlers 1. Art sowie zur Güte durchgeführt. Neben den Modellen mit festen und zufälligen Effekten wurde ein weiteres Modell mit Skalenheterogenität betrachtet. Hierbei steht die Frage, ob Heterogenität allein bedingt durch die falsche Skala (Relatives Risiko und Risikodifferenz) dazu führen kann, dass ein Therapieeffekt nicht entdeckt wird, im Vordergrund. Dazu wurde ein weiteres Verteilungsmodell (Gleichverteilung) der Basisrisiken angenommen. Unter diesen verschiedenen Modellannahmen wurden die statistischen Eigenschaften (Güte und Fehler 1. Art) der drei genannten Methoden verglichen.

References:

[1] EMA, 2001. Points to consider on applications with 1. Meta-analyses; 2. one pivotal study.

[2] FDA, 1998. Guidance for Industry - Providing Clinical Evidence of Effectiveness for Human Drug Use and Biological Products.

[3] Janatzek, S., 2011. Untersuchung von und Umgang mit Heterogenität in Nutzenbewertungen - ein Problemaufriss.

[4] Hartung, J. & Knapp, G., 2001. A refined method for the meta-analysis of controlled clinical trials with binary outcome. Statist. Med., 20: 3875-3889.

245: 3

Inference in the random effects meta regression model

Thomas Friedrich, Guido Knapp

Technische Universität Dortmund, Germany

The explanation of heterogeneity when combining different studies is an important issue in meta analysis. Besides including a heterogeneity parameter in the analysis, it is also important to understand the possible causes of heterogeneity. A possibility is to incorporate study-specific covariates in the model that account for between-trial variability. This leads to the random effects meta regression model.

Commonly used methods for constructing confidence intervals for the regression coefficients are examined and two new methods based on generalised inference principles are proposed. The different methods are compared by an extensive simulation study with respect to coverage probability and average length in three different and common settings.

Some pitfalls and suggestions are being made for how to choose the right methods for the right task.

245: 4

Random zero: Meta-analyses with three, two and one studies

Theodor Framke, Andrea Gonnermann, Armin Koch

Institut für Biometrie, Medizinische Hochschule Hannover, Germany

Zero cell counts may occur in studies with binary endpoints. Frequently they can be observed in small studies and/or in the case of rare events. Safety data from clinical trials may be prone to 'empty cells'. Problems also arise in studies in which all subjects have been successful. Several approaches (e.g. excluding studies, corrections such as adding 0.5 to all cells of a fourfold table) have been proposed to deal with the estimation of treatment effects in meta-analyses that include study arms with no events. Further difficulties may arise from the estimation of a variance.

This presentation highlights crucial situations and summarizes different methods that have been proposed to handle this problem. In addition, the impact of selected methods on different strategies to combine studies will be assessed with a simulation study. Results for the type I error will be presented for fixed and random effects models.

References:

Sweeting MJ, Sutton AJ, Lambert PC (2004): What to add to nothing? Use and Avoidance of continuity corrections in meta-analysis of sparse data, Statistics in Medicine 23, 1351-1375.

Bradburn MJ, Deeks JJ, Berlin JA, Localio AR (2007): Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events, Statistics in Medicine 26, 53-77.

Lane PW (2012): Meta-Analysis of incidence of rare events, Stat Methods Med Res, Epub ahead of print.

Gart JJ, Zweifel JR (1967): On the bias of various estimators of the logit and its variance with application to quantal bioassay, Biometrika 54(1and2), 181-187.

246: Time Series Analysis

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1015
Session Chair: Jens-Peter Kreiß

246: 1

A model-specification test for GARCH(1,1) processes

Anne Leucht[1], Jens-Peter Kreiß[2], Michael H. Neumann[3]

[1]Universität Mannheim, Germany; [2]Technische Universität Braunschweig, Germany; [3]Friedrich-Schiller-Universität Jena, Germany

There is already an overwhelming amount of model specification tests in the econometric literature. These methods typically rely on the assumption that the information variables as well as the response variables are observable. However, this condition is violated in the case of GARCH models, where unobserved quantities enter the information variable.

We establish a consistent model-specification test for GARCH(1,1) models based on an L2-type test statistic. The latter can be approximated by a degenerate V-statistic and its asymptotics are then derived invoking results of Leucht and Neumann (2012). Since the test statistic as well as its limit distribution depend on unknown parameters in a complicated way, critical values cannot be derived directly. We present a bootstrap-based testing procedure that overcomes these difficulties. The approach presented in this talk can be carried over to many other test problems.

References:

Leucht, A. and Neumann, M. H. (2012). Degenerate U- and V-statistics under ergodicity: Asymptotics, bootstrap and applications in statistics. (Forthcoming in Annals of the Institute of Statistical Mathematics).

246: 2

Financial modelling with generalized moving average processes

Jeannette H.C. Woerner

Technische Universität Dortmund, Germany

Ornstein-Uhlenbeck processes, especially Lévy driven Ornstein-Uhlenbeck processes build a popular class of stochastic processes, since they are analytically tractable and reasonably well fit empirical features of some financial models, e.g. volatility processes or electricity prices. Ornstein-Uhlenbeck processes may be defined as moving average processes. We take this as a starting point and consider modifications of the kernel function, which lead to new features in the paths properties and the autocorrelation function. We examine the properties of the corresponding processes and show how they fit empirical data in a more realistic way. One feature is for example that we might include seasonalities or other oscillating behaviour directly into the model.

246: 3

Consistency and asymptotic normality of M-estimators for augmented GARCH models

Fabian Michael Tinkl

University Erlangen-Nuremberg, Germany

We prove consistency and asymptotic normality of M-estimators for augmented generalized autoregressive conditional heteroscedasticity (GARCH) models that have been introduced by Duan (1997). This broad class of models contains many important GARCH models as special cases, for instance the linear GARCH, the asymmetric GARCH and the threshold GARCH model. Using this general setting we focus on the quasi maximum likelihood estimator (QMLE) and the least absolute deviation estimator (LADE). We show that under mild assumptions both estimators are consistent and asymptotically normally distributed for augmented GARCH models that are strictly stationary and ergodic. In particular we do not impose any moment assumptions on the GARCH process and we derive consistency and asymptotic normality of the LADE under very weak assumptions. A simulation study supports these findings by showing the good finite sample properties of the two estimators in different GARCH settings. From the simulation study we deduce that LADE is superior to the QMLE when the error distribution is heavy tailed.

246: 4

Diagnosing and Modelling Extra-Binomial Variation for Time-Dependent Counts

Christian Weiß[1], Hee-Young Kim[2]

[1]Department of Mathematics, TU Darmstadt, Germany; [2]Institute of Economics, Korea University, Seoul, Korea

The analysis and modelling of time series of counts has become a popular research area during the last decades. Meanwhile, more and more articles consider the problem of count data time series with a finite range, and also more and more such applications have been

reported in the literature. A quite popular approach for modelling such data is the binomial AR(1) model, which is based on the binomial thinning operation. The stationary marginal distribution of this model is the binomial distribution, which is characterized by a fixed relation between mean and variance. In many applications, however, we observe variances being larger than allowed by the binomial model.

We consider the modelling of count data time series with a finite range having extra binomial variation. We propose a beta-binomial AR(1) model using the concept of random coefficient thinning. We discuss the stationarity conditions, derive the moments and autocovariance functions, and consider approaches for parameter estimation. Furthermore, we develop two new tests for detecting extra-binomial variation, and we derive the asymptotic distributions of the test statistics under the null hypothesis of a binomial AR(1) model. The size and power performance of the two tests are analyzed under various alternatives taken from a beta-binomial AR(1) model with Monte Carlo experiments.

Finally, we conclude with an application being related to the current Euro crisis. Our real-data example, using the Harmonised Index of Consumer Prices (HICP) of the European Union to express price stability in the Euro countries, demonstrates the practical value of the developed tests.

247: Non- and Semi-Parametric Statistics

Time: Tuesday, 19th Mar 2013: 1:00pm - 2:20pm · Location: KG III, HS 3042
Session Chair: Markus Pauly

247: 1

Flexible Pair-Copula Estimation in D-vines using Bivariate Penalized Splines

Göran Kauermann[1], Christian Schellhase[2]

[1]Department of Statistics, Ludwig-Maximilians-University Munich, Germany; [2]Centre for Statistics, Department for Business Administration and Economics, Bielefeld University, Germany

The paper presents a new method for flexible fitting of D-vines. It is the idea to model a multivariate copula by a collection of pairwise, that is two dimensional copulas. The pair-copula uses conditional distributions as arguments but the copula itself is independent of any conditioning variables. We estimate each pair-copula density in a flexible, that is semi-parametric way by refraining from any strong parametric assumptions on the structure of the pairs, using penalized Bernstein polynomials or linear B-splines, respectively, as spline bases in each knot of the D-vine throughout each level.

A penalty on the spline basis coefficients induce smoothness of the fit while the high dimensional spline basis guarantees flexibility. To ensure uniform univariate margins of each pair-copula, linear constraints are placed on the spline coefficients and quadratic programming is used to fit the model. The amount of penalizations for each pair-copula is driven by a penalty parameter which is selected in a numerically efficient way, using the link between penalized splines and linear mixed models.

The order of variables in the first tree level of the D-vine completely specifies the vine. We specify this order by solving a traveler salesman problem interpreting the pairwise corrected Akaike Information Criterion as distance measure between two variables.

Simulations from different copula families and practical examples accompany the presentation.

247: 2

Nonparametric Low Frequency Lévy Copula Estimation

Christian Palmes

Technische Universität Dortmund, Germany

We intend to estimate the jump dependence structure of a multidimensional Lévy process under a discrete low frequency sample scheme. The Lévy measure in the Lévy Khintchine representation describes the average jump behaviour in a time unit. Our aim is to estimate the dependence structure of the Lévy measure by estimating its Lévy copula, a concept introduced by Kallsen and Tankov (2006).

We use the low frequency techniques presented in Neumann and Reiss (2009) and Nickl and Reiss (2012) in a one dimensional setting to construct two Lévy copula estimators based on low frequency observations. We prove that one estimator converges in a logarithmic convergence rate to the true Lévy copula. This convergence holds under quite general assumptions, which include also Lévy triplets with a non vanishing Brownian motion part and arbitrary Blumenthal-Getoor index. Note that in a low frequence observation scheme it is statistically hard to distinguish between infinitely many small jumps and a Brownian motion part. Hence such a slow logarithmic convergence rate is not surprising. In the complementary case of a compound Poisson process we establish further the convergence of a second estimator to the true Lévy copula with a square root convergence rate. Both convergence rates are optimal in the sense of the paper of Neumann and Reiss (2009).

247: 3

On the iterative plug-in algorithm for estimating diurnal patterns of financial trading durations

Yuanhua Feng, Sarah Anna-Eva Forstinger, Christian Peitz

University of Paderborn, Germany

This paper discusses the practical performance of the iterative plug-in bandwidth selector in a recently proposed semi-parametric autoregressive conditional duration (semi-ACD) model, which is defined by introducing a smooth nonparametric diurnal pattern into the standard ACD model. We focus particularly on studying the impact of different factors on the quality of the bandwidth selection in detail. A large simulation study with different ACD models, a typical and a non-typical diurnal pattern, a few methods for calculating the bandwidth for estimating the second derivative and a few methods for estimating the sum of innovation autocovariances is carried out for n = 8000, 16000 and 32000, respectively, with 400 replications in each case. To each of the simulated data, a semi-ACD model is fitted in order to identify which combination of the selected factors of the algorithm works best in practice. Based on the simulation results, the practical behavior of the proposed bandwidth selector is assessed in different ways. The results of the bandwidth selection, the scale function estimation as well as the ACD model parameter estimation are discussed. It is shown that the proposal works very well in practice, but some combinations of the above mentioned impact factors are superior to others. A further application to real data examples shows that the data-driven diurnal estimation works well and that the quality of the estimated ACD model is improved.

247: 4

Double-conditional smoothing of high-frequency volatility surface under a spatial model

Yuanhua Feng, Christian Peitz

Universität Paderborn, Germany

We want to investigate a spatial model for analyzing high frequency returns non-parametrically. The introduced model allows for examining and analyzing the slowly change of volatility over a long period of time as well as the daily volatility patterns. A double conditional kernel regression is proposed in order to estimate the mean surface and the volatility surface at the same time. The idea is to smooth the data over the time of day on a given day in a first step. The results obtained in the first step are then smoothed over all observed days. It is shown that our proposal is equivalent to a common two dimensional kernel regression. The running time, however, is much less than for the traditional approach. Moreover, step one -- the first conditional smoothing -- also already provides useful intermediate results. This idea works for both, Ultra High Frequency Data and adapted equidistant, i.e. regularly time spaced High Frequency data. Asymptotic results for the proposed estimators in the latter case are obtained under suitable stationarity conditions. Selected examples show that the proposal works very well in practice and the impact of the 2008 financial crisis on the volatility surface is further discussed briefly.

250: Invited session: Treatment effects in observational studies (Hernan, Vytacil)

Time: Tuesday, 19th Mar 2013: 2:50pm - 4:10pm · Location: KG II, Audimax
Session Chair: Claudia Schmoor
Session Chair: Bernd Fitzenberger

250: 1

Observational studies analyzed like randomized trials, and vice versa

Miguel Hernan

Harvard School of Public Health, United States of America

Randomized experiments and observational studies are often used to answer the same causal questions. For example, the effects of statins and of postmenopausal hormone therapy on the risk of coronary heart disease have been estimated both using randomized clinical trials and observational studies. Often the design of open-label randomized trials and of observational follow-up studies is similar, except that baseline treatment assignment is not randomly assigned in the latter. Because treatment choices and participation decisions after baseline are not randomized in neither randomized trials nor observational follow-up studies, time-varying confounding and selection bias may arise under both designs. Therefore randomized trials are just follow-up studies with baseline randomization. Leaving aside the adjustment for baseline confounders, which is generally necessary in observational follow-up studies, there are no reasons why the analysis of follow-up studies with and without randomization should differ. This talk reviews a framework for the analysis of both randomized trials and observational studies in the presence of time-varying confounding, and presents several applications.

250: 2

Instrumental Variables and the Sign of the Average Treatment Effect

Edward Vytlacil, Cecilia Machado, Azeem Shaikh

Yale University, United States of America

We establish conditions under which one can use linear instrumental variables to infer the sign of the Average Treatment Effect without imposing a linear model with constant coefficients. In order to obtain simple closed form results and for ease of exposition, we consider the case of a binary outcome variable, binary endogenous regressor, and binary instrument. We explore alternative sets of conditions that impose monotonicity either on the outcome equation, the equation for the endogenous regressor, or both. We use these conditions to analyze when linear instrumental variables identifies the sign of the Average Treatment Effect, and how to use the linear instrumental variables estimator to test monotonicity restrictions.

251: Statistics in Clinical and Preclinical Research -- Adaptive Designs

Time: Tuesday, 19th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1199
Session Chair: Ekkehard Glimm

251: 1

An approach for unplanned sample size changes in one-armed phase II cancer clinical trials

Stefan Englert, Meinhard Kieser

University of Heidelberg, Germany

Due to ethical considerations, phase II trials in oncology are typically performed with planned interim analyses. The sample sizes and decision boundaries are determined in the planning stage such that the type I and II error rates are controlled and have to be followed strictly later on. In practice, however, attaining the pre-specified sample size in each stage can be problematic.

In the talk, we focus on the open problem of unplanned changes in the sample size of the first stage. Green and Dahlberg (1992), Chen and Ng (1998) and Li and Heitjan (2012) developed methodology to allow a certain degree of flexibility. They impose, either by frequentist or Bayesian procedures, assumptions on the distribution of possible scenarios of over- or underrunning and develop designs with appropriate characteristics. From a methodical point of view, these designs control the type I error rate averaged over a range of first stage sample sizes, but not for each specific sample size that may occur in the course of the trial. Thereby, these methods allow reacting for unintentional sample size changes only.

We present a methodology that allows complete flexibility in first stage sample size with full control of the type I error rate. Underrunning or unplanned looks at the data can be easily handled with our approach. Overrunning of the sample size, however, requires conservative designs. The approach cannot be misused to hunt for a desirable result, as the initially planned design has superior characteristics as compared to all designs with unplanned changes.

251: 2

Blinded sample size reestimation in clinical trials with time-to-event endpoint

Tim Friede[1], Harald Pohlmann[2], Heinz Schmidli[2]

[1]Universitätsmedizin Göttingen, Germany; [2]Novartis Pharma AG, Switzerland

Two-stage designs are popular means to maintain the power of a hypothesis tests at a prespecified level independent of the size of the nuisance parameters by estimating the nuisance parameters from data of the first design stage and adjusting the sample size of the second stage accordingly. International guidelines emphasize the importance of controlling the type I error rate and maintaining of blinding, which might be interpreted as maintenance of trial integrity. The focus in the field of nuisance parameter based sample size reestimation (BSSR) so far has mainly been on continuous and binary endpoints [1, 2]. In the setting of time-to-event endpoints the power is driven by the number of events [3]. To predict the duration to achieve the necessary number of events from interim data the time to event process as well as the processes of time to dropout and recruitment have to be estimated [4-6]. In this presentation we present new flexible parametric procedures controlling the total duration of a trial, which are motivated and illustrated by trials in secondary progressive multiple sclerosis.

References:

[1] Friede T, Kieser M (2006) Biometrical Journal 48: 537-555.

[2] Proschan MA (2009) Biometrical Journal 51: 348–357.

[3] Schoenfeld DA (1983) Biometrics 39: 499-503.

[4] Whitehead J, Whitehead A, Todd S, Bolland K, Sooriyarachchi MR (2001) Statistics in Medicine 20: 165-176.

[5] Hade EM, Jarjoura D, Wei L (2010) Clinical Trials 7: 219-226.

[6] Todd S, Valdes-Marquez E, West J (2012) Pharmaceutical Statistics 11: 141–148.

251: 3

Entscheidungsregeln für eine datengesteuerte Wahl der Zielpopulation zur Untersuchung der Wirksamkeit neuer Therapien

Johannes Krisam, Meinhard Kieser

Institute of Medical Biometry and Informatics, University Medicine Heidelberg, Germany

Bei der Untersuchung einer neuen Therapie besteht oftmals Anlass zur Vermutung, dass die Behandlung in einer Subgruppe effektiver im Vergleich zur gesamten Patientenpopulation oder nur dort effektiv ist. Die Zielpopulation für den Wirksamkeitsnachweis wird üblicherweise auf Basis vorhandener Daten ausgewählt, beispielsweise Ergebnisse von Pilotstudien oder geplanten Zwischenauswertungen (siehe z.B. Brannath et al.,2009; Jenkins et al.,2011; Friede et al.,2012). Die Güte der angewandten Selektionsregel hat maßgeblichen Einfluss auf den

Erfolg einer klinischen Studie oder einer Arzneimittelentwicklung. Wir betrachten den Fall, dass die Selektion der Patientenpopulation anhand eines Biomarkers erfolgt, der die Subgruppe nicht notwendigerweise perfekt identifiziert, d.h. eine Spezifität und/oder Sensitivität unter 100 % aufweisen kann.

Wir entwickeln Methoden, die eine Evaluation der Charakteristika der Selektionsregeln für die Zielpopulation erlauben und somit eine adäquate Wahl einer geeigneten Strategie unterstützen. Insbesondere können die Verfahren zur Berechnung der Fallzahl, die notwendig ist, um eine vorgegebene Selektionswahrscheinlichkeit sicherzustellen, eingesetzt werden. Für die Situation von Unsicherheit über Parameterwerte, modelliert mittels a-priori-Verteilungen, werden optimale Selektionsregeln hergeleitet. Die Ergebnisse werden anhand von Beispielen demonstriert, und Konsequenzen für die Planung werden diskutiert.

References:

Brannath W, Zuber E, Branson M, Bretz F, Gallo P, Posch M, Racine-Poon A. Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy in oncology. Statistics in Medicine 2009;28:1445–1463.

Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharmaceutical Statistics 2011;10:347–356.

Friede T, Parsons N, Stallard N. A conditional error function approach for subgroup selection in adaptive clinical trials. Statistics in Medicine 2012.Epublished ahead of print(DOI:10.1002/sim.5541).

251: 4

Alternative views on setting clinical trial futility criteria

Paul Gallo[1], Lu Mao[2]

[1]Novartis Pharmaceuticals, United States of America; [2]University of North Carolina, United States of America

A feature increasingly utilized in clinical trial design is to allow a trial to stop early when it seems as if it will not likely achieve its primary efficacy objectives. This is commonly referred to as stopping a trial for futility, and can be motivated by ethical and financial considerations. The possibility of stopping for futility has implications for the operating characteristics of the trial design. A number of approaches have been described in the literature to set criteria for futility stopping, including rules based upon conditional power, predictive probability, beta spending functions, and others. We consider futility stopping from the point of view of quantifying and providing an objective sensible balance between risks of making incorrect decisions (e.g., stopping trials which should stop, and continuing trials which should continue), and make proposals for how specific considerations in individual trials can lead to choice of a sensible scheme. This approach is not specific to any of the particular scales in the literature such as those mentioned above, and we describe the interrelationship between criteria as they might be expressed on different scales. As futility may be evaluated multiple times in a long-term trial and the amount of information available at scheduled interim analyses may be difficult to predict in advance, we present some optimality criteria and discuss which of the well-known scales tend to produce schemes simple to implement and with better and more robust behavior across different timepoints at which futility might be evaluated within a given trial.

252: Statistics for High Dimensional Data

Time: Tuesday, 19th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1098
Session Chair: Arnold Janssen

252: 1

Estimation of the Global Minimum Variance Portfolio in High Dimensions

Nestor Parolya[1], Taras Bodnar[2], Wolfgang Schmid[1]

[1]European University Viadrina (FFO), Germany; [2]Humboldt-University of Berlin, Germany

We estimate the global minimum variance (GMV) portfolio in the high dimensional case using the results from the random matrix theory. This approach leads to the shrinkage-type estimators which are distribution-free and thus do not depend on the structure of the market. The first estimator is build using nonlinear shrinkage technique for large-dimensional covariance matrices. The second estimator, which we develop, is the optimal shrinkage estimator for GMV portfolio. We study in detail its asymptotic properties and show its theoretical advantages. The case we consider includes both the number of assets [fn]p[/fn] and the sample size [fn]n[/fn] tend to infinity so that [fn]p/n[/fn] tends to some [fn]c[/fn] which satisfies [fn]0<c<1[/fn]. Further, we provide the simulation study in order to compare the small- and large-sample behavior of derived estimator within the traditional sample estimator and other existent in the literature estimators of the GMV portfolio. We observe the significant improvements and robustness of the resulting estimator even for non-normally distributed data. At the end, we provide a brief discussion on the case when [fn]c>1[/fn].

252: 2

Copula-based models in multiple hypothesis testing

Thorsten Dickhaus, Taras Bodnar, Jakob Gierl

Humboldt-University Berlin, Germany

We are considered with simultaneous testing of a family of null hypotheses under the framework of a single statistical model. In this, we assume that the individual tests are carried out by means of (marginal) p-values and that these p-values, regarded as random variables, are dependent. Two popular type I error measures in multiple testing are the family-wise error rate (FWER) and the false discovery rate (FDR).

In the first part of the presentation, we express the threshold of an FWER-controlling simultaneous test procedure (STP) in the sense of Gabriel (1969) in terms of the copula function of the family of p-values, assuming that each of these p-values is marginally uniformly distributed on the unit interval under the corresponding null hypothesis. This offers the opportunity to exploit the rich and growing literature on copula-based modeling of multivariate dependency structures for the construction of STPs in non-Gaussian situations.

The second part is concerned with the role of Archimedean copulae in multiple testing in the case that p-values are elements of an infinite sequence of exchangeable random variables. We utilize analytic properties of Archimedean copulae to prove FDR control of the linear step-up test by Benjamini and Hochberg (1995) for such p-values.

References:

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc., Ser. B, Stat. Methodol. 57, 1, 289-300.

Gabriel, K. R. (1969). Simultaneous test procedures - some theory of multiple comparisons. Ann. Math. Stat. 40, 224-250.

252: 3

The Affinely Invariant Distance Correlation

Johannes Dueck[1], Dominic Edelmann[1], Tilmann Gneiting[1], Donald Richards[2]

[1]University of Heidelberg, Germany; [2]Pennsylvania State University, United States of America

Székely, Rizzo and Bakirov (2007) and Székely and Rizzo (2009), in two seminal papers, introduced the powerful concept of the distance correlation as a measure of dependence between sets of random variables. We study in this talk an affinely invariant version of the distance correlation and propose a consistent sample quantity. In the case of subvectors of a multivariate normally distributed random vector, we provide exact expressions for the distance correlation. For high-dimensional Gaussian data, as the dimensions of the subvectors converge to infinity, the direct calculation of affine distance correlation measures will return values which are virtually zero. Therefore, in high-dimensional settings distance correlation measures need to be rescaled properly.

To illustrate our results, we consider time series of wind vectors at the Stateline wind energy center in Oregon and Washington, and we derive the empirical auto and cross distance correlation functions between wind vectors at distinct meteorological stations.

252: 4

Verwendung von symmetrischen und asymmetrischen Abstandsmaßen für hochdimensionale Tests auf Unterschied, Äquivalenz, Überlegenheit und Nichtunterlegenheit

Kai Antweiler, Siegfried Kropf

Otto-von-Guericke-Universität Magdeburg, Germany

Tests auf der Basis von (üblicherweise symmetrischen) Ähnlichkeits- oder Abstandsmaßen sind ein effektives Instrument für Vergleiche mit hochdimensionalen Beobachtungsvektoren. Im einfachen Test auf Unterschiede können Permutationstests und im parametrischen Fall Rotationstests durchgeführt werden. Beides sind exakte Tests, die sich auch bei erstaunlich kleinen Stichroben noch anwenden lassen (Kropf et al., 2004; Kropf, Adolf, 2009).

In neueren Arbeiten nutzen wir die abstandsbasierten Tests für multivariate Äquivalenzfragestellungen. Da die Nullhypothese hier keine Punkthypothese mehr ist, sind Permutations- und Rotationstechniken nicht mehr einfach zu benutzen. Jedoch konnte in Leave-one-out-Techniken eine gute Approximation für hochdimensionale, gleichskalierte Variablen festgestellt werden.

Für einseitige Fragestellungen im Überlegenheitstest müssen unsymmetrische Abstandsmaße verwendet werden, da man sonst nicht zwischen „besser“ und „schlechter“ unterscheiden kann. Es werden Vorschläge zur Nutzung solcher Teststatistiken unterbreitet und demonstriert und auftretende Schwierigkeiten diskutiert.

Bei Nichtunterlegenheitsstudien werden ebenfalls unsymmetrische Abstandsmaße benötigt, aber hier zusätzlich unter der Bedingung einer Nullhypothese, die ganze Bereiche einschließt, so dass hier wieder Leave-one-out-Techniken zum Einsatz kommen.

Für die verschiedenen Fragestellungen werden Simulationsuntersuchungen zur Kontrolle des Fehlers erster Art vorgestellt, da bis auf den zweiseitigen Test auf Unterschied alle anderen Testverfahren approximativ sind.

References:

Kropf, S., Heuer, H., Grüning, M., Smalla, K. (2004). Significance test for comparing complex microbial community fingerprints using pairwise similarity measures. Journal of Microbiological Methods 57/2, 187-195.

Kropf, S., Adolf, D. (2009). Rotation test with pairwise distance measures of sample vectors in a GLM. Journal of Statistical Planning and Inference 11, 3857-3864.

253: Penalized and Regularized Regression Analysis

Time: Tuesday, 19th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1010
Session Chair: Göran Kauermann

253: 1

Corrected Confidence Bands for Functional Data Using Principal Components

Jeff Goldsmith[1], Sonja Greven[2], Ciprian Crainiceanu[3]

[1]Columbia University, United States of America; [2]Ludwig-Maximilians-Universität München, Germany; [3]Johns Hopkins University, United States of America

Functional principal components (FPC) analysis is widely used to decompose and express functional observations. Curve estimates implicitly condition on basis functions and other quantities derived from FPC decompositions; however these objects are unknown in practice. In this paper, we propose a method for obtaining correct curve estimates by accounting for uncertainty in FPC decompositions. Additionally, pointwise and simultaneous confidence intervals that account for both model-based and decomposition-based variability are constructed. Standard mixed-model representations of functional expansions are used to construct curve estimates and variances conditional on a specific decomposition. A bootstrap procedure is implemented to understand the uncertainty in principal component decomposition quantities. Iterated expectation and variance formulas combine both sources of uncertainty by combining model-based conditional estimates across the distribution of decompositions. Our method compares favorably to competing approaches in simulation studies that include both densely- and sparsely-observed functions. We apply our method to sparse observations of CD4 cell counts and to dense white-matter tract profiles. Code for the analyses and simulations is publicly available, and our method is implemented as the IVfpca() function in the R package refund on CRAN.

253: 2

[fn]L_0[/fn]-Penalization for Categorical Effects

Margret-Ruth Oelker, Gerhard Tutz

Ludwig-Maximilians-Universität München, Germany

In regression modeling, one categorical covariate with [fn]k+1[/fn] levels results in [fn]k[/fn] coefficients. Hence, the number of coefficients can become huge if several covariates are available. To reduce model complexity, coefficients of similar categories should be fused and coefficients of non-influential categories should be set to zero. To this end, Lasso-type penalties on the differences of coefficients are a standard approach. However, the clustering/selection performance of this approach is sometimes poor - especially when adaptive weights are bad conditioned or not existing. To overcome this, a [fn]L_0[/fn]-norm penalty on the differences of coefficients is proposed. The proposed penalty shows nearly no shrinkage effects and the clustering performance is enhanced. Applied to pure coefficients and with according tuning, the proposed penalty corresponds to model selection with information criteria AIC and BIC. Allowing for the fusion of categories, the proposed penalty extends model selection with AIC/BIC. For estimation, the optimization problem is approximated quadratically. Solutions seem to be stable -- even though the penalty is neither concave nor convex. Numerical experiments in the framework of generalized linear models are promising. For illustration, data on the unemployment rates in Germany between 2005 and 2010 is analyzed.

The contribution has been withdrawn.

Degrees of Freedom Test for Penalized B-Splines in Additive Models

Helene Roth, Stefan Lang

University of Innsbruck, Austria

253: 4

Penalized Approach for Estimating Party Positions with Time-Varying Word Weights from Political Texts

Carsten Jentsch, Julian Koltes, Eun Ryung Lee, Enno Mammen, Carsten Trenkler

University of Mannheim, Germany

Political texts offer a lot of information and provide a cheap and freely accessible data source for policy positions of political actors. As the structure of raw text is sophisticated, statistical inference is commonly based on word count data obtained from party manifestos of different parties at different time points.

We use combinations of LASSO and fused LASSO to make inference on the evolution of party positions. Relevant words describing the differences between manifestos are selected and their weights are estimated. The approach relies on Slapin and Proksch (2008), who model word counts via Poisson distributions with parameters depending parametrically on word, party and time.

Our main contribution is to allow for time dependency of the words' discriminating effects, i.e. the political lexicon can change over time. This

extends the approach of Slapin and Proksch (2008) and allows us to identify changes in political debates and to get more insights into the use of words by left and right-wing parties.

We apply our approach to German party manifestos of five parties over five federal elections after German reunification. Although the relaxation of the model assumptions introduces a large amount of additional parameters, penalization makes it possible to deal with the resulting high-dimensional setup. This allows us to fit a model of about 54000 parameters to a (9000 [fn]\times[/fn] 25) data matrix. Simulation studies confirm that our procedure is robust, runs stable and leads to meaningful and interpretable results. Non-convexity of the models makes the implementation rather challenging.

254: Classification

Time: Tuesday, 19th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1015
Session Chair: Hans Kestler

254: 1

The Blackwell Prediction Algorithm for 0-1 Sequences and a Generalization

Hans Rudolf Lerche

Universität Freiburg , Germany

Let [fn]x^{1}, x^{2}, ...[/fn] be a not necessary random but arbitrary long 0-1 sequence. We wish to sequentially predict the sequence. This means that, for each [fn]n>1[/fn] we will guess the value of [fn]x^{n+1}[/fn], basing our guess on knowledge of [fn]x^{1},x^{2},..x^{n}[/fn]. The Blackwell-Algorithm is a procedure which predicts well for all 0-1 sequences. We explain this algorithm and extend it to more than two categories. The three-category algorithm will be explained using a geometric model (the so called prediction prism). It has some special geometric properties, which also hold for an arbitrary number of categories. The results show nicely a curse of dimension.

254: 2

On Instance Selection in Multi Classifier Systems

Friedhelm Schwenker

Ulm University, Germany

In any data mining application the training set design is the most important part of the overall data mining process. Designing a training set means pre-processing the raw data, selecting the relevant features, selecting the representative instances (samples), and labeling the instances for the classification or regression application at hand. Labeling data is usually time consuming, expensive (e.g. in cases where more than one expert must be asked), and error-prone. Instance selection deals with searching for a subset [fn]S[/fn] of the original training set [fn]T[/fn], such that a classifier trained on [fn]S[/fn] shows similar, or even better classification performance than a classifier trained on the full data set [fn]T[/fn] (Olvera-López et al 2010). We will present confidence-based instance selection criteria for [fn]k[/fn]-nearest-neighbor classifiers and probabilistic support vector machines. In particular we propose criteria for multi classifier systems and discuss them in the context of classifier diversity. The statistical evaluation of the proposed selection methods has been performed on affect recognition from speech and facial expressions. Classes are not defined very well in this type of application leading to data sets with high label noise. Numerical evaluations on these data sets show that classifiers can benefit form instance selection not only in terms of computational costs, but even in terms of classification accuracy.

References:

Olvera-López, J. A. and Carrasco-Ochoa, J. A. and Martinez-Trinidad, J. F. and Kittler, J. (2010): A review of instance selection methods. Artif. Intell. Rev. 34(2), 133-143

254: 3

Data Integration via Network Smoothed T-Statistics

Holger Fröhlich

Universität Bonn, Germany

Predictive, stable and interpretable gene signatures are generally seen as an important step towards a better personalized medicine. During the last decade various methods have been proposed for that purpose. However, one important obstacle for making gene signatures a standard tool in clinics is the typical low reproducibility of these signatures combined with the difficulty to achieve a clear biological interpretation. For that purpose in the last years there has been a growing interest in approaches that try to integrate information from molecular interaction networks.

We here propose to a technique to smooth gene-wise t-statistics over the structure of a literature derived protein-protein interaction network. Compared to several other competing methods our algorithm reveals a significantly better prediction performance and higher signature stability. Moreover, obtained gene lists are highly enriched with known disease genes and drug targets. We show, how our approach can be extended further by a-priori integration of information on candidate disease genes and known somatic and germline mutations, which can lead to additional stabilization of the gene signature. Finally, we demonstrate that our method also allows for integrating gene and miRNA expression data for classification in a natural manner. The latter can improve classification performance additionally in some cases.

254: 4

Reduction of Feature and Instrument Space - Robustness of Chord Recognition

Markus Eichhoff[1], Igor Vatolkin[2], Claus Weihs[1]

[1]Computational Statistics, Technische Universität Dortmund, Germany, Germany; [2]Algorithm Engineering, Technische Universität Dortmund,

Germany, Germany

Instrument recognition is a challenging task especially for polyphonic recordings. This study focuses on chords of three single tones being played simultaneously. The single tones are derived by the three databases Iowa (2011), McGill (2010) and RWC (2003). A subset [fn]S[/fn] of 18 musical instruments from five classes (acoustic-pluck, electric-pluck, piano, strings, wind) is used to calculate high-level audio features (see Eichhoff and Weihs (2010)). Classification tasks are carried out by different supervised classification methods.Variable Selection and a pre-filtering method is further integrated to reduce dimension of feature space. Target variable per instrument class is 0/1 (instrument class inside or not) and hamming-loss is calculated to aggregate the five error rates.

The target of our study is to estimate a reasonable instrument subset [fn]S[/fn] which is large enough to represent the complete set of instruments and small enough to keep the classification performance acceptable. A variant of sequential forward selection is carried out to find a minimal subset S and a minimal classification error. In each step of the forward selection 2000 chords for training and 1000 chords for testing are created.

References:

M. Eichhoff, C. Weihs. Musical Instrument Recognition by High-Level Features, Proc. of the 34th Annual Conf. of the German Classification Society, pp. 373-381, 2010

University of Iowa. Electronic Music Studios. Musical instrument samples, 2011

McGill University. Mcgill master samples collection on DVD, 2010

M. Goto, H. Hashiguchi, T. Nishimura, R. Oka. RWC music database: Music genre database and musical instrument sound database, ISMIR 2003 Proceedings, 2003

261: Statistics in Clinical and Preclinical Research -- Non-clinical Statistics

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG II, Audimax
Session Chair: Hannes-Friedrich Ulbrich

261: 1

A Random Effects Binomial Model for the Distribution of Counts in Developmental Toxicity Experiments

Mehdi Razzaghi

Bloomsburg University, United States of America

To assess the developmental effects of a toxic substance, bioassay experiments are conducted on laboratory animals in controlled environments. In developmental toxicity experiments, pregnant female animals are exposed to a dose of a toxic chemical during a critical time of the gestation period. The animals are generally sacrificed just before term and the content of the uterus is examined for a variety of outcomes. One such outcome may be the number of live fetuses or the number of fetuses with a certain effect such as malformation. Traditionally, binomial models with fixed cluster sizes have been used to analyze these experiments. Since the number of fetuses in a litter is variable, more appropriate approaches with random cluster sizes have become popular in recent years. Here, we introduce a new approach for analyzing count data from developmental toxicity studies. Our approach is based on using the Lindley distribution to describe the random effects. We show that if the number of fetuses is modeled by a Poisson distribution, then the count data follows an extension of the Poisson-Lindley distribution. The properties of this new distribution are discussed and maximum likelihood estimates of model parameters are derived. An example is used for illustration.

261: 2

Remarks on the interpretation of limits of agreement based on small samples

Michael Vock

University of Bern, Switzerland

Limits of agreement (LOA) are widely used to describe the differences between two methods of measuring the same quantity. The usual interpretation of, e.g., 95% LOA is that 95% of the differences between the two methods should lie within these LOA. For large sample sizes and normally distributed differences, this interpretation of LOA is approximately correct for all commonly used definitions of LOA.

However, caution should be used in the case of small sample sizes since different definitions of LOA may disagree -- and they may have markedly different interpretations: Simple definitions just provide estimates for quantiles, while more complicated definitions provide prediction intervals for a future difference. In addition, even for the latter type of LOA, an interpretation in the way mentioned above is dangerous since, for small samples, the probability mass effectively covered by a single realization of a prediction interval may be markedly different from the targeted coverage probability.

These problems are pointed out from a theoretical perspective and illustrated using examples and simulations.

261: 3

Bead-basierte Liquidarray-Technologie: Vergleich unterschiedlicher Vorverarbeitungsmethoden

Andrea Senft[1], Daniel C. Chamrad[2], Andreas Ziegler[1]

[1]IMBS, Universität zu Lübeck, Germany; [2]Protagen AG, Dortmund, Germany

In der Immunproteomik erlauben Hochdurchsatzverfahren inzwischen die parallele Analyse von mehreren tausend unterschiedlichen Proteinen. Um systematische Verzerrungen in den Daten, wie z.B. Batch-Effekte, vor der statistischen Analyse zu verringern, muss eine Vorverarbeitung der experimentellen Daten erfolgen. Elementare Bestandteile der Vorverarbeitung sind insbesondere die Transformation und die Normalisierung der gemessenen Werte. Dabei ist es entscheidend, aus der Fülle vorhandener Transformations- und Normalisierungsmethoden anhand objektiver Kriterien eine für die Daten geeignete Kombination auszuwählen. Während für anderen „Omics“-Technologien bereits entsprechende Standardvorgehensweisen der Vorverarbeitung etabliert sind, existiert bislang kein standardisiertes Vorgehen für Bead-basierte Immunarrays. Deshalb vergleichen wir anhand von Realdaten zahlreiche Kombinationen bekannter Transformations- und Normalisierungsmethoden. Als Transformationsmethoden werden z.B. die Log-Transformation oder die Box-Cox-Transformation herangezogen. Betrachtete Normalisierungsmethoden sind u.a. LOESS sowie die Quantilnormalisierung. Die verschiedenen Kombinationen werden anhand unterschiedlicher Bewertungskriterien beurteilt, um die für die Datensituation am besten geeignete Vorverarbeitungsmethode zu identifizieren.

261: 4

Stochastic models for brain aggregate cell cultures: Quantitative description of neurotoxicity

Maria Renner[1], Marie-Gabrielle Zurich[2], Annette Kopp-Schneider[1]

[1]Division of Biostatistics, DKFZ, Germany; [2]Departement de Physiologie, Universite de Lausanne, Switzerland

Brain aggregate cell cultures are three-dimensional in vitro cultures which offer a unique system for neurotoxicity testing. The cultures contain the different brain cell types which are able to interact in a physiological manner. The populations are present in proportions similar to those in the brain of an adult rat in vivo. For toxicity testing the cultures are exposed to different concentrations of a compound at several time points. Cell-type specific toxicity of a compound is assessed by routine protocols for qRT-PCR analyses and enzyme activity measurements at several time points.

We have developed a Markov model for the size and behavior of a single brain cell population which incorporates cellular stress and death in continuous time with concentration-dependent transition rates. Since cell numbers are not directly measurable, intracellular lactate dehydrogenase (LDH) activity is used as a surrogate resulting in a stochastic activity model. A functional relationship between fold changes of genes and the Markov model is established.

Maximum likelihood and least squares regression techniques are applied for estimation of the transition rates. Likelihood ratio tests are performed to test hypotheses about the transition rates. Simulation studies are used to investigate the performance of the transition rate estimators and to analyze the error rates of the likelihood ratio tests. The activity model is applied to experimental data and describes transitions from healthy to stressed cells and from stressed cells to cellular death. The estimated transition rates are interpreted in a mechanistic and quantitative manner.

262: Survival and Event History Analysis -- Frailties, stratifizierte Analysen

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG III, HS 3044
Session Chair: Steffen Unkel

262: 1

Interpretation stratifizierter Überlebenszeitanalysen unter Verwendung der verbleibenden Lebenserwartung

Benjamin Mayer[1], Silvia Sander[1], Marko Kornmann[2]

[1]Institut für Epidemiologie und Medizinische Biometrie, Universität Ulm, Germany; [2]Klinik für Allgemein- und Viszeralchirurgie, Universität Ulm, Germany

Hintergrund: FOGT1 war eine randomisierte Phase III Studie zur adjuvanten Therapie bei Kolonkarzinom (1992-1999), deren Patienten seither nachbeobachtet werden. Die dreiarmige Studie untersuchte den adjuvanten chemotherapeutischen Effekt der Modulatoren Folinsäure (Arm B) bzw. Interferon α-2a (Arm C) im Vergleich zur adjuvanten Standardbehandlung (Arm A). Eine stratifizierte Überlebenszeitanalyse nach Therapie und Alter der Patienten (< 70 Jahre vs. ≥ 70 Jahre) zeigte für die günstigste Therapie (B) nahezu gleich große 5-Jahres-Überlebensraten für beide Altersgruppen, obwohl sich die Patienten bezüglich ihrer restlichen Lebenserwartung teilweise deutlich unterscheiden.

Methoden: Zur Klärung der Frage, ob ältere Patienten von Therapie B besonders profitieren, wurde für alle Patienten (n=855) der Quotient QSL aus Überlebenszeit und noch verbleibender Lebenserwartung (Daten des Statistischen Bundesamtes) berechnet. Neben einer deskriptiven und teststatistischen Auswertung dieses Quotienten wurden zudem entsprechende Korrelationsanalysen auf Basis der verbleibenden Lebenserwartung durchgeführt.

Ergebnisse: Eine nach Altersgruppe und Therapie stratifizierte Analyse des Quotienten aus Überlebenszeit und verbleibender Lebenserwartung zeigte eine mediane Rate von 0.37 für ältere Patienten in Therapiearm B gegenüber medianen Raten von jeweils 0.34 für ältere Patienten in den Therapiearmen A und C (p=0.57). Der lineare Zusammenhang zwischen QSL und Alter (r=0.46) war für Patienten in Arm B am höchsten. Die Analysen stützen die implizit beobachtete Vermutung, dass insbesondere ältere Patienten des Kollektivs von Behandlung mit Therapie B profitieren.

Fazit: Der aus Überlebenszeit und verbleibender Lebenserwartung ermittelte Quotient kann mit zu diskutierenden Einschränkungen verwendet werden, um den in der stratifizierten Überlebenszeitanalyse implizit beobachteten, positiven Zusammenhang von Therapie B und höheren Lebensalters anschaulich zu quantifizieren.

262: 2

Frailty-Modelle für Lebensdauerdaten der Carla-Studie

Diana Pietzner, Beatrice Herzog, Andreas Wienke

IMEBI Halle, Germany

Für die Carla-Studie [1] wurde eine Zufallsstichprobe von 1779 Personen aus der 45- bis 80-jährigen Bevölkerung der Stadt Halle gezogen mit dem Ziel Risikofaktoren für kardiovaskuläre Erkrankungen in der Bevölkerung zu untersuchen. Von 2002 bis 2005 wurden die Baseline-Untersuchungen durchgeführt. Ca. 6.5% der Probanden sind während der Nachbeobachtungszeit bis Ende 2009 verstorben. Es wird eine Lebensdaueranalyse durchgeführt, dabei werden verschiedene Risikofaktoren wie Blutdruck, BMI, Diabetes und Rauchgewohnheiten eingeschlossen. Sozioökonomische Einflüsse des Stadtteils können mithilfe eines log-normalen Frailty-Models berücksichtigt werden. Für nichterhobene Kovariablen kann mithilfe eines univariaten Frailty-Models adjustiert werden. Die Ergebnisse werden mit denen eines traditionellen Cox-Modells verglichen.

References:

[1] Greiser KH, Kluttig A, Schumann B, Kors JA, Swenne CA, Kuss O, Werdan K, Haerting J: Cardiovascular disease, risk factors and heart rate variability in the elderly general population: Design and objectives of the CARdiovascular disease, Living and Ageing in Halle (CARLA) Study. BMC Cardiovascular Disorders 2005, 5:33

262: 3

Vergleich von Konfidenzintervallen für zufällige Effekte im log-normalen Frailty-Modell

Katharina Hirsch, Andreas Wienke, Oliver Kuß

IMEBI Halle, Germany

Oft ist bei der Analyse von multivariaten Ereigniszeiten die Assoziationsstruktur der Daten von primärem Interesse. Wegen ihrer Parallelen zu gemischten Modellen bieten sich hier log-normale Frailty-Modelle mit abschnittsweise konstanten Baseline-Hazardfunktionen zur Modellierung an [1]. Jedoch ist die Schätzung von zufälligen Effekten generell mit erheblichen Herausforderungen verbunden. So ist die Verteilung der Schätzwerte im Allgemeinen nicht symmetrisch da der Parameterraum auf die positive Halbachse eingeschränkt ist. Damit ergeben sich zusätzliche Fragen für die Konstruktion von Konfidenzintervallen. Hier soll deshalb untersucht werden, welche Art von Konfidenzintervallen für die Schätzung der zufälligen Effekte in log-normalen Frailty-Modellen am besten geeignet sind. Betrachtet werden

das symmetrische- und asymmetrische Wald-Konfidenzintervall sowie das Profile Likelihood Konfidenzintervall. Dabei wird ihre Präzision und Überdeckungsrate [2] untersucht, wobei große Unterschiede erwartet werden.

References:

[1] Hirsch, K.;Wienke, A.; Kuss, O.: Log-normal frailty models fitted as Poisson generalized linear mixed models; Statistics in Medicine (submitted)

[2] Burton, A.; Altman, D. G.; Royston, P. & Holder, R. L.: The design of simulation studies in medical statistics; Statistics in Medicine, Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, 2006, 25, 4279 - 4292

263: Statistics for High Dimensional Data

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1098
Session Chair: Harald Binder

263: 1

Assigning statistical significance in high-dimensional problems

Peter Bühlmann

ETH Zurich, Switzerland

High-dimensional data, where the number of variables is much larger than sample size, occur in many applications nowadays. During the last decade, remarkable progress has been achieved in terms of point estimation and computation. However, one of the core statistical tasks, namely to quantify uncertainty or to assign statistical significance, is still in its infancy for many problems and models. After a brief review of some main concepts for high-dimensional statistical inference, we present procedures and corresponding theory for quantifying significance in single and multiple testing problems. Illustration on various examples highlights the methods' user-friendly nature for high-dimensional data analysis.

263: 2

Two-Sample Testing in High-Dimensional Models

Nicolas Staedler, Sach Mukherjee

Netherlands Cancer Institute, Amsterdam, The Netherlands

We propose novel methodology for testing equality of model parameters between two high-dimensional populations. The technique is very general and applicable to a wide range of models. The method is based on sample splitting: the data is split into two parts; on the first part we reduce the dimensionality of the model to a manageable size; on the second part we perform significance testing (p-value calculation) based on a restricted likelihood ratio statistic. Assuming that both populations arise from the same distribution, we show that the restricted likelihood ratio statistic is asymptotically distributed as a weighted sum of chi-squares with weights which can be efficiently estimated from the data. Arbitrary splitting of the data results in a "p-value lottery". In order to get reproducible results we repeat the splitting procedure various times and aggregate the resulting p-values. This multi-split approach provides improved p-values. We illustrate the use of our general two-sample approach in two-sample comparison of high-dimensional regression models ("differential regression") and graphical models ("differential network"). In both cases we show results on simulated data as well as real data from recent, high-throughput cancer studies.

263: 3

Building multivariable models for RNA-Seq data

Isabella Zwiener[1,2], Barbara Frisch[1], Harald Binder[1]

[1]Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Germany; [2]Center for Thrombosis and Hemostasis, University Medical Center Mainz, Germany

RNA-Seq is a new technology to measure gene expressions. It can be used to build signatures that predict the clinical outcome of patients, e.g. group membership or survival time. Prognostic signatures can be derived by multivariable model building. Expression data arising from RNA-Seq experiments are count data; data derived from genes with large mean count values also have large variances. The state-of-the-art statistical analysis to select differentially expressed genes from RNA-Seq data is a 2-step process including normalization and univariable testing. Existing normalization methods transform the data to a scale independent of the total number of counts per sample. The usual mean-variance dependence of count data has not been considered to date. In multivariable models the selection probability of a gene may depend on the variance, so that further transformation could be necessary. In a simulated two-group setting we compare multivariable models using different transformations of RNA-Seq data to account for the mean-variance dependence. As we do not know the true underlying scale of the effects, we simulate different scales for the true effects. Different transformations lead to different sensitivities and different lengths of selected genes. Transformation performance depends on the scenario for the scale of the effect. We illustrate the application using RNA-Seq data from kidney renal clear cell carcinoma patients, with survival time as the clinical endpoint. In this application, the selected genes depend on the transformation, too. This illustrates the need of appropriately handling the scale and distribution of RNA-Seq data in building multivariable models.

264: Penalized and Regularized Regression Analysis

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1010
Session Chair: Sonja Greven

264: 1

Copula based approach for the estimation of sample selection models involving semi-parametric predictors

Malgorzata Wojtys, Giampiero Marra

University College London, United Kingdom

Sample selection is a problem frequently encountered in applied research, arising when a response variable of interest is observed only for a restricted, non random sample. We consider a sample selection model defined by a system of two regression models: one for the outcome equation and the other for the selection equation, each of which contains a parametric linear part, a non-parametric component in the form of a sum of smooth functions of continuous covariates and a random error term. Selectivity arises if error terms for these two equations are mutually dependent.

We propose a fitting method based on penalized likelihood simultaneous equation estimation where the smooth functions are approximated using the regression spline approach and the bivariate distribution of errors is represented in the form of an Archimedean copula. This approach allows to capture a wide range of bivariate dependence and is a generalization of the existing method implemented in R package SemiParSampleSel where the errors are assumed to follow the bivariate normal distribution.

The procedure will be illustrated with numerical examples for which a number of popular copulas and marginal distributions of the outcome will be considered.

264: 2

Flexible Distributed Lags

Viola Obermeier[1], Fabian Scheipl[1], Christian Heumann[1], Joachim Wassermann[2], Helmut Küchenhoff[1]

[1]Department of Statistics, Munich, Germany; [2]Department of Earth-and Environmental Sciences, Munich, Germany

Regression models with effects of lagged covariates are often used in biostatistical and geophysical data analysis. In this talk, we present a novel penalty structure for interpretable and flexible estimates of lag coefficients based on spline representations. We provide a user-friendly implementation of our flexible distributed lag approach that can be used directly in the established R package mgcv for estimation of generalized additive models. This allows our approach to be immediately included in complex additive models for generalized responses even in hierarchical or longitudinal data settings, making use of established robust and well-tested inference algorithms. The benefit of flexible distributed lag modeling is shown in a simulation study. We demonstrate the performance and utility of the proposed flexible distributed lag model in a case study on earthquake data from Mount Hochstaufen, Bavaria with focus on the specific shape of the lagged rain influence on the occurrence of earthquakes in different depths.

264: 3

Adaptive Semiparametric M-Quantile Regression

Fabian Sobotka[1], Nicola Salvati[2], Giovanna Ranalli[3], Thomas Kneib[1]

[1]University Göttingen, Germany; [2]University Pisa, Italy; [3]University Perugia, Italy

M-Quantiles combine the robustness and interpretability of quantiles with the flexibility and intuitive estimation of expectiles. They allow for an iteratively weighted least squares estimation including quadratic penalties to incorporate a semiparametric model. The inclusion of p-splines and spatial effects, like Markov random fields, is possible. And by definition their estimate is still robust against outliers. However, this is only true for homoscedastic scenarios. In heteroscedastic cases the distinction between outliers and "trustable" observations is likely to fail. Here, we introduce adaptive M-Quantile regression models to overcome this problem by replacing the tuning constant of the M-Quantile estimation with a robustness curve. The latter is constructed from the scale part of a location-scale model. Our findings will be analysed in a simulation study and made available as R-package "Mreg".

264: 4

Boosted beta regression -- an analysis tool for bounded response variables

Matthias Schmid[1], Nora Fenske[2], Florian Wickler[2], Andreas Mayr[1]

[1]University of Erlangen-Nuremberg, Germany; [2]University of Munich, Germany

Regression analysis with a bounded outcome is a common problem in applied statistics. Typical examples include prediction models for percentage outcomes and ratings that are measured on a bounded scale. The focus of the talk is on beta regression [1], which is a convenient tool for analyzing bounded responses. The classical approach to fit a beta regression model is to use maximum likelihood

estimation [2,3] with subsequent AIC-based variable selection. As an alternative to this established - yet unstable - approach, we propose a new estimation technique called boosted beta regression. Boosted beta regression is based on the gamboostLSS algorithm [4], which implies that both estimation and variable selection are carried out simultaneously in a highly efficient way. Additionally, both the mean and the variance of a bounded response can be modelled using flexible nonlinear covariate effects. As a consequence, the new method accounts for common problems such as overdispersion and non-binomial variance structures.

References:

[1] Ferrari S., Cribari-Neto F. (2004). "Beta regression for modelling rates and proportions." Journal of Applied Statistics 31(7), 799-815.

[2] Cribari-Neto F., Zeileis A. (2010). "Beta regression in R". Journal of Statistical Software 34(2).

[3] Stasinopoulos M., Rigby B. (2012). "gamlss.dist – distributions to be used for GAMLSS modelling." R package version 4.2-0.

[4] Mayr A., Fenske N., Hofner B., Kneib T., Schmid M. (2012): "Generalized additive models for location, scale and shape for high dimensional data". JRSS C 61(3), 403-427.

265: Observational Studies -- Specific Modelling in Theory and Practice

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1199
Session Chair: Peggy Sekula

265: 1

Estimating Heterogeneous Growth Curve Models

Holger Brandt, Andreas Klein, Jana Gäde

Goethe-Universität Frankfurt, Germany

Analyzing longitudinal data with latent growth curve models (LGCM) has become a common procedure in social and behavioral sciences. In LGCMs, individual growth trajectories are modeled within the Structural Equation Modeling framework. While for ordinary LGCMs homogeneous slope variances are assumed, in many empirical situations this assumption is violated, because growth trajectories can be predicted more precisely for some subjects than for others. Here, we present an ML estimator for a heterogeneous growth curve model (HGCM; Klein & Muthén, 2006) that allows a specification of heterogeneous growth trajectories. The HGCM is a parsimonious complementary model to growth mixture models. In contrast to growth mixture models that assume distinct latent classes, the HGCM assumes a continuous change of the variance of the growth trajectories for subjects with different start points. We illustrate the model by an empirical data set from the HIV prevention sciences.

265: 2

Bias correction in linear mixed models with a deficient covariate – an application to effects of ultrafine particles on human health

Veronika Fensterer[1], Helmut Küchenhoff[1], Josef Cyrys[2,3], Susanne Breitner[2], Alexandra Schneider[2], Annette Peters[2]

[1]Statistical Consulting Unit, Department of Statistics, Ludwig-Maximilians-Universität (LMU), Munich, Germany; [2]German Research Center for Environmental Health (HMGU), Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany; [3]Environment Science Centre (WZU), University of Augsburg, Germany

In order to examine short-term effects of ultrafine particles (UFP) on human health personal measurements of particle number concentration (PNC) were collected within the “Augsburger Umweltstudie”. In parallel we recorded PNC at a fixed monitoring site located in an urban background. Apart from classical measurement error of the mobile devices they occasionally break down. Therefore the combination of personal exposure measurements with data obtained at the fixed monitoring site was investigated and bias due to measurement error was corrected using validation data collected within a DFG-funded project.

The method-of-moments allowing for classical measurement error is extended to the case of linear mixed models with autocorrelated residuals. Additionally, more complex error types are examined involving a mixture between Berkson and classical measurement error, random effects in the measurement error and autocorrelated measurement error. Whereas random effects in the measurement error have little influence on the bias of classical error, they become relevant regarding mixture error. Autoregressive measurement error of order one affects the well-known form of the bias through weighting the variance of the respective component; this may lead to changes of the bias in both directions. A correction method in the presence of additional covariates, which are associated with the deficient measurement, is also proposed. Distributional properties of the corrected estimators are used for the construction of confidence intervals with the delta method and Bootstrap.

Besides the theoretical findings, asymptotic characteristics of the correction factor are presented with results from simulations and the methods are applied to the “Augsburger Umweltstudie”.

265: 3

Analysis of low birth weight -- a new approach using a competing risk model with a continuous marker

Reinhard Meister[1], Arthur Allignol[2]

[1]Mathematics Department, Beuth Hochschule für Technik Berlin, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Birth weight is an important indicator and predictor of well-being of the newborn at time of delivery and for the future. As the embryo undergoes a dramatic growth during pregnancy, birth weight depends strongly on gestational age at delivery. In contrary, most approaches for assessing the proportion of low birth weight infants are based on birth weight solely. The WHO defines (very) low birth weight as values below 2500g (1500 g). Using percentile curves for birth weight, we can define pregnancy outcome more specifically. Intra uterine growth reduction (IUGR) is specified as birth weight below a p-percentile, assessing birth weight relative to gestational age. This defines a new category of competing risks for pregnancy outcome, and can be treated similar to other final states in a multi-state mode. In addition, we will show how standardized birth weights -- the continuous marker -- can be analyzed appropriately. Application to real cohort data of pregnancy outcome will be presented for illustration of the approach.

Our proposed analysis improves the common approach substantially: unbiased estimation of cumulative incidences, differentiation between preterm birth and IUGR, and analysis of time dependent covariates is enabled.

265: 4

Diagnostic accuracy measures for the relationship between carotid intima-media thickness and the Metabolic syndrome in adolescents

André Scherag[1], Carolin Pütter[1], Henrike Berg[1], Rainer Wunsch[2], Thomas Reinehr[3]

[1]Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University of Duisburg-Essen, Germany; [2]Department of Pediatric Radiology and Ultrasound, Vestische Hospital for Children and Adolescents Datteln, University of Witten/Herdecke, Germany; [3]Department of Pediatric Endocrinology, Diabetes and Nutrition Medicine, Vestische Hospital for Children and Adolescents Datteln, University of Witten/Herdecke, Germany

Objectives: The concept of Metabolic Syndrome (MetS) is discussed controversially in adolescence. It is unclear whether fulfilling the definition is more predictive for morbidity and mortality than the sum of the components of MetS.

Study design: This study included 461 overweight adolescents aged 10-18 years (median BMI 28.6kg/m2). We analyzed the relationships between two definitions of the MetS (IDF and Weiss) and carotid intima-media thickness (IMT), which is predictive for later cardiovascular diseases. We used regression models, areas under the receiver operating characteristics curve (AUCs) and new measures of diagnostic performance (e.g. Pencina et al., 2012) for increased IMT as gold standard (defined by ≥0.7mm).

Results: At the group level, quantitative IMT was associated with BMI, blood pressure, glucose levels at 2 h in oral glucose tolerance test, and with each of the MetS states (all $p<0.05$). At an individual level, using the MetS definitions alone as diagnostic test for the presence of increased IMT (AUCs: 0.60-0.66) was inferior when compare to the sum of all individual components (AUCs: 0.65-0.85). Adding presence or absence of MetS to the components did not improve the accuracy.

Conclusions: Overweight adolescents with MetS demonstrated increased IMT values as compared to overweight adolescents without MetS. However, the best model to diagnose increased IMT was the sum of the quantitative components. The use of dichotomized variables reduced the diagnostic accuracy. These AUC findings are discussed in relationship to the new measures of diagnostic performance using our real data example.

266: Time Series Analysis

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1015
Session Chair: Anne Leucht

266: 1

Using multiple filters to detect rate changes in point processes

Michael Messer[1], Marietta Kirchner[1], Julia Schiemann[2], Ralph Neininger[1], Jochen Roeper[2], Gaby Schneider[1]

[1]Institute of Mathematics, Goethe University, Frankfurt, Germany; [2]Institute for Neurophysiology, Goethe University, Frankfurt, Germany

Non-stationarity of the event intensity is a persistent problem in modelling time series of events, such as neuronal spike trains. In many cases it seems reasonable to approximate the intensity over time as a step function. We propose a new technique to detect the change points in this case. Our technique uses multiple filters and can thus detect rate changes on fast and slow time scales at the same time.

To capture the great heterogeneity of many experimental time series of events, we introduce a generalization of renewal processes which allows for variability in the variance of increments.

In a sample data set of dopaminergic spike trains of anesthetized mice, the detected change points agreed closely with visual inspection, and in over 70% of all spike trains which were classified as non-stationary, different change points were detected by different filters.

266: 2

Non-parametric analysis of resting state fMRI data

Daniela Adolf, Snezhana Weston, Siegfried Kropf

Otto-von-Guericke-Universität Magdeburg, Germany

Via functional magnetic resonance imaging (fMRI) an indirect measurement of neuronal activation in the human brain is possible. Since the response signal is a temporal series in single three-dimensional regions (voxels), a challenge in fMRI analysis is that measurements are temporally correlated, which goes beyond the scope of classical statistics.

In standard fMRI analysis software as SPM (statistical parametric mapping), data are analyzed voxel-wise using a general linear model, where the temporal correlation of measurements is taken into account by pre-whitening. There, usually a first-order autoregressive process is assumed and its correlation coefficient is estimated using a restricted maximum likelihood approach.

A recently published empirical study on 1484 resting state data sets (Eklund et al., 2012) demonstrated, that this AR(1) autocorrelation correction is not sufficient to avoid high family-wise error rates up to 70%.

We used these data sets to show that it is possible to yield valid results in fMRI analysis using a special permutation approach (Adolf et al., 2011). Here we account for temporal autocorrelation by permuting blocks of adjacent elements instead of single measurements and thus we can keep the nominal family-wise test level.

References:

A. Eklund, M. Andersson, C. Josephson, M. Johannesson, H. Knutsson (2012). Does parametric fMRI analysis with SPM yield valid results? - An empirical study of 1484 rest datasets. NeuroImage 61, 565–578.

D. Adolf, S. Baecke, W. Kahle, J. Bernarding, S. Kropf (2011). Applying multivariate techniques to high-dimensional temporally correlated fMRI data. Journal of Statistical Planning and Inference 141. 3760–3770.

266: 3

Order Determination for Functional Autoregressive Processes

Thorsten Fink[1], Alexander Aue[2], Thomas Lee[2]

[1]TU Braunschweig, Germany; [2]UC Davis, United States of America

We consider functional time series that are assumed to follow an autoregressive scheme of unknown order and show how to estimate this order consistently. We precisely establish the connection between functional autoregressive processes and multivariate autoregressive processes and show how to obtain a multivariate process if we are given a functional AR process. The resulting process follows an autoregressive scheme, but is not a standard AR process anymore. The coefficient matrices are random and the residuals are correlated with each other and with the observations of the process. Following earlier contributions for AR processes, we introduce a general loss function and show that the estimated order obtained by a minimization of this function converges to the correct order of the multivariate non-standard AR process and therefore of the functional AR process in probability. We also evaluate the small to medium sample size performance of this estimator by a simulation study and compare it with an existing sequential testing method. Finally, we apply the method to real data sets.

Trends and projections of vehicle accident related injuries in Northwest Gondar, Ethiopia: a time series analysis

Solomon Meseret Woldeyohannes, Haimanot Ghiwot Moges, Berihun Megabiaw Zeleke, Sisay Yifru Malefiya, Yigzaw Kebede Gete

University of Gondar, Ethiopia

Background: Road traffic injuries are an emerging challenge to public health in the world. Forecasts suggest that road traffic injuries become the third leading cause of death and disability worldwide by 2020. These projections highlight the essential need to address road-traffic injuries as a public-health priority.

Objective: The study aimed at estimating trends and projections of fatal injuries attributed to vehicle accidents in North Gondar, Ethiopia.

Methodology: Data on fatalities, total and partial permanent injuries, and lost workday attributable to vehicle accident were collected from North Gondar Traffic Offices from 1996 to 2011. Holt and Brown exponential smoothing techniques were used to model the number of fatalities and other injuries due to vehicle accidents.

Results: There were 2,300 vehicle accidents that occurred from 1996 to 2011 causing an estimated 968 fatalities, 1,665 lost workday and 1,185 permanent total and partial injuries, and 1,899,950.60$ losses. Future forecasts showed that by 2015, there could be 414 fatalities, 1,123 lost workday and 438 permanent total and partial injuries, and 955,249.12$ losses.

Conclusion: The numbers of lives lost and disabilities due to vehicle accidents indicated an upward trend in the last decade showing future burden in terms of societal and economic costs threatening the lives of many individuals. Surveillance systems that could enable to monitor patterns of Vehicle accidents with preventive strategies must be established.

Key words: Vehicle accident, fatal injuries, economic loss, trends, forecasts.

267: Non- and Semi-Parametric Statistics

Time: Tuesday, 19th Mar 2013: 4:40pm - 6:00pm · Location: KG III, HS 3042
Session Chair: Enno Mammen

267: 1

Permutation based Confidence Intervals for the Area under the ROC Curve

Thomas Andreas Asendorf

University of Göttingen, Germany

In diagnostic trials, the performance of a medical test is most frequently measured in terms such as sensitivity, specificity and the area under the ROC-curve (AUC), making the use of accurate estimation methods for the latter inevitable. When confronted with small sample sizes, classical approaches using the central limit theorem, and relying on its approximative properties, are not very effective in estimating the AUC. Instead, we will study various permutation methods for calculating confidence intervals for the AUC of different types of data. We will see, that using these permutation methods leads to more accurate confidence intervals, especially when confronted by data with few observations. Using this method gives the experimentor better possibilities of quantifying the performance of medical tests in diagnostic trials, using the AUC.

267: 2

Nichtparametrische Alternativen zur MANOVA

Arne Bathke[1], Solomon Harrar[2], Woody Burchett[3], Amanda Ellis[3]

[1]Universität Salzburg, Austria; [2]University of Montana, United States of America; [3]University of Kentucky, United States of America

Wir präsentieren nichtparametrische Methoden zur Analyse von multivariaten Daten in faktoriellen Designs. Die Zielgrößen werden nicht als normalverteilt angenommen. Es werden sogar unterschiedliche Skalenniveaus der unterschiedlichen Zielgrößen zugelassen. Das heißt, es kann zum Beispiel eine Variable metrisch sein, eine weitere ordinal, eine weitere binär. Neben Theorie und Simulationsergebnissen demonstrieren wir anhand eines Beispiels, wie die neue Methodik mit Hilfe eines R-Paketes angewendet werden kann.

267: 3

Asymptotic permutation tests in heteroscedastic factorial models

Markus Pauly[1], Edgar Brunner[2], Frank Konietschke[2]

[1]Heinrich-Heine University Duesseldorf, Germany; [2]Georg-August University Goettingen, Germany

Classical inference methods for factorial designs postulate a homosscedastic model assumption. However, especially in higher factorial designs, this assumption is rather unrealistic.

In this talk we study inference methods based on quadratic forms for analyzing unbalanced heteroscedastic models. Existing examples are given by ANOVA-, Welch-James or Wald- type tests. While the latter is asymptotically exact, both the ANOVA- and the Welch-James-type tests are derived from asymptotic approximations and their actual significance levels are not known.

Based on theoretical results we propose a permutation tests which is not only fintely exact under exchangeability of the data but also asymptotic exact in general unbalanced heteroscedastic models.

Intensive simulation studies show that the test keeps the prescribed level satisfactorily.

A real data example illustrates the application of the methods.

References:

Johansen, S. (1980). The Welch-James Approximation to the Distribution of the Residual Sum of Squares in a Weighted Linear Regression. Biometrika 67, 85 - 92.

Brunner, E., Dette, H., Munk, A. (1997). Box-Type Approximations in Nonparametric Factorial Designs. Journal of the American Statistical Association 92, 1494 - 1502.

267: 4

A studentized permutation test for the nonparametric Behrens-Fisher problem in paired data

Frank Konietschke

University Medical Center Göttingen, Germany

In this talk we consider nonparametric ranking methods for matched pairs, whose distributions can have different shapes even under the null hypothesis of no treatment effect. Although the data may not be exchangeable under the null, we investigate a permutation approach as a valid procedure for finite sample sizes. In particular, we derive the limit of the studentized permutation distribution under alternatives, which can be used for the con-struction of [fn](1 − \alpha)[/fn]-confidence intervals. Simulation studies show that the new approach is more accurate than its competitors. The procedures are illustrated using a real data set.

311: Meth Börse

Time: Wednesday, 20th Mar 2013: 8:00am - 8:30am · Location: KG I, HS 1199

311: 1

Modellierungsstrategien bei hochdimensionalen hierarchischen Daten in der tiermedizinischen Anwendung

Lothar Kreienbrock

University of Veterinary Medicine Hannover, Germany

tba

321: Statistics in Clinical and Preclinical Research -- Data Monitoring Committees for confirmatory adaptive clinical trials

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG II, Audimax
Session Chair: Joachim Röhmel

321: 1

Statistical Considerations for Data Monitoring Committees in Adaptive Clinical Trials

Lisa LaVange

FDA, United States of America

Data Monitoring Committees (DMCs) are called upon to make increasingly more complex decisions for pharmaceutical sponsors as adaptive trials become more prevalent. Rather than limiting decisions to continuing or discontinuing a clinical trial at an interim time point for safety, futility, or early evidence of efficacy, DMCs are asked to provide input on major design changes, including dose selection, enrichment of certain subgroups, etc. A number of issues accompany the increased responsibilities of DMCs, including the amount and quality of data upon which the decisions are based and the sponsors' need to act on the DMC's recommendations while still being masked to the analysis details. Also at issue are the sponsor's comfort with the DMC's new role and the DMC members' comfort with the role themselves. Regulatory acceptance of adaptive design decisions depends on the ability to pre-specify the possible adaptations at the study planning phase. Sponsors may seek regulatory approval of substantial changes in an ongoing trial, but such approval puts the regulatory agency in an awkward position of serving as another DMC. In this presentation, we will discuss these and other issues arisin in the monitoring of adaptive trials, paying particular attention to regulatory interactions and concerns.

321: 2

Controlled unblinding and information sharing in adaptive design trials from a vendor perspective

Reinhard Eisebitt

Aptiv Solutions Inc., United States of America

In today's world the conduct of a multinational clinical trial is a highly regulated, complex process that requires coordination of multiple vendors. Adaptive designs add further complexity to the task, because they allow for design modifications after examination of unblinded interim data. Flaws in the process and the hand-offs among the many parties involved can introduce serious bias. The presentation addresses some common issues of potential information leakage and process control from a service provider perspective.

321: 3

Bias in Adaptive Clinical Trials and the Role of Data Monitoring Committees

Martin Posch

Section for Medical Statistics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria

In adaptive clinical trials Data Monitoring Committees have the task to implement adaptations based on unblinded interim data. It is well known that without appropriate adjustment, modifications to the trial designs based on unblinded interim data may introduce a severe type I error rate inflation and biased estimates. Explicit adaptations as, for example, sample size reassessment, can be accounted for with appropriate statistical methodology as combination tests or the conditional error rate principle. These methods guarantee the integrity of the statistical analysis even if the exact adaptation rules are not pre-specified in detail a priori. However, these tests provide no adjustment for “implicit adaptations” that may result from leakage of interim data to patients or investigators. Examples of such implicit adaptations are changes to the study population, to the evaluation of response or a change of the placebo effect. Therefore, appropriate fire walls to prevent the dissemination of unblinded interim results are essential to preserve the integrity of the trial.

While it is generally accepted that unblinded interim data must not be leaked, the issue of bias introduced by the leakage of blinded interim data has achieved less attention: If a treatment has an effect on a secondary or safety endpoint, this may lead to a partial unblinding which in turn may introduce bias.

Therefore, in trials with (adaptive) interim analyses, a definition of appropriate information flows as well as a stage-wise examination of the trial data to detect implicit adaptations is important to demonstrate the integrity of the trial.

322: Survival and Event History Analysis -- Biometrical Problems

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1015
Session Chair: Antje Jahn

322: 1

Analysis of time-to-event data from randomized trials for personalized treatment selection

Martin Grupp

Albert-Ludwigs-Universität Freiburg, Germany

In medical statistics the so called subgroup analysis attracts more and more attention. This thesis develops a two-stage-prediction-model to estimate individual treatment effects on survival data.

The structure of the model proposed is based on a model of Cai et al. (2010) and predicts the outcome in a two-stage-model of a (semi-)parametric and a nonparametric model for time-to-event data since this is a significant part in medical practice. It is aimed to build the model as intuitive as possible so that it can easily be used in clinical studies.

Further, the original structure of Cai et al. has been upgraded by a variable selection to identify important biomarkers which are interacting with either outcome or treatment and therefore not to over-specify the model. In the end a further treatment selection has been added via a 'Selection Impact' (SI) curve to make the results more comparable. Moreover the SI-curve is an intuitively and demonstrative way of finding the optimal treatment decision for each patient, which should help to understand the treatment selection. This could especially be helpful in clinical practice.

The model has been used and tested in two randomized studies, the GLIOMA study and the MRC RE01, before a simulation study has been constructed. It should be noted that the variable selection as well as a similar model of the SI-curve can in addition be used for non-survival-data. As a consequence one could adopt these improvements on the original model of Cai et al. as well.

322: 2

Modelling the prevalence and incidence of rare chronic diseases via stochastic differential equations

Sandra Landwehr[1], Ralph Brinks[1], Rebecca Fischer-Betz[2], Matthias Schneider[2], Guido Giani[1]

[1]German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Institute for Biometrics and Epidemiology, Germany; [2]Policlinics for Rheumatology, Heinrich Heine University Düsseldorf, Germany

Epidemiological data on rare diseases are usually rather rare themselves, hence why it is often very difficult to obtain reliable estimates of quantities such as prevalence or incidence of these diseases. For the characterisation of these quantities it has been proved useful to consider so-called compartment models, and stochastic approaches are used to estimate them. On the basis of an illness-death model we describe a system of stochastic differential equations (SDEs), which will be solved numerically to derive the age-specific prevalence from given incidence and mortality rates.

We apply this SDE method to data on the rare rheumatic disease Systemic Lupus Erythematosus (SLE) to obtain simulated age-profiles of SLE prevalences from given incidence and mortality rates taken from appropriate UK register. Moreover, we show how to reconstruct the age-specific incidences with the help of spline-based smoothing of the prevalence paths and a non-parameteric estimation technique. It will become apparent that the reconstructed incidence is in good accordance with the input data.

322: 3

On deriving incidences from current status data

Ralph Brinks, Sandra Landwehr, Guido Giani

German Diabetes Center, Düsseldorf, Germany

Background: In his seminal work [1] Keiding uses the illness-death model (IDM) of Fix and Neyman [2] to formulate a general relationships between the prevalence (p), incidence (i) and mortalities in chronic diseases. Considered covariates are age (a), calendar-time (t) and duration (d) of the disease [1]. In certain situations i can be estimated from p [1,3], which is valuable when only current status data are available.

Methods: In the case p and mortalities in the IDM depend on t, a and d, but i just depends on a, we prove that the dependence on d does not have to be known in estimating i from p. The proof is illustrated by an analytical example and an associated simulation. A population of 300000 (born in 60 consecutive years at constant birth rate of 5000/year) is followed from birth to death. Eventual diagnosis of the chronic disease and death are modeled as competing risk. Theoretical and simulated p at t = 100 years are used to reconstruct i.

Results: Using the theoretical p, i can perfectly be reconstructed. In the simulated p, due to sampling small deviations from the true i occur.

Conclusion: In epidemiology, duration dependency is difficult to measure. In this work we show that sometimes this is not necessary to estimate the incidence from prevalence data.

References

[1] Keiding N, J Roy Stat Soc A, 1991

[2] Fix E & Neyman J, Human Biology, 1951

[3] Brinks R et al, Stat Med, 2012

322: 4

Robust estimation of basic reproduction numbers using social contact and serological data

Steffen Unkel[1], C. Paddy Farrington[2], Karim Anaya-Izquierdo[3]

[1]Justus-Liebig-Universität Giessen, Germany; [2]The Open University, United Kingdom; [3]London School of Hygiene and Tropical Medicine, United Kingdom

The basic reproduction number of an infection in a given population, R0, is inflated by individual heterogeneity in contact rates. Recently, new methods for estimating R0 using social contact data and serological survey data have been proposed. These methods, like most of their predecessors, ignore individual heterogeneity, and are sensitive to mis-specification of the contact function.

Using a frailty framework, we derive expressions for R0 in the presence of age-varying heterogeneity. In this case, R0 involves the variance function of the age-dependent frailty. This variance can be estimated within a shared frailty framework from paired data on two infections transmitted by the same route. We propose two estimators of R0 for infections in endemic equilibrium. We investigate their performance by simulation, and find that one is generally less efficient but more robust than the other to mis-specification of the effective contact function.

These methods are applied to data on varicella zoster virus infection from two European countries.

323: Statistics for High Dimensional Data

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1098
Session Chair: Isabella Zwiener

323: 1

Detection of Significantly Differentially Methylated Regions in Reduced Representation Bisulfite Sequencing Data

Katja Hebestreit, Martin Dugas, Hans-Ulrich Klein

University of Münster, Germany

DNA methylation plays an important role for epigenetic gene regulation in development and disease. A DNA sequence is called methylated, if methyl groups are attached to the C bases within CpG dinucleotides (CpG sites). Reduced representation bisulfite sequencing (RRBS) is a cost-efficient method for DNA methylation profiling at single-nucleotide resolution in CpG rich regions (e.g. promoter regions). After sequencing, the number of methylated and unmethylated reads are obtained for each covered CpG. So far, only few approaches to process and analyze this kind of data are published. We propose an algorithm to detect significantly differentially methylated regions (DMRs) in cancer versus control samples.

To reduce the number of tests, we implemented a hierarchical testing procedure that first tests clusters and then tests locations within rejected clusters (Benjamini and Heller, 2007).

The algorithm is divided into four steps: 1.) Find clusters of covered CpGs considering all samples. 2.) Smooth methylation levels in these CpG clusters for each sample. 3.) Estimate the group effect along the CpG sites via beta regression. 4.) Test for significant group effects in each CpG cluster and controlling a cluster-wise FDR. 5.) Trimming of rejected CpG clusters to DMRs and controlling a location-wise FDR.

We tested the algorithm on simulated data with 5,000 simulated DMRs of different lengths and differences and could show that control of the FDRs was achieved. Additionally, we analyzed real data from 18 patients with acute promyelocytic leukemia and 16 controls and could show that the approach results in biological meaningful results.

323: 2

Accurate Metagenomic Abundance Estimation on Species Level and Influences on Sequence Coverages

Martin S Lindner, Maximilian Kollock, Bernhard Y Renard

Robert Koch-Institut, Germany

One particular goal of sequencing based metagenomic community analysis is the quantitative taxonomic assessment of microbial community compositions. Reliable, relative quantification of taxons is of high relevance for metagenomic diagnostics or microbial community comparison. However, the majority of existing approaches either quantify at low resolution (e.g. at phylum level), or rely on the existence of specific genes, or have severe problems discerning species with highly similar genome sequences. We developed Genome Abundance Similarity Correction (GASiC), a versatile method to estimate true genomic abundances in metagenomic datasets on the species.

Within our approach, metagenomic sequence reads are first mapped against a set of reference genomes of species potentially present in the dataset. Then, we esestimate the pairwise similarities of the reference genomes using a simulation approach. The similarities are then used to correct the mapping results from the first step and to obtain estimates of the true genomic abundances in the dataset. To this end, we formulate the problem as a non-negative LASSO. Further, we study how changes in the sequence coverage of a specific genome can serve as indicator for incorrect metagenomic assignments and propose a mixture model approach for detecting these artifacts.

We applied our method to the metagenomic FAMeS benchmark dataset and compared its performance to existing methods, showing that reduces the quantitative error by up to 60% even in the presence of sequence similarities of 95% and above.

323: 3

Comparison of methods for the detection of gene set activation profiles in gene expression time series

Andre König, Jörg Rahnenführer

TU Dortmund, Germany

Time series of gene expression measurements are used to study various biological processes, e.g. immune response, cancer progression or embryonic development. Due to the costs of microarray experiments in many research projects only a few times are analyzed. Moreover, due to limited biological material or money, often none or just few replicates are considered.

Summarizing the expression information for a set of genes is a promising and widely used approach in the field of analyzing gene expression data. The use of gene sets defined by Gene Ontology (GO) or KEGG is established, but it is not yet common for the analysis of time series experiments.

Within an extensive simulation study we compare several algorithms for analyzing gene sets in time series experiments, including three methods from the literature and two new own methods. The maSigFun approach (Nueda et al., 2009) needs a clear signal fitting well to a

regression model. The Short Time-series Expression Miner (STEM) by Ernst et al. (2006) uses clustering around pre-chosen model profiles. In contrast, our methods proceed per time point without such model assumptions. We first compare the expression values at every time point with a reference distribution, determine differentially expressed genes, and then identify gene groups overrepresented among the differentially expressed genes. This yields a characteristic and well interpretable time profile for all considered gene groups. In distinction from Al-Shahrour et al. (2006) we consider the significance of the differential gene expression per time point and improve the obtained profiles by smoothing.

323: 4

Modelle für hochdimensionale Repeated Measures Daten

Edgar Brunner[1], Markus Pauly[2], David Ellenberger[1]

[1]Universität Göttingen, Germany; [2]Universität Düsseldorf, Germany

Hochdimensionale Repeated Measures Daten liegen vor, wenn bei Zeitverläufen die Anzahl der Messzeitpunkte [fn]d[/fn] größer ist als die Anzahl [fn]n[/fn] der unabhängigen Individuen. Dabei kann dem Zeitverlauf eine faktorielle Struktur unterliegen und es können mehrere (unterschiedlich behandelte / konditionierte) Gruppen von Individuen vorhanden sein. Die Herleitung asymptotischer Verfahren zum Testen globaler Hypothesen beruht auf der Annahme, dass [fn]n[/fn] und [fn]d[/fn] (mit [fn]n \leq d[/fn]) [fn]\to \infty[/fn] gehen. Bei Zeitverläufen bedeutet die Bedingung [fn]d \to \infty[/fn], dass entweder der Zeitverlauf weiter fortgesetzt wird oder dass die Beobachtungen immer dichter werden. Dies hat unterschiedliche Auswirkungen auf die Struktur der Kovarianzmatrizen und damit auf die Grenzverteilungen. Es werden verschiedene asymptotische Verfahren und Approximationen für diese Situationen vorgestellt und die Voraussetzungen werden anhand von Beispielen diskutiert.

324: Penalized and Regularized Regression Analysis

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1010
Session Chair: Fabian Sobotka

324: 1

Estimation of Non- and Semiparametric Additive Models by Regularized Kernel Methods

Robert Hable

Universität Bayreuth, Germany

Regularized kernel methods such as support vector machines (SVM) are among the most important and commonly used methods in machine learning. They are defined by minimizing a regularized empirical risk over a function space defined by a kernel function. In this way, they provide a broad and flexible class of statistical methods: by use of suitable loss functions, they can be used for many different tasks such as classification, mean regression, quantile regression, and estimation of scale functions in heteroscedastic regression problems; by use of suitable kernels, they can be used as nonparametric and also as parametric methods.

In the talk, it is shown how these methods can also be used for estimating additive models. Additive models are often a good compromise between nonparametric and linear models in practice: the additive component functions can be estimated in a nonparametric way but, due to the additive structure, additive models retain much of the good interpretability properties of linear models. Regularized kernel methods are more flexible than many other methods for additive models because they can also easily estimate models with multivariate component functions.

References:

A. Christmann & R. Hable (2012): Consistency of support vector machines using additive kernels for additive models. Computational Statistics & Data Analysis, 56:854-873, 2012.

324: 2

Spatially regularized estimation for the analysis of DCE-MRI data

Julia C. Sommer[1], Jan Gertheiss[1,2], Volker J. Schmid[1]

[1]Department of Statistics, Ludwig-Maximilians-Universtität München, Germany; [2]Department of Animal Sciences, Georg-August-Universität Göttingen, Germany

Compartment models are typically used for the quantitative analysis of Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) data. Such models link the observed concentration uptake over time to physiologic properties like blood flow and capillary permeability. However, models of different architecture and complexity have been used and it remains unclear which model to use. The adequate number of compartments might be different in different types of tissue; and in heterogeneous tissue, like for example cancerous tissue, it might even vary over a field of voxels. Therefore, we propose a spatial Elastic Net approach that allows to estimate the number of compartments for each voxel such that the model complexity is not fixed a priori.

A multi-compartment approach is considered, which is translated into a restricted least square model selection problem. This is done by using a set of basis functions for a given set of candidate rate constants. The form of the basis functions is derived from a kinetic model and thus describes the contribution of a specific compartment. We choose a sparse set of basis functions per voxel, and hence, rate constants of compartments.

The proposed spatial Elastic Net estimator uses the intrinsic spatial information given by the voxel structure of the image. Spatial smoothness of the parameters is incorporated by penalizing the differences of neighboring coefficients. The spatial penalty performs better than a penalty treating voxels independently. The proposed estimation method is evaluated for simulated images and applied to an in-vivo data set.

324: 3

Extracting interactions from high-dimensional data using random forests and regularized regression for clinical endpoints

Isabell Hoffmann, Murat Sariyar, Harald Binder

Institute for Medical Biostatistics, Epidemiology and Informatics, University Medical Center, Johannes Gutenberg-University Mainz, Germany

When linking high-dimensional molecular covariates to some clinical endpoint, e.g., when using gene expression measurements for prognosis, sparse regression techniques are destined to provide a short list of marginal or main effects. While interactions are highly likely to be present in molecular applications, there is no established method for determining interactions that should be considered together with potential main effects. For this task, we present a general strategy based on a combination of a regularized regression approach and random forests. We specifically consider likelihood based componentwise boosting for time-to-event settings in order to select and estimate main effects. Random survival forests are used as a screening technique for obtaining interaction information. The screening is built on permutation accuracy importance and pairwise inclusion frequencies. We integrate these techniques into a procedure that extracts interactions from the forests for enriching the Cox model obtained from the boosting technique. The benefits and limits of this new approach

are evaluated on different scenarios of simulated data sets with respect to prediction performance and sensitivity as well as positive predictive values concerning main effects and interactions. Our approach is seen to find relevant interactions, even without strong marginal components, but not as stable as interactions with such components. This leads to improvements of prediction performances in many of the investigated scenarios. We also illustrate the approach with gene expression data from patients with diffuse large B-cell lymphoma.

324: 4

Adapting the sparsity of risk prediction models for different levels of molecular data

Murat Sariyar[1], Martin Schumacher[2], Harald Binder[1]

[1]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center, Johannes Gutenberg-University Mainz, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Risk prediction models can connect high-dimensional molecular measurements to clinical endpoints, such as survival. Often, selection of a small number of important variables, i.e., sparsity, and good prediction performance are wanted at the same time. The appropriate degree of sparsity will depend on the specific molecular level under consideration. For example, analysis for DNA methylation data could be performed at various levels ranging from a single CpG site to a chromosome. The goals related to the aggregation of molecular information could be reduction of the number of variables to be screened and enhanced biological interpretation of the model, without seriously sacrificing prediction performance. Currently, there is no unified and simple way for automatically adapting the degree of sparsity to the data at hand. To address this, we extend a componentwise boosting algorithm such that it can direct sparsity towards a larger or smaller level, by means of a single tuning parameter. The new approach is investigated on different aggregation levels of DNA methylation data of kidney cancer patients with time to death as response. On all aggregation levels, data-based selection of the sparsity tuning parameter improves at least one of the two considered measures, stability of variable selection or prediction performance. The results confirm that automatic adaptation of sparsity can provide deeper insight into the required model complexity for specific molecular levels when building risk prediction models.

325: Clustering

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3043
Session Chair: Berthold Lausen

325: 1

Finite Mixture Model Clustering of SNP Data

John Hinde[1], Norma Coffey[2], Augusto Franco Garcia[3]

[1]NUI Galway, Ireland; [2]UCD, Dublin, Ireland; [3]ESALQ/USP, Piracicaba, Brazil

Sugarcane is polypoid and it is important to develop methods that identify the many different alleles and associated genotypes. Single nucleotide polymorphisms (SNPs) can give an indication as to the number of allele haplotypes present for a gene and such information could have implications in sugarcane breeding since high yield potential may be due to the presence of and/or different number of copies of, a specific allele(s) present at a gene locus. Clustering these data provides a means of identifying different genotypes and therefore it is necessary to develop a technique that can determine the number of clusters present, determine the angles between the clusters to identify different genotypes, and provide a probabilistic clustering to identify points that have high probability of belonging to a particular cluster (have a particular genotype) and those that are regarded as an unclear genotype. Standard clustering methods, such as mclust do not perform well because of the radial nature of the data, although the performance can be improved by moving to polar coordinates, as previously proposed in the literature. Here we propose the use of finite mixtures of orthogonal regression lines to cluster the data. We implement this technique in R, show its usefulness in clustering these data and compare the performance with the other methods.

325: 2

Merging States for Cluster Analysis in hidden Markov models

Hajo Holzmann

Marburg University, Germany

We analyse clustering problems in case of dependent data. Specifically we consider the observable part of a finite state hidden Markov model (HMM), where the stationary distribution is a finite mixture of parametric distributions (e.g. of multivariate normal distributions) and the hidden state process has a Markov chain structure. In contrast to independent finite mixtures, the dependence structure of the model plays an important role for estimating the non-observable states of the HMM. Baudry et al. (2010) model i.i.d. samples with finite mixtures and merge certain components to clusters according to an entropy based criterion. Similarly, for a HMM it is not always suitable to assign to each state of the Markov chain its own cluster. Therefore, we analyze the merging of states to clusters for HMMs. In contrast to independent finite mixtures, where merging states does not affect the probabilistic structure of the model, merging states of a HMM changes the transition probability matrix and the state dependent distributions become mixtures themselves. We develop and analyze merging algorithms which take into account this loss of information after merging.

325: 3

Some empirical evidence on the effectiveness of Factor PD-clustering

Mireille Gettler Summa[1], Francesco Palumbo[2], Cristina Tortora[3]

[1]Université Paris Dauphine CEREMADE CNRS, France; [2]Università di Napoli Federico II, Italy; [3]Stazione Zoologia Anton Dohrn, Italy

The ever-increasing number of big datasets brings to the necessity to cluster dataset with a large number of units and variables. When the number of variables increases, clustering methods can become unstable and factorial-clustering methods can be used instead. Factorial clustering methods aim at finding a clustering structure and a factorial space that best represent the partition. To obtain this result the clustering method and the factorial decomposition must optimize the same function; consequently factorial clustering methods consist in a linear transformation of data and a clustering on transformed data optimizing a common criterion. Among these methods we have proposed Factor PD-clustering. It is based on Probabilistic Distance clustering (PD-clustering) that is an iterative, distribution free, probabilistic, clustering method. Factor PD-clustering makes a linear transformation of original variables into a reduced number of orthogonal ones using a common criterion with PD-Clustering. It has been demonstrated that Tucker3 decomposition permits to obtain this transformation. Factor PD-clustering exploits alternatively Tucker3 decomposition and PD-clustering on transformed data until convergence is achieved. This method can significantly improve the algorithm performance. The factorial step makes the method more stable, more robust and allows us to work with dataset with large number of variables. The aim of this talk is to show how Factor PD-clustering improves PD-clustering performance dealing with large datasets. The two methods have been applied on some real dataset with a large number of variables and on some simulated datasets.

325: 4

Assessing Answer Patterns in Questionnaire/Item Response Data Using Mixtures of Rasch Models

Hannah Frick[1], Carolin Strobl[2], Friedrich Leisch[3], Achim Zeileis[1]

[1]Universität Innsbruck, Austria; [2]Universität Zürich, Switzerland; [3]Universität für Bodenkultur Wien, Austria

The Rasch model is a standard model for binary item response data, e.g., as generated by respondents "solving" items in an aptitude test or by giving yes/no answers in questionnaires. The Rasch model logistically links the probability of solving an item (or answering "yes") to the difference between the respondent's ability and the item's difficulty. Based on observed item response data, conditional maximum likelihood estimation can then be used to infer item difficulties and respondent abilities. However, comparisons of respondents' abilities obtained this way are only fair if the same item difficulties hold for all respondents. Mixture models are flexible means to assess this assumption and uncover heterogeneity if there is any. Furthermore, if different patterns of item difficulty can be found in a mixture of Rasch models, covariates may be available to explain these: either ex-post or in a concomitant variable model within the mixture model. Here, we present a general framework for such Rasch mixture models along with its implementation in the R package "psychomix". It includes the possibility of concomitant variable models as well as various options for modeling the distribution of respondents' abilities. The methods are illustrated using questionnaire data on the extrinsic motivation of R package authors. In this application, the mixture model can uncover different patterns of item difficulty for groups of the R package authors which can be partially attributed to the covariates education and occupation.

326: Robust and Distribution Free Statistics

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1199
Session Chair: Andreas Christmann

326: 1

Sparse and Robust Principal Component Analysis

Peter Filzmoser

Vienna University of Technology, Austria

A method for principal component analysis is proposed that is sparse and robust at the same time. The sparsity delivers principal components that have loadings on a smaller number of variables, making them easier to interpret. The robustness makes the analysis more resistant to outlying observations. The principal components correspond to directions that maximize a robust measure of the variance, with an additional penalty term to take sparseness into account. An algorithm to compute the sparse and robust principal components has been implemented in the R package pcaPP. The method is applied on several real data examples, and diagnostic plots for detecting outliers and for selecting the degree of sparsity are provided. A simulation experiment studies the loss in statistical efficiency by requiring both robustness and sparsity.

326: 2

On the Reliable Analysis of Interval Data: Unbiased Sets of Estimating Functions and Partial Identification

Thomas Augustin[1], Ulrich Pötter[2], Michael Seitz[1]

[1]Ludwig-Maximilians-Universität München (LMU), Germany; [2]German Youth Institute, Munich

Survey data for variables like income or age are often grouped, heaped or roughly rounded and thus only provide the information that the precise value of the variable lies in some interval. While traditional modelling ignores the imprecision in the data or has to rely on other optimistic assumptions on the coarsening process, an alternative methodology is currently getting strong momentum in the literature. Questioning the implicit paradigm that even very imprecise data should nevertheless always produce single-valued point estimates, the novel methods look for an optimal set of observationally equivalent models, naturally reflecting the extent of imprecision and telling exactly what can reliably be learned from the data. Typically, the results are still informative enough to provide valuable insights into the underlying subject matter questions.

Current work in this spirit, like approaches of partial identification (e.g., Beresteanu, Molchanov, Molinari 2011, Econometrica; Tamer 2011, AnnRevEcon) but also (profile-) likelihood-based imprecise regression (e.g., Cattaneo, Wiencierz 2012, IntJApproxReason), mainly confines consideration to linear models. To handle general models, we suggest a notion of unbiasedness for sets [fn]\Psi[/fn] of estimating functions and report approaches to make this idea operational in the context of GLMs under interval data. We first discuss the special, but instructive one-parameter case, where Fisher-scoring can be applied to the lower and upper envelope of [fn]\Psi[/fn], and then present different formulations of the estimation task as constrained optimization problems. We compare, by simulations and some illustrative applications, estimated regions determined with the L-BFGS-B algorithm and a new heuristic optimization approach.

326: 3

On Bootstrap and Kernel Based Methods

Andreas Christmann, Robert Hable

University of Bayreuth, Germany

The talk will focus on the mainly open question how to draw statistical decisions based on support vector machines and related kernel based methods. This question has to my knowledge not yet attracted much consideration in the literature.

Some recent results on asymptotic properties of bootstrap approximations for the distribution of SVMs and related kernel based methods will be given.

327: Non- and Semi-Parametric Statistics

Time: Wednesday, 20th Mar 2013: 8:50am - 10:10am · Location: KG I, Aula
Session Chair: Frank Konietschke

327: 1

Proper Local Scoring Rules on [fn]\R^n[/fn]

Evgeni Ovcharov

University of Heidelberg, Germany

Scoring rules provide measures for the accuracy of probabilistic forecasts. We give a complete characterisation of the proper [fn]m[/fn]-local scoring rules on [fn]\R^n[/fn] in terms of convex functionals. The class of convex functionals that generate proper [fn]m[/fn]-local scoring rules is defined precisely and various representation formulas for their kernels are derived. We show the important property that the scoring rules of orders higher than zero are 0-homogeneous and that they exist for even orders only. Every scoring rule of order [fn]2m[/fn] is shown to be generated by a kernel of minimal order [fn]m[/fn].

327: 2

Data-based optimal stoppping via forecasting of time series

Daniel Jones

TU Darmstadt, Germany

The problem of optimal stopping in discrete time is considered. The algorithm proposed uses techniques of forecasting of time series and is completely nonparametric in the sense that it is solely based on observations. It is shown that the expected gain of the corresponding stopping rule converges to the optimal value whenever the observations are drawn from a stationary and ergodic sequence. The algorithm is illustrated by applying it to the problem of optimal exercising an American option.

327: 3

Expectiles and Expectile Regression: A Comparison with Quantiles towards Efficiency

Linda Schulze Waltrup[1], Fabian Sobotka[2], Thomas Kneib[2], Göran Kauermann[1]

[1]Ludwig-Maximilians-Universität München, Germany; [2]Georg-August-Universität Göttingen, Germany

Modern methods in regression are becoming more and more general. They allow for heteroscedasticity, correlation, multiple observations and many more. In this context, models were proposed which not only estimate the trend, but also other features like the tail of the data. Examples for that are quantile and expectile regression. Similarly as quantiles generalize the median, expectiles generalize the mean. While quantiles are obtained by minimizing a sum of weighted absolute errors, expectiles result from weighted squared errors. Both, expectiles and quantiles, can be seen as special cases of so called M-quantiles.

In this talk, we explore the connection between expectiles and quantiles and note that this connection can be used to estimate the expected shortfall. The transformation also allows to investigate the performance of quantile and expectile estimators on a comparable level. We start with sample expectiles and quantiles and then switch to regressions estimates. An extension to semiparametric regression is given. We observe that, like quantiles, neighboring expectiles may cross each other, although crossing seems less likely for expectiles than for quantiles, and derive non-crossing estimates. We conclude the talk with an extension to allow for longitudinal data.

327: 4

Asymptotische Verteilung von Strukturbruchschätzer für nichtlineare AR-Prozesse

Stefanie Schwaar

University of Kaiserslautern, Germany

Damit wir Aussagen über Konfidenzbereiche des Strukturbruches treffen können, sind wir an der asymptotischen Verteilung des Schätzers interessiert. In unserem Modell nehmen wir an, dass genau ein Strukturwechsel auftritt und die Regressionsfunktion des zugrunde liegenden AR Prozesses mittels eines Neuronalen Netzes mit einer verdeckten Schicht approximiert wird. Ausgehend von Strukturbruchtests erhalten wir einen Schätzer für den Zeitpunkt der Änderung, dessen asymptotische Eigenschaften wir genauer untersuchen. Es kann gezeigt werden, dass dieser für eine große Klasse von Alternativen asymptotisch konsistent ist mit Rate [fn]O_P(1/n)[/fn]. Diese Rate kann nicht verbessert werden, wie die asymptotische Verteilung des entsprechend standardisierten Schätzers zeigt. Simulationen illustrieren das Verhalten des Schätzers mit einem besonderen Fokus auf den misspezifierten Fall, bei dem die Regressionsfunktion tatsächlich nicht durch ein neuronales Netz gegeben ist. Schließlich wenden wir den Schätzer auf Datensätze an. Der Vortrag basiert auf einer gemeinsamen Arbeit mit Claudia Kirch.

328: Industry Day

Time: Wednesday, 20th Mar 2013: 9:00am - 10:00am · Location: KG III, HS 3042

328: 1

Requirements of an SPC System -- shown alongside the example of InfinityQS

Boris Kulig

Statcon, Germany

The requirements for any SPC system are briefly introduced and will be shown alongside a practical example using InfinityQS software.

1.) Flexible dataentry: We differentiate between manufacturing master data and measurements. Masterdata are e.g. specification limits or control limits which have to be entered only once and the measurements are compared to them. Measurement data might be entered manually or read directly from the gauge using an interface.

2.) Online reactions: on the basis of both data typs the SPC-systems takes "online" or realtime decisions and actions, e.g. corrective actions.

3.) Speed: SPC systems gather data from many positions across a company, which implies that it must be guaranteed that no value gets lost and that data will not accumulate somewhere without being processed. The system speed is a very critical parameter for many SPC implementations. Some key data and database systems, which fulfill different requirements, will be discussed.

4.) Data analysis: The typical analysis within SPC systems focuses on the recent or timewise actual data values. Thus it must be possible without complex IT knowledge to query these data and directly analyze them using statistical or graphical procedures. Any arbitrary data selection and/or stratification should also easily be possible, because this might limit the use of the data to support the continuous learning process.

5.) Data integrity: Editing and changing data is only possible when there is an audit trail.

338: Industry Day

Time: Wednesday, 20th Mar 2013: 10:30am - 11:30am · Location: KG III, HS 3042

338: 1

Active quality management with LineWorks SPACE

Steffen Meyer

camLine, Germany

The Advanced SPC and Quality Management System LineWorks SPACE is the backbone of modern high-tech mass production lines worldwide. Current conditions in manufacturing, such as limited resources, time pressure, high data volume and high product mix require automation of the QM decision-making chain - from data collection and evaluation to initiating corrective actions in the MES. Audits require traceability of all decisions. Engineers must be able to transfer the SPC methodology simply in their process environment

Based on the requirements of highly automated FABs their implementation is presented with LineWorks SPACE. Focus of the presentation is the area of Real-Time-SPC, the automated detection of process deviations and their treatment by OCAP workflows and Corrective Actions with feedback to MES and APC scenarios. An important factor for the acceptance of SPC systems is the easy maintainability with low resources. With more than 10,000 active rules or control charts, the automation of the SPC administration is required. SPACE offers features such as pre-run and automated calculation of control limits. This reduces the effort required to maintain the system in large production environments dramatically. SPACE Examples are practically demonstrated.

330: Plenary session: Event history (Aalen, van den Berg)

Time: Wednesday, 20th Mar 2013: 10:40am - 12:00pm · Location: KG II, Audimax
Session Chair: Clarice Garcia Borges Demetrio

330: 1

Causal inference and survival analysis

Odd O. Aalen

University of Oslo, Norway

Modern causal inference represents a useful way of thinking about the ever-present causal issues in medical data. We shall cover two aspects:
- The marginal structural model of James M. Robins is an important tool for handling the thorny issue of time-dependent confounding. We shall present another approach, based on a sequential Cox regression.
- Mediation analysis of survival data has been sorely missing. In many settings one runs Cox analyses with baseline covariates while internal time-dependent covariates are not included in the analysis due to perceived difficulties of interpreting results. At the same time it is clear that the time-dependent covariates may contain information about the mechanism of the treatment effects. We shall discuss the use of dynamic path analysis for studying this issue.

330: 2

Testing for unobserved heterogeneity in survival analysis

Gerard J. van den Berg

University of Mannheim, Germany

tba

341: Statistics in Clinical and Preclinical Research -- Exposure-response Modelling and Dose Finding

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG II, Audimax
Session Chair: Günter Heimann

341: 1

Designing a Trial to find the Minimum Inhibitory Concentration of an Anti-Infective drug

Roland Fisch[1], Baldur Magnusson[1], Yi Cheng[2]

[1]Novartis Pharma AG, Switzerland; [2]Novartis Pharma, Shanghai, China

In the early clinical development of an anti-infective compound, one crucial aspect is to characterize the exposure-response relationship, where the response is a count of infective individuals (e.g. viral load, parasite or bacteria count). Ideally, the relationship should be explored not only for high concentrations, where the drug effect is saturated, but down to concentrations close to the minimum inhibitory concentration (MIC), i.e. the exposure for which the counts' time course is flat. This is particularly important when combination treatments should be developed later.

We describe an adaptive trial design to characterize exposure-response, taking into account restrictions in several respects: Firstly, there is a lower limit of quantitation (LLOQ) of the outcome variable; all count data below LLOQ will be of little use; high doses may therefore contribute little information, since most measured outcomes will be below LLOQ. Secondly, very small doses may have insufficient clinical efficacy, and may therefore not be ethically acceptable. Thirdly, interim decisions may have to be taken based on partial data; e.g. exposure (i.e. PK) may not be available instantly, such that interim decisions may have to be taken based on response data alone. Trial simulations are done in order to optimize the adaptive design, under such constraints.

341: 2

Analysis of clinical trials with biologics using dose-time-response models

Markus Reiner Lange, Heinz Schmidli

Statistical Methodology, Novartis Pharma AG, Switzerland

Biologics such as monoclonal antibodies are increasingly and successfully used for the treatment of many chronic diseases. Unlike conventional small drug molecules, which are commonly given as tablets once daily, biologics are typically injected at much longer time intervals, i.e. weeks or months. Hence both the dose and the time interval have to be optimized during the drug development process for biologics. To identify an adequate regimen for the investigated biologic, the dose-time-response relationship must be well characterized, based on clinical trial data. The proposed approach uses semi-mechanistic nonlinear models to describe the dose-time-response relationship. Both likelihood-based and Bayesian methods for inference and prediction are discussed. The methodology is illustrated with data from a clinical study in an auto-immune disease.

341: 3

Analysis of dose-finding studies with active control

Hans-Joachim Helms[1], Norbert Benda[2], Jörg Zinserling[2], Tim Friede[1]

[1]Department of Medical Statistics, Universitätsmedizin Göttingen, Germany; [2]Federal Institute for Drugs and Medical Devices, Bonn, Germany

In a dose finding study with an active control a new drug with several dose levels is compared with an established drug (the active control). The aim of such studies is to find the target dose concentration d which leads to the same efficacy as the active control (Källén and Larsson, 1999; Benda et al., 2009). For this purpose the intersection point of the dose response function with the efficacy of the active control has to be derived. Also a confidence interval of the dose d is of interest. Linear (log-)dose response curves were investigated in Helms et al. (2012). In this talk we will focus on nonlinear dose-response curves and will introduce a flexible approach using polynomial interpolation (Stoer and Bulirsch, 2007). This new approach will be compared with parametric standard methods as delta approximation or a profile likelihood approach (Pawitan, 2001). The methods will be illustrated with an example of a phase 2 dose-finding study in diabetes mellitus.

References

Benda N, Branson M, Maurer W, Friede T (2009). Drug Development 4: 84-88.

Freund R W, Hoppe R (2007). Stoer/Bulirsch: Numerische Mathematik 1, 10. Auflage, Springer Verlag, Berlin. S. 39-57.

Helms HJ, Benda N, Friede T (2012). Point and interval estimators of the target dose in clinical dose-finding studies with active control. (Submitted)

Källén A, Larsson P (1999). Statistics in Medicine 18 (6), 629–641.

Pawitan Y (2001). Oxford University Press.

341: 4

Sample size for Phase II dose-response trials

Björn Bornkamp[1], Jan Rekowski[2]

[1]Novartis, Switzerland; [2]TU Dortmund, Germany

Late-phase II trials are of central importance in pharmaceutical development as they generate the information that is ultimately used to plan large-scale confirmatory trials. The two major questions at end of Phase II are (i) whether one should move into confirmatory development and (ii) what dose(s) to use. The sample size for Phase II, however is typically calculated based on establishing statistical significance of comparing active doses to placebo, which addresses the real questions at end of Phase II only partially. In this talk I provide some considerations on alternative approaches for sample size calculation, for example when dose-response modelling is used for analysis.

342: Survival and Event History Analysis – Multistate models and time-varying covariates

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1015
Session Chair: Odd O. Aalen

342: 1

A Regression Model for the Excess Length-of-Stay in Hospital due to Nosocomial Infections

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

[1]Freiburg Center for Data Analysis and Modeling, University of Freiburg, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

The occurrence of a nosocomial infection (NI) constitutes a major complication that can be severe both in terms of mortality and morbidity. NI often leads to a prolonged length of stay (LoS) in the hospital which is one of the main driver for extra costs due to NI. Excess LoS due to NI is often used in cost-benefit studies that weigh the costs of infection control measures like isolation rooms against the costs raised by NI.

Estimation of extra LoS is complicated by the fact that the occurrence of NI is time-dependent. Cox proportional hazards models including NI as a time-dependent covariate could be used but do not allow to quantify the extra days spent in hospital following an infection. Using the multistate model framework, Schulgen and Schumacher (1996) devised a way to quantify extra LoS comparing the mean LoS given current NI status and averaging this quantity over time. It has also been extended to the competing risks setting in order to distinguish between discharge alive and death.

However, a way of studying the impact of covariates on the excess LoS is so far limited to stratified analyses. We propose to use the pseudo value regression technique (Andersen, Klein, Rosthøj, 2003). The idea is to use a generalized estimating equation model on the pseudo-values of the extra LoS. Motivated by a study on hospital-acquired infection, we investigate the use of pseudo values regression for identifying additional risk-factors that influence the extra LoS.

342: 2

Analysis strategies for complex treatment patterns in observational data from a clinical cancer registry

Irene Schmidtmann[1], Arndt Weinmann[2], Harald Binder[1]

[1]IMBEI, Universitätsmedizin Mainz, Germany; [2]Medizinische Klinik, Universitätsmedizin Mainz, Germany

In contrast to randomised clinical trials in which well-defined groups of patients are treated according to standardised protocols, clinical reality is more complex. Treatment strategies are modified over time, often depending on outcomes of previous interventions or the availability of new treatment modalities. Therefore it is necessary to develop statistical analysis strategies which can take such complexity into account in a time-to-event setting. Competing risks approaches and models with time-dependent covariates provide different perspectives on the data. We specifically investigate the usefulness of such tools as parts of an analysis strategy when considering the survival of patients from the beginning of transarterial chemoembolization (TACE) treatment and the transition to other treatment options such as Sorafenib, as documented in the Mainz hepatocellular carcinoma registry. We model the transition to death or to other treatments as competing risks. Survival after beginning of TACE is modelled using Cox regression with time-dependent covariates, including further treatment and time-varying laboratory values. Cumulative incidence functions prove to be a useful tool for distinguishing factors with weak or strong influence and for determining the period of influence. Cox regression models are found to be useful for disentangling the effect of many potentially influential baseline factors and intermediate observations. The joint interpretation of the results from these different tools provides considerable insight into treatment decisions and patient status. However, such a summary over different techniques also is rather difficult to perform, and we indicate several potential pitfalls.

342: 3

Estimation of stochastic employment and unemployment intensities

Andreas Groll, Francesca Biagini, Jan Widenmann

Ludwig-Maximilians University, Munich, Germany

One result of the current European debt crisis are high unemployment rates, burdening both public and private unemployment insurance systems.

An adequate class for the underlying stochastic processes are F-doubly stochastic Markov chains (Jakubowski and Niewkeglowski, 2010), allowing for matrix-valued stochastic intensity processes and incorporating covariate processes. Such models have already been extensively studied, see e.g. Kaplan and Meier (1958), Cox (1972, 1975) and Andersen et al. (1993). They base on the theory of multivariate counting processes and their compensators. For the simple case of Markov chains with deterministic intensity matrices, Andersen et al. (1993) describe the connection between multivariate counting processes and their underlying multi-state processes. The connection of counting processes with stochastic compensators to their underlying multi-state processes is considered in the present talk. We show that the class of F-doubly stochastic Markov chains is a natural candidate for the underlying multi-state processes corresponding to multivariate counting processes with stochastic compensators.

The intensity processes are estimated using the R-package mstate (de Wreede et al., 2011), which is based on the theoretical findings in Andersen et al. (1993). In cooperation with the IAB, a dataset of integrated labour market biographies (SIAB) is provided, including the duration of employment and unemployment periods between 1975-2008 of more than 1.5 million individuals together with socio-demographic and macro-economic covariates.

An application of the estimated intensity processes is Monte-Carlo-simulation-based calculation of fair insurance premiums of unemployment insurance products. As general pricing rule for the simulations we adopt the benchmark approach (Platen and Heath, 2007).

342: 4

First in First out? Working Status of Students and Time to Degree at German Universities

Katja Theune

Universität Duisburg-Essen, Germany

This paper analyzes time to first degree at German universities. The data base is the “Absolventenpanel” 2001, a panel study conducted by the “Hochschul-Informations-System” (HIS). In international comparison the German university system is characterized by a long duration of study. One potential reason might be the high proportion of students reporting part-time work during their studies. This paper focuses on the relationship between the working status of German students and their time to degree. Besides that, additional individual characteristics, parental background and fields of study are included in the analysis. The descriptive analysis reveals a positive correlation between the amount of part-time work and the duration of study. In the empirical analysis the Cox Proportional Hazards model is applied. The results confirm that part-time work has an increasing effect on time to degree. This finding is further substantiated by the prolonging effect of the time spent at part-time work. As mainly socially underprivileged students are forced to engage in part-time working during their studies, the politics of introducing university fees can be expected to result in even longer times to degree in Germany.

343: Statistics for High Dimensional Data

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1098
Session Chair: Jörg Rahnenführer

343: 1

On the false discovery rate (FDR) control of adaptive multiple tests

Arnold Janssen, Philipp Heesen

Heinrich-Heine Universitaet Duesseldorf, Germany

A multiple test is able to reject a portion of high dimensional hypotheses. Valid procedures which control the FDR were introduced by Benjamini and Hochberg (1995) and by Storey, Taylor and Siegmund (2004) who proposed an adaptive multiple test. In this talk new adaptive tests are proposed which include the method of Storey et al. as special case.

343: 2

Exceedance control of the number of false rejections in multiple testing

Marsel Scheer[1,2,3]

[1]German Diabetes Center, Düsseldorf, Germany; [2]Leibniz Center for Diabetes Research, Heinrich Heine University Düsseldorf, Germany; [3]Institute for Biometrics and Epidemiology, Düsseldorf, Germany

Controlling the k-FWER at level alpha in multiple testing problems with [fn]m[/fn] null hypotheses means that the probability of rejecting at least [fn]k[/fn] true null hypotheses is bounded by alpha, cf. e.g. [1,2,3]. Considering [fn]k[/fn] as fixed may be viewed as unsatisfactory. In this talk we propose a new and more flexible approach, NFRX control for short, where [fn]k[/fn] is allowed to depend on the unknown number [fn]m1[/fn] of false null hypotheses (NFRX = Number of False Rejections eXeedance). For example, it seems more appropriate to require [fn]k = k(m1)[/fn] to be small (large) if the number of false null hypotheses is small (large). We present new methods for the construction of suitable multiple tests and sufficient conditions such that the NFRX is asymptotically ([fn]m[/fn] tends to infinity) controlled, that is, the probability of rejecting at least [fn]k(m1)[/fn] true null hypotheses is asymptotically bounded by alpha. The material presented here is part of [4].

References:

[1] Victor, N. (1982). Exploratory data analysis and clinical research. Meth. Inform. Med., 21, 53-54.

[2] Hommel, G. and Hoffmann, T. (1988). Controlled uncertainty. In: Bauer, P. et al. (Eds.): Multiple Hypothesenprüfung. Springer, Berlin, 154-161.

[3] Lehmann, E. L. and Romano, J. P. (2005). Generalizations of the familywise error rate. Ann. Statist., 33, 1138-1154.

[4] Scheer, M. (2012). Controlling the Number of False Rejections in Multiple Hypotheses Testing. Dissertation, Heinrich-Heine Universität, Düsseldorf.

343: 3

Some improvements for multiple testing procedures that control generalised error rates

Sebastian Döhler

Hochschule Darmstadt, Germany

We describe a method for improving some multiple testing procedures. The goal is control of some generalized error rates under arbitrary dependence of the [fn]p[/fn]-values. The main idea is to view the involved critical constants as solutions of certain sets of linear inequalities. This property can be used to find optimal solutions to an associated linear programming problem. This approach is based on the work of Lehmann and Romano (Ann. Stat. 33 (2005), 1138 - 1154) and Romano and Shaikh (Ann. Stat., 34 (2006) 1850-1873; and The second Erich L. Lehmann symposium (2006), 33-50).

343: 4

Goodness-of-fit tests in terms of local levels and related processes

Veronika Gontscharuk[1,2,3], Sandra Landwehr[1,2,3], Helmut Finner[1,2,3]

[1]German Diabetes Center, Düsseldorf, Germany; [2]Leibniz Center for Diabetes Research, Heinrich Heine University Düsseldorf, Germany; [3]Institute for Biometrics and Epidemiology, Düsseldorf, Germany

Many goodness-of-fit (GOF) tests can be viewed as simultaneous multiple test procedures based on p-values for n hypotheses, where

[fn]p[/fn]-values under the null hypotheses are iid uniformly distributed on [fn][0,1][/fn]. Typically, such GOF tests are sensitive only for some special kind of alternatives. For example, the classical Kolmogorov-Smirnov test has higher power for alternatives that differ from the null distribution in the central range, while so-called Higher Criticism (HC) tests are sensitive only in the moderate tails. In order to compare various competing GOF procedures we introduce a new measure indicating local power, so-called local levels, which are defined as the probability that the ith smallest [fn]p[/fn]-value exceeds the corresponding critical value. Then we investigate finite and asymptotic local levels corresponding to Kolmogorov-Smirnov and HC tests. We study their properties and provide appropriate formulas for asymptotic local levels. Based on the theory of stochastic processes we show why HC tests are sensitive only in tails. Moreover, by means of local levels we demonstrate extremely slow convergence when applying asymptotic results. As a consequence, the calculation of suitable critical values for HC tests remains a challenging problem if [fn]n[/fn] is fixed. We propose a new approximation for the distribution of the HC test statistic in this case. With respect to family wise error rate control, critical values based on the new approximation seem to be much better than critical values related to the limiting distribution of the standardised HC test statistic, that is, the Gumbel distribution.

344: Penalized and Regularized Regression Analysis

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1010
Session Chair: Göran Kauermann

344: 1

Smooth additive models for large datasets

Simon N Wood

University of Bath, United Kingdom

Datasets of with hundreds of thousands or millions of observations are increasingly common, and smooth regression models sometimes have something useful to offer in their analysis. This talk reviews some of the options for making generalized additive model fitting methods suitable for use with such large datasets, and shows how some relatively straightforward approaches can lead to well founded methods that are surprisingly efficient, and make good use of existing methods for less large datasets.

344: 2

Objective Bayesian model selection with the deviance statistic

Daniel Sabanés Bové, Leonhard Held

University of Zurich, Switzerland

In medical research and elsewhere, variable and model selection in regression is a common problem. The Bayesian approach to this problem is conceptually straightforward, since it melts prior assumptions with information from the data into the posterior distribution. The posterior distribution can then be used to quantify the importance of selecting variables (as posterior variable inclusion probabilities), ranking the models (by their posterior model probabilities) and estimating the effects of variables while taking the model uncertainty into account (via Bayesian model averaging of within-model posterior distributions). However, the results depend critically on the choice of the within-model prior distributions on the parameters. Often no prior information is available and hence default and objective prior distributions are required. For this situation, Johnson (2005) proposed to transform frequentist test statistics into Bayes factors between models, eliminating the need to directly specify prior distributions on model parameters at all. The method is applicable to generalised linear models and the Cox model using the deviance statistic, and only depends on a single hyperparameter [fn]g[/fn]. In this talk we will discuss the role of [fn]g[/fn] and introduce a conjugate hyperprior for it. Furthermore, we will carve out the strong connections of the test-based Bayes factors to shrinkage estimates for prediction (Copas, 1983) and to objective Bayesian model selection based on generalised hyper-[fn]g[/fn] priors (Sabanés Bové and Held, 2011). The methodology is then applied in selected biomedical applications, where simultaneous selection and transformation of covariates is required for developing prognostic models.

344: 3

Regularization in Finite Mixtures of Generalized Linear Models

Wolfgang Pößnecker, Gerhard Tutz

Ludwig-Maximilians-Universität München, Germany

Finite mixture models are a popular tool to deal with unobserved heterogeneity, which frequently occurs if the population under consideration consists of heterogeneous subpopulations with differing effect of the predictor variables on the response. Finite mixtures of generalized linear models use a seperate GLM for each potential subpopulation, so that they require substantially more parameters than ordinary GLMs. To tackle the consequential difficulties with stability and interpretation, we propose new regularization techniques for finite mixtures of GLMs. Existing approaches based on Lasso penalties are extended to yield selection of both coefficients and variables. Additionally, a novel reparameterization is introduced that allows to decide whether a variable is affected by the latent heterogeneity or if it has a fixed, homogeneous effect. The usefulness of our approach is demonstrated on real and simulated data.

345: Clustering

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG III, HS 3043
Session Chair: Hajo Holzmann

345: 1

From Local to Global Search in Modularity Clustering

Andreas Geyer-Schulz[1], Michael Ovelgönne[2]

[1]KIT, Germany; [2]University of Maryland, United States of America

The modularity clustering criterium of Newman and Girvan is a popular criterium for graph-clustering. In this contribution we analyze a randomized greedy algorithm and the core groups graph clustering scheme. For this purpose we describe the search space of the algorithm and its properties. We establish the structure of the Morse graph on the lattice of the algorithm's search space. The Morse graph can be partially constructed from a sample of locally optimal algorithm runs with the help of the core groups graph clustering scheme. The core groups graph clustering scheme is shown to be an efficient method to identify saddle points on the Morse graph which are promising restart points of the algorithm.

In the 2010 DIMACS clustering challenge, the family of randomized greedy algorithms with the core groups graph clustering scheme have dominated the modularity clustering challenge.

345: 2

Information content in clustering algorithms

Morteza Haghir Chehreghani, Ludwig Busse, Joachim M. Buhmann

ETH Zurich, Switzerland

Algorithms are usually analyzed according to their time and space requirements, i.e., speed and memory. Machine learning algorithms are additionally required to show robustness to input fluctuations. This work elaborates a new framework to measure the informativeness of algorithmic procedures and their stability against noise.

An algorithm is considered to be a noisy channel which is characterized by a generalization capacity (GC). Informative algorithms yield a high generalization capacity, whereas fragile algorithms suffer from a low capacity. GC optimally measures the generalization ability of algorithms to extract context-sensitive information w.r.t. a given output space. GC objectively ranks different algorithms for the same data processing task based on the bit rate of their respective capacities. The problem of grouping data is used to demonstrate this validation principle for clustering algorithms, e.g. K-means, pairwise clustering, normalized cut, DBSCAN and dominant set clustering. Our new validation approach selects the most informative clustering algorithm, i.e. the procedure which filters out the maximal number of stable, task-related bits relative to the underlying hypothesis class of data groupings. Particularly, the GC principle addresses the three relevant fundamental learning questions (1.) finding the optimal number of clusters for a given method, (2.) ranking different distance measures used in clustering methods, and (3.) validating different clustering mechanisms.

In an application example, we employ the generalization capacity principle to analyze clustering algorithms for experimental gene expression data. This analysis provides a worked-through procedure for ranking and validating different clustering algorithms.

345: 3

Benchmarking different clustering algorithms on functional data

Christina Yassouridis, Friedrich Leisch

Universität für Bodenkultur Wien, Austria

The theoretical knowledge of clustering functions is still scarce and despite an increasing demand in many disciplines only a few methods are available in form of applicable code. Similar as in the non-functional case a division into model based, partitioning or hierarchical clustering can be made but the distance measure has to be redefined since curve points are not independent any longer. This can become a difficult task especially when the measurement grid is sparse. In literature it is often solved by projecting the functions on a spline basis and building a fixed or random effects model of the basis coefficients or by treating them as outcome of a stochastic process represented by a time series for instance a Karhunen-Loeve expansion (see Gareth James 2003, Peng Jie 2008). All concepts usually involve various parameters, among them number of basis functions, projection dimension, maximum iteration etc. Published algorithms usually work well on the data presented in the article but their performance has never been tested objectively on other data sets nor against each other. The purpose of the presentation is to give an overview of several existing methods for clustering functional data. An outline of their theoretic concepts is given and practicability is verified. This includes the efficiency and sensitivity to hyperparameters tested on different data sets registered on a regular and irregular basis. First data sets with known class membership are generated and finally the effectiveness of the methods is demonstrated on a real life situation.

345: 4

Ensemble learning for density estimation

Friedhelm Schwenker

Ulm University, Germany

Estimation of probability density functions (PDF) is a fundamental concept in statistics and machine learning and has various applications in pattern recognition. In this contribution ensemble learning approaches will be discussed in the context of density estimation, particularly these methods will be applied to two PDF estimation methods: Kernel density estimation (or Parzen window approach) and Gaussian mixture models (GMM) (Fukunaga, 1990). The idea of ensemble learning is to combine a set of [fn]L[/fn] pre-trained models [fn]g_1,\ldots,g_L[/fn] into an overall ensemble estimate [fn]g[/fn]. Combining multiple models is a natural step to overcome shortcomings and problems appearing in the design of single models. Along with the design of single models [fn]g_i[/fn] an aggregation mapping must be realized in order to achieve a final combined estimate, usually this mapping has to be fixed a priori but trainable fusion mappings can be applied as well. Examples of fixed fusion schemes are median or (weighted) average of the predicted models (Kuncheva, 2004), e.g. weighted average is defined through

[fn]g_\mathbf{w} (\mathbf{x}) = \sum_{l=1}^L w_l g_l (\mathbf{x})[/fn]

with [fn]w_l \geq 0[/fn] and [fn]\sum_{l=1}^L w_l = 1[/fn].

For example, weighted averaging of kernel density estimates leads to a representation with a new kernel function. The proposed ensemble PDF approach will be analyzed by statistical evaluations on benchmark data sets. The behavior of these algorithms in classification and cluster analysis applications will be presented as well.

References:

KUNCHEVA, L. (2004): Combining pattern classifiers: Methods and algorithms, Wiley.

346: Robust and Distribution Free Statistics

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1199
Session Chair: Robert Hable

346: 1

Constrained simultaneous nonlinear quantile regression

Luca Sartore

University of Padua, Italy

Quantile regression provides a more informative description of the phenomenon under study without making any assumption on its underlying distribution; in fact, it provides a conditional quantile response for each probability level.

Several solutions were presented in the literature to provide simultaneous estimation methods in order to solve the problem of crossing quantile curves, such as the restricted quantile regression, the stepwise multiple quantile regression and the rearranged quantile regression.

When we deal with parametric nonlinear models, most of the current non crossing quantile regression techniques require adaptation case by case. Instead, constrained simultaneous nonlinear quantile regression can be considered as a flexible tool which allows to satisfy the non crossing constraints and the theoretical properties of the model under study. In fact, this technique estimates the parameters of a nonlinear model where some of the parameters are fixed for each probability level and others are allowed to vary.

A practical example is presented which involves the analysis of the data related to a diffusion process. In particular, the Bass model and its generalisations are considered.

346: 2

Empirical assessment of robust principal components analysis to investigate genetic population substructure

Carine Legrand, Justo Lorenzo Bermejo

Institute of Medical Biometry and Informatics, University of Heidelberg, Germany

Population stratification plays a major role in genetic association studies. For example, hidden genetic structure may result in false positive associations, and the identification of genetically similar populations may guide the selection of suitable reference panels for genotype imputation. Principal component analysis (PCA) is often used to detect population stratification, and the estimated principal components are frequently integrated into the subsequent regression models used to identify genetic associations. However, standard PCA is very sensitive to individuals with a departing genotype, who are commonplace in association studies. The first estimated components are often attracted towards these outliers instead of reflecting the ancestry of the majority of individuals.

We therefore explored the potential of several robust PCA methods - including projection pursuit approaches, ROBPCA and spherical PCA - applied to a subpopulation phenotype and single nucleotide variants. Real data from the HapMap project were used to assess inter-group and intra-group variances, as well as the classification accuracies based on standard and robust PCA. In order to compare the ability of standard and robust methods to mirror population evolution, we conducted simulations relying on the population genetics simulation environment simuPOP, complemented with our own scripts. The investigated scenarios included recombination, subpopulation isolation, bottlenecks and selection pressure on a few loci. Preliminary results show an extreme down-weighting of complete outlying groups of individuals by some robust approaches.

346: 3

Comparative Performance of Robust Logistic Regression in the Framework of Genetic Risk Prediction

Miriam Kesselmeier, Justo Lorenzo Bermejo

Universität Heidelberg, Germany

The relationship between a binary disease phenotype and inherited genetic variants is often investigated by logistic regression. The strong influence of few observations deviating from the majority of the data (outliers), which is particularly relevant in the context of high-dimensional genetic data, is a major limitation of standard procedures used to estimate regression parameters, e.g. iteratively reweighted least squares. We have investigated the possible benefit of robust logistic regression taking advantage of the class of M-estimators for generalized linear models proposed by Cantoni and Ronchetti. Using simulated and real data, we assessed the influence of outliers on the estimated odds ratios, areas under receiver operating characteristic curves, clinical net benefits and integrative discrimination indexes relying on both standard and robust parameter estimates.

Our results confirmed a substantial effect of single outliers on the estimated genotype relative risks. Standard and robust logistic regression differed in the ranking of genetic variants by their strength of association. Standard and robust model selection resulted in different prediction models. The clinical net benefit estimated by leave-one-out cross-validation was larger under robust logistic regression for several scenarios.

Preliminary findings indicate some advantage of robust statistics in the context of genetic association studies. The theoretical background and technical details will be discussed during the meeting.

346: 4

Comparison of Genotype Imputation Accuracy Between Different Approaches to Identify the Population of Reference

Barbara Peil[1], Maria Kabisch[2], Christine Fischer[3], Carine Legrand[1], Miriam Kesselmeier[1], Ute Hamann[2], Justo Lorenzo Bermejo[1]

[1]Institute of Medical Biometry and Informatics, University of Heidelberg, Germany; [2]Molecular Genetics of Breast Cancer (B072), Deutsches Krebsforschungszentrum (DFKZ), Heidelberg, Germany; [3]Institute of Human Genetics, University Hospital Heidelberg, Germany

Genotypes are imputed in association studies to improve statistical power and to enhance fine-mapping resolution. A population of reference, usually denominated ‘reference panel’, is generally required to impute genotypes that have not been directly measured in the association study. The selection of an ‘optimal’ panel of reference is therefore crucial and different selection strategies have been proposed.

The standard approach for panel selection is to identify the subpopulation in a public data repository, e.g. the HapMap Project, which is genetically most similar to the study population. The genetic similarity between the study population and possible reference panels is often quantified by Wright’s FST statistic.

The aim of the present work is to develop alternative strategies in order to select a reference panel for genotype imputation. The investigated strategies rely on a genetic principal component analysis (PCA) of the study population and subpopulations in HapMap. Based on this PCA, a reference panel is selected according to a bivariate generalization of the nonparametric univariate boxplot denominated bagplot, and according to the univariate depth of the first principal components. In contrast to the standard approach, alternative strategies for panel selection are expected to be robust against individuals with an outlying genetic background.

Preliminary results suggest that an alternative strategy based on a bagplot of the first two principal components outperforms the standard approach. The first genetic principal component seems to be particularly relevant to imputation accuracy. Detailed results will be provided at the conference.

354: Benefit Assessment in Public Health -- The Clincial Benefit of Diagnostic Methods: General Issues: Common Workshop of IBS-DR, GMDS and IQWiG

Time: Wednesday, 20th Mar 2013: 1:00pm - 2:20pm · Location: KG I, Aula
Session Chair: Werner Vach
Session Chair: Ralf Bender

354: 1

Evidence-based testing: the architecture of medical test evaluations

Patrick Bossuyt

Universiteit van Amsterdam, The Netherlands

The unraveling of the human genome and the expansion of genomics, proteomics and metabolomics have fuelled the development of novel biomarkers. New in-vitro diagnostics other forms of medical testing and can improve health care, bringing us closer to stratified and personalized medicine. At the same time, society is concerned about the never-ending increase in health care expenditure while other groups lament a creeping medicalization through the use of novel forms of testing. Increasingly, decision-makers, physicians and other users request evidence that using medical tests are beneficial. Like any other intervention in health care, testing should prevent premature death, restore functioning, or simplify health care as it is. Health care policymakers are calling on manufacturers to shift from a narrow technical or biomedical perspective to a wider one, one that considers whether the diagnostic technology improves final outcomes in typical patient populations. Before recommending the use of diagnostic tests and markers, and before deciding on their reimbursement, decision-makers and users now want to see data that testing actually improves outcomes in relevant patient populations, or that it enhances patient outcome, health care quality, efficiency and cost-effectiveness. In this presentation we will highlight the state of the art in evaluations of medical tests. We will distinguish between technical performance, clinical performance, and clinical effectiveness. Problems with designing and evaluating randomized trials of testing will be highlighted. Way to improve the relevance of evaluations of technical performance and clinical performance will be demonstrated.

354: 2

Assessing Diagnostic Technologies for Coverage – What would be "Proper Reasons" to Relinquish Patient Outcome Studies?

Philipp Storz-Pfennig

GKV-Spitzenverband, Germany

Traditionally diagnostic technologies have received less attention and less methodological rigour was exercised in their assessment compared to e. g. drugs or devices in coverage decisions. Increasing use and innovations regarding diagnostics (e. g. advanced imaging, biomarkers and patient monitoring systems) often outpaces those of therapeutics, with accompanied cost growth. More reliable frameworks of how to put tests to the test regarding their contribution to, ultimately, patient benefit is called for. From the point of view of a coverage decision making body (the German Federal Joint Committee – FJC, composed of statutory health insurance, doctor, hospital and patient representatives deciding on the benefits catalogue for about 70 million beneficiaries) a major problem point is highlighted and respective methodological questions discussed. The FJC procedural guidelines determine that, also for diagnostics, outcome studies of reliable designs (e. g. randomized trials) should be preferred, while decisions based on less reliable or more indirect evidence (e. g. diagnostic accuracy studies) may be possible – if "proper reasons" can be given to relinquish outcome studies. Using examples from past and present assessments/appraisals, to infer what has be considered proper in the past – and what might be in the future, given more rigorous methodological thinking. While it will be in the remit of the FJC to determine what will be considered proper reasoning, support from methodological research might be crucial to understand decision uncertainties involved and might contribute to transparency and propriety of decisions

354: 3

From comparative accuracy studies to diagnostic RCTs: A long way?

Werner Vach[1], Bettina Schnitter[2]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]University Medical Center Freiburg, Germany

To assess the additional clinical benefit of a new diagnostic method, well conducted RCTs with patient relevant outcomes may provide information at a high evidence level. They can overcome limitations of comparative accuracy studies, which can only study the improvement in sensitivity and specificity, but do not allow directly to assess the benefit of improved patient management due to better diagnoses. However, comparative diagnostic RCTs are still rare. We discuss some potential issues which have to be taken into account in planning comparative diagnostic RCTs: Choice of patient relevant outcome, sample size, standardization of management procedures, standardization of population and patient recruitment, center differences, speed of modernisation, and limitations in generalizability. We approach this using results from a systematic review of comparative diagnostic RCTs involving PET/CT as one diagnostic modality, allowing to look at how these issues are currently tackled. All issues demonstrate, that comparative, diagnostics RCTs require a careful planning and cannot solve all problems in diagnostic research. Hence prior to a comparative, diagnostic RCT the evidence from comparative accuracy studies combined with external evidence on the impact of potential management options should be carefully aluated to avoid unnecessary RCTs and unnecessary delays in providing new diagnostic modalities to all patients.

348: Industry Day

Time: Wednesday, 20th Mar 2013: 1:30pm - 2:30pm · Location: KG III, HS 3042

348: 1

Statgraphics Centurion

Helga Meinhardt

DPC, Germany

STATGRAPHICS Centurion is one of the few available statistical software applications that is flexible enough to provide entry-level access for virtually everyone within a company, while still ensuring that high-level statistical algorithms are available for addressing complex issues and calculations. It is a statistical software package that will serve a range of users from machine operators and shop floor supervisors to design and process engineers. Since its development in the early 1980s, STATGRAPHICS has concentrated on providing statistical tools that can be used both to design quality into products, and ensure that acceptable quality is maintained throughout the production process. This approach fits well with the broad Six Sigma mandate to develop an in-depth understanding of the philosophy as well as the theory, tactics, strategy, and application tools. STATGRAPHICS Centurion is designed for anyone who wishes to do serious data analysis without investing weeks learning how to use a statistical package. Everything is completely menu-driven, and there are tools such as the StatWizard and StatAdvisor to help you run the program most effectively. During the presentation we will address topics as data import, analysis capabilities, automatic run e.g. of control charts or reporting features.

351: Statistics in Clinical and Preclinical Research -- Missing Data Analyses

Time: Wednesday, 20th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1199
Session Chair: Michael Kunz

351: 1

Sensitivity Analysis for Recurrent Event Data subject to Informative Dropout

Mouna Akacha[1], Emmanuel Ogundimu[2]

[1]Statistical Methodology, IIS, Novartis Pharma AG, Switzerland; [2]Statistics Department, University of Warwick, United Kingdom

The studies that motivated this work seek to analyze processes which generate events repeatedly over time. Such processes are referred to as recurrent event processes. Examples include seizures in epileptic studies, flares in gout studies or occurrence of cancer tumors.

Interest lies in understanding the underlying event occurrence process. This includes the investigation of the rate at which events occur, the inter-individual variation, and most importantly, the relationship between the event occurrence and covariates such as treatment.

One considerable challenge in analyzing recurrent event data arises when a large proportion of patients discontinues before the end of the study, e.g. due to adverse events, leading to partially observed data. Any analysis of such data relies on untestable assumptions regarding the post-discontinuation behaviour of patients that drop out early. Regulatory agencies are therefore increasingly asking for sensitivity analyses which assess the robustness of conclusions across a range of different assumptions.

Sophisticated sensitivity analyses for continuous data, e.g. using the pattern-mixture model approach, are being increasingly performed. However, this is less the case for recurrent event or discrete data.

In this talk, we will present an approach to perform sensitivity analyses for recurrent event data. This approach bases on a pattern mixture model framework. Different assumptions about the future behavior of dropouts dependent on reason for dropout and received treatment can be incorporated.

We illustrate our approach with a clinical trial in patients who suffer from bladder cancer.

351: 2

Treatment of Non-ignorable Missing Data under Heterogeneity

Thomas Lehmann, Peter Schlattmann

University Hospital Jena, Friedrich Schiller University Jena, Germany

Statistical analysis of clinical data is often complicated by missing values. Excluding patients with missing data from analysis not only reduces statistical power, but also often leads to biased estimates. Multiple imputation (MI) is a statistical technique to deal with incomplete data sets, and if the data are missing at random (MAR, i.e. observed data contain useful information for predicting the missing values), MI yields unbiased estimates of the unknown parameters.

Although the theoretical background of multiple imputation is well-established, there is still a lack of practical applications, in particular if imputation and parameter estimation have to be realized simultaneously. One example is the statistical analysis of heterogeneous data: Here observations can be viewed as arising from a mixture of different populations, finite mixture models provide a natural framework to account for heterogeneity. The proportions of the mixture distribution are usually unknown, and the EM algorithm is often applied to estimate mixing weights as well as the parameters of the component densities.

We present an approach to incorporate both problems of unknown indicator variables and missing values in mixture models. The focus is on data which are missing not at random (MNAR), i.e. the probability of missingness depends on the unobserved value itself. Since the observed data contain no information to distinguish MNAR from MAR, different assumptions have to be made about the mechanism to reflect the uncertainty about the missing values (sensitivity analyses). We provide the practical implementation of this approach to analyse a clinical incomplete dataset.

351: 3

Effect estimation in clinical trials accounting for treatment discontinuation

Gerd Karl Rosenkranz

Novartis Pharma AG, Switzerland

To establish the effect of an investigational treatment, randomized, controlled clinical trials are designed where patients are expected to follow a pre-defined treatment schedule for both the new treatment and a control. During the conduct of the trial some patients may decide that continuing treatment as per protocol is no longer warranted for various reasons or an investigator may conclude that further treatment is no longer in the best interest of the patient. These situations result in discontinuation of treatment and as a consequence in missing data. Even if patients are further observed according to the original schedule to avoid missing data -- as expected by Health Authorities -- it is not obvious how to make best use of the data after treatment discontinuation for estimating a treatment effect. This presentation discusses a

proposal based on the concept of principal stratification (Frangakis and Rubin, 2002).

351: 4

Incorporating Age- and Sex-dependent Reference Ranges in Joint Longitudinal and Time-to-Event Modelling for Bone Marrow Transplantation

Markus Christian Elze[1], Annekathrin Heinze[2], Stephan Klöss[3], Oana Ciocarlie[4], Melanie Bremm[2], Ulrike Köhl[3], Jane Luise Hutton[1]

[1]Department of Statistics, University of Warwick, United Kingdom; [2]Department of Medicine, Goethe-University Frankfurt, Germany; [3]Integrated Research and Treatment Center Transplantation, Hannover Medical School, Germany; [4]Pediatric Department, Timisoara University of Medicine and Pharmacy, Romania

The incorporation of time-varying data in survival models is a common objective in areas where longitudinal measurements are collected at arbitrary time points, such as clinical trials or the social sciences. Joint modelling of longitudinal measurements and time-to-event data is a natural solution to this problem, but the amount of available data may limit the use of joint models. Here, we show that transforming the longitudinal data using additional information from external sources may increase the amount of information gained from the data.

‘Bone marrow transplantation’ is a potentially curative treatment option for different hematologic disorders, such as severe leukaemia. However, it is still associated with high mortality rates due to complications after transplantation. Early identification of high-risk patients is crucial for successful intervention. Thus, predictive models are needed to assist clinical decision making. The development of longitudinal immune measurements is relevant for complication prediction and should be considered in modelling. As studies are often faced with limited patient numbers, assessing the immune recovery may be difficult. Here, we demonstrate how the use of reference data from healthy persons may assist the model fitting.

Age- and sex-dependent reference ranges are created from 100 healthy children for several immune subpopulations using the LMS method (Cole et al.). These are then employed to assess the immune recovery for 67 paediatric patients who underwent bone marrow transplantation. The performances of joint models with and without the use of reference ranges are compared. The use of these models in clinical practice is discussed.

352: Survival and Event History Analysis -- General Models and Methods

Time: Wednesday, 20th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1015
Session Chair: Marco Burkschat

352: 1

Incomplete Repair in Degradation Processes

Waltraud Kahle

Otto-von-Guericke University, Magdeburg, Germany

In the last years, there is an increasing number of papers concerning with incomplete (or general) repair models where the degree of repair is somewhere between a renewal (as good as new) and a minimal repair (as bad as old).

If a failure is the result of an underlying degradation process, then we have a natural interpretation of an incomplete repair: the repair action decreases the degradation level.

We consider the Wiener process with drift as a model of damage and degradation. A failure occurs when the degradation reaches a given level h first time. For preventive maintenance, inspections of the degradation are regularly carried out. If at inspection time the degradation is larger than a predefined level a, then the damage level is decreased. The costs of a maintenance depends on the degree of repair.

In the talk, we consider the problem of defining cost optimal maintenance policies.

352: 2

Maximum likelihood estimation for left- and right-censored survival times with time-dependent covariates

Alexander Kremer, Rafael Weißbach, Friedrich Liese

University of Rostock, Germany

There is a large body of papers on parameter estimation for right censored data. The proofs for the asymptotically normality of the MLE are typically based on techniques from martingale theory. This is not applicable for left censored data that appear in the estimation of the default probability.

Motivated by this application we study left censored data that come from duration times with time-dependent covariates in the hazard rate and unknown censoring times. As the log-likelihood is concave in the parameter we are able to give a short proof for the asymptotic normality by applying results on convex processes due to Hjord and Pollard (1993). This new technique is also applied to give a new short and direct proof for right censored data, without reverting to martingales.

Monte carlo simulations are carried out to check the actual level of the confidence intervals that are based on the asymptotic normality of the MLE. The application to data from practice concerns the estimation of the intensity of credit rating transitions. Left-censored observations, as compared to right-censored, yield a remarkable difference for the one-year probability of default, an important parameter in finance. By time-dependent covariates we find that the credit cycle has a relevant impact on the default probability.

352: 3

New Estimation Methods for the Hazard Ratio in Clinical Trials with Survival Endpoints

Sandra Ligges

Institut für Biometrie und Klinische Forschung (IBKF), Universität Münster, Germany

Point and interval estimation of the hazard ratio in clinical trials with survival endpoints is usually attained by the point estimator and confidence interval resulting from a Cox proportional hazards model. Wassmer (2006) constructed a mathematically closed point estimator based on the approximate expected value of the Logrank test statistic of Schoenfeld (1981). Trading variance for bias this estimator is advantageous to the Cox estimator in the MSE if the hazard ratio is not extreme. A bias correction of the Wassmer estimator leads to a new approximately unbiased estimator with smaller variance than the Cox estimator in the same settings.

Besides by making use of some ideas of Schoenfeld (1981) another new point estimator as well as a new confidence interval have been constructed. Up to a moderate treatment difference and if the sample size is not too large the new point estimator performs best in the MSE among all considered point estimators, and the new confidence interval reduces the expected interval width compared to the Cox model interval while keeping the aspired confidence level.

In this talk, all these new estimation methods are presented and compared to each other as well as to the already established estimators by showing some results of an extensive simulation study composed of many practical relevant scenarios.

References:

[1] Schoenfeld, D. (1981): The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions. Biometrika, 68 (1), 316-

319.

[2]Wassmer, G. (2006): Planning and Analyzing Adaptive Group Sequential Survival Trials. Biometrical Journal, 48, 714-729.

352: 4

Type-I censored sequential order statistics

Marco Burkschat[1], Erhard Cramer[2]

[1]Otto-von-Guericke University, Magdeburg, Germany; [2]RWTH Aachen University, Germany

Sequential order statistics can be used to describe the successive failure times of components in a system, where failures may affect the performance of remaining components. If the observation of such failure times is stopped at a fixed threshold time, a Type-I censored sample is obtained. In the talk, conditional distributions of Type-I censored sequential order statistics are considered. In the particular case of an underlying exponential distribution, the distribution of corresponding spacings is examined. The results are used to derive distributions of maximum likelihood estimators in this setting.

353: Statistics for High Dimensional Data

Time: Wednesday, 20th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1098
Session Chair: Anne-Laure Boulesteix

353: 1

Combining multiple cancer data sets for survival models with high-dimensional genetic measurements

Jörg Rahnenführer, Michel Lang

TU Dortmund, Germany

Cancer survival analysis is often suffering from too small patient numbers in individual cohorts from single study centers. In the presence of high-dimensional genetic measurements, for example obtained from arrayCGH or gene expression experiments, a meaningful inference is even more challenging. Both for increasing prediction accuracy and for validation, the use of multiple data sets, especially also from different medical centers, is necessary. Our goal is to assess the heterogeneity between such data sets and investigate the potential of integrative analyses. We have gathered several breast cancer and lung cancer data sets with survival times, clinical covariates, and genetic high-dimensional covariates.

Two aspects are discussed. First, we present Cox models with clinical covariates and a single genetic covariate. In a genome-wide analysis we compare different strategies for combining the information from different studies, including simple pooling of data sets, stratified Cox models, and meta analysis. We point out and explain differences for selected examples. Second, we present results from a validation analysis. Cox models are estimated using the CoxBoost algorithm where clinical variables are included in the models as mandatory covariates and high-dimensional genetic variables as penalized covariates. We evaluate the performance of models trained and tested on different data sets. It is shown that even a cross-validated performance measure is often too optimistic regarding a different test set. Finally, we discuss a potential solution to this problem where in the model building process different data sets enter with different weights, accounting for their similarity to the test data set.

353: 2

Imputation approach as a basis for integrating multiple genome wide data with partial overlap

Stefanie Hieke[1,2], Richard F. Schlenk[3], Martin Schumacher[1], Axel Benner[4], Lars Bullinger[3], Harald Binder[1,5]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Freiburg Center for Data Analysis and Modeling, University Freiburg, Germany; [3]Department of Internal Medicine III, University Hospital of Ulm, Germany; [4]Division of Biostatistics, German Cancer Research Center, Heidelberg, Germany; [5]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Johannes Gutenberg University Mainz, Germany

Integration of multiple genome-wide datasets in risk prediction models relating to clinical endpoints might help to improve therapy management for patients. We consider a stepwise procedure analyzing first one molecular source and using the results for guiding another molecular source. A typical problem in integrative analysis is the small number of cases with parallel measurements from several molecular sources. Most patients will have measurements from a single source only. However, the extracted results from one source can only guide the model for another source for these samples with measurements form both sources. We investigate imputation approaches to allow the use of all available data. The proposed imputation approach forms the basis of the integration of different molecular sources with partial overlap. Imputation is based on the linear predictor, the signature comprising the most important information from the first source, as imputation at the measurement level given the multiple source of missing information would not be feasible. We investigate componentwise likelihood-based boosting and aggregated random forest techniques as imputation approaches. We illustrate the strategies in an application to survival data considering at least two different molecular sources with partial overlap in the samples. The competing approaches are judged with regard to prediction performance and the effect of the different strategies on the selected molecular entities. The results indicate that the integration of several molecular sources can identify molecular entities that could not be seen from the analysis of one single source alone.

353: 3

Strategies to derive combined prediction models using both clinical predictors and high-throughput molecular data

Riccardo De Bin[1], Willi Sauerbrei[2], Anne-Laure Boulesteix[1]

[1]University of Munich (LMU), Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In biomedical research, the goal of a prediction model is to provide a function useful to predict a specific disease outcome. In recent years, a lot of attention has been devoted to taking advantage of the information contained in high-throughput molecular data (omics data), as evidenced by the proliferation of studies which provide prediction models based on gene expressions or other predictors derived from omics data. Nonetheless, in medical practice, several clinical variables are often available, which have a predictive value well validated in the literature. Unfortunately, most of the studies consider only the omics or only the clinical variables. Recent researches, however, suggest that prediction models based on both these kinds of variables may outperform those models based only on one of them. In this regard, the main challenge derives from the different nature of the data, belonging to the high-dimensional and low-dimensional world, respectively. In order to combine clinical and molecular information, several strategies have been proposed and used in literature. An important difference is often the way in which the clinical variables are considered in the high-dimensional model fitting. Here we briefly review these strategies and we show how to apply them in practice, exploiting and adapting some existing methods, such as univariate variable selection, boosting regression,

lasso, etc. We implement these methods and compare the results in publicly available datasets, with a particular focus on outcome prediction in breast cancer patients.

353: 4

Developing subgroup signatures by weighted regression for time-to-event methylation data

Veronika Weyer[1], Harald Binder[1,2]

[1]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Mainz, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In the analysis of high throughput DNA methylation data with a time-to-event endpoint, often a small selection of CpG sites is wanted that can predict the risk for events such as death. For considering heterogeneity from known subgroups, such as cancer stages, we introduce a new weighted regression technique. When developing a risk prediction signature for each of the groups, the technique utilizes some information from the other groups. Variable selection for signature development is performed using componentwise boosting with a weighted likelihood. For choosing the weight, we propose two approaches based on resampling. The first approach chooses the weight that results in the largest CpG inclusion frequencies, when repeatedly building the signatures in resampling data sets. In a second approach, we select the weight such that inclusion frequencies are large for different signature sizes on average. We investigate model stability and resulting prediction performance in a simulation study and a real data example for comparing the two approaches. Choosing appropriate weights for the respective other groups is seen to be crucial for optimizing the risk prediction signatures. There is a systematic difference in resampling inclusion frequencies depending on the weight. We illustrate how different CpG inclusion frequency profiles can be obtained depending on the approach for weight selection. Weighted regression is seen to be useful when examining subgroups with the same sign of CpG effects. Thus weighted regression seems promising when subgroups need to be taken into account for signature development in a time-to-event setting.

355: Advanced Statistical Methodology and Software

Time: Wednesday, 20th Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1010
Session Chair: Willi Sauerbrei

355: 1

Survey estimation and marginal effects in Stata

Ben Jann

University of Bern, Switzerland

Stata is a general purpose software package that has become popular among various disciplines such as epidemiology, economics, or social sciences. Users like Stata for its scientific approach, its robustness and reliability, and the ease with which its functionality can be extended by user written programs. In this talk I will first give a brief overview of the functionality of Stata and then discuss two specific features: survey estimation and predictive margins/marginal effects. Most surveys are based on complex samples that contain multiple sampling stages, are stratified or clustered, and feature unequal selection probabilities. Standard estimators can produce misleading results in such samples unless the peculiarities of the sampling plan are taken into account. Stata offers survey statistics for complex samples for a wide variety of estimators and supports several variance estimation procedures such as linearization, jackknife, and balanced repeated replication (see Kreuter and Valliant, 2007, Stata Journal 7: 123-129). In the talk I will illustrate these features using applied examples and I will also show how user written commands can be adapted to support complex samples. Complex can also be the models we fit to our data, making it difficult to interpret them, especially in case of nonlinear or non-additive models (Mood, 2010, European Sociological Review 26: 67-82). Stata provides a number of highly useful commands to make results of such models accessible by computing and displaying predictive margins and marginal effects. In my talk I will discuss these commands provide various examples demonstrating their use.

355: 2

Multivariable model building when some variables are continuous – The MFP approach and extensions for interactions

Willi Sauerbrei[1], Carolin Jenkner[1], Patrick Royston[2]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]MRC Clinical Trials Unit, London, United Kingdom

The multivariable fractional polynomial (MFP) procedure (Sauerbrei et al 2007, Royston and Sauerbrei 2008) is a suitable approach for the selection of variables and selection of functional form for continuous variables. It combines backward elimination with a systematic search for a good fitting FP function. Extensions to investigate for interactions between continuous variables, between a continuous variable and a categorical variable (relevant to identify differential treatment effects in randomized trials) or between a variable and time in a Cox model for survival data (time-varying effect) are available. When developing the MFP approach and the extensions all programs were written in Stata by Patrick Royston. In SAS and R MFP is also available, but none of the extensions.

By analyzing medical data in the framework of regression models key issues of multivariable model building with continuous variables will be discussed and the analysis will be illustrated by using Stata. Concerning statistical criteria it will become obvious that MFP usually selects suitable models. It is also important that usually results from these models are easily understandable, interpretable and transferable. The key components of the approach and the Stata programs are easy to understand and to use.

363: Benefit Assessment in Public Health -- The Clincial Benefit of Diagnostic Methods: Methods and Examples: Common Workshop of IBS-DR, GMDS and IQWiG

Time: Wednesday, 20th Mar 2013: 2:50pm - 4:10pm · Location: KG I, Aula
Session Chair: Claudia Schmoor
Session Chair: Carsten Schwenke

363: 1

How to demonstrate the added value of diagnostic tests

Carsten Schwenke

SCOSSiS, Germany

Diagnostic tests are a major part in the diagnostic cascade to come to a diagnosis in patients with unknown disease or to confirm a disease in cases of doubt. The aim of the diagnostic test is to change the post-test probability for a disease compared to the pre-test probability in the physician's mind.

A reasonable reliability is the basic requirement for new test to qualify for clinical routine. The second part is the validity of the diagnostic test, shown by technical parameters like sensitivity and specificity. In case of sufficient technical efficacy, the add-on value needs to be assessed compared to the standard procedures. This add-on value can be quantified by assessing the technical efficacy (e.g. Sensitivity) or effectiveness (e.g. change of diagnosis) or efficiency (consequence for the patient).

In several cases, even the technical efficacy is difficult to obtain, e.g. if the standard of reference is difficult to optain like in Alzheimer's disease. Therefore, the clinical development of new diagnostic tests needs to be planned carefully with regard to the type of assessment, choice of standard of reference, control group, endpoint (technical vs. others), etc.

In this presentation, I will discuss the challenges of diagnostic test development in the light of FDA and EMA guidelines and the requirements to show an add-on value compared to existing tests.

363: 2

Diagnostic tests and randomized controlled trials: Design and Interpretation

Lars Beckmann[1], Johanna Buncke[2], Ralf Bender[1], Fülöp Scheibler[1]

[1]Institute for Quality and Efficiency in Healthcare (IQWiG), Köln, Germany; [2]Johannes Gutenberg-Universität Mainz, Germany

The therapeutic benefit of a diagnostic test cannot be evaluated solely based upon its diagnostic accuracy. As in the assessment of drug benefit, for proving a causal relationship between the use of a diagnostic test and a benefit for patients in general randomized controlled trials (RCTs) involving the test in combination with a specific therapy are required. The scientific literature describes three basic study designs to evaluate the benefit of a diagnostic test: the interaction design, the marker-based strategy-design, and the enrichment design. The advantages and disadvantages of each design and its role in particular diagnostic pathways are discussed. RCTs based on the marker-based strategy-design and the enrichment design cannot a priori be considered as adequate. Rather, the interpretation of the results depends upon assumptions about the relationship between the attribute to be tested and the therapies. The interaction design theoretically facilitates assessment of benefit without requiring further assumptions as long as there is clinical equipoise between the two arms. In contrast to the marker-based strategy, it is not possible within the interaction design to evaluate the diagnostic pathway. It can be concluded that it is feasible to conduct RCTs to assess the benefit of a diagnostic test. The choice of study design depends upon both ethical considerations and the specific role in a diagnostic pathway. Alternative approaches, such as decision modelling or linked evidence, in which diagnostic accuracy is linked with external information from controlled trials, must be judged critically when used as the basis for evidence-based decisions.

363: 3

Combining diagnostic studies – a practical approach

Wiebke Sieben, Lars Beckmann, Ralf Bender

Institute for Quality and Efficiency in Healthcare (IQWiG), Köln, Germany

Combining evidence from multiple diagnostic studies in a meta-analysis is confronted with the dependency between sensitivity and specificity. Recently, several approaches have been proposed for bivariate meta-analysis which account for this dependency in random effects models. The resulting pooled estimates for sensitivity and specificity and corresponding confidence regions for both diagnostic technologies can be compared to answer the question whether one test is preferable. A Likelihood-Ratio-test can be applied to test for differences in the expected sensitivities and specificities.

To account for sources of heterogeneity, diagnostic studies using a paired design, in which two tests are applied to the same patient population, are preferable. In projects to evaluate the PET technology for several types of cancer we have found that often only few diagnostic studies with a paired design exist. However, own simulations suggest that even if all model assumptions hold, the coverage of confidence regions in a bivariate meta-analysis can be extremely low when only few studies are entered into the model.

In this presentation we discuss aspects of model choice, the implementation in SAS and propose a strategy for the presentation of the results, i.e. the visualization and interpretation with respect to the number of available studies and their risk of bias.

363: 4

The evaluation of diagnostic methods and the need for randomized controlled trials: Medical topics evaluated by IQWiG

Stefan Sauerland, Fülöp Scheibler

Institute for Quality and Efficiency in Healthcare (IQWiG), Köln, Germany

The Institute for Quality and Efficiency in Health Care (IQWiG) is an independent scientific institute that investigates the benefits and harms of medical interventions, including screening and diagnostic interventions. Examples of diagnostic tests evaluated by IQWiG include osteodensitometry, fetal ultrasonography and positron emission tomography (PET). Most assessments, including all those regarding PET and PET/CT, focused on randomized controlled trials (RCT) but also included diagnostic test accuracy (DTA) studies. This presentation will address the question how much can be learned about the effectiveness of diagnostic test from studying only DTA data (e.g. in the absence of RCTs).

Test accuracy has been called a surrogate endpoint, because ruling a diagnosis in or out does not necessarily imply that a specific therapy can be applied and will lead to beneficial effects for the patient. However, the exact role of the new test has also to be considered. Some new tests are aimed at replacing existing tests, because the new tests offer the same test accuracy but with less invasiveness or radiation exposure. In these cases, test accuracy data may suffice for a proper evaluation.

In all other situations, using DTA data is hampered by the fact that it usually is impossible to link the results of diagnostic to those of therapeutic interventions (“linked evidence”). We will illustrate this problem with some examples. Additional problems of DTA studies may include a wrong selection of unit of analysis (lesions rather than patients) and the impossibility to obtain a valid reference standard.

358: Industry Day

Time: Wednesday, 20th Mar 2013: 3:00pm - 4:00pm · Location: KG III, HS 3042

358: 1

Advanced process control with STATISTICA Enterprise

Michael Busch

StatSoft, Hamburg, Germany

Modern statistical process control goes beyond the use of traditional univariate and multivariate control charts. STATISTICA Enterprise combines comprehensive graphic and statistical methods with an enterprise-wide infrastructure. This includes the integration with existing data sources as well as automated, server-driven identification and monitoring of process-relevant parameters. After a general description of the architecture of the software solution the lecture will give an example scenario with traditional control charts, and the use of modern data mining techniques for process monitoring where the traditional methods are not sufficient.

368: Industry Day

Time: Wednesday, 20th Mar 2013: 4:30pm - 5:30pm · Location: KG III, HS 3042

368: 1

Standardization Aspects in the SPC Environment

Michael Radeck, Markus Pfirsching

Q-DAS, Germany

1) Data format and data transfer

Due to a standardized data format it is much easier to record, save and transfer data. Important aspects are

* the integration of measuring instruments

* the automated process control by means of SPC systems

* a uniform data format for the exchange of data.

2) Standardized calculations of quality indicators lead to transparent production processes Global companies have their production sites in many different countries. The easy exchange of data and the reproducibility of evaluation results are of particular importance.

* Uniform evaluation methods due to a corporate evaluation strategy

* Operation and output of results in the respective national language

* Automated data processing by creating statistics, graphics and reports

3) Process improvement and Six Sigma offensives

Due to the standardized data format it is easy to isolate existing perturbations and to determine models describing the interaction between product characteristics and process parameters for processes with an optimization potential. This is e.g. based on graphical analyses and correlation methods, the analysis of variance and the regression analysis.

361: Statistics in Clinical and Preclinical Research -- Others

Time: Wednesday, 20th Mar 2013: 4:40pm - 6:00pm · Location: KG II, Audimax
Session Chair: Stephanie Roll

361: 1

Analyzing Blinded Phase 3 Interim Safety Data

Jouni Kerman, Achim Güttner, Elodie Incera

Novartis Pharma AG, Switzerland

According to a new rule by the FDA, sponsors are requested to review safety data of ongoing trials in an unblinded fashion in order to identify suspected unexpected serious adverse reactions. Such reviews should be based on the entire project database of completed and ongoing studies. However, this rule has not been broadly implemented due to sponsors' concerns of data and trial integrity.

Taking these concerns into account, we set out to produce some exploratory analyses based on fully blinded (pooled) adverse event count data from an ongoing Phase 3 study, with the goal to characterize the imbalance of expected and observed data, even though direct estimation of active-to-comparator odds ratios was not possible. Based on unblinded data from completed Phase 2 studies, we used Bayesian models to estimate the occurrence rates of adverse events for each safety category of interest, and to predict the aggregate adverse event counts of the Phase 3 study. Comparing the posterior predictive distribution of the event counts to the observed counts, we were able to characterize the level of statistical significance of the interim findings by a "probability of imbalance" for each adverse event category. This probability served then as a measure to rank the categories so that those with the largest imbalances could be easily identified.

We discuss the background of the case study, details of the statistical analyses, the graphical visualization of the results, and their interpretability. We also address some concerns related to internal processes and to possible operational bias.

361: 2

Combining Rank Preserving Structural Failure Time Models and Multistage Modelling in the Context of Progression-Free Survival

Karin Schiefele[1], Frank Fleischer[2]

[1]Universität Ulm, Germany; [2]Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach/Riss, Germany

In oncological trials confounding effects due to subsequent treatments after progression or non-compliance are frequently observed. These effects may lead to biased estimates and wrong test decisions in an intention-to-treat analysis. A model that has been developed to adjust for confounders during treatment is the rank preserving structural failure time model (RPSFTM). On the other hand a model often used in the context of the two main oncological endpoints progression-free survival and overall survival is the multistate or illness-death model. We show how to combine the two approaches mathematically. By this new approach it is possible to estimate treatment effects before and after progression based on the multistage model when simultaneously adjusting for confounding effects in the sense of the RPSFTM. The combined approach is applied to simulated data as well as data of a non-small-cell lung cancer trial.

REFERENCES:

Fleischer F, Gaschler-Markefski B, Bluhmki E. 2009. A statistical model for the dependence between progression-free survival and overall survival. Statistics in Medicine 28 (21): 2669-2686.

Korhonen P A, Laird N M, Palmgren J. 1999. Correcting for non-compliance in randomized trials: an application to the ATBC study. Statistics in Medicine 18: 2879-2897

Robins J M, Tsiatis A A. 1991. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics 20: 2609-2631.

Robins J M, Tsiatis A A. 1992. Semiparametric estimation of an accelerated failure time model with time-dependent covariates. Biometrika 79 (2): 311-319.

361: 3

Reconstruction of the joint distribution when only the marginal distributions are known with application to the estimation of heterogeneous treatment effects in randomized controlled trials based on the parallel group design

Ruediger P. Laubender, Ulrich R. Mansmann

Ludwig-Maximilians-Universität München, Germany

One of the aims of "individualized medicine" is to allocate the optimal treatment to a patient when it is known that patients differently react to several treatments. The presence of heterogeneous treatment effects can be estimated by study designs which allow for replication of the treatment effects like crossover designs (Senn 2001). However, such randomized controlled trials (RCTs) are not possible when treatments irreversibly change the patients (e.g. treatments used in oncology). In such situations only RCTs based on the parallel group designs can be

used. Here, we only observe the marginal distributions of the outcomes of the considered treatments but we cannot observe the joint distribution of these outcomes. For normally distributed outcomes, it is possible to reconstruct the joint distribution and hence to estimate heterogeneous treatment effects from such a RCT by re-defining a predictive baseline covariate as an indicator for the individual treatment effect (ITE) and by assuming a multivariate normal distribution. Given these assumptions we derive estimators for reconstructing the joint distribution, for the ITE and for the probability that an individual will profit from treatment A compared to treatment B given his or her value of the predictive covariate. This is an estimator easily understandable by physicians and patients and cannot be derived from a usual regression model with a treatment-covariate interaction. The conditions for judging whether heterogeneous treatment effects are present at all within that approach are shown. Our methods are applied to three data examples. Possible extensions of the proposed approach are discussed.

361: 4

Multiple testing in group sequential trials using graphical approaches

Willi Maurer, Frank Bretz

Novartis Pharma AG, Switzerland

Confirmatory clinical trials are becoming increasingly more complex and often involve multiple statistical hypotheses that reflect the structured objectives. Examples include the investigation of multiple doses or regimens of a new treatment, several clinical endpoints and populations, non-inferiority and superiority, or any combination thereof. O'Neill (1997) introduced the framework of structuring the experimental questions to best reflect the clinical study's objectives. This task comprises the identification of the study's primary, secondary and exploratory objective(s) and the requirements as to when the corresponding statistical tests are considered meaningful. A topic, which has been considered much less in the literature until very recently, is the application of group sequential trial designs to multiple endpoints. Hung et al. (2007) showed that borrowing testing strategies naively from trial designs without interim analyses may not maintain the familywise Type I error rate at level alpha. The authors gave examples in the context of testing two hierarchically ordered endpoints in two-armed group sequential trials.

In this talk we consider the general situation of testing multiple hypotheses repeatedly in time using recently developed graphical approaches. We focus on closed testing procedures using weighted group sequential Bonferroni tests for the intersection hypotheses. Under mild monotonicity conditions on the error spending functions, this allows one to implement sequentially rejective graphical procedures in group sequential trials. The methodology is illustrated with a numerical example for a three-armed trial comparing two doses against control for two hierarchically ordered endpoints.

362: Survival and Event History Analysis -- Complex data

Time: Wednesday, 20th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1199
Session Chair: Gerard J. van den Berg

362: 1

Comparison of approaches to assess time-varying effects in high-dimensional survival studies

Anika Buchholz[1], Willi Sauerbrei[1], Harald Binder[2]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Johannes Gutenberg University Mainz, Germany

Survival studies with microarray data often focus on identifying a set of genes with significant influence on a time-to-event outcome. Typically, a gene expression signature (i.e. predictor) is derived using the Cox proportional hazards (PH) model. However, the Cox model assumes that effects are constant over time which might not hold for the predictor and/or some of the genes. Ignoring this may lead to false conclusions about the influence of genes. Hence, it is important to account for possible non-PH in selection of important genes and investigate the shape of time-varying effects.

Recently we have compared several strategies for identifying, selecting and modelling time-varying effects in low-dimensional settings [1]. Some of them can also be applied to high-dimensional data. We will illustrate and compare three different approaches using publicly available gene expression data with time-to-event outcome from cancer patients, for which predictors have been derived [2]. In addition, we will extend the investigation to individual genes to assess whether for some of them the effect varies in time, which usually means that they only have a short term effect.

References:

[1] A. Buchholz, W. Sauerbrei. Comparison of procedures to assess non-linear and time-varying effects in multivariable models for survival data. Biometrical Journal, 53:308-331, 2011.

[2] H. Binder and M. Schumacher. Incorporating pathway information into boosting estimation of high-dimensional risk prediction models. BMC Bioinformatics, 10:18, 2009.

362: 2

Gene-longevity association studies using a modification of the relative risk model.

Alexander Begun[1], Andrea Icks[1,2], Guido Giani[1]

[1]German Diabetes-Center at the Heinrich-Heine-University, Düsseldorf, Germany; [2]Department of Public Health, Faculty of Medicine, Heinrich-Heine-University, Düsseldorf, Germany

A significant difference in the gene and allele frequencies in distinct age groups can indicate the presence of a genetic influence on longevity. Some extensions of the basic gene frequency method involve the use of demographic information about the population under study and allow the estimation of genotype specific hazards for candidate genes. We discuss a modification of the relative risk method, which allows us to separate the aging effect from the cohort effect and include the effect of antagonistic pleiotropy. We illustrate our method with numerical example based on simulated data.

362: 3

Simple correction methods for task completion times contaminated by errors, outliers and omissions

Matthias Gondan[1], Bruno Fimm[2]

[1]University of Heidelberg, Germany; [2]University Hospital Aachen, Germany

In attention diagnostics and applied evaluative research, inference on behavioral performance is often based on task completion times, for example, the time it takes to respond to a visual signal, or to choose among a number of alternatives in a multiple choice task. When the intensity of the signal is low and/or the task is difficult, participants usually do not respond with 100% accuracy, however. In such a situation, estimation of response speed is difficult because the task completion times are contaminated by outliers, omitted and erroneous responses. Estimation of response speed is often based to the subset of correct responses only, which is known to yield biased estimates and does not make full use of the available information. I present simple ad hoc data cleaning methods for task completion times contaminated by outliers, omissions and erroneous responses, and I show that these methods can substantially reduce bias and uncertainty of estimates of response speed. The method is illustrated using data from an evaluation of integrated displays in intensive care units.

362: 4

Unemployment duration and sports participation

Charlotte Cabane

SEW-St.Gallen University, Switzerland

In this study I use the German Socio-Economic Panel to evaluate the impact of practising sport -while being unemployed- on the unemployment duration. The empirical literature on sport participation has focused on labour market outcomes and job quality while the impact of this activity on job search has not been studied. However, sports participation fosters socialization which, through the networking effect, accelerates the exit from unemployment to employment. Moreover, sporty people develop specific non-cognitive skills which are highly valued on the labour market. Last, there is a well-known positive impact of sports participation on individuals' health status which should also favour job-search efficiency and applicant's attractiveness. Since other activities could have similar impacts on unemployment duration -through social connections and non-cognitive skills- I compare a set of activities to sporting activities. I use a duration discrete time model and take seriously into account the endogeneity issue due to unobservables. I find encouraging results with respect to participation in sporting activities and also results outlining the specificity of this activity with respect to others.

364: Advanced Statistical Methodology and Software

Time: Wednesday, 20th Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1010
Session Chair: Matthias Schmid

364: 1

Applied and Reproducible Econometrics with R

Achim Zeileis

University of Innsbruck, Austria

Recently, computational methods and software have been receiving more attention in the econometrics literature, emphasizing that they are integral components of modern econometric research. This has also promoted the usage of the R system for statistical computing which is starting to be adopted for empirical research in economics (amongst many other systems such as Stata, EViews, RATS, or programming environments such as GAUSS or Ox). With its flexibility, object orientation, superior graphics, and tools for reproducible research, R still has much more potential in econometrics. However, some obstacles for greater popularity of R so far have been that it was mainly developed by and for statisticians, and that it thus sometimes uses unfamiliar terminology, follows a different workflow than other software packages, and that many econometric methods became available only relatively recently. The talk will give an overview of the econometric methods already available in R, demonstrate the look-and-feel of R in typical econometric analyses (using examples from Kleiber and Zeileis, 2008), and highlight some of the special strengths of R. In particular, it will be shown in a case study of "forensic econometrics" (Koenker and Zeileis, 2009) how reproducible econometric research can be greatly facilitated in R.

References:

Christian Kleiber, Achim Zeileis (2008). Applied Econometrics with R. New York: Springer-Verlag. ISBN 978-0-387-77316-2. Code/data/examples/demos at http://CRAN.R-project.org/package=AER

Roger Koenker, Achim Zeileis (2009). On Reproducible Econometric Research. Journal of Applied Econometrics, 24(5), 833-847.

364: 2

A General Framework for Function-on-Function Regression with R

Fabian Scheipl, Ana-Maria Staicu, Sonja Greven

Ludwig-Maximilians-University Munich, Germany

The pffr() function in the refund package (Crainiceanu et al. 2011) for R offers a general framework for regression of (correlated, incomplete) functional responses on scalar and functional covariates, with the full flexibility of generalized additive mixed models for scalar responses made available for the functional data paradigm. We describe inference based on the mixed model framework that allows us

to bring robust and highly performant algorithms developed for scalar regression to bear in this new paradigm, discuss extensions for models with non-standard errors based on ideas from the marginal models literature, and show how this approach can easily be extended for non-Gaussian data. Results from extensive empirical evaluations of our approach on simulated data and real case studies are presented.

Keywords: functional data, penalized splines, mixed models, generalized additive models

References:

C.M. Crainiceanu, P.T. Reiss (Coordinating authors), J. Goldsmith, S. Greven,

L. Huang, and F. Scheipl (Contributors). refund: Regression with Functional

Data, 2011. URL http://CRAN.R-project.org/package=refund. R package version 0.1-6.

A.E. Ivanescu, A.-M. Staicu, S. Greven, F. Scheipl, and C.M. Crainiceanu. Penalized function-on-function regression. 2012. URL biostats.bepress.com/jhubiostat/paper240/. under revision.

F. Scheipl, A.-M. Staicu, and S. Greven. Additive mixed models for correlated functional data. 2012. URL http://arxiv.org/abs/1207.5947. under revision.

364: 3

More than just the mean: Flexible statistical modelling via GAMLSS

Andreas Mayr, Benjamin Hofner, Matthias Schmid

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Generalized additive models for location, scale and shape (GAMLSS, Rigby and Stasinopoulos, 2005) are a very flexible model class

extending the popular GAM framework beyond mean regression. In contrast to conventional GAMs, the idea of GAMLSS is to regress not only the expected mean but every distribution parameter (e.g. location, scale and shape – hence the extension LSS) to a set of covariates via an additive predictor and associated link function. As a result, common data phenomenons like heteroscedasticity and severe skewness can be explicitly modeled. Furthermore, the common exponential family distribution assumption as in GAMs is relaxed. We present two different fitting algorithms that are available in R: The package gamlss is provided by the inventors of the model class and is based on backfitting; the gamboostLSS package is a recent proposal based on multi-dimensional boosting (Mayr et al., 2012). We discuss the general properties of both algorithms and present some examples to demonstrate their practical usage.

References:

Mayr, A., N. Fenske, B. Hofner, T. Kneib and M. Schmid (2012). Generalized additive models for location, scale and shape for high-dimensional data - a flexible approach based on boosting. Applied Statistics 61(3): 403-427.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). Applied Statistics 54(3), 507–554.

365: Young Statisticians

Time: Wednesday, 20th Mar 2013: 4:40pm - 6:00pm · Location: KG III, HS 3043
Session Chair: Benjamin Hofner
Session Chair: Tina Müller

365: 1

Comparing adaptive versus fixed p-values in the Promising Zone Design

Julia Katharina Krzykalla[1], Frank Fleischer[2]

[1]Universität Ulm, Germany; [2]Boehringer Ingelheim Pharma GmbH & Co. KG, Germany

Das Promising Zone Design ist ein spezielles adaptives Studiendesign, das es ermöglicht, im Zuge einer Zwischenauswertung eine Neuberechnung der Sample Size durchzuführen, ohne in der finalen Analyse Adaptionen der Teststatistiken vornehmen zu müssen. Die Neuberechnung der Sample Size geschieht dabei auf Basis der Conditional Power.

Um trotz der Verwendung von „einfachen“ Testmethoden, wie sie auch für fixe Studiendesigns eingesetzt werden, die Einhaltung des Typ-I Fehlers zu sichern, wird ein Wertebereich der Conditional Power festgelegt, für den eine Erhöhung der Sample Size in Frage kommt.

Vor diesem Hintergrund sollte nun untersucht werden, ob und inwiefern sich ein adjustierter (adaptiver) p-Wert von dem aus dem oben beschriebenen Design resultierenden (fixen) p-Wert unterscheidet. Dazu wurden zunächst die p-Werte formal hergeleitet und anschließend für verschiedene Szenarien abhängig von den Ergebnissen aus der Zwischen- und Endauswertung betrachtet.

Dabei stellte sich heraus, dass sich, entgegen der weitverbreiteten Annahme, dass der adaptive p-Wert grundsätzlich ungünstiger (größer) als ein fixer p-Wert ist, Situationen identifizieren lassen, für die ein adaptives Design vorteilhaft ist, d.h. in denen der adaptive p-Wert kleiner als der fixe p-Wert wird.

365: 2

Locally weighted Cox-regression for analyzing treatment-by-subgroup interactions

Anne-Sophie Stöhlker

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

A main task in survival analysis is to detect differential treatment effects for a sample. Since this may become problematic in case of heterogeneous populations it is important to determine possibly internally homogeneous subgroups and the respective treatment effects.

In this thesis, the primary idea to achieve the latter is to conduct a locally weighted regression for several subgroups. Initially, these subgroups are to be determined via an interaction-tree procedure detecting treatment-by-subgroup interactions similar to the method by Su et al. (2011). The standard procedure at that point would be to perform a subgroup analysis respectively to carry out a Cox-regression for each subgroup. Therefore, perfectly homogeneous subgroup-allocation is assumed implicitly. Considering that this does not hold naturally, an alternative way of analysis is suggested: In contrast to the standard subgroup analysis, treatment effects for one subgroup are not only based on fitting a Cox model to the data of the respective group. Instead, they relate to a Cox model being fit to the data of the whole population by additionally incorporating specific weights. These weights express the influences of all subgroups on one of them and are obtained by adjusting and applying the weighting scheme by Simon (2002). The entire procedure is called locally weighted Cox-regression.

To examine the developed model, it is applied to the data of two randomized trials in patients with brain tumors respectively kidney cancer. Additionally, a simulation study covering various hypothetical scenarios is conducted in order to investigate the performance of the method.

365: 3

An AUC-based Permutation Variable Importance Measure for Random Forests for Unbalanced Data

Silke Janitza, Anne-Laure Boulesteix

Ludwig-Maximilians-Universität München, Germany

The random forest method is a commonly used tool for classification with high dimensional data as well as for ranking candidate predictors based on the so-called variable importance measures. However the classification performance of random forest is known to be suboptimal in case of strongly unbalanced data, i.e. data where response class sizes differ considerably. In this case it tends to predict the majority class, yielding a minimal error rate. The standard random forest permutation variable importance measure which is based on the error rate is directly affected by this problem and loses its ability to discriminate between important and unimportant predictors in case of class imbalance. This effect is more pronounced for small effects and small sample sizes.

The area under the curve (AUC) is a promising alternative to the error rate for class imbalance since it puts the same weight on both classes. A novel permutation variable importance measure in which the error rate is replaced by the AUC is therefore a promising alternative for unbalanced data settings. It can be shown in simulation studies that this measure outperforms the standard error-rate-based permutation variable importance measure for unbalanced data settings while both measures have equal performance for balanced data settings.

365: 4

Change point test for tail index for dependent data

Yannick Peer Hoga

Universität Duisburg-Essen, Germany

Kim and Lee (Metrika 74:297-311, 2011) proposed a change point test for the tail index of stationary [fn]\beta[/fn]-mixing random variables. In my master's thesis I was able to show that their test, which is based on comparing Hill's estimator from different time periods, can be successfully extended to [fn]\alpha[/fn]-mixing random variables and even [fn]\mathcal{L}_{2}[/fn]-E-NED sequences of random variables. The limiting distibution of the test statistic in the first extension to [fn]\alpha[/fn]-mixing random variables is the same as in Kim an Lee, whereas it is different in the second. The [fn]\mathcal{L}_{2}[/fn]-E-NED concept is due to Hill (Econometric Theory 26:1398-1436, 2010) and includes a wide range of processes, including among others [fn]\mathcal{L}_{2}[/fn]-NED, ARFIMA, FIGARCH, explosive GARCH, nonlinear ARMA-GARCH, bilinear processes, and nonlinear distributed lags.

A simulation study is implemented to investigate the empirical sizes and powers of the test in finite samples. As an application, the data set of daily DJIA returns is examined for a possible change point in the tail index.

370: Poster & Wine

Time: Wednesday, 20th Mar 2013: 6:00pm - 10:00pm · Location: KG I, Eingangshalle EG

370: 1

Boosting functional regression models [A1]

Sarah Brockhaus, Fabian Scheipl, Sonja Greven

Ludwig-Maximilians-Universität München, Germany

The employment of boosting allows us to estimate a broad class of functional regression models accommodating models with functional response and multiple functional and scalar covariates. Using different base learners linear effects of functional covariates can be combined arbitrarily with linear and smooth effects of scalar covariates and functional random effects.

Furthermore it is possible to minimize diverse loss functions yielding e.g. generalized linear models or robust models like quantile regression for functional data. Another benefit of boosting is its capability for variable selection even in datasets with more covariates than observations.

To estimate our models we use the fast and stable R package mboost extended by some base learners meeting the requirements of functional variables.

370: 2

Using the lasso to identify early predictors of Alzheimer's disease [A2]

Manfred Berres[1], Nicole S. Schmid[2], Kirsten I. Taylor[2,3], Nancy S. Foldi[4], Andreas U. Monsch[2]

[1]RheinAhrCampus Remagen, Germany; [2]Memory Clinic, University Hospital, Basel, Switzerland; [3]University of Cambridge, United Kingdom; [4]City University of New York, United States of America

Recent findings suggest that very subtle changes in cognition may precede clinical Alzheimer's disease (AD) by several years. To identify which cognitive, mood and physical status variables influence conversion to AD and estimate the strength of their effects, we followed 825 participants from the longitudinal BASEL study (BAsel Study of the ELderly). Over 13 years, 29 individuals were diagnosed with probable AD within an average of 8 years. Each diagnosed individual was pairwise matched by age, education, demographic status, observation period and ApoE genotype with another BASEL participant who had remained healthy. A logistic regression model with a lasso penalty (Tibshirani, 1996) was applied to 115 potential predictors. The lasso operator shrinks coefficients of a regression model and at the same time selects predictors by shrinking coefficients to zero.

The penalty parameter was determined by cross-validation. With the optimal parameter, 11 predictors were selected for the regression model which achieved a correct classification rate of 60%. Errors and repetitions in verbal fluency, intrusions in memory tests, visual-spatial deficiencies, subjective complaints and cardiac problems were the most reliable very early predictors for conversion to AD.

In this high dimensional setting with twice as many variables as observations, the lasso technique generated a model which is in line with the neuropathological development of AD and shows high consistency within the selected variables.

References:

Schmid et al., Journal of Alzheimer's Disease (in press).

Tibshirani, JRSS B 58 (1996) 267-288.

370: 3

Comparison between two cubic spline models and a seasonal model. [A3]

Joaquim Fernando Costa, Ana Rita Gaio

Department of Mathematics, Faculty of Sciences, University of Porto, Portugal

We have recently introduced a novel methodology based on cubic splines for the case of constrained time-dependent data. For instance, one of the independent variables can be the hour of the day; when it reaches 24h, it immediately goes back to 0h again. In cubic splines the prediction function is continuous everywhere and has continuous first and second derivatives. In addition, these constraints should also be imposed at the turning points, as the [fn]24 \rightarrow 0[/fn] above. These constraints introduce new basis functions for the regression and allow for the inclusion of a greater number of knots than in the cubic spline scenario, for the same number of degrees of freedom. We apply our methodology to simulated data and to a real dataset consisting of the number of births, in the North of Portugal, for different hours of the day and different days of the week. We have compared the results of our constrained cubic spline model with those of a usual cubic spline and also with a harmonic sesonal model using sines and cosines, which also satisfies the boundary continuity constraints. Using the same number of degrees of freedom, our model shows superiority, with respect to error fitting terms. This positions our approach as a new and relevant spline model for constrained data and also as a strong competitor to seasonal models.

370: 4

Conditional AIC for generalized additive mixed models [A4]

Benjamin Säfken, Thomas Kneib

Georg-August University Göttingen, Germany

In simple regression settings the number of parameters is used as bias correction to the log-likelihood when comparing different nested models with the Akaike information criterion. In the mixed model setting the question arises whether the random effects should be treated as parameters in the bias correction. Integrating out the random effects leads to a biased criterion. Conditioning on the random effects the effective degrees of freedom, defined as the trace of the hat matrix, are an appropriate bias correction, if the random effects covariance is known. Since with unknown random effects covariance the fitting process is non-linear a generalization of the trace of the hat matrix is used, i.e. the derivative of the prediction with respect to the data. This bias correction can be interpreted as the sensitivity of the estimator to small changes in the data and therefore as the flexibility of the fitting process. In generalized linear mixed models similar flexibility measures for other exponential family distributions can be found. They enable AIC based model selection not only for generalized linear mixed models but also for generalized additive models.

370: 5

Influence diagnostics in LASSO regression [A5]

Choongrak Kim[1], Whasoo Bae[2]

[1]Pusan National University, South Korea; [2]Inje University, South Korea

One or few observations may be unduly influential on regression estimates, and LASSO regression also suffers from this phenomenon. Unfortunately, the analytic expression for LASSO regression estimates are not available. Here, we suggest a one-step version of Cook's distance in LASSO regression using the result of ridge regression, and compare it with the exact Cook's distance, which is only possible numerically. Also, we study the effect of few observations on LASSO regression in selecting variables since the model selection via the LASSO would give different results when we delete few influential observations.

370: 6

Functional Linear Mixed Models for Sparsely and Irregularly Sampled Data [A6]

Jona Cederbaum, Sonja Greven

Ludwig-Maximilians-Universität München, Germany

Recent technological advances allow today's scientists to collect an increasing amount of data consisting of functional observations. In many practical applications these functional observations are correlated, e.g. due to repeated measurements or spatial proximity.

Scalar correlated data are commonly analyzed using linear mixed models. For densely and regularly sampled data, a functional analogue, the functional linear mixed model (FLMM), has been proposed (e.g. Guo, 2002, Morris and Carroll, 2006, Greven et al., 2010).

Although in practice the observed curves are frequently evaluated on - possibly only few - irregularly spaced points, irregularly observed and sparse functional data with additional correlation structure have received little attention in the literature so far.

We propose an estimation procedure for FLMMs which is applicable to sparsely and irregularly sampled data. Inference is based on dimension reduction via functional principle component analysis, adapted to the additional correlation structure and to the irregularity of the sampling design. The procedure is motivated by and applied to data from speech production research.

370: 7

Autoregressive model of Sweet orange (Citru sciensis L. osbeck) productivity in Ibadan, Nigeria. [F4]

Oludare Ariyo[1], Taofeek Dauda[2], Abayomi Olaniyan[1], Bernard, Okafor[1]

[1]National Horticultural Reserach Institutes, Ibadan, Nigeria; [2]Institutes of Agricultural Research OAU, Ibadan, Nigeria

This study was carried out to evaluate citrus productivity through an autoregressive model using data from the citrus orchard established by the National Horticultural Research Institute, Ibadanin 1977. The citrus orchard covered 20ha of alfisoil, composes of twelve varieties of sweet oranges which represents the blocks while the annuals yield represents the treatments. The results of the analysis of variance showed that the annual yields of citrus regardless of the variety are significantly different from one another because the F- statistics 1578120 returned for the year is greater than F (23, 576: 0.01) = 2.26. Also, there exist significant difference in the mean yield of the variety irrespective of the year because, the 369479 returned for the variety is greater than F (11, 576: 0.01) = 3.60. The interaction of the year by variety of the yield of citrus clearly indicated a significantly different result since the mean (9095.06) returned for the interaction was greater than F (253, 576: 0.01) = 1.00. Yield extension rate (YER) of the citrus yield does not follow a regular pattern and it differs across the different period with no two periods ([fn]x_{ij}[/fn]) having the same mean YER. The auto regression analysis of the citrus yield gave a linear relationship between the current yield and preceding year’s yield of citrus with a very high coefficient of determination (0.993) and a very low residual (Σ = 0.0022 r Y)

370: 8

Qualitative Robustness of SVMs and Other Regularized Kernel Methods for Dependent Data [A8]

Katharina Strohriegl, Robert Hable

Universität Bayreuth, Germany

In nonparametric classification and regression problems, support vector machines (SVMs) and related regularized kernel methods currently attract much attention in theoretical and in applied statistics. Though these methods are successfully used for dependent data in many applications, almost all theoretical investigations assume i.i.d. pairs of random variables. For the case of i.i.d. observations, qualitative robustness of a broad class of regularized kernel was shown in Hable & Christmann (2011). This was done in the ordinary way by showing that the corresponding statistical functional is continuous. In case of stochastic processes with non-i.i.d. observations, we demonstrate that continuity of the statistical functional can also be used to show qualitative robustness of regularized kernel methods for many processes.

References:

G. Boente, R. Fraiman, and V. J. Yohai. Qualitative robustness for stochastic processes. Ann. of Stat., 15:1293-1312, 1987.

R. Hable and A. Christmann. On qualitative robustness of support vector machines. J. Multivariate Anal., 102:993-1007, 2011.

F. R. Hampel. A general qualitative definition of robustness. Ann. Math. Statist., 42:1887-1896, 1971.

370: 9

nparcomp: An R Software Package for Nonparametric Multiple Comparisons [E1]

Marius Placzek

Georg-August-Universität Göttingen, Germany

When analyzing one-way layouts, i.e. a single factor with several levels and multiple observations at each level, usually not only a global hypothesis is of interest but also multiple comparisons between the different treatment levels. Parametric and semiparametric procedures impose quite restrictive distributional assumptions which are often not justifiable or impossible to validate. Hence, most practical situations demand statistical procedures that enable us to accurately and reliably analyze one-way layouts with minimal conditions on the assumptions required. Nonparametric methods offer such a possibility and thus become of particular practical importance. In this talk, rank-based methods for the analysis of unbalanced one-way layouts will be presented. A single-step procedure for multiple comparisons controlling the FWE in the strong sense as well as simultaneous confidence intervals for the estimated effects will be presented along with the analysis of an example using the corresponding R package nparcomp.

370: 10

Multiavariate Bayesian Quantile Regression [E2]

Elisabeth Waldmann, Thomas Kneib

Georg-August-Universität Göttingen, Germany

Quantile regression for conditional random variables has become a widely used tool to analyse relations within data. It provides a detailed description of the conditional distribution, without assuming a distribution type for the conditional distribution. The Bayesian version, which can be implemented by considering the asymmetric Laplace distribution (ALD) as an error distribution is an attractive alternative to other methods, because it returns knowledge on the whole parameter distribution instead of solely point estimations. While for the univariate case there has been a lot of development in the last few years, multivariate responses have only been treated to little extend in the literature, especially in the Bayesian case. By using a multivariate version of the location scale mixture representation for the ALD we are able to apply inference techniques developed for multivariate Gaussian models on multivariate quantile regression and make thus the impact of covariates on the quantiles of more than one dependent variables accessible.

370: 11

Comparison of permutation approaches for testing hypotheses from a theory of causal effects [E3]

Sonja Hahn

Universität Jena, Germany

When analyzing randomized experiments or observational studies, it is often feasible to control for covariates in order to enhance power or to remove bias. Based on a theory of causal effects (Steyer, Mayer & Fiege, submitted) a hypothesis about the total causal effect of a treatment can be formulated. There has been wide research on parametric procedures testing these hypotheses. Although Rosenbaum (1984) already developed a permutation approach in the 1980s, there are few investigations on nonparametric procedures in this area comparing these procedures by means of simulation studies.

For the present simulation study datasets were created containing a dichotomous treatment variable and a dichotomous covariate. Datasets varied in size of the average treatment effect, size of the interaction, total sample size and in the presence of confounding. Different permutation approaches were compared with respect to robustness and power.

Results suggest that many of the existing permutation approaches do not work well when testing the hypothesis of interest, especially when there is confounding and/or interaction. Nonetheless these approaches can be adopted to test the hypothesis about the average effect by altering the procedures.

References:

Rosenbaum, P.R. (1984). Conditional Permutation Tests and the Propensity Score in Observational Studies. Journal of the American Statistical Association, 79, 565-574.

Steyer, R., Mayer, A. & Fiege, C. (submitted) Causal Inference on Total, Direct, and Indirect Effects.

370: 12

On Adjusted Method of Moments [E4]

Ahmad Reza Soltani[1], Kamel Abdollahnezhad[2]

[1]Kuwait University; [2]Golestan University

The adjusted method of moments estimator, amme, is introduced by Soltani, and is appeard in Soltani and Homei (2009), Statistics, 43(6), 611-620. A closed formula for the amme for the parameter a, whenever the population has uniform distribution on (0,a), is derived by Soltani and Abdollahnezhad, Metron 2012, and its asymptotic properties are studied. In this article we prove that an amme estimator is consistent, then derive the amme for the location parameter of the power distribution and exhibit its superiority to the maximum likelihood estimator. We also discuss the limiting distribution of the an amme estimators and provide numerical inference procedures based on amme.

370: 13

An approach to constructing testing procedures by means of eye-tracking technology [B12]

Pavel Alekseevich Marmalyuk, Grigory Aleksandrovich Yuryev, Lev Semenovich Kuravsky, Valery Ivanovich Alkhimov

Department of Information Technologies of Moscow State University of Psychology and Education, Russian Federation

A brand-new approach to constructing intellectual and competence-based tests, which is based on a representation of subjects gaze movement on a stimulus surface with aid of a Markov stochastic process, which transition probabilities distribution satisfies the Fokker-Planck-Kolmogorov equation, is proposed. The relevant diagnostic procedure is afforded by identification of gaze probability distributions using eye-tracking observation results for different diagnosed subject groups and subsequent evaluation of subject's gaze trajectory conditional likelihood estimates. An illustrative example of practical application of the approach to determine mathematical skill level of school and university students is given.

370: 14

Multivariate Resistenzprofile als Zielgröße in Risikofaktor-Modellen [B7]

Inga Ruddat[1], Erhard Tietze[2], Dagmar Ziehm[3], Lothar Kreienbrock[1]

[1]University of Veterinary Medicine Hannover, Germany; [2]Robert Koch-Institute, Germany; [3]Niedersächsisches Landesgesundheitsamt, Germany

Um einer wachsenden Bedrohung für erfolgreiche Therapien durch Zunahme von antibiotikaresistenten pathogenen Bakterien entgegenzuwirken, ist es wichtig sämtliche Faktoren, welche mit der Entstehung und Verbreitung von Resistenzen assoziiert sind, zu kennen und zu verstehen.

Aufgrund der Tatsache, dass viele Resistenzmechanismen genetisch verknüpft sind und somit Resistenzen oft parallel auftreten, schlagen wir in dieser Arbeit die Verwendung eines Resistenzscores, gebildet aus der multivariaten quantitativen Information des Resistenzprofils, als Zielgröße in Ursache-Wirkungs-Modellen vor. Dieser Score basiert auf den minimalen Hemmkonzentrationen (MHK) und wird definiert als die Manhattan Distanz zwischen dem betreffenden Isolat und einem fiktiven Isolat, welches die kleinst möglichen MHK-Werte für alle Resistenztests ausprägt. Somit kann der Score als Maß für die globale Resistenz interpretiert werden.

Zur Evaluierung der Scorebildung wird eine multivariate Analyse der MHK-Werte basierend auf distanzbasierten Permutationstests durchgeführt (Anderson 2001).

Die Methoden werden auf Resistenzdaten von 383 Salmonella Typhimurium Isolaten von sporadischen Salmonellose-Fällen in Niedersachsen angewendet, welche ein hohes Maß an Multiresistenz zeigen (Ruddat et al. (2012)). Beide Verfahren zeigen ähnliche Ergebnisse in den Testentscheidungen bzgl. 70 Faktoren (epidemiologische Informationen über die ursächlichen Träger der Salmonellen). Somit ist die Auswertung des 1-dimensionalen Resistenzscores bei gegebener Datenlage eine gute Alternative zur Betrachtung der vollständigen multivariaten Resistenzinformation.

References:

Anderson, M. J. (2001). "A new method for non-parametric multivariate analysis of variance" Austral Ecology 26: 32-46.

Ruddat, I.; Schwarz, S.; Tietze, E.; Ziehm, D.; Kreienbrock, L. (2012). "A quantitative approach to analyse linkages between antimicrobial resistance properties in Salmonella Typhimurium isolates" Epidemiology and Infection 140, 1: 157-167.

370: 15

Modelling the hierarchical structure in datasets with very small clusters: Is a general linear mixed model for binomial outcome always best? [B9]

Odile Sauzet[1], Janet L Peacock[2]

[1]Universität Bielefeld, Germany; [2]King's College London, United Kingdom

Datasets of preterm infants incorporate clusters of size two (twins), size three (triplets) and so on, with the majority of infants being in ‘clusters’ of size one. It is unclear whether adjustment for clustering is needed or even possible. We compared analyses allowing for clustering (GLMM and GEE) with analyses ignoring clustering (GLM) in order to inform on the magnitude of the error introduced by not adjusting for non independence, and we documented the difficulties encountered when trying to do so. Using a strategy similar to the one in our previously published work for continuous outcomes (Sauzet et al, SIM 2012), through simulations based on a real dataset, we explored estimation bias of parameters, standard errors and random effect variance in predictors of binomial outcomes in different size datasets, with varying percentages of twins.

The Gauß-Hermite and Laplacian approximation methods provide the least biased estimates while under-estimating the random effect variance. However the difference with GLM estimates is small. PQL and GEE methods fail to estimate the parameters accurately when the number of clusters is insufficient. GLM is the model which gives the best measure of the parameter estimate variability.

Contrary to the continuous case where standard errors obtained from mixed models were a much more accurate measure of the type I error than those obtained from the linear regression, it seems that there is little benefit in fitting a GLMM model for datasets with a large percentage of twins unless an estimate of the ICC is needed.

370: 16

Modelling change of carotid intima-media thickness in a population-based prospective cohort - preliminary results from the Heinz Nixdorf Recall Study [B10]

Marie Henrike Berg[1], Frauke Hennig[2], Marcus Bauer[3], Stefan Möhlenkamp[4], Susanne Moebus[1], Raimund Erbel[3], Karl-Heinz Jöckel[1], Barbara Hoffmann[2], André Scherag[1]

[1]Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Germany; [2]IUF-Leibniz Research Institute for Environmental Medicine, Düsseldorf, Germany; [3]West German Heart Center, Department of cardiology, University Hospital of Essen, Germany; [4]Department of Cardiology, Hospital Bethanien Moers, Germany

Background: Intima-media thickness (IMT) is often used as surrogate marker of subclinical atherosclerosis and predicts vascular events [1]. Changes of IMT might provide additional information although this view has recently been challenged [2].

Materials and Methods: We analysed baseline and follow-up data (2000-2003 and 2006-2008) of the population-based Heinz Nixdorf Recall Study (n=4.814 subjects, aged 45-75 years). Side-specific IMT was assessed for the common carotid artery by B-mode ultrasonography and measured in the first 10 mm proximal to the bulb using the semi-automatic software “Artery Measurement System II“. Differences between sides, sex and age groups (under / over 60) were evaluated. Analyses were limited to subjects without prior notice to carotid plaque.

Results: Results are based on data from 2.024 subjects. Mean yearly change (±SD) was 0.020±0.026 (min, max: -0.085, 0.153) mm/year for the right side and 0.019 ± 0.025 (min, max: -0.113, 0.159) mm/year for the left. We detected no evidence for sex differences regarding change (right: p=0.47, left: p=0.07), but there was some evidence for side-dependencies (p = 0.04). The change was smaller in older subjects: left 0.016mm/year vs. 0.020 mm/year (p=0.001) and right 0.019 mm/year vs. 0.022 mm/year (p=0.001).

Discussion: The progression is comparable to other findings [2]. To address the problem of measurement error related to IMT measurements we discuss the results of regression calibration [3] and covariate measurement error models [4] on our descriptive findings.

This project is supported by the German Research Foundation (DFG) within the program HO 3314_2-1.

370: 17

Randomization does not help much [B8]

Uwe Saint-Mont

FH Nordhausen, Germany

Following Fisher, it is widely believed that randomization "relieves the experimenter from the anxiety of considering innumerable causes by

which the data may be disturbed." In particular, it is said to control for known and unknown nuisance factors that may considerably challenge the validity of a result. Looking for quantitative advice, I will study a number of straightforward, mathematically simple models. However, they all demonstrate that the optimism with respect to randomization is wishful thinking rather than based on fact: In small to medium-sized samples, random allocation of units to treatments typically yields a considerable imbalance between the groups, i.e., confounding due to randomization is the rule rather than the exception. For this, and some further reasons, classical experimentation based on sound background theory and the systematic construction of comparable groups, seems to be preferable.

370: 18

Validity of one-day cross-sectional assessments – the NutritionDay study [B6]

Sophie Frantal, Peter Bauer, Karin Schindler, Michael Hiesmayr

Medizinische Universität Wien, Austria

The nutritionDay is a one-day multinational cross-sectional audit capturing data on disease and nutrition of patients admitted to hospital wards. Its aim is to improve knowledge and awareness of malnutrition in hospital. Until now the nutritionDay has been repeated 7 times since the first data collection in 2006, including about 70000 patients from 46 countries [1,2]. Participation is open to any clinical ward via the web (www.nutritionday.org). Patient outcome is evaluated 30 days after nutritionDay.

A main problem occurring in cross-sectional studies is sampling bias. Length of hospital stay (LOS) differs strongly between patients and causes a length bias as patients with longer LOS are more likely to be included and are therefore overrepresented in this type of study. To compensate for the length bias it was suggested to give higher weight to patients with shorter LOS and to account for censoring [3].

The adjusted analysis dramatically changed the results. While the median (lower quartile; upper quartile) LOS in the original sampled patients was 16(8;31), it decreased to 7(4;12) in the adjusted sample. As this result shows it is very important to statistically adjust cross-sectional studies for this length bias to achieve reliable information. Adjustment can be interpreted as a change in the point of view; while the crude data belongs to the view of the caregivers, the adjusted data belongs to a patient population view. Results are validated including all patients admitted to the General Hospital of Vienna between 2005 and 2009, where the distribution of LOS is known.

370: 19

Mortality after myocardial infarction in non-diabetic and diabetic people between 1985 and 2009. The MONICA/KORA registry [B2]

Heiner Claessen[1], Andrea Icks[1,2], Inge Kirchberger[3,4], Margit Heier[3,4], Annette Peters[3], Ines Trentinaglia[3], Giudo Giani[1], Wolfgang von Scheidt[5], Christa Meisinger[3,4]

[1]German Diabetes Centre at the Heinrich-Heine-University, Leibniz-Centre for Diabetes Research, Düsseldorf, Germany; [2]Department of Public Health, Medical Faculty, Heinrich-Heine University Düsseldorf, Germany; [3]Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Institute of Epidemiology II, Neuherberg, Germany; [4]Central Hospital of Augsburg, MONICA/KORA Myocardial Infarction Registry, Augsburg, Germany; [5]Central Hospital of Augsburg, Department of Internal Medicine I - Cardiology, Augsburg, Germany

Objective:

To analyse mortality after myocardial infarction (MI) in people with compared to those without diabetes in Southern Germany, 1985-2009.

Methods:

Using data of the population-based MONICA/KORA Myocardial Infarction Registry, we ascertained all patients with a first fatal or non-fatal MI between 1985 and 2009 (n=16,478, age 25 to 74 years, 71% male, 29% with diabetes). The impact of diabetes on mortality was examined using multiple logistic and Cox regression. All analyses were conducted for the total population and stratified by sex and time period of first MI.

Results:

The cumulative 25 year survival was 7.8% and 3.7% in diabetic men and women, and 16.2% and 10.2% in their non-diabetic counterparts, respectively. The proportion of fatal MIs was significantly higher in diabetic compared to non-diabetic patients (multiple adjusted OR: 1.27; 95% confidence interval 1.18-1.36) with no differences between sexes and time periods. Likewise, mortality after non-fatal MI was significantly higher among both diabetic men and women after adjustment for age, year of first MI, history of angina pectoris, hypertension, stroke, smoking status and highest degree of education (HR 1.63; 1.46-1.81 and 1.83; 1.56-2.14) with no significant changes between time periods. Time dependency of HR comparing diabetic vs. non-diabetic subjects was not statistically significant (men: p=0.65, women: p=0.66).

Conclusions:

The probability of fatal MI and mortality in non-fatal MI is significantly higher in the diabetic population, in particular in women. However, during the past 25 years, survival has improved in both diabetic and non-diabetic patients in a similar manner.

370: 20

Application of competing risks regression models to evaluate risk stratification for cardiac death after myocardial infarction [B3]

Bernhard Haller[1], Georg Schmidt[2], Kurt Ulm[1]

[1]Institut für Medizinische Statistik und Epidemiologie der Technischen Universität München, Munich, Germany; [2]I. Medizinische Klinik und Poliklinik der Technischen Universität München, Munich, Germany

Background: In many clinical research applications the time to occurrence of one event of interest, that may be obscured by another - so called competing - event, is investigated. Specific interventions can only have an effect on the endpoint they address or research questions might focus on risk factors for a certain outcome. Different approaches for the analysis of time-to-event data in the presence of competing risks were introduced in the last decades including some new methodologies, which are not yet frequently used in the analysis of competing risks data.

Methods: Cause-specific hazard regression, subdistribution hazard regression, mixture models and vertical modeling were applied to a dataset of a cohort study including 2,341 patients intending to establish risk stratification for cardiac death after myocardial infarction. Death from cardiac reasons was the event of interest and death from other causes was treated as competing event. Results obtained from different competing risks regression approaches are presented and discussed.

Aim: Different methods for analysis of competing risks data were applied to the same dataset. Data analysts are encouraged to use the appropriate methods for their specific research questions.

The results of the analyses are published in [1].

References:

[1] Haller B, Schmidt G, Ulm K. (2012) Applying competing risks regression models: An overview. Lifetime Data Analysis. DOI 10.1007/s10985-012-9230-8

370: 21

Tree-based identification of subgroups with enhanced treatment effect for survival data [B1]

Harriet Sommer

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In medical statistics it is a challenging task to determine the effects of different therapy alternatives, however, one therapy can have enormously diverse implications for different patients. This thesis deals with the question which method can be used to solve this kind of problem: to find the right therapy for an individuals' disease.

With different subgroup identification procedures we will extract subgroups of patients out of a population which will benefit significantly from the new treatment and as well subgroups for which the old treatment seems to be the better alternative - if this situation exists. We will especially focus on a tree-building subgroup identification procedure called interaction tree proposed by Su et al. in 2011 for longitudinal trials and translate it in order to analyse clinical studies that involve censored survival times. Our findings are compared to other techniques and applied to different data sets in the application part including the Glioma and the MRC RE01 trial. Simulated data were designed to investigate the procedure and explore possible constraints and weaknesses.

We will also introduce the selection impact curve (Song and Pepe), a method which seeks at finding the optimal cutpoint for the treatment allocation, and illustrate our results. Additionally we apply a measure for the heterogeneity of the treatment effect between the resulting subgroups, Cochran's Q and the popular [fn]I^2[/fn] (Higgins and Thompson). With the help of these measures we can asses our resulting subgroups and justify to plot a selection impact curve if a sufficient heterogeneity is given.

370: 22

Adaptive designs with arbitrary dependence structure [D15]

Rene Schmidt, Andreas Faldum, Joachim Gerss

University of Münster, Germany

Adaptive designs were originally developed for independent and uniformly distributed p-values. Sometimes, independence is not satisfied or it may not be possible to check whether it is satisfied. We start with examples of trial settings, where adaptive designs with complex dependence structure between the p-values of the stages may arise. The probability of a type I error for a fixed adaptive design depends on the true dependence structure between the p-values of the stages. If the dependence structure is not taken into account adequately, control of type I error might be endangered. In particular, the worst case among all dependence structures and the type I error rate in the worst case may be considered. We discuss the type I error rate in the worst case for the class of inverse normal designs [1] as well as for designs based on Fisher's combination test [2]. Thereby, the focus is on two-stage designs without futility stop. For this purpose, we pursue a copula approach to adaptive designs. A comparison of the worst case level with the reference level for independent, uniformly distributed p-values is performed. On this basis the decision boundary for the second stage can be modified so that the type I error is controlled in the worst case and thus for any dependence structure.

References:

[1] Lehmacher, W. and Wassmer, G. (1999). Adaptive sample size calculations in group sequential trials. Biometrics 55, 1286–1290.

[2] Bauer, P. and Köhne, K. (1994). Evaluation of experiments with adaptive interim analyses. Biometrics 50, 1029–1041.

370: 23

Statistical Process Modeling and Simulation of Inhomogeneous Mineral Subsoil Machining [E5]

Swetlana Herbrandt, Nils Raabe, Manuel Ferreira, Christian Rautert, Claus Weihs

TU Dortmund University, Germany

The target of optimizing the concrete drilling process is the reduction of tool wear and machining time, which can be achieved by adapting tools to the particular machining process. To determine optimal parameter settings for specific conditions models are developed to explain the influences of the process parameters on the forces affecting the workpiece, the chip removal rate and the diamond wear rate. The obtained findings are used to derive a geometrically motivated simulation model describing cutting forces, tool and workpiece wear. In this simulation model workpiece and diamond grain are subdivided into Delaunay tessellations where the resulting micropart connections are the predefined breaking points. In addition the specific material heterogeneity of the used workpiece is modeled by Gaussian Random Fields. Finally we extend the simulation model for multiple diamonds randomly distributed in the grinding segment.

370: 24

Minimization and extensions in Diagnostic Development: subject/sample allocation for training/test sets and cross-validation [E6]

Holger Hoefling, Anthony Rossini

Novartis, Switzerland

Minimization is a deterministic dynamic allocation technique used in clinical trials to ensure that important factors are balanced across treatment groups. In the setting of Diagnostic Development, we often would like to use available samples from prior studies in a manner ideally comparable to a prospective study. These retrospective samples are often distributed between 2 groups, a training set where the diagnostic algorithm will be developed and a test set for performance evaluation. In order to make this technique applicable to the assignment into training and test sets, we investigate an extension of Minimization - Randomized Minimization. Randomized Minimization consists of permuting the order of the samples/patients in a design-reasonable way, to an ordering which differs from the original accrual order, and then applying Minimization. This approach balances assignment on key factors between the training and the test set while avoiding the criticism that Minimization is a deterministic allocation procedure. The same approach can be used for cross-validation of the classification algorithm so that we can characterize the diagnostic. We use simulations to compare Minimization s performance to other techniques including such as standard randomization and stratification in terms of bias as well as variability.

References:

[1] Minimization: a new method of assigning patients to treatment and control groups. Taves DR. Clin Pharmacol Ther. 1974 May;15(5):443-53.

370: 25

Simultane Konfidenzintervalle mit gleichen marginalen Überdeckungswahrscheinlichkeiten [E7]

Benjamin Markus Becker

Universitätsmedizin Göttingen, Germany

Um simultane Konfidenzintervalle zu Tests bei unbalancierten heteroskedastischen normalverteilten Stichproben auszurechnen, greifen die bekannten Verfahren auf die Verteilung des Maximums der einzelnen Teststatistiken zurück. Die Konfidenzintervalle beruhen dann auf demselben Quantil der Maximumverteilung der Statistiken, auch wenn die Freiheitsgrade dort unterschiedlich sind. Auf diese Weise ist die Überdeckungswahrscheinlichkeit für den wahren Parameter in jedem Konfidenzintervall unterschiedlich, so dass die Längen der Konfidenzintervalle zuungunsten der Parameter nivelliert werden, deren Schätzung ein größerer Stichprobenumfang zugrunde liegt. In Bezug auf die dualen Tests bedeutet das, dass in der globalen Testentscheidung ausgerechnet die Vergleiche stärker eingehen, die für sich genommen mit einer geringeren Power getestet würden.

In diesem Konferenzbeitrag wird eine Berechnung der Konfidenzintervalle auf der Grundlage eines parametrischen Bootstraps vorgestellt, bei der dieser Effekt nicht auftritt.

370: 26

Subscan - a cluster algorithm for identifying statistically dense subspaces with application to biomedical data [A9]

Johann M. Kraus, Hans A. Kestler

Ulm University, Germany

Cluster analysis is an important technique of initial explorative data mining. Recent approaches in clustering aim at detecting groups of data points that exist in arbitrary, possibly overlapping subspaces. Generally, subspace clusters are neither exclusive nor exhaustive, i.e.

subspace clusters can overlap as well as data points are not forced to participate in clusters. In this context subspace clustering supports the search for meaningful clusters by including dimensionality reduction in the clustering process. Subspace clustering can overcome drawbacks from searching groups in high-dimensional data sets, as often observed in clustering biological or medical data. In the context of microarray or next-generation sequencing data this refers to the hypothesis that only a small number of genes is responsible for different tumor subgroups. We generalize the notion of scan statistics to multi-dimensional space and introduce a new formulation of subspace clusters as aggregated structures from dense intervals reported by single axis scans. We present a bottom-up algorithm to grow high-dimensional subspace clusters from one-dimensional statistically dense seed regions. Our approach objectifies the search for subspace clusters as the reported clusters are of statistical relevance and are not artifacts observed by chance. Our experiments demonstrate the applicability of the approach to both low-dimensional as well as high-dimensional data.

370: 27

An efficient subspace KNN [A13]

Ludwig Lausser, Hans A. Kestler

RG Bioinformatics & Systems Biology, Ulm University, Germany

Feature selection algorithms are the main ingredients for constructing interpretable and sparse classification models. By removing features they help to identify key components and to neglect distracting measurements. Designed for high-dimensional spaces modern feature selection algorithms are based on search heuristics, which only evaluate a small fraction of all possible feature combinations. It can seldom be guaranteed that a gained score corresponds to the optimal achievable value for the chosen objective.

In this work we present a feature selection algorithm based on the cross-validation results of a k-Nearest Neighbor classifier. The structure of this classifier allows a very fast evaluation if done for ascending sequences of feature sets. It can be used to exhaustively evaluate all subspaces of a datasets of small or medium dimensionality in a reasonable time span. A first implementation of the algorithm showed a performance of about 200.000 10x10 cross-validations per minute on a dataset of 100 samples (CPU: 2.66 GHz).

The method is based on an in-order enumeration of all possible feature combinations and the corresponding distance matrices. It is a prototype of a fast exhaustive search algorithm suitable for a variety distance based feature selection methods. Indexing each feature combination in the correct order allows parallelization of this scheme.

The performance of the new algorithm is investigated in both real and artificial datasets. The feature selections are tested on their suitability for state-of-the-art classification algorithms. The method is also analyzed on its descriptive capabilities.

370: 28

Suitable loss functions for cost sensitive classification [A12]

Lydia Rybizki

FAU Erlangen-Nürnberg, Germany

In this work, we investigate the influence of different criteria on construction and assessment of classifiers. We focus on score based classification rules, i.e. classification depends on whether a score, defined as a class probability estimate or some monotonic transformation of it, exceeds a certain threshold. If criteria are chosen according to the specific problem, as often demanded, we argue that three criteria are most relevant in practical classification problems with two classes: criteria based on (1) generalized 0-1-loss when misclassification costs are given, (2) ranking loss (relating to area under the ROC curve) when they are completely unknown or (3) a special class of proper scoring rules (beta-losses) proposed by Buja et al (2005) when assumptions about the ratio of misclassification costs can be made. In a simulation study, we first compare estimated scores resulting from beta-, ranking and standard (e.g. log) losses in terms of classification performance. Thereby, we identify estimation methods eligible for certain performance criteria, especially when underlying models are misspecified. Second, we employ 0-1-, beta- and ranking loss based criteria to select variables for score construction in a cross-validatory framework. The results are compared to standard variable selection criteria in terms of size and classification performance of selected subsets. In addition, we combine cost based criteria with costs for observing variables to derive an overall cost sensitive selection of variables.

370: 29

Classifying real-world data with the DD[fn]\alpha[/fn]-procedure [A11]

Pavlo Mozharovskyi[1], Karl Mosler[1], Tatjana Lange[2]

[1]University of Cologne, Germany; [2]Hochschule Merseburg, Germany

The DD[fn]\alpha[/fn]-procedure is applied to fifty binary classification problems regarding real-world benchmark data. The procedure first transforms the data from their original feature space into a depth space, which is the unit square, and then separates them by a projective invariant procedure, called [fn]\alpha[/fn]-procedure. To each data point the transformation assigns its depth values with respect to the two classes. Here the random Tukey depth is employed, which approximates the Tukey depth by minimizing univariate Tukey depths over a finite number of directions. An important practical question is how to treat `outsiders', that is data points having zero depth in both classes. The talk introduces several possible treatments and provides broad numerical experience about their usefulness. Also evidence is given that in almost all cases two features in the (possibly extended) depth space are enough to satisfactorily separate the data.

370: 30

A new measure of the impact of incomplete cross validation, with applications to various steps of prediction rule construction [A10]

Roman Hornung, Christoph Bernau, Anne-Laure Boulesteix

Ludwig-Maximilians-Universität München, Germany

In supervised statistical learning, it is widely recognized that prediction models should not be constructed and evaluated using the same data set. In practice, however, the available data set is often too small to be split into training and validation sets. It is then standard practice to estimate the prediction error using cross-validation (CV) or a related procedure.

By “incomplete CV”, we mean a CV procedure in which for one or several analysis steps the excluded folds are taken into account before or during the construction of the prediction rules. This is known to yield strongly biased error estimates resulting in over-optimistic conclusions in some cases, e.g. supervised variable selection in high-dimensional prediction problems. However, other aspects are far less acknowledged although also potentially dangerous with respect to over-optimistic error estimation.

In this poster, we present a new measure quantifying the impact of incomplete CV and, using numerous real-life low- or high-dimensional data sets, we apply this new measure to various data preparation steps including preliminary variable selection, choice of tuning parameters, normalization and batch effect removal (in the case of high-throughput omics data), dichotomization of metric predictors, or imputation of missing values. The new measure is based on the ratio between the prediction errors estimated when these procedures are applied only once to the whole data set and when they are repeated for each CV iteration, i.e. for each considered training set anew, and subsequently applied to the excluded fold via addon-procedures.

370: 31

Fold change classifiers for high-dimensional data [A14]

Ludwig Lausser, Hans A. Kestler

RG Bioinformatics & Systems Biology, Ulm University, Germany

The classification of biomolecular data is based on high-dimensional profiles consisting of thousands of measurements. Both for reducing the model complexity and increasing insights into the processes underlying the classification it is assumed that the used models are to be interpretable in terms of a small signature of measurements. Types of classification systems suitable for this scenario are fusion strategies of low dimensional base classifiers. While a large variety of fusion strategies exist, the base ensemble classifier is commonly chosen to be a single threshold classifier (e.g., [fn]x^{(i)}>t[/fn]). Although this base classifier has some beneficial characteristics, it is not guaranteed to be the best choice in many settings.

In this work we discuss an alternative base classifier of type [fn]x^{(i)} > tx^{(j)}[/fn]. Its decision criteria can be interpreted as evaluating the fold-change of two features of a query sample. This type of decision boundary is independent from a global threshold and therefore invariant against global multiplicative effects. An ensemble of such fold-change classifiers inherits this property. We give bounds on the generalization ability of these types of classifiers and show their impact on analyzing gene expression microarray data.

370: 32

Which antigens are really important? – An interdisciplinary and pragmatic statistical approach within diagnostic biomarker panel discovery for rheumatoid arthritis [A15]

Carmen Theek, Anna Telaar

Protagen AG, Germany

High dimensional data analysis plays an important role in the context of biomarker discovery as the number of parameters exceeds the number of observations immensely. Within an exploratory discovery study a pragmatic and interdisciplinary strategy is shown to extract informative antigens for classification of rheumatoid arthritis (RA) patients.

Using the Luminex® xMAP technology providing median fluorescence intensity values it was intended to identify a marker panel.

Based on the results of a pilot study measuring a total of 3051 antigens the number of antigens was reduced to 145 for this study, observed in 74 RA patients and 71 healthy controls. Following pre-processing procedures univariate non-parametric testing for group comparison including correction of multiplicity was applied. Univariate classification performance was evaluated by ROC analysis. Additionally, multivariate analyses were embedded within a 10-fold cross-validation (training and test set) within a 200-fold loop. For feature selection two different approaches were used: 1) Powered Partial Least Squares Discriminant Analysis and 2) Random Forest, both followed by fitting a logistic regression model using a stepwise procedure. The classification performance of this model was evaluated on the test set.

Intersections of different statistical approaches passed expert review and were investigated further applying correlation analyses and considering classification performance. Final panel definition was based on biological evidence and statistical relevance.

As an independent external test set 30 samples of patients with RA were used to determine classification characteristics of the panel defined. A correct classification rate of above 90% could be achieved within this external test set.

370: 33

Penalized estimation in high-dimensional hidden Markov models with state-specific graphical models [A17]

Nicolas Staedler, Sach Mukherjee

Netherlands Cancer Institute, The Netherlands

We consider penalized estimation in hidden Markov models (HMMs) with multivariate Normal observations. In the moderate-to-large dimensional setting, estimation for HMMs remains challenging in practice, due to several concerns arising from the hidden nature of the states. We address these concerns by l1-penalization of state-specific inverse covariance matrices. Penalized estimation leads to sparse inverse covariance matrices which can be interpreted as state-specific conditional independence graphs. Penalization is non-trivial in this latent variable setting; we propose a penalty that automatically adapts to number of states K and state-specific sample size and can cope with scaling issues arising from the unknown states. The methodology is adaptive and very general, applying in particular to both low- and high-dimensional settings without requiring hand tuning. Furthermore, our approach facilitates exploration of the number of states K by coupling estimation for successive candidate values K. Empirical results on simulated examples demonstrate the effectiveness of the proposed approach. In a challenging real data example from genome biology, we demonstrate the ability of our approach to yields gains in predictive power and deliver richer estimates than existing methods.

370: 34

Integrative analysis of histone ChIP-seq and gene expression microarray data using Bayesian mixture models [A20]

Hans-Ulrich Klein[1], Martin Schäfer[2], Bo Porse[3], Martin Dugas[1], Katja Ickstadt[2]

[1]University of Münster, Germany; [2]TU Dortmund University, Germany; [3]University of Copenhagen, Denmark

Histone modifications are an epigenetic mechanism to activate or repress the expression of genes. Several data sets of matched microarray expression data and histone modification data measured by ChIP-seq have been published. We present a novel approach to detect genes that are differentially expressed between two conditions due to an altered histone modification, useful for small sample sizes.

Both data types are matched by assigning the number of ChIP-seq reads aligning within the promoter region of a gene to the normalized expression value of that gene. Quantile normalization is applied to ChIP-seq values. Then, a score is calculated for each gene by multiplying the standardized difference of ChIP-seq values by the standardized difference of expression values. A Bayesian mixture model with a standard Dirichlet prior is fitted to these scores, allocating each gene to one of six components to finally classify genes into three groups: Genes with (i) equally directed differences in both data sets, (ii) reversely directed differences in both data sets and (iii) no differences in at least one of the data sets. Group (iii) is represented by four centered normal components whereas an exponential component is used for group (i) and a mirrored exponential component for group (ii).

We applied the method to ChIP-seq measurements and matched gene expression data from three mice. Overall, 450 (29) genes out of 21236 genes were classified into group (i) ((ii)), indicating that the investigated chromatin signature, H3K4me3, is an activating histone mark. Stability of the classification was studied using replicates.

370: 35

Estimating the Mahalanobis distance in High-dimensional Data [A16]

Deliang Dai

Linnaeus University, Sweden

The Mahalanobis distance is a fundamental statistic in many fields such as Outlier detection, Normality testing and Cluster analysis. However, the standard estimator developed by Mahalanobis (1936) and Wilks (1963) is not well behaved in cases when the dimension (p) of the parent variable increases proportional to the sample size (n). This case is frequently referred to as Increasing Dimension Asymptotics (IDA). Specifically, the sample covariance matrix on which the Mahalanobis distance depends becomes degenerate under IDA settings, which in turn produce stochastically unstable Mahalanobis distances. This research project consists of several parts. It (a) shows that a previously suggested family of “improved” shrinkage estimators of the covariance matrix produce inoperable Mahalanobis distances, both under classical and increasing dimension asymptotics. It (b) develops a risk function specifically designed to assess the Mahalanobis distance and identifies good estimators thereof and (c) develops a family of resolvent-type estimators of the Mahalanobis distance. This family of estimators is shown to remain well behaved even under IDA settings. Sufficient conditions for the proposed estimator to outperform the traditional estimator are also supplied. The proposed estimator is argued to be a useful tool for descriptive statistics, such as Assessment of influential values or Cluster analysis, in cases when the dimension of data is proportional to the sample size.

370: 36

Challenges for integrating prior biological knowledge in high-dimensional survival models [A18]

Kai Kammers, Jörg Rahnenführer

TU Dortmund University, Department of Statistics, Germany

First, we present an approach for improving performance and interpretability of survival prediction models by integrating gene expression data with prior biological knowledge. To increase the interpretability of prognostic models, we combine genes to gene groups and use these

groups as covariates in the survival models. Since expression profiles within gene groups are often heterogeneous, we present a new method for obtaining subgroups with coherent patterns. We apply preclustering to genes within predefined groups according to the correlation of their gene expression measurements. For evaluating the fitted survival models, we show prediction error curves revealing that models with preclustered gene groups have improved prediction performance compared to models built with single genes or gene groups.

Besides gene expression data, there exist other data sources containing genomic, epigenetic and proteomic information. Thus, in a next step, we discuss challenges in extending the preclustering approach to other high-dimensional data. How can we make use of different types of biological prior knowledge (GO, KEGG, MSigDB and PANTHER databases)? How can we integrate genetic data whose measurements are categorical and not numerical? What are competing statistical models to be compared with the preclusting approach (Sparse Lasso, Group Lasso, and CoxBoost)? How can we provide an efficient computational framework, e.g. based on the R package BatchExperiments?

370: 37

A Hidden Markov Model Integrating Genotype and Copy Number Information from Exome Sequencing Data [A19]

Christian Ruckert[1], Michael Gombert[2], Arndt Borkhardt[2], Martin Dugas[1]

[1]University of Münster, Germany; [2]Heinrich Heine University, Düsseldorf, Germany

In routine diagnostics targeted sequencing is preferred over whole genome approaches due to its lower costs. Various articles in the last months showed that, in addition to the detection of structural aberrations, exome sequencing data could also be used for copy number (CN) estimation. Several methods are based on hidden markov models (HMM) using the underlying CNs as hidden states. Curiously none of these incorporates data from genotype (GT) calling, although this could provide valuable additional information.

The observed data sequence along genomic coordinates consists of two components: read counts in targeted regions and allele frequencies at known SNP locations. To integrate these two we developed a HMM with four hidden states:

1.) deletion -- decreased read depth and loss of heterozygosity for GT calls,

2.) normal - no CN change and retention of zygosity,

3.) copy neutral loss of heterozygosity -- normal CN paired with an unusually long sequence of homozygote calls, and

4.) amplification -- gain in CN with no information about influence on GT.

Emission probabilities were calculated for CN and GT separately and then combined to estimate the most probable sequence of hidden states. The model for CN estimation accounts for GC-content and background read depth.

Using simulated data as well as whole genome and exome sequencing data from a sample obtained at three different times (initial, remission, relapse) we compared our method to other algorithms for copy number estimation from exome sequencing data.

370: 38

A nonparametrically improved Sharpe ratio guided by prior knowledge [C6]

Michael Scholz[1], Jens P. Nielsen[2]

[1]University of Graz, Austria; [2]Cass Business School, United Kingdom

We propose a two step approach to improve one of the most used measures of performance evaluation for investment strategies, the Sharpe ratio. Expected returns and their variability in financial markets vary over time and contain a significant predictable component. Thus, we first predict excess stock returns in a heteroscedastic regression model applying a local-linear kernel smoother improved by prior knowledge, and second we repeat the approach for the conditional variance function. To measure the quality of the prediction model, we apply the validated R-sq., where the prediction of our cross-validated model is compared to the simple cross-validated mean. A ratio of the estimates of both steps gives our new estimator for which we also provide statistical properties. Based on this Sharpe ratio estimator, we propose an advanced trading strategy that is superior to different often applied simple investment plans. In an applied part, we show the efficiency of our estimator and strategy using monthly data from the S&P500 in a period from 1953 to 2010.

370: 39

Quality Leisure Time and Youth Development: What Are the Effects of Sports? [C4]

Benjamin Fuchs[1], Aderonke Osikominu[2,1,3]

[1]University of Freiburg, Germany; [2]University of Zurich, Switzerland; [3]IZA, Bonn, Germany

Parents and policymakers alike worry about what activities may provide valuable learning experiences to youth beyond the domain of schooling. Sports has been singled out as a popular pastime that is positively related with educational and labor market outcomes at later ages. While existing research supports the hypothesis that athletic participation may have a positive causal effect on skills we know little about the underlying mechanisms. Does athletic involvement crowd out leisure activities with potentially adverse effects on skill formation? What additional skills do athletes acquire that lead to better educational and labor market outcomes? Psychological research shows that interpersonal and executive skills develop strongly during adolescence and are likely sensitive to particular experiences during that time. A

growing body of economic research documents the importance of interpersonal and executive skills for economic success. This paper sets out to explore what youth do in their leisure and whether athletic participation affects behavioral outcomes reflecting interpersonal and executive skills. Our empirical analysis exploits data from the German Socio-Economic Panel that offers the unique advantage of both a large, representative sample and experimentally validated survey questions on behavioral measures. Our results are consistent with the hypothesis that informal learning activities during adolescence influence the development of interpersonal and executive skills. Athletes also show better educational outcomes than comparable youth who do not play sports. The benefits from participating in sports tend to be higher for youth from less advantaged families suggesting that athletic involvement may contribute to reducing socioeconomic disparities.

370: 40

Gold as an inflation hedge in a time-varying coefficient framework [C3]

Robert Czudaj, Joscha Beckmann

University of Duisburg-Essen, Germany

This study analyzes the question whether gold provides the ability of hedging against inflation from a new perspective. Using data for four major economies, namely the USA, the UK, the Euro Area, and Japan, we allow for nonlinearity and discriminate between long-run and time-varying short-run dynamics. Thus, we conduct a Markov-switching vector error correction model (MS-VECM) approach for a sample period ranging from January 1970 to December 2011. Our main findings are threefold: First, we show that gold is partially able to hedge future inflation in the long-run and this ability is stronger for the USA and the UK compared to Japan and the Euro Area. In addition, the adjustment of the general price level is characterized by regime-dependence, implying that the usefulness of gold as an inflation hedge for investors crucially depends on the time horizon. Finally, one regime approximately accounts for times of turbulences while the other roughly corresponds to `normal times'.

370: 41

Adverse Selection in Secondary Insurance Markets: Evidence from the Life Settlement Market [C5]

Daniel Bauer[1], Jochen Russ[2], Nan Zhu[3]

[1]Georgia State University, USA; [2]Institut für Finanz- und Aktuarwissenschaften, Germany; [3]Illinois State University, USA

Life expectancy providers specialize on compiling individualized mortality forecasts for underwriting life insurance contracts in the primary and secondary insurance market. We use a dataset of estimated life expectancies for one of the largest US providers and corresponding realized death times to analyze the adequacy of the individualized mortality forecasts. In doing so, we develop an array of metrics and statistical tests to discern the goodness of the provided estimates.

We find that while the provided distributions do not perfectly match the observation, the average error in the “level” of the life expectancies -- while statistically significant -- is small. By deriving non-parametric estimates for an adjustment factor, we demonstrate that the baseline “shape” of the mortality intensities varies across different partitions of the sample. Nonetheless, even after adjusting the shapes, we find that goodness-of-fit tests are rejected within many of these partitions even for the (in-sample) adjusted individualized mortality curves. We interpret this finding as evidence for residual information in the death times not captured by the compiled estimates.

This residual information creates the possibility for adverse selection in the secondary life insurance. To test for it, we compare the quality of the estimates of a sub-sample of cases that are known to be underwritten in the secondary life insurance market to the rest of the sample. We find significant positive difference suggesting the existence and the significance of asymmetric information -- a finding that has important economic consequences. We use survival regressions to check the robustness of this result.

370: 42

Variance estimators for Brier scores [B5]

Sarah Julia Marzi

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Assessment of predictive accuracy is crucial to clinical research and applications. Survival time predictions of disease outcomes are of interest to researchers, clinicians and patients alike. The Brier score (Graf et al., 1999) has been proposed and studied as a useful measure of predictive accuracy. However, the variance that stems from estimating it in patient samples has not been addressed in detail yet.

A recent paper by Antolini et al. (2012) provides a subgroup weighting scheme for censored survival data. Using this scheme an estimator for the variance of the Brier score is derived for discrete prediction models. These provide a finite classification scheme of with equal survival probability for patients in the same category. Since one of the major applications of the Brier score is the comparison of different prognostic models a second variance estimator is presented: it quantifies the variance which results when estimating the difference of two Brier scores for two discrete prognostic models that are evaluated on the same data set.

A study in patients with breast cancer will be used to illustrate the results of the variance estimators comparing two established prognostic classification schemes. Additionally, a simulation study will show the performance of the estimators in small samples and in presence of censoring.

References:

[1] Antolini et al. Stat.Med. 2012; 31: 1113-28.

[2] Graf et al. Stat.Med. 1999; 18: 2529-45.

370: 43

Is it beneficial to be included in a sustainability stock index? A comparative short-term event study for the US, Europe and Japan [C2]

Franziska Gückelhorn, Gunnar Gutsche, Andreas Ziegler

University of Kassel, Germany

An increasing awareness of climate change, modifications in the consumer behavior as well as a regular appearance of the subject in the political debate accelerated the interest of the financial community in socially responsible investing (SRI). Perhaps the most controversial question in the context of SRI still asks: Is the responsible investor doing well by doing good? This study attempts to shed some more light on this issue by specifically looking on the effects of the inclusion in respectively the exclusion from two widely recognized global sustainability stock indices - the DJSI World and the FTSE4Good Global Index – on the stock prices of the affected firms. Methodically we apply an event study approach. The basic idea behind this methodology rests on the estimation of so called abnormal returns for each firm experiencing this event. Abnormal returns are defined as difference between the actual observed stock returns on the event day and the expected returns given the event did not take place. The estimation of the expected returns is in former studies mostly based on the common market model. However, a well-known stylized fact of financial data are volatility clusters meaning that high (low) amplitudes of asset returns are followed by high (low) amplitudes. This information can be incorporated by using a GARCH approach developed by Bollerslev (1986) to obtain more efficient estimations of the abnormal returns. Therefore, we additionally consider a GARCH (1,1) approach. To test the significance of the abnormal returns we use several parametric and non-parametric tests.

370: 44

Exposure-adjusted binary response models in a large portfolio of car insurance policies [C1]

Linda Elin Walter

University of Basel, Switzerland

This paper models the claim incidence in a Swiss motor third-party liability insurance portfolio by applying extensions of generalized linear models (GLM). Our unique data set is characterized by a large number of contracts, different lengths of risk exposure and changes made to contracts during the insurance period. In order to investigate the relationship between reporting a claim and various covariates we employ an exposure-adjusted logit model and, additionally, keep track of the changes of the covariates over one calendar year. This results in clustering effects with a highly skewed distribution of cluster sizes. Among others, we find that leasing a car strongly increases the probability of reporting a claim, whereas the gender of the main driver shows no statistically significant effect. We can show that the benefit from keeping track of the changes in the contract is limited, whereas it is important to adjust for the exposure time. Furthermore, we conduct various robustness checks and provide alternatives – modified noncanonical link functions and weighted likelihood procedures – to correct for exposure in a binary response setting.

370: 45

Statistische Evaluation des Projekts Rauchzeichen der Deutschen Herzstiftung e.V. [D20]

Mario Schwarzer[1], Jule Thomas[2], Martin Kaltenbach[2], Michaela Nedela-Morales[2], Eva Herrmann[1]

[1]Institut für Biostatistik und mathematische Modellierung. Wolfgang Goethe Universität Frankfurt, Germany; [2]Deutsche Herzstiftung e.V., Projekt Rauchzeichen, Frankfurt am Main, Germany

Der Tabakkonsum ist an der Entstehung vieler schwerer Volkserkrankungen beteiligt. Das Primärpräventionsprogramm Rauchzeichen hat zum Ziel den Einstieg in eine Raucherkarriere bei Jugendlichen zu verhindern. Dazu werden Schulbesuche mit Ärzten an den 7. Klassen aller Schulformen durchgeführt.

In einer kontrollierten und anonymisierten Studie wurden nun die mittelfristigen Effekte des Programms mit Hilfe eines eigens konzipierten Fragebogen exemplarisch an 23 Schulen und über 6000 Schülern der Jahrgänge 8, 9 und 10 evaluiert. Dabei wurden neben SchülerInnen von Schulen, die an dem Rauchzeichenprojekt teilgenommen hatten auch solche vergleichbarer Schulen der Region, die bisher nicht an dem Projekt teilgenommen hatten befragt. Der statistische Vergleich zwischen Interventions- und Kontrollgruppe erfolgte nach Jahrgängen und Schulart stratifiziert. Primäre Zielgrößen waren neben dem Anteil der regelmäßig oder gelegentlich rauchenden SchülerInnen ein Score zum Gesundheits- und Problembewusstsein im Bezug auf das Rauchen (Score 1), sowie ein Score zur Stärke des Rauchens der Schüler (Score 2).

Der Anteil der SchülerInnen, die als Raucher, sowohl gelegentlich wie regelmäßig, kategorisiert wurden, betrug in den 9-ten Klassen zwischen 25.1% (Gymnasien) und 41.6% (Hauptschulen), nahm insgesamt mit dem Alter zu und unterschied sich nach Schultyp. Es gab keinen signifikanten Unterschied zwischen Interventions- und Kontrollgruppe für den Anteil der Raucher und für Score 2 ($p=0,15$). Allerdings traten Unterschiede in Score 1 auf, auch bis zu zwei Jahre nach der Intervention ($p<0.001$).

Im Hinblick auf die meist sehr ungünstigen Ergebnisse in der Evaluierung von Rauchpräventionsprogrammen ist das positive Ergebnis, dass die Interventionsgruppe die gesundheitlichen Risiken höher einschätzt oder kritischeren Umgang mit Tabakerzeugnissen zeigt, besonders

erfreulich.

370: 46

The influence of respondent incentives on item nonresponse and measurement error in a web survey [C7]

Barbara Felderer[1], Frauke Kreuter[1], Joachim Winter[2]

[1]Institut für Arbeitsmarkt- und Berufsforschung, Germany; [2]Ludwig-Maximilians-Universität München, Germany

It is well known that incentives can effectively be used to decrease unit nonresponse. The question we are analyzing here is whether incentives are able to decrease item nonresponse and measurement error as well.

To study the effect of incentives on item nonresponse and measurement error, an experiment was conducted with participants of a web survey. In addition to an incentive for participation, an extra prepaid incentive ranging from 0.50 Euro to 4.50 Euro was given to some respondents towards the end of the questionnaire in the form of an Amazon-voucher. At the same time, respondents were requested to think hard about the answers to the next questions and be as precise as possible. In this experiment there are two reference groups: one group received the request but no incentive and the other did not receive any request or incentive.

We approach our research questions in three steps: Our first analysis focuses on the effect of incentives on the proportion of "don't know's" and "no answer's". In a second step, we look at the amount of rounding and heaping as an indicator for measurement error. In the third step, we examine measurement error directly for two variables (income, unemployment benefit recipiency) by linking the survey data to German administrative records and computing the difference between survey response and administrative records.

Comparisons across the different incentive groups will allow for an assessment of the effectiveness of incentives on item nonresponse and measurement error.

370: 47

Estimation of non-linear indicators on regional level using parametric distributions [C8]

Jan Seger, Ralf Münnich

Universität Trier, Germany

In economics and social sciences the measurement of welfare and progress has been a topic of main interest for decades due to its high policy relevance. Beyond GDP further raises the needs for sophisticated indicators to measure progress. From a methodological point of view, the measurement of complex phenomena, like for example income inequality, leads to non-linear indicators.

In practice, these indicators have to be estimated from survey samples. In case indicator estimates are needed on regional level, classical design-based methods suffer from very low area-specific sub-sample sizes which hardly allow deriving reliable indicator estimates. Alternatively model-based methods such as small area estimators may be applied. In order to derive accurate synthetic estimates for non-linear indicators, the non-sampled part of the universe for which full auxiliary information is available, can be estimated using model-based predictions. We extend this approach by fitting an adequate parametric distribution to the estimated non-sampled elements of the variable of interest. Afterwards we calculate estimators for the indicators of interest applying the fitted distributions in each area. However, the appropriateness of this method is based on the prediction power of the underlying model and the fit of the chosen parametric distributions.

The aim of the paper is to compare design- and model-based methods for estimating non-linear indicators in light of Beyond GDP. The evaluation is based on a large model-based Monte-Carlo study including various realistic scenarios.

370: 48

Common trend smoothing with extended Hodrick-Prescott (HP) filters [C14]

Wolfgang Polasek

Institute for Advanced Studies, Austria

The Hodrick-Prescott (HP) method was originally developed to smooth time series, i.e. to get a smooth (long-term) component. We show that the HP smoother can be viewed as a Bayesian linear model with a strong prior for the smoothness component. Extending this Bayesian version of he HP approach for a common trend model is possible using

a non-conjugate prior and the MCMC estimation technique. The results are more complex since the assumption of a common trend requires additional restrictions between the 2 series than the univariate approach in Polasek (2012).

This extended Hodrick-Prescott approach decomposes the smoother into additive components and can be used for a new class of model-based smoothers for time series and spatial models.

The common trend smoothing approach are applied to two macro-economic time series and a regional smoothing of macro-economic variables in Europe.

370: 49

Physically coherent probabilistic weather forecasts via ensemble copula coupling (ECC) [C11]

Roman Schefzik[1], Thordis Thorarinsdottir[2], Tilmann Gneiting[1]

[1]Institute for Applied Mathematics, Heidelberg University, Germany; [2]Norwegian Computing Center, Oslo, Norway

State of the art weather forecasts depend on ensemble prediction systems, which consist of multiple runs of dynamical numerical weather prediction models differing in the initial conditions and/or the parameterized numerical representation of the atmosphere. Statistical postprocessing of the ensemble forecasts is required to realize their full potential, in the sense that biases and dispersion errors need to be addressed. Current postprocessing approaches mostly apply to a single weather quantity at a single location and for a single prediction horizon only. However, in many applications there is a critical need to account for spatial, temporal and inter-variable dependencies.

To address this, we propose a tool called ensemble copula coupling (ECC), in which existing univariate postprocessing methods are employed to obtain calibrated and sharp forecasts for each location, variable and look-ahead time separately. Then, the univariate distributions are aggregated in a discrete copula approach. The key idea is that the postprocessed ECC ensemble inherits the multivariate rank dependence pattern from the unprocessed ensemble, thereby capturing the flow dependence.

In this talk, we present the ECC approach, study its relationships to discrete copulas, and assess the predictive performance in an application to ensemble data over Germany.

370: 50

Temperature forecasts over Germany using spatial ensemble model output statistics [C13]

Kira Feldmann[1], Michael Scheuerer[1], Thordis L. Thorarinsdottir[2]

[1]Heidelberg University, Germany; [2]Norwegian Computing Center, Norway

Over the past two decades, meteorologists have developed ensemble prediction systems as the state of the art in weather forecasting. Instead of a single point forecast, multiple forecasts are generated, by running a numerical weather prediction model with different initial conditions and/or perturbated model parameters. However, as ensemble systems suffer from model biases and dispersion errors, statistical post-processing is essential to make use of their full potential. A number of approaches, such as ensemble model output statistics (EMOS), have been proposed and successfully applied to statistically correct ensemble output at single locations. However, when forecasting composite quantities, such as minima, maxima, totals or averages, spatial dependency structures are of great importance. To take this into consideration, we merge EMOS with the geostatistical output perturbation (GOP) method, which models spatial correlations of temperature forecast errors via Gaussian random fields. In this way, we obtain probabilistic forecasts for whole spatial weather fields, taking the form of multivariate normal distributions, where the mean vector derives from the ensemble output according to EMOS, and GOP is used to specify the covariance matrix. The proposed spatial EMOS technique was applied to ensemble forecasts of surface temperature with the COSMO-DE ensemble prediction system, run by the German Weather Service. The post-processed forecasts captured the spatial structure of observed temperature fields and showed considerable improvement over the raw ensemble and alternative post-processing techniques.

370: 51

Testing for spatial structural change at unknown positions with application to inhomogeneous mineral subsoil [C12]

Matthias Arnold, Nils Raabe, Dominik Wied

TU Dortmund, Germany

We propose a fluctuation-type procedure for detecting breaks in spatial regions. While such tests are common in the context of time series, it is not a priori clear how to apply them to spatial data as there is no natural order of the observations. We demonstrate how this order can be constructed from a spatial autoregressive model. Once such an order is derived, standard time series results apply and break points can be consistently identified.

In the application, the resulting spatial fluctuation test distinguishes between hard and soft areas of inhomogeneous mineral subsoil. Both simulation evidence and the application on the subsoil data yield favorable results.

370: 52

Eine systematische Übersicht unter Verwendung von Meta-ANOVA für zwei Therapiemodi bei der konservativen Behandlung analer Inkontinenz [D16]

Reinhard Vonthein[1], Thilo Schwandner[2], Tankred Heimerl[2], Andreas Ziegler[1,3]

[1]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Germany; [2]Klinik für Allgemein-, Viszeral-, Thorax-, Transplantations- und Kinderchirurgie, Zentrum für minimalinvasive Chirurgie, Universitätsklinikum Gießen und Marburg GmbH, Standort Gießen, Germany; [3]Zentrum für Klinische Studien Lübeck, Universität zu Lübeck, Germany

Hintergrund: Unsere systematische Übersicht bestimmt die beste Biofeedback- (BF) oder Elektrostimulationsbehandlung (ES) bei analer Inkontinenz bei Erwachsenen und beurteilt den Grad der Evidenz gemäß GRADE. Berücksichtigt werden Wirksamkeit und Sicherheit. Die

Heterogenität der Interventionen erlaubt nur die Meta-Analyse einzelner Effekte mit einzelnen Studien. Die direkten Vergleiche könnten um indirekte bereichert werden.

Methode: Nach Cochrane-Handbuch wurden randomisierte Parallelgruppen-Studien selektiert (PROSPERO: CRD42011001334). Direkte und indirekte Vergleiche quantifiziert die zweifaktorielle Meta-Varianzanalyse mit den Faktoren BF und ES mittels R-Paket metafor.

Ergebnisse: BF und/oder ES wurde in 13 randomisierten Parallelgruppen-Studien untersucht, drei von hoher und vier von moderater Qualität. Keine Studie zeigte die Überlegenheit einer Kontrollbehandlung oder einer Monotherapie über eine BF+ES-Kombinationstherapie. Die Überlegenheit von BF+ES über Monotherapien ergaben verschiedene Studien. Amplitudenmodulierte Mittelfrequenz- (AM-MF) Stimulation kombiniert mit BF war sowohl Niederfrequenz-ES als auch alleinigem BF überlegen. Die quantitative Zusammenfassung per Meta-ANOVA zeigte übereinstimmend BF-RR (relatives Kontinenz-Risiko) 2 (95%CI 1.2-3.1) und Synergie-RR von BF+ES über BF hinaus 2 (95%CI 1.4-2.8), bei ES-RR 0.9 (95%CI 0.3–3.0). Diese Aussagen resultieren bei Analyse aller Studien und änderten sich kaum bei Betrachtung der Untergruppe der Studien mindestens moderater Qualität sowie bei Berücksichtigung der Behandlungsdauer oder der Kontinenzdauer, die als Erfolg gewertet wurde.

Sicherheitsaspekte wurden unvollständig berichtet, obwohl Niederfrequenz-ES unerwünschte Wirkungen wie Elektrolyse haben kann. Schwerwiegende unerwünschte Ereignisse wurden z.T. offensichtlich übersehen.

Schlussfolgerung: Es gibt ausreichende Evidenz für eine Wirksamkeit von AM-MF plus BF zur Behandlung analer Inkontinenz. Die Adjustierung der Meta-Analysen mit den Faktoren, die die Heterogenität bewirken, ermöglicht die quantitative Evidenzsynthese.

370: 53

Factors Affecting the Intubation Conditions Created by Mivacurium. A Meta-Analysis and Meta-Regression Analysis [D17]

Samson Mesfin

Johannes Gutenberg Universität Mainz, Germany

Intubation is the process of inserting a flexible tube anywhere in the human body. It is used in emergency medicine to help when a patient have difficulty in breathing. Mivacurium is a drug used to facilitate intubation. The objective of the paper was to identify the factors that affect the probability of excellent intubation condition of Mivacurium (EIC). A total of 1029 patients from 51 randomized clinical trials were studied using meta-analysis methods.In meta-analysis fixed effect and random effects models can be used to combine results from the different studies included in the meta-analysis. Fixed effect model assumes a common true effect underlying all the studies and all difference in the effect size is due to sampling error. In contrast, the random effects model allows the true effect size to vary between studies. Results from fixed effect and random effects meta-analysis showed lack of significant effect of mivacurium on the probability of EIC. Graphical and statistical methods for heterogeneity showed substantial heterogeneity in the effect size across the different studies included in the meta-analysis.To explore the sources of heterogeneity fixed effect and random effects meta-regression models were fitted. Results from the classical meta-regression models showed dose, average age, time to intubation(tstart) and age by tstart interaction are the variables that significantly affect EIC. The Bayesian approach on the other hand showed EIC varies with dose and age by tstart interaction. However, interpretation of the results should be done with caution since meta-analysis as an observational study is subject to confounding and ecological bias.

370: 54

Will the increasing interest in Individualized Health Care concepts question Randomized Controlled Clinical Trials? [D18]

Frank Krummenauer, Thomas Ostermann

University of Witten/Herdecke, Germany

The “conformity paradigm” of RCTs – the rigid standardisation of treatment and diagnostic circumstances in comparative evaluations – is considered as a basic precondition for the valid evaluation of therapeutic offers. The concept of Individualized Health Care, however, inverts this “conformity paradigm” and is directed into maximum possible individual adjustment of treatment circumstances. As a consequence, the evaluation and comparison of kernel therapeutic concepts under varying individual application settings has to be addressed by Individualized Health Care Researchers, who increasingly question the rigid concept of RCTs and their underlying conformity paradigm.

Nevertheless, from the methodological perspective, RCTs can still serve as a gold standard research tool in Individualized Health Care Research as well – but they may look different! The choice of primary endpoints will have to relate to the concept of individual cost effectiveness rather than to efficacy: Note that the introduction of Individualized Health Care will usually mean cost increases in the first place, but result in medium and long term cost savings due to increasing the number of patients positively responding to their respective treatment offer (and thereby reduce long term costs due to lacking rehabilitation of patients). As a consequence, RCTs in Individualized Health Care research have to quantify individual clinical outcome from the patient’s perspective and individual cost investment from the health care providers’ perspective simultaneously. The concept of individual cost effectiveness estimation alongside clinical pathways may therefore become essential in implementing of RCTs in Individualized Health Care research.

370: 55

A framework to assess the added value of subgroup analyses when the overall treatment effect (TE) was significant [D19]

Hong Sun, Werner Vach

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Background: Although statisticians have already noticed and discussed the problem, it is still not clear whether we should perform and how to perform such subgroup analyses when the overall TE is significant in both single clinical trials and systematic reviews. So a framework is needed to assess and compute the long term effect of different strategies to perform subgroup analysis.

Methods: A hierarchical structured simulation study is set up for evaluating the average post-study TE for patients in all studies (E) and fraction of patients with a negative TE in the successful studies (P). Seven decision rules are applied to optimistic, moderate and pessimistic scenarios. The particular case of binary outcomes with TE presented as risk difference is considered. Due to the fixed power of 90% and an assumed constant TE for each study, a constant sample size is required.

Discussion: We demonstrate that there are decision rules for subgroup analysis which decrease P and increase E simultaneously comparing to the situation of no subgroup analysis. These rules are much more liberal than the usual significance testing, since there is a high risk to decrease E using the latter.

370: 56

Erhöht gesteigerte Hautirritabilität das Risiko einer Kontaktallergie? Eine Analyse mit dem Hurdle Modell [D11]

Judith Schwitulla[1], Jochen Brasch[2], Wolfgang Uter[1]

[1]Universität Erlangen Nürnberg, Germany; [2]Universitätsklinikum Schleswig-Holstein, Germany

Natriumlaurylsulfat (NLS) ist ein Irritans und wird zusammen mit den wichtigsten Kontaktallergenen getestet, um eine gesteigerte Irritabilität der Haut zu diagnostizieren. Es soll helfen, die Interpretation von Testreaktionen auf Allergene zu unterstützen. Im Folgenden soll untersucht werden, ob die NLS-Reaktivität ein Risikofaktor für das Auftreten von allergischen Reaktionen ist. Verwendet werden Daten des Informationsverbundes dermatologischer Kliniken (IVDK, www.ivdk.org), die zwischen 2008 und 2011 (n=34864) erhoben wurden. Die Zielgröße gibt Aufschluss über die Anzahl allergischer Reaktionen pro Patient und besteht zu über 50% aus Nullen. Der Vergleich von Zählmodellen, anhand von AIC, log Likelihood und den vorhergesagten Anzahlen, soll eine geeignete Analyse ermöglichen, welche sowohl sogenannte ”excess zeros” als auch Überdispersion adäquat modellieren kann. Die besten Ergebnisse liefert das Hurdle Modell. Gleichzeitig bietet es, durch zwei separate Schätzer, eine sinnvolle Interpretation der Effekte. Auf Grund der getrennten Schätzung des ”Zero” und ”Count” Parts ermöglicht das Hurdle Modell eine differenzierte Quantifizierung der Assoziation sowohl bezüglich der Odds, überhaupt sensibilisiert zu sein, als auch hinsichtlich des Risikos zusätzlicher Allergien. Sowohl eine größere Chance für das Vorhandensein einer allergischen Reaktion bei irritabler Haut (NLS) als auch eine erhöhte Anzahl positiver Reaktionen werden bestätigt. Neben der Irritabilität sind auch andere Einflussgrößen signifikant mit der Zielgröße assoziiert.

370: 57

Feasibility of propensity score matching for observational studies on drug effects on pregnancy outcome [D12]

Evelin Wacker[1], Corinna Weber-Schöndorfer[1], Reinhard Meister[2]

[1]Institute for Clinical Teratology and Drug Risk Assessment in Pregnancy, Charité University Medicine Berlin, Germany; [2]Department of Mathematics, Beuth Hochschule für Technik – University of Applied Sciences, Berlin, Germany

For ethical reasons, studies on drug risk assessment in pregnancy are observational. In observational studies, baseline characteristics of treated subjects often differ from those of untreated subjects, leading to biased estimates of drug effects on pregnancy outcome. Therefore, the choice of an appropriate comparison cohort is essential, and selection bias has to be taken into account. Propensity score based analysis plays a prominent role among the approaches for counteracting bias. When the number of potential controls is limited, stratification or regression adjustment using the propensity score are the methods of choice. These approaches run into problems, when event rates to be compared are small, and adjustment may become unfeasible. This occurs regularly in studies on drug effects on pregnancy outcome with malformation rates in the magnitude of 3% to 5%. We took the chance of replacing regression adjustment by propensity score matching, using a large pool of potential controls. After matching, no further adjustment is necessary. However, when estimating the propensity score, imputation of missing covariates is mandatory.

We show how propensity score matching can be implemented, how multiple imputation of missing values is handled, and compare the matching approach with regression adjustment. We demonstrate our findings using a real study cohort (n=113) and a control pool of more than 25,000 prospectively ascertained pregnancies with complete data regarding the covariates. Propensity score estimation was based on few, but essential covariates (maternal age, previous parities, previous spontaneous abortions, smoking habits, alcohol consumption).

370: 58

Strukturgleichungsmodelle mit kontinuierlichen Variablen versus Modellen mit geordneten kategorialen Indikatoren – ein praxisorientierter Vergleich bei der Entwicklung klinischer Fragebögen [D3]

Rainer Leonhart

University of Freiburg, Germany

Bei der Entwicklung von Fragebögen im Bereich der klinischen Forschung wird immer von intervallskalierten, normalverteilten Prädiktoren für die latenten Variablen (Faktoren) ausgegangen. Die im Rahmen einer Konfirmatorischen Faktorenanalyse eingesetzten Schätzalgorithmen

basieren hierbei meist auf einer Maximum-Likelihood-Schätzung, welche nur bei erfüllten Voraussetzungen eine gute Modellpassung erreichen. Verteilungsfreie Verfahren führen hingegen oft zu schlechteren Passungsindizes und benötigen größere Stichproben. Wird hingegen davon ausgegangen, dass die Indikatoren „nur“ kategorial und geordnet vorliegen, so können mit einer weighted-least-square Schätzung über die kategorialen Prädiktoren die zugrundeliegenden latenten, normalverteilten Faktoren geschätzt werden. Es kann gezeigt werden, dass sich die Passung eines guten Modells (z.B. Gönner, Leonhart & Ecker, 2008) noch verbessern lässt. Auch liegen Beispiele anderer Messinstrumente aus der psychologisch-medizinischen Forschung vor, bei welchen die Annahme von kategorialen Indikatoren eine Parameterschätzung erst möglich macht. Vergleichende Berechnungen mit Hilfe von AMOS (Arbuckle, 2012) und MPlus (Muthén & Muthén, 2010) können den Vorteil dieses Vorgehens praxisnah belegen.

References:

[1] Arbuckle, J. L. (2012). Amos (Version 20.0) [Computer Program]. Chicago: SPSS.

[2] Gönner, S., Leonhart, R. & Ecker, W. (2008). The Obsessive-Compulsive Inventory-Revised (OCI-R): Validation of the German Version in a Sample of Patients with OCD, Anxiety Disorders, and Depressive Disorders. Journal of Anxiety Disorders, 22, 734-749.

370: 59

The use of R in a biomedical setting [D4]

Daniel Chamrad, Klaus Marquart, Carmen Theek

Protagen AG, Germany

Currently, SAS is the dominant software tool for statistical analyses in a biomedical environment. The R environment for statistical computing is well known for its excellent graphical capabilities, flexibility, novel methods and parallelization capabilities. However, despite of immense cost advantage compared to SAS, R is applied rarely in clinical trials, as its acceptance by authorities is questionable.

Here, we demonstrate the use of R in a qualified fashion, supporting some regulatory requirements for validated systems, especially focusing on reproducibility and standardization.

For this purpose we have developed a process in which statistical programming in R is carried out in a standardized way providing standard statistical reports in PDF-format. Work within projects follows a pre-defined structure of files and folders, being generated and maintained in automated manner. Standardized statistical analyses use R source code from an internal library, which has been developed following software development standards.

Based on the statistical analysis plan, a configuration file controls hierarchy, numbering, sections and headings of planned analyses. This file is automatically processed within the R environment, generating R-objects as well as CSV-files with results in tabular format and figures in a standard graphic format. Using the Sweave technology in the background all output is collected and transformed into a statistical report in PDF-format, automatically, including information about programmer, date and time stamp. As result, the cumbersome and error-prone workload of the statistical programmer to generate reports is minimized.

The contribution has been withdrawn.

Use of Bayesian approach to design and evaluation of bridging studies [D9]

Chin-Fu Hsiao[1], Wei-Hsiang Hsu[1], Hsien-Ming Hung[2]

[1]National Health Research Institutes, Taiwan, Republic of China; [2]FDA, United States of America

370: 61

Choosing non-inferiority margins to protect against degradation of treatment effects on average: A proposal for M3 [D5]

Primrose Beryl, Werner Vach

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Concern about risk of degradation of treatment effects in non-inferiority(NI) trials have been raised by researchers. To control this risk, NI margins have to be chosen appropriately. ICH-E10 as well as FDA provide guidelines for choice of margin based on: preservation of control effect and exclusion of clinically meaningful difference. However, it is possible that NI margins are set relatively wide to avoid time and cost of performing huge trials. We suggest a new additional hurdle for the choice of margin to protect against a degradation of treatment effect on an average.

The principle behind our proposal is: the probability of having declared a treatment with negative true effect as non-inferior in a successful NI trial is called likelihood of harm. If more than half of the NI trials are associated with a likelihood of harm above 50%, we have to fear degradation. Hence, assuming a pre-trial distribution of true effects characterised by 69% negative effects, the margin should be ideally chosen stringent enough to limit the likelihood of harm to 50%.

The optimal margin associated with lesser than 50% likelihood of harm was estimated using simulations. It is a fixed constant in case of continuous and survival outcomes while it is dependent on prevalence for binary outcomes. We propose that choice of margin in NI trials consider the protective margin M3, in addition to the preserve-effect margin, M1 and the clinical margin, M2. The proposed margin requires

reasonable sample sizes and can be implemented in future NI trials.

370: 62

Blinded sample size reestimation for recurrent event data with time trends [D2]

Simon Schneider[1], Heinz Schmidli[2], Tim Friede[1]

[1]University Medical Center Goettingen, Germany; [2]Statistical Methodology, Novartis Pharma AG, Basel, Switzerland

The use of an internal pilot study for blinded sample size reestimation allows to reduce uncertainty on the appropriate sample size compared to conventional fixed sample size designs. Recently blinded sample size reestimation procedures for recurrent event data were proposed and investigated (Friede and Schmidli, 2010a, b). These approaches assume treatment-specific constant event rates which might not always be appropriate as found in relapsing multiple sclerosis (RMS) (Nicholas et al, 2012). Based on a proportional intensity frailty model we propose methods for blinded sample size reestimation in situations where a time trend of the event rates is present. For the sample size planning and the final analysis standard negative binomial methods can be used, as long as the patient follow-up time is approximately equal in the treatment groups. To reestimate the sample size at interim, however, a full likelihood analysis is necessary. Operating characteristics such as rejection probalities and sample size distribution are evaluated in a simulation study motivated by a systematic review in RMS (Nicholas et al., 2012). The key factors affecting the operating characteristics are the study duration and the length of the recruitment period. The proposed procedure for blinded sample size reestimation controls the type I error rate and maintains the desired power against misspecifications of the nuisance parameters.

References

Friede T, Schmidli H. Statistics in Medicine 2010; 29: 1145–1156.

Friede T, Schmidli H. Methods of Information in Medicine 2010; 49: 618–624.

Nicholas R et al. Multiple Sclerosis Journal 2012; 18: 1290–1296.

370: 63

Vergleichende Analyse adaptiver Designs nach der α-spending- und der Inversen-Normal-Methode [D6]

Hans Christian Kuhl[1], Sandra Ligges[2], Andreas Faldum[2]

[1]Meda, Germany; [2]Institut für Biometrie und Klinische Forschung Münster, Germany

Klassische adaptive Designs basieren auf Kombinationsregeln bzw. Conditional Error Funktionen (CEFs, z.B. Inverse-Normal Methode, INM). Da hier die a-priori festzulegenden Gewichte im Verlauf einer Studie nicht mehr geändert werden dürfen, sind die zu verwendenden Teststatistiken nicht suffizient, wenn der Stichprobenumfang in einer Zwischenauswertung von der Planung abweicht. Die α-spending-Methode (ALSPM) hingegen bietet die Möglichkeit, aus einer Familie von CEFs eine CEF im Verlauf der Studie zu bestimmen: Hier werden die kritischen Schranken stufenweise nach dem tatsächlich aufgetretenen Stichprobenumfang und dem zu verbrauchenden Signifikanzniveau bestimmt.

In dieser Studie werden zweistufige adaptive Designs nach der ALSPM und nach der klassischen INM vergleichend untersucht. Dabei werden die Power und der mittlere zu erwartende Stichprobenumfang (ASN) der Designs für systematisch variierte Rekrutierungsraten berechnet und graphisch veranschaulicht. Neben einer Fallzahlrekalkulation zur Einhaltung einer vorgegebenen bedingten Power in der zweiten Stufe wird in weiteren Untersuchungen zur besseren Vergleichbarkeit der Designs die Fallzahl in der zweiten Stufe so gewählt, dass entweder die Power oder der ASN für beide Verfahren gleich ist und somit der jeweils andere Parameter direkt verglichen werden kann.

Aufgrund ihres flexibleren Charakters hinsichtlich der stufenweisen Gewichte ist die ALSPM der INM erwartungsgemäß dann überlegen, wenn die Anzahl rekrutierter Patienten der ersten Stufe deutlich nach unten vom geplanten Stichprobenumfang abweicht und in der zweiten Stufe nahezu der geplante Gesamtstichprobenumfang erreicht wird. Dies entspricht einer Situation in der Praxis, wenn die Auswertungen zu festen Terminen geplant sind und nach anfänglich schwacher Rekrutierungsrate weitere Prüfzentren nach der Zwischenauswertung hinzugenommen werden.

370: 64

Welche Konsequenz kann eine unverbundene Auswertung verbundener Daten haben? [D1]

Jessica Hirsch, Frank Krummenauer

Institut für Medizinische Biometrie und Epidemiologie, Universität Witten / Herdecke, Germany

Zielsetzung: Was ist die Konsequenz einer unverbundenen Auswertung verbundener Daten. Beispielsweise, wenn verbundene Daten mit dem unverbundenen statt verbundenen t-Test ausgewertet werden? Analog im nonparametrischen Fall, wenn symmetrisch-verteilte Daten mit dem unverbundenen statt verbundenen Wilcoxon-Test oder Vorzeichentest analysiert werden. Intuitiv zu erwarten ist, dass die Ergebnisse der Tests konservativ ausfallen-im Zweifelsfall ein wahrer Effekt nicht aufgedeckt wird.

Methode: Für diese Fragestellung wurden die genannten fünf etabliertesten Tests miteinander verglichen. Hierzu wurden bivariat-

normalverteilte Daten(Mittelwerte: μ1,μ2 aus [-1,1]) für verschiedene Fallzahlen (1≤n≤80), verschiedene Korrelation (ρ aus [-1,1]) und fest angenommener Varianz (σ1=σ2=1) simuliert. Entlang der resultierenden p-Werte werden die Tests miteinander numerisch und grafisch gegenüber gestellt.

Ergebnis: Die Simulationsanalyse zeigte, dass bei unverbundener Auswertung negativ-korrelierter Daten beide Tests (t-Test und Wilcoxon-Test) antikonservativ werden gegenüber den entsprechenden verbundenen Testversionen. Konkret ergab sich für negative Korrelation ρ=-0.4, Mittelwerte μ1=0.0, μ2=-0.4 bei einer Fallzahl von n=60 folgende p-Werte: p(t-Test)=0.030, p(Wilcoxon-Test)=0.039, p(verb.t-Test)=0.069, p(verb.Wilcoxon-Test)=0.079 und p(Vorzeichentest)=0.245. Grundsätzlich zeigte der Vorzeichentest bei moderat positiven bzw. negativen Korrelationen (-1≤ρ≤0.6) hochgradig konservative Ergebnisse gegenüber den anderen verbundenen Tests. Bei positiver Korrelation zeigten sich die unverbundenen Testversionen weniger konservativ gegenüber den verbundenen Testversionen als der Vorzeichentest.

Schlussfolgerung: In vielen tutoriellen Darstellungen gilt der Vorzeichentest als verteilungsunabhängige Alternative zum verbundenen t-Test bzw. Wilcoxon-Test. Aufgrund der Diskretheit seiner Teststatistik ist die Verwendung eines „suboptimalen"-Tests (unverbundener Wilcoxon-Test) bei positiver Korrelation (ρ≤0.6) empfehlenswert. Lediglich bei stark-positiver Korrelation legitimiert sich die Verwendung des Vorzeichentests. Bei negativer Korrelation zeigt sich die „falsche" Auswertungsstrategie mittels unverbundener Testversionen durchweg antikonservativ-insbesondere sollte bei symmetrisch-verteilten Daten mit dem verbundenen Wilcoxon-Test oder im parametrischen Fall verbundenen t-Test ausgewertet werden.

370: 65

Analyzing Health Care Costs of Multimorbid Patients: a Statistical Learning Approach [D10]

Matthias Eckardt[1,2,3]

[1]Universitätsklinikum Hamburg-Eppendorf, Germany; [2]Hamburg Center for Health Economics, Germany; [3]for the Multicare study group

Individuals with multiple chronic diseases (multimorbidity) consume a disproportionally large share of total health care resources. Classical health economic modelling techniques can hardly deal with the heterogeneity and high complexity of multimorbidity data caused by individual specific disease combinations and socio-economic covariates.

We therefore applied statistical learning techniques including generalized finite mixture models and conditional inference trees / forests to deal with the inherent heterogeneity and complexity of the data, and to detect components of patients suffering from multimorbidity as well as the most influential covariates within the different disease combinations with regard to health care costs. Multimorbidity was defined as co-occurrence of three or more conditions from a list of 29 chronic diseases reported by the patient's general practitioner. Data come from a multicentre prospective cohort study (MultiCare) of randomly selected multimorbid primary care patients aged 65 to 85 years in Germany. As a result certain most influential single diseases and multimorbidity components have been detected.

370: 66

Analyzing factorial diagnostic trials: The R-Package facROC [D8]

Katharina Lange, Edgar Brunner

Universitätsmedizin Göttingen, Germany

For some years diagnostic studies have developed into an important mainstay for clinical research. Just as other clinical trials diagnostic studies come up with data of various structures such as factorial designs. Furthermore in diagnostic trials we face another complex problem: there is a great number of possible diagnostic accuracy measurements (e.g. the area under the ROC-curve, sensitivity, specificity and predictive values).

In this talk we present an approach to analyze factorial diagnostic trials with paired and unpaired observations. Originally, this nonparametric method of analysis was developed for the evaluation of the area under the ROC-curve (AUC) [1], by applying the theory of the multivariate nonparametric Behrens-Fisher-Problem [2]. But this approach can be extended to sensitivity and specificity [3] and even to predictive values [4]. Therefore, a unified nonparametric approach for analyzing the most important diagnostic accuracy measurements in factorial set-ups is introduced. This method of analysis forms the basis of the R-Package facROC which allows evaluating the AUC, sensitivity and specificity in arbitrary factorial diagnostic trials with paired and unpaired observations.

References

[1] Lange K . (2008). Nichtparametrische Modelle für faktorielle Diagnosestudien, Georg-August- Universität Göttingen, Diplomarbeit

[2] Brunner E, Munzel U, Puri ML (2002). The multivariate nonparametric Behrens-Fisher problem. Journal of Statistical Planning and Inference 108:37–53

[3] Lange K, Brunner E (2012). Sensitivity, specificity and ROC-curves in multiple reader diagnostic trials — A unified, nonparametric approach. Statistical Methodology 9:490–500.

[4] Lange K . (2011). Nichtparametrische Analyse diagnostischer Gütemaße bei Clusterdaten, Georg-August-Universität Göttingen, Dissertation

370: 67

Mixed and marginal beta regression models to analyze longitudinal quality of life scores bounded at 0 and 1 [D7]

Matthias Hunger[1], Angela Döring[2], Rolf Holle[1]

[1]Helmholtz Zentrum München, Institute of Health Economics and Health Care Management, Germany; [2]Helmholtz Zentrum München, Institute of Epidemiology II, Germany

Background: Beta regression has been recommended for the analysis of variables observed on the unit interval and has recently been extended to the mixed model framework. However, there is only little empirical evidence about whether mixed beta regression (beta GLMM) has any benefits over linear mixed models (LMM) in the analysis of longitudinal data in practice. An important field of application are quality of life (QoL) scores, which are a crucial outcome parameter in clinical trials and health economics. Our objective was to compare beta regression models with the LMM to analyze longitudinal QoL in two typical empirical examples.

Methods: We used generic QoL data from a cohort study and disease-specific QoL data from a clinical trial. We described the conceptual differences between mixed and marginal beta (GEE) regression models and compared both models to the LMM in terms of overall fit and predictive accuracy.

Results: The beta distribution fitted the highly skewed empirical distribution of the QoL scores better than the normal distribution. Measures of overall fit suggested that the beta GLMM was superior to the LMM (AIC: -2723 vs. -2441 and -904.1 vs. -376.4). However, it underestimated the mean at the upper part of the distribution. There were nearly no differences between adjusted group mean scores from the beta GEE and the LMM.

Conclusion: Beta regression provided a slightly better overall fit to our data, however, if focus is on estimating group mean scores ant not individual predictions, results from beta models and LMM did not differ substantially.

370: 68

A Sequential Trial with a Binary Endpoint When the Number of Cases Fixed is Fixed by Design [D13]

Günter Heimann

Novartis Pharma AG, Switzerland

In this presentation I will discuss the design and analysis of a clinical study with a binary endpoint and very small success rates. Patients will be randomized to the two treatment groups until a fixed number of cases r is observed. In this case, the overall sample size Nis a random variables.

With small success rates, one usually performs Fisher’s exact test to compare the two groups. When the overall number of casesis fixed by design rather than the sample sizes, there is a simpler exact alternative. Given r and the overall sample size N, the test statistic hasa binomial distribution with parameters r and success rate q. Under H0, the parameter q equals the proportion of patients q randomized to the treatment group. Iderive the (conditional) distributions of all sufficient statistics under both H0 and alternatives, which allows me to simplify simulations, and to derive simple stopping rules using stochastic curtailment. I will also provide asymptotic power functions and show that the proposed test is close to the optimum.

370: 69

Modelling Survival following Bone Marrow Transplantation using Longitudinal Immune Measurements at Arbitrary Time Points [B4]

Markus Christian Elze[1], Stephan Klöss[2], Annekathrin Heinze[3], Oana Ciocarlie[4], Melanie Bremm[3], Ulrike Köhl[2], Jane Luise Hutton[1]

[1]Department of Statistics, University of Warwick, United Kingdom; [2]Integrated Research and Treatment Center Transplantation, Hannover Medical School, Germany; [3]Department of Medicine, Goethe-University Frankfurt, Germany; [4]Pediatric Department, Timisoara University of Medicine and Pharmacy, Romania

Severe leukaemias in paediatric patients may be treated by bone marrow transplantation. However, complications after transplantation may arise and early identification of high-risk patients is crucial for successful intervention. Thus, predictive models are needed for clinical decision making. These models need to incorporate longitudinal immune measurements taken at arbitrary time points. Also, measurement time points and frequency cannot be assumed to be independent of the survival process.

Joint modelling of longitudinal and time-to-event data is a natural approach to data of this nature. Implementations of joint modelling techniques are readily available. However, no consensus has been reached yet on how to best assess goodness of fit and predictive ability of such models.

For use in clinical practice simpler models may be preferred. A simple comparison of measurements at a single time point is sometimes found in biological publications to group patients into ‘high-risk’ and ‘low-risk’ categories. However, bias may be associated with such an approach. Bias can be reduced by using appropriate summary measurements and accelerated failure time models can be fitted. This has the advantage that many measures of the prognostic ability of survival models are readily available, e.g. explained variation (Stare et al.). These simpler models may be used to specify a joint model and joint modelling may in turn inform the simpler models.

A possible extension of joint modelling is the inclusion of a proportion cured component. Also, the incorporation of the measurement

frequency in the longitudinal model may inform on the risk assessment by the clinicians.

370: 70

A generalised method for the exact computation of partial inbreeding coefficients [F2]

Manfred Mayer, Reinsch Norbert

Leibniz Institute for Farm Animal Biology, Germany

Partial inbreeding coefficients assign portions of the overall inbreeding coefficient to founders. These are individuals of a certain generation, from which all alleles of descendants are derived. According to a broader understanding the term founder may include inbred and/or related individuals and even individuals from different generations. First it is demonstrated by example that the usual additive matrix method for computing partial inbreeding coefficients fails when founders are inbred. A more general method obtains correct results by using an appropriately weighted sum of inbreeding-contributions from all common ancestors in the pedigree. When contributions of common ancestors are adequately positioned as elements of a matrix C summation is over elements of columns. Finally it is shown how repeated recalculation of modified coancestry matrices can be employed in order to fill the matrix C with all the necessary contributions of common ancestors, which is an alternative to existing path-searching methods, which can be very intricate for complex pedigrees.

370: 71

Distribution of SNP effects in genomic selection [F1]

Angela Maria Bernal-Vasquez, Torben Schulz-Streeck, Andres Gordillo, Hans-Peter Piepho

Universität Hohenheim, Germany

In plant breeding, genomic selection (GS) is a powerful tool to reduce the time to obtain new hybrids. Based on marker information (i.e. SNP) and phenotypic data of a training population, GS allows predicting genetic estimated breeding values without preselecting procedures of markers. In general, genotypic values are regressed on SNP marker data. The methods to perform GS vary on the assumptions regarding the distribution of the SNP effects, relevant to understand the quantitative genetics underlying a trait.

Our objective was to investigate in empirical data, which distribution fit well to observed SNP effect estimates. We fitted symmetric distributions around zero to the marker effects for the kernel dry weight of a maize double haploid population.

The estimated marker effects can be written as the sum of a true additive effect plus and error. The latter can be assumed normal with zero mean and variance equal to the squared standard error of each SNP estimate, and the former can be modelled as the estimated effect minus the error.

The best fitting model was obtained by the t-distribution, followed by the SU distribution, then the normal and mixture of normal distributions and finally, the Lapace distribution. The algorithm of the mixture of two normal distributions converged to a single component.

Additionally, expected order statistics of SNP effects estimates were produced by simulation for each distribution and plotted against the observed estimates using Q-Q plots. The graphics agreed with the AICs, showing that the best fitting were accomplished by the t-distribution.

370: 72

Penalized scalar on function regression with interaction term as sensor signal evaluation technique [F6]

Karen Fuchs[1,2], Sonja Greven[1], Fabian Scheipl[1], Evamaria Stütz[2]

[1]LMU Munich, Germany; [2]Siemens AG Munich, Germany

In all fields of technology, the deployment of sensors is growing for years. While the demand on the explanatory power of a sensor increases, size as well as costs are expected to decrease. This necessitates the usage of sophisticated analysing methods to evaluate the obtained signals and maximize the information drawn from the respective results. We use a generalized regression model with functional covariates, including a functional interaction term, to predict scalar responses of interest. Coefficient functions are estimated using basis expansions, supposing a regular grid for the observation points of the covariates, and maximization of the respective log-likelihood, which is penalized to impose smoothness. The respective smoothing parameters for the penalties are estimated from the data, e.g. via generalized cross-validation. The add-on package mgcv for R can be used to do this. We test the goodness of estimation for our approach in simulations. The predictive power of our approach will be shown for some exemplary applications. The first uses cell chip sensor data as functional covariates, the second uses data of gas sensors. Both applications provide independent but concurrently measured signals. The main aim is to predict the concentration of either a chemical substance in the cell chip medium or else the concentration of certain gases in synthetic air.

370: 73

Low Frequency Oscillations in Electrical Transmission Systems [F5]

Dirk Surmann, Sebastian Krey, Uwe Ligges, Claus Weihs

Technische Universität Dortmund, Germany

The European electrical transmission system is operated increasingly close to its operational limits due to market integration, energy trading and the increased feed-in by renewable energies. For this reason it is necessary to analyse that part of energy that permanently oscillates through the electrical transmission system with a low frequency. These so called Low Frequency Oscillations are described and analysed within a smaller electrical system, the New England Test System, which guarantees a convenient handling. The analysis results in a new model which describes each node of the transmission system over partly excited mechanical harmonic oscillators. As in a true transmission system, the harmonic oscillators are connected over mechanical components according to the transmission lines of the electrical system. This model which bases on a system of differential equations, is compared with a well established and much more complex simulation system used at the Institute of Energy Systems, Energy Efficiency and Energy Economics of TU Dortmund University. The future aim is to optimise the parameters of the connected mechanical harmonic oscillators (mass, damping, stiffness and excitation) to get the same behaviour for the Low Frequency Oscillations as in the complex simulation. In addition stochastic elements will be implemented in the differential equations to analyse their impact on the Low Frequency Oscillations.

370: 74

Bayes factors for quantifying the evidence of nested effect models' structural features [F7]

Giusi Moffa, Rainer Spang

University of Regensburg, Germany

We study Bayes factors as a means to quantify evidence of certain structural features of nested effect models. For such models we find that the ratio between the marginal likelihoods of the highest scoring models with or without the feature of interest provides a good approximation to the Bayes factor and can therefore be used in practice. Moreover we show how the full network reconstruction becomes more quickly unstable as the noise level increases, with respect to feature recognition. Finally we apply the method to a real problem in colorectal cancer modelling in order to support the idea that contrary to the established literature belief WNT is actually involved in the pathway signalling.

370: 75

Bayesian Analysis of FRAP-Images [F8]

Martina Feilke, Volker J. Schmid

Ludwig-Maximilians-University Munich, Germany

FRAP (Fluorescence Recovery after Photobleaching) is a frequently used microscopy technique, with which the binding behaviour of molecules in various cellular compartments in vivo, in our case in cell nuclei, can be investigated. The molecules of interest are therefore fluorescently tagged, a part of the cell nucleus of the cell of interest is bleached, and the recovery of the bleached part of the nucleus is observed by taking images of the nucleus in predefined time intervals.

For quantitative analysis of FRAP data, we propose to use a Bayesian non-linear mixed effects model. The non-linear model is derived from a compartment model approach frequently used for FRAP data. We suggest a Bayesian approach allowing for mixed effect priors on the parameters of the regression model in order to analyse the recovery curves of several cell nuclei together. With that we get joint parameter estimates for all curves as well as a measure for the variation of the parameters due to the biological variation between cell nuclei.

We also suggest a method for spatial FRAP data. Here we are able to perform an analysis of the entire cell nucleus at the pixel level. Therefore, the concentration of the molecules of interest in the neighbouring pixels is incorporated into the fit of the recovery curve per pixel to account for diffusion processes. This is done using a Bayesian non-linear model derived from a set of differential equations. The results of both methods allow a deeper insight into the behaviour of protein binding in the cell nucleus.

370: 76

When biologists meet statisticians: A workshop concept to foster interdisciplinary team work [G3]

Benjamin Hofner[1], Lea Vaas[2,3], John-Philip Lawo[4], Tina Müller[5], Johannes Sikorski[3], Dirk Repsilber[6]

[1]FAU Erlangen-Nürnberg, Germany; [2]CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands; [3]Leibniz Institut DSMZ, Braunschweig, Germany; [4]CSL Behring, Marburg, Germany; [5]Bayer Pharma AG, Berlin, Germany; [6]Leibniz Institute for Farm Animal Biology, Genetics and Biometry, Dummerstorf, Germany

Today, life science and statistics have become essential partners. The need to plan complex, structured experiments, involving elaborated designs, and the need to analyse datasets in the era of systems biology and high throughput technologies has to build upon professional statistical expertise. On the other hand, conducting such analyses and also developing improved or new methods, also for novel kinds of data, has to build upon solid biological understanding and practise. However, the meeting of scientists of both fields is often hampered by a variety of communicative hurdles -- which are based on field-specific working languages and cultural differences.

As a step towards a better mutual understanding, we developed a workshop concept bringing together young experimental biologists and statisticians, to work as pairs and learn to value each others competences and practise interdisciplinary communication in a casual atmosphere. The first implementation of our concept was a cooperation of the German Region of the International Biometrical Society and the Leibnitz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (short: DSMZ), Braunschweig, Germany. We collected feedback in form of three questionnaires. Here, we will present the concept of the workshop, the results of the questionnaires and discuss how the workshop could be further improved.

370: 77

Grundvorlesung Biometrie und Epidemiologie - Welche Anforderungen stellen Lehrende der Veterinärmedizin an das Curriculum? [G1]

Ramona Zeimet[1], Marcus Doherr[2], Lothar Kreienbrock[1]

[1]IBEI, Stiftung Tierärztliche Hochschule Hannover, Deutschland; [2]VPHI, Vetsuisse Fakultät, Universität Bern, Schweiz

Angesichts des heterogenen mathematischen Grundkenntnisstands und Interesses bei den Studierenden der Tiermedizin, stellt die Lehre statistischer Konzepte eine (für das Fach Biometrie und Epidemiologie nicht ungewöhnliche) Herausforderung dar. Infolgedessen begann in den tiermedizinischen Bildungsstätten eine anhaltende Debatte über den besten Vermittlungszeitpunkt im Studienverlauf, konzeptionelle Vermittlungsansätze und die thematischen Inhalte der biostatistischen Grundvorlesung, wobei weltweit an den Fakultäten verschiedene Vermittlungs-Modelle etabliert wurden.

Während eines Treffens der Biostatistik-Dozierenden der österreichischen (1), deutschen (5) und schweizerischen (2) veterinärmedizinischen Bildungsstätten wurde die (zusätzliche) Frage aufgeworfen, ob durch die Themen, die in der Grundvorlesung Biometrie und Epidemiologie gelehrt werden, die Studierenden mit den entsprechenden, für andere Lehrveranstaltungen relevanten, Kenntnissen ausgestattet werden. Ende des Jahres 2011 wurde daher eine Online-Befragung entwickelt und an allen acht veterinärmedizinischen Bildungsstätten in Deutschland, Österreich und der Schweiz durchgeführt, um dieser Frage nachzugehen.

Über 250 Dozenten (70% zwischen 40 und 60 Jahren) aus den acht Bildungsstätten nahmen daran teil. Dabei gaben über 50% der Befragten an, dass statistische Konzepte generell früher gelehrt werden sollten (1. oder 2. Jahr). Für epidemiologische Konzepte war dieser Anteil geringer (27%). Außerdem gab es deutliche Unterschiede in der beurteilten Relevanz der 44 biometrischen und epidemiologischen Kernthemen, die in der Umfrage vorgestellt wurden. Hier war es für die Dozierenden der verschiedenen Fachbereiche möglich, durch eine Wertung von 0 (keine Relevanz für die eigene Vorlesung) bis 4 (sehr hohe Bedeutung) die individuelle Themenrelevanz zu beurteilen. Weitere Ergebnisse dieser Umfrage, als auch mögliche Einflüsse auf das Antwortverhalten der Befragten, sollen während der Konferenz vorgestellt werden.

370: 78

Regression Models as a Tool in Medical Research -- Lessons learned from writing a book [G4]

Werner Vach

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

In December 2012 my book "Regression Models as a Tool in Medical Research" was published by Chapman & Hall. Although written by a statistician, the book tries to focus on those issues of regression models, which are most relevant when applied in medical research by medical researchers. So when writing the book, my main task was to identify these topics. Two major sources was used for this task: My experience from working together with medical researcher for more than 20 years, and the aim to write a self-consistent book, avoiding any unmotivated recommendation and presenting all topics in a logical sequence. As a result, the book starts with explaining the classical regression models, the logistic regression model, and the Cox regression model with focus on the interpretation of regression parameters, confidence intervals, p-values and presentation of results. In addition, the typical problems in verbalizing the results from regression analyses are discussed. In the second part, some topics of relevance across different types of regression models like modelling non-linear relations are discussed. In preparing the book, it turned out that it is important to consider also the power of regression analyses in some depth, as this is the key issue to understand why questions like variable selection or choice of complexity of models is so difficult. The topic of the construction of risk scores and predictors is covered only briefly, as in my impression this field is still rather immature.

370: 79

Einsatz elektronischer Medien in der Biometrieausbildung. E-Klausur - technische Spielerei oder innovatives Arbeitsmittel [G2]

Friedrich-Wilhelm Röhl, Sebastan Baecke, Frank Hertel, Siegfried Kropf

Medizinische Fakultät Magdeburg, Germany

In regelmäßigen Abständen gibt es Gründe zur Überarbeitung der Ausbildungsabschnitte. Neben inhaltlichen Neuorientierungen gehört dazu der Einsatz technischer Hilfsmittel. Seit dem vergangenem Jahr werden unsere Abschlussprüfungen als E-Klausur mit dem Lernmanagementsystem MOODLE durchgeführt.

Bei den schriftlichen Klausuren lag der Schwerpunkt bei den statistischen Tests. Als Ergänzung gab es einzelne Alternativ- bzw. MC-Fragen. Für die E-Klausur haben wir uns bisher auf 3 Aufgabentypen beschränkt: MC-Fragen, Lückentext und Berechnung von Kenngrößen. Bei MOODLE können die Aufgaben in zufälliger Reihenfolge präsentiert und die Antworten bei MC-Fragen zufällig gemischt werden.

Durch den höheren Anteil an MC-Fragen, für die unsere ursprünglichen Aufgabenstellungen in mehrere Fragen aufgeteilt werden müssen, gibt es einen erheblichen Initialaufwand. Wenn ein größerer Aufgabenpool besteht, lässt sich eine neue Klausur schnell zusammenstellen. In unserem Fach haben wir darüber hinaus den Vorteil, dass Aufgaben durch leichte Änderungen der Zahlen variiert werden können. Der größte Vorteil dieser Art von Klausuren besteht aber in der schnellen und objektiven Bewertung der Klausuren.

Die Reaktionen bei Ankündigung, dass eine E-Klausur geschrieben wird, waren überwiegend positiv. An einer ersten Umfrage beteiligten sich 50 Studenten, bei der sich Befürworter und Skeptiker die Waage hielten. Nach der Präsentation erster Fragen stieg die Zahl der

Skeptiker und nach Durchführung der Klausur sank der Anteil der Befürworter auf 1/3. Die Notenverteilung der E-Klausur war mit denen der Papierklausuren aus den Vorjahren vergleichbar. Mit besserer Vorbereitung der Studenten auf die zu erwartenden Fragen und die Möglichkeit, eine Übungsklausur zu testen, erwarten wir in den geplanten aktuellen Umfragen eine deutliche Verschiebung der Meinung.

370: 80

On individual control treatments in designed genetical and agricultural experiments [F3]

Stanislaw Mejza, Iwona Mejza

Poznan University of Life Sciences, Poland

A common aim of genetical and agricultural experiments is to compare the test treatments with an individual control (standard) treatment. Two kinds of experiments are usually considered, namely;

1) nonreplicated genetical experiments performed at early stage breeding program and

2) the factorial experiments with crossed and nested structures of factors.

Our paper deals with first type of designs only.

By nonreplicated genetical experiment we mean one in which examined genotypes are replicated only once. The use of nonreplicated design is only one possible way to carry out an evaluation (inference) of the lines. Additionally, to control the real or potential heterogeneity of experimental units, control (check) plots are arranged in the trial. Some plots (check plots) with a control variety are usually placed between the plots with the lines. There are two main problems that have to be considered in the experiment, i.e. density of check plots and arrangement of them, random or systematic

In the article a response surface methodology is proposed for the analysis of nonreplicated breeding experiments. First, estimates of the yield response surface based on check plots as supporting points are obtained. Then the treatment (genotype, hybrid) effect is estimated as the difference between the observation obtained for the treatment and the response surface forecast. The consequences of density and arrangements of controls (check plots) on statistical inference using both simulation and uniformity trials are investigated.

370: 81

Effizienzmessung und Steuerung des RCS-Designs - Eine Betrachtung im Hinblick auf die Bundestagswahl 2013 [C9]

Christian Hoops, Tobias Michael

Ipsos Public Affairs

Das Rolling Cross-Section-Survey Design (Johnston & Brady 2002) setzt die Idee um, eine Querschnittsuntersuchung über einen vorab definierten Zeitraum derart zu verteilen, dass nicht nur alle Befragte zusammen einer Zufallsstichprobe der Grundgesamtheit entsprechen, sondern auch die Erhebung jedes einzelnen Tages (Schmitt-Beck et al. 2006). Die ideale Struktur eines RCS-Surveys sieht demnach vor, dass eine Tagesstichprobe eine identische Mischung aus mehreren Replikaten darstellt. Dadurch besitzen RCS-Daten ein hohes Analysepotential zur Auswertung von dynamischer Prozessen wie unvorhersehbaren Naturkatastrophen oder TV-Duellen in Wahlkampfphasen.

Der Vortrag beschreibt die Durchführung einer solchen RCS-Studie in der Praxis und die Möglichkeit diese anhand verschiedener Effizienz-Maße entsprechend zu steuern. Vorab werden mit einer Hazard-Regression die Effekte erklärender Variablen auf die Interview-Realisierung untersucht und die optimale Replikatsgröße geschätzt, die während der Feldzeit aufgrund der noch zu erwartenden Interviews reguliert werden kann.

References:

Johnston, R. und H. E. Brady, 2002: The Rolling Cross-Section Design. Electoral Studes 21(2): 283-295.

Schmitt-Beck, R. / T. Faas / C. Holst, 2006: Der Rolling Cross-Section Survey – ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung. ZUMA-Nachrichten 58: 13-49.

370: 82

Anforderungen an transparente dienstliche Beurteilungsverfahren, insbesondere im Wissenschaftsbereich [C10]

Udo Rempe

Universität Kiel, Germany

Nach Rechtsvorschriften und Rechtsprechung sind “Rang und Reihenfolge” von Eignung, Befähigung und fachlicher Leistung einer zu beurteilenden Person innerhalb einer hinreichend großen homogenen Gruppe von Mitkonkurrentinnen und Mitkonkurrenten um Gehaltsverbesserungen festzustellen. Damit die Aufgabe mit einem vertretbaren Aufwand erledigt werden kann, sind die Begutachtungsaufgaben der Beurteilenden auf das Reihen der Leistungen von höchstens vier Wissenschaftlerinnen oder Wissenschaftlern zu begrenzen. – Man kann eine Kohorte aus n fachnahen Personen zunächst mit Indizes [fn]i=0, 1, ..., n-1[/fn] durchnummerieren und dann

“per Los” eine zufällige Anordnung mit den Indizes [fn]k=0, 1, ..., n-1[/fn] erzeugen. Für jedes [fn]k[/fn] muss eine begutachtende Person mit dem Index [fn](k+3)\mod n[/fn] drei andere Personen mit den Indizes [fn](k+1)\mod n[/fn], [fn](k+2)\mod n[/fn] und [fn](k+4)\mod n[/fn] in einer Dreierliste anordnen. Jede zu beurteilende Person mit dem Index [fn](k+2) \mod n[/fn] wird von drei beurteilenden Personen mit den Indizes [fn](k+1)\mod n[/fn], [fn](k+3)\mod n[/fn] und [fn](k+4)\mod n[/fn] mit insgesamt sechs anderen Personen verglichen. So kann für jede Person ein Rang [fn]r[/fn] durch Einordnen in eine Zufallsstichprobe aus sechs Personen der Kohorte festgestellt werden. Zu klären sind Fragen wie jene, ob bei zwei Personen mit den so ermittelten Rängen [fn]r[/fn] und [fn]s[/fn] und [fn]r>s[/fn] die erste wirklich die bessere Leistung hat. Das Verfahren hat beispielsweise Bedeutung bei politischen Entscheidungen zur Korrektur der W-Besoldung.

370: 83

1000G imputations in a family-based genome-wide association study for genomic imprinting of early onset extreme obesity [A21]

Carolin Pütter[1], Ivonne Jarick[2], Marina Arendt[1], Johannes Hebebrand[3], Anke Hinney[3], André Scherag[1]

[1]Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Germany; [2]Institute of Medical Biometry and Epidemiology, Phillips-University of Marburg, Germany; [3]Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Germany

Introduction: Genome-wide association studies (GWAS) have had a dramatic impact on our understanding of genetic factors involved in common complex disorders such as obesity (Speliotes et al., 2010; Bradfield et al., 2012). Typically, GWAS focus on allelic effects of “single-nucleotide polymorphisms” (SNPs). Genetic phenomena like genetic imprinting in which alleles are expressed differentially conditional on parental origin are largely ignored. Family-based GWAS provide the option to assess this epigenetic process of genomic imprinting.

Materials and Methods: We extended our previous work using the 1000 Genomes data set for imputations to also access rare variants (http://www.1000genomes.org). We performed genome-wide imprinting analyses in 705 German nuclear families with extremely obese offspring originally genotyped by the Affymetrix Genome-Wide Human SNP Array 6.0 data. We analysed the imputed data set by stratified transmission-disequilibrium-tests (TDT) as standard software option in PLINK 1.07 (Purcell et al., 2007) and by the parental-asymmetry tests (PAT, Weinberg, 1999).

Results: Even upon extending the marker density using 1000 genomes imputation, we observed no genome-wide significant imprinting signal at a level α of 5x10-8. Comparing the results of the stratified TDT and the PAT the single SNP results were in some cases very different.

Discussion: Our analyses demonstrate the 1000 genomes imputation had a very limited effect on our results at least when compared to our previous findings. In contrast, the choice of the test statistic for addressing parent-of-origin effects in GWAS is important – at least for moderate sample sizes. We discuss our ongoing replication efforts.

370: 84

Analyse von Signalintensitätsabbildungen für die seltenen Varianten des Illumina HumanExome Beadchips [A22]

Arne Schillert[1], Jochen Kruppa[1], Jeanette Erdmann[2], Inke R. König[1], Andreas Ziegler[1]

[1]Universität zu Lübeck, Institut für Medizinische Biometrie und Statistik, Germany; [2]Medizinische Klinik II, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Germany

In der genetischen Epidemiologie werden Next-Generation-Sequencing-Verfahren zur Genotypbestimmung seltener Varianten eingesetzt. Mit der Entwicklung von Exom-Chips versprechen die Hersteller von Genotypisierungs-Mikroarrays diese Genotypisierung seltener Varianten mit Hilfe von Mikroarrays. Allerdings betritt man bei der Verwendung von Exom-Chips Neuland: Die Algorithmen verwenden pro Variante die Informationen aller Personen simultan, um die Genotypen zu bestimmen. Für seltene Varianten sind jedoch zwei der drei Genotypgruppen nur spärlich besetzt, was eine sorgfältige Qualitätssicherung erforderlich macht.

Wir haben die Genotypdaten für 15000 Personen aus einer Studie zur koronaren Herzkrankheit mit Hilfe des Illumina HumanExome BeadChip bestimmt. Anschließend wurden die Signalintensitätsdaten mit unserem R-Paket ATOMIC analysiert. Dabei wurde untersucht, welche der verwendeten Kennzahlen sich für seltene Varianten eignen, um die Qualität der Daten zu verbessern.

370: 85

Selecting B-spline regression models with shape constraints with the generalized order restricted information criterion [A7]

Anton Garder, Daniel Gerhard

Leibniz Universität Hannover, Germany

In many practical situations, when analysing a dependence of an explanatory variable on a response variable, it is essential to assume that the relationship of interest obeys certain shape constraints, since unconstrained models might be too flexible and give implausible results.

Model selection with information criteria can be used to identify a single best model, or to make inference based on weighted support from a set of competing models, incorporating model selection uncertainty into parameter estimates and estimates of precision. The generalized order restricted information criterion (GORIC) can be used to select one-way analysis of variance models when the population means are subject to a mixture of linear equality and inequality constraints (Kuiper et al. 2011).

Instead of assuming order constraints on population means in ANOVA settings, the GORIC approach can be extended to select from a set of B-spline regression models, specifying the shape restrictions by matrix constraints on the B-spline parameters.

The method is illustrated using the open source environment R with the add-on package goric (Gerhard and Kuiper 2012).

References:

Kuiper, Hoijtink, Silvapulle (2011). An Akaike-type Information Criterion for Model Selection Under Inequality Constraints. Biometrika, 98, 495-501.

Gerhard and Kuiper (2012). goric: Generalized Order-Restricted Information Criterion. R package version 0.0-6.

370: 86

Second order admissibility of confidence intervals [E8]

Yoshiji Takagi

Nara University of Education, Japan

We are interested in second order admissibility of confidence intervals in parameter estimation problems. Let us define a class of estimators for a single parameter [fn]\theta[/fn] as

[fn]{\cal C} = \{ \tilde{\theta}_{c} = \hat{\theta}_{ML} + c(\hat{\theta}_{ML})/n \} [/fn]

where [fn]\hat{\theta}_{ML}[/fn] is the MLE and [fn]c(\cdot)[/fn] is any continuous and differentiable function. In point estimation, Takagi (1999) derived a necessary and sufficient condition on the function [fn]c(\cdot)[/fn] for an estimator [fn]\tilde{\theta}_{c}[/fn] to be second order admissible in the class [fn]{\cal C}[/fn] under a given loss function. In this paper, we obtain a necessary and sufficient condition for a confidence interval constructed based on [fn]\tilde{\theta}_{c}[/fn] to be second order admissible in the class [fn]{\cal C}[/fn] under the criterion of the length.

In addition, by comparing the two conditions, we find a set of the estimators with second order admissibility in both point estimation and interval estimation.

370: 87

Relaxing monotonicity in the identification of local average treatment effects [B11]

Giovanni Mellace

University of St. Gallen, Switzerland

In heterogeneous treatment effect models with endogeneity, the identification of the local average treatment effect (LATE) typically relies on an instrument that satisfies two conditions: (i) joint independence of the potential post-instrument variables and the instrument and (ii) monotonicity of the treatment in the instrument, see Imbens and Angrist (1994). The contribution of this paper is to show that LATEs can still be identified and consistently estimated when relying on a condition that is strictly weaker than global monotonicity, while maintaining joint independence. We will refer to this condition as ``local monotonicity'' (LM). Crudely speaking and in contrast to (global) monotonicity, LM allows for the existence of both compliers and defiers, but requires that they do not occur at the same time at any support point of the outcome conditional on a particular treatment state. I.e., monotonicity is assumed to hold locally in subregions of the potential outcome distribution (and may switch the sign across subregions), rather than over the entire support. Under LM we can identify the LATEs on the (i) compliers (whose treatment reacts to the instrument in the intended way), (ii) defiers (who react counter-intuitively), and (iii) both populations jointly. Furthermore, (i) and (iii) coincides with standard LATE if monotonicity holds. We also present an application to the quarter of birth instrument of Angrist and Krueger (1991).

370: 88

Multiple treatment comparisons in presence of a treatment-covariate interaction [D14]

Frank Schaarschmidt

Leibniz Universität Hannover, Germany

When multiple treatments are analyzed together with a covariate, a treatment-covariate interaction complicates the interpretation of the treatment effects. The construction of simultaneous confidence bands for differences of the treatment specific regression lines allows a detailed interpretation of the effects, but the practical application of these methods is difficult because they are described as a collection of many special cases and the available implementations require additional programming or rely on non-standard or proprietary software. A flexible alternative is to compute simultaneous confidence intervals for multiple contrasts of the treatment effects over a grid of covariate values. This approach is freely available for the R software using the package multcomp. The underlying methodology is briefly reviewed and hints to a number of extensions are given. The application to examples in the linear and generalized linear model is illustrated.

The contribution has been withdrawn.

Nonparametric hazard rate estimation for relative survival models

Sophie Frantal[1], Werner Brannath[2]

[1]Medizinische Universität Wien, Austria; [2]Universität Bremen, Deutschland

411: Meth Börse

Time: Thursday, 21st Mar 2013: 8:00am - 8:30am · Location: KG I, HS 1199

411: 1

Bootstrap für abhängige Daten

Jens-Peter Kreiß

TU Braunschweig, Germany

tba

421: Statistics in Clinical and Preclinical Research -- Biomarker Development

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG II, Audimax
Session Chair: Annette Kopp-Schneider

421: 1

Combining mRNA and miRNA Measurements for Prediction of Clinical Endpoints

Moritz Gronbach[1], Harald Binder[1,3], Rolf Backofen[2], Sita Lange[2], Martin Schumacher[1]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Bioinformatics, University of Freiburg, Germany; [3]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center Johannes Gutenberg University Mainz, Germany

In the last decade, microRNA has been linked to several diseases. MicroRNA is believed to act by binding to mRNA, thereby affecting the stability of the mRNA or suppressing translation into proteins. While several estimated mappings from microRNA to mRNA (i.e., which microRNAs target which mRNAs) are available, there is no widely accepted general approach. We investigate how such estimated mappings can be used and evaluated to improving risk predictions for patients when linking simultaneous measurements of microRNA and mRNA expression to a clinical endpoint. We present an iterative method for incorporating mappings into multivariable risk prediction models for time-to-event endpoints. First, a set of mRNAs relevant to the endpoint is selected by a componentwise likelihood-based boosting algorithm. Next, the microRNA-mRNA mapping for these mRNAs is determined. The mapping information then is incorporated into a modified version of the boosting algorithm, and the multivariable model for the clinical endpoint is re-fitted. We evaluate the method on a prostate cancer data set containing both microRNA and mRNA measurements, where time to biochemical relapse is the endpoint. For the mapping estimates, we use a method based on correlation of expression profiles and two sequence-based methods. Comparison of prediction performance then allows to evaluate different mapping estimates. We also gain information about the importance of single microRNA-mRNA pairs, enabling further investigation of the underlying biological processes.

421: 2

Challenges for the Planning and Conduct of Diagnostic Studies With Molecular Biomarkers

Andreas Ziegler[1], Inke R. König[1], Peter Schulz-Knappe[2]

[1]Universität zu Lübeck, Germany; [2]Protagen AG, Dortmund, Germany

Biomarkers are of increasing importance for personalized medicine in many areas of application, such as diagnosis, prognosis, or the selection of targeted therapies. In many molecular biomarker studies, intensity values are obtained from large scale omics experiments. These intensity values, such as protein concentrations, are often compared between at least two groups of subjects to determine the diagnostic ability of the molecular biomarker. Various prospective or retrospective study designs are available for molecular biomarker studies, and the biomarker used may be univariate or even consist in a multimarker rule. In this work, several challenges are discussed for the planning and conduct of biomarker studies. The phases of diagnostic biomarker studies are closely related to levels of evidence in diagnosis, and they are therefore discussed upfront. Different study designs for molecular biomarker studies are discussed, and they primarily differ in the way subjects are selected. Using two systematic reviews from the literature, common sources of bias of molecular diagnostic studies are illustrated. The extreme selection of patients and controls and verification bias are specifically discussed. The pre-analytical and technical variability of biomarker measurements is usually expressed in terms of the coefficient of variation, and is of great importance for subsequent validation studies for molecular biomarkers. It is finally shown that the required sample size for biomarker validation quadratically increases with the coefficient of variation, and the effect is illustrated using real data from different laboratory technologies.

421: 3

Identification of predictive biomarkers in clinical drug development using maximally selected statistics and objective function maximization.

Natalja Strelkowa, Frank Fleischer

Boehringer Ingelheim Pharma GmbH & Co. KG, Germany

Cutpoints for continuous predictive biomarkers can be necessary and attractive decision rules used in clinical practice for treatment assignment of patients. Examples include patient classification according to biomarkers, e.g. tumor sizes or protein levels, in order to decide for more or less aggressive treatment.

Using the frameworks of generalized maximally selected statistics and objective function maximization we develop strategies for cutpoint determination of potentially predictive biomarkers based on the efficacy outcomes of early phase clinical trials. We analyze different study designs and patient recruitment scenarios with continuous and binary primary endpoints, where the treatment success is assumed to be dependent on the biomarker levels. The outcomes of our analysis provide estimates for reasonable sample sizes in clinical trials as well as the precision of estimated cutpoint values for biomarker levels at which the experimental treatment becomes more effective than the standard therapy. Analytical results are obtained for some special cases.Our strategies provide thereby an alternative for heuristic rules of thumb as for example 'dichotomization at median' or 'educated guess of an expert'.

421: 4

Sample Size Considerations for Safety Biomarker Development

Noemi Hummel, Carina Ittrich, Frank Fleischer

Boehringer Ingelheim Pharma GmbH & Co. KG, Germany

Safety biomarkers can play an important role to differentiate the underlying disease process from a potential iatrogenic adverse event (AE) during the development of a new compound in clinical research. Especially for severe diseases the treatment related benefits and risks have to be balanced. Safety biomarkers can be used to detect early and noninvasively patients at risk to develop certain AEs in the course of the study. Ideally biomarker-based dose reduction or treatment withdrawal criteria would allow reversal or even avoidance of the observed symptoms.

As outlined by Pepe (2003), biomarker development can be divided into five different phases. We provide proposals on how a safety biomarker development strategy can be incorporated into a specific Clinical Development.

For the practical implementation of clinical tests based on biomarkers, it is crucial to have a sufficiently high sample size to ensure a desired specificity and sensitivity of the test, given predefined significance level and power. We investigate the influence of the AE prevalence on the sample size and compare these sample sizes with those of usual Phase II/III trials. Moreover, we derive the minimal AE prevalence needed for usual Phase II/III sample sizes and estimate the power that can be reached with such sample sizes via simulations. In case of too low prevalence, references ranges of the biomarker derived specifically for the patient population of interest may be an alternative for safety monitoring during clinical development and in clinical practice. Finally we present a case study to illustrate our results.

422: Statistics in Agriculture and Ecology -- Modelling and Estimation

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3043
Session Chair: Arne Nothdurft

422: 1

Maximum Likelihood Estimation of Mark-Recapture-Recovery Models in the Presence of Continuous Covariates

Roland Langrock, Ruth King

University of St Andrews, United Kingdom

We consider mark-recapture-recovery (MRR) data of animals where the model parameters are a function of individual time-varying continuous covariates. For example, the survival probability of an individual may be a function of condition, with weight used as a proxy for this underlying condition. For time-varying individual covariates, the covariate value is unobserved if the corresponding individual is unobserved, in which case the survival probability cannot be evaluated. For continuous-valued covariates, the corresponding likelihood can only be expressed in the form of an integral that is analytically intractable, and, to date, no maximum likelihood approach that uses all the information in the data has been developed. We accomplish this task by formulating the MRR setting in a state-space framework and considering an approximate likelihood approach which essentially discretises the range of covariate values, reducing the integral to a summation. Assuming a first-order Markov structure for the covariate values, the likelihood can be efficiently calculated and maximized using standard techniques for hidden Markov models. We initially assess the approach using simulated data before applying to real data relating to Soay sheep.

422: 2

Meta-analysis: A need for well-defined usage in ecology and conservation biology

Daniela Vetter[1], Gerta Rücker[2], Ilse Storch[1]

[1]University of Freiburg, Germany; [2]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Meta-analysis is a powerful research summarization technique. The advantages of meta-analysis have also been recognized in the fields of ecology and conservation biology with the method becoming increasingly popular since the 1990s. “Meta-analysis”, however, is not well-defined in these fields, but is regularly confused with other summary analysis techniques, such as multiple regression methods, vote-counting or other quantitative analyses. We argue that this vague and inconsistent utilization of the term is problematic, because a meta-analysis typically provides scientifically rigorous results. We therefore advocate a consistent and well-defined usage of the term in our disciplines, based on the standardized definition applied in the medical sciences. We conducted a literature search for meta-analyses in the Web of Knowledge, determined steps that in our opinion are mandatory when performing meta-analysis and rated articles according to these steps. Of the 133 evaluated articles 25% did not fulfill any of the requisite steps for a meta-analysis. Our findings highlight the ambiguous and vague usage of the term “meta-analysis” in ecology and conservation biology and underline the importance of a consistent and clear definition. We conclude with recommendations on how the term should be applied in the future.

422: 3

Using a-priori information on networks and pathways for finding relations between molecular profiles and behavior factors from cattle

Nina Melzer[1], K. L. Graunke[1], J. Langbein[1], S. Saebo[2], R. Zimmer[3], G. Nürnberg[1], M. Schwerin[1], D. Repsilber[1]

[1]Leibniz Institute for Farm Animal Biology, Germany; [2]Norwegian University of Life Sciences, Norway; [3]Ludwig-Maximilian-University Munich, Germany

Is it possible to benefit from integrating a-priori knowledge about interactions of molecular variables for predicting behavioral scores from gene expression data? We seek relations between behavioral scores and gene expression data collected from cows of a crossing experiment of Holstein Friesian and Charolais cattle. Partial least squares regression offers a possibility to model and predict the assumed relationship, including opportunities to report important variables as well as to estimate prediction accuracy in a cross-validation approach. Recently L-PLS was proposed to enable the use of a-priori information about the molecular variables in the form of pathways or networks from public databases. We assess the feasibility of this approach in our dataset and expand it to a wide-ranging comparison of different sources of a-priori information, different choices regarding degree of belief in this information, and also towards optimizing a hyper-parameter specifying how much of the a-priori information should be incorporated to build the PLS-model. In our presentation, we will also show in more detail the special form of the underlying NIPALS algorithm in L-PLS, which enables this embedded integrative bioinformatics approach.

422: 4

Modelling the risk of deer-vehicle collisions depending on time of day and time of the year

Lisa Möst[1], Jörg Müller[2], Torsten Hothorn[1]

[1]Ludwig-Maximilians-Universität München, Germany; [2]Nationalparkverwaltung Bayerischer Wald, Departement Zoology, Grafenau, Germany

Deer-vehicle collisions are an enduring and serious threat to road safety. In Germany, approximately 240.000 vehicle collisions caused by

roe deer occur each year leading to a number of about 3.000 (partly heavy) injured people and costs of 520 Mio. € (www.dgv.de). Municipal authorities are constantly looking for new strategies to decrease the number of these collisions. Hothorn et al. (2012) model the risk of deer-vehicle collisions depending on environmental conditions and browsing intensities.

This talk focuses on the risk of deer-vehicle collisions depending on time of day and time of the year. By means of our model, we identify temporally limited high-risk times which can be used for alternative safety strategies like temporally bounded speed limits. The analysis is based on data from the Bavarian State Home Office comprising more than 74.000 vehicle collisions with roe deer, red deer and fallow deer in Bavaria in 2006 and 2009.

For estimating the number of deer-vehicle collisions depending on a smooth bivariate surface depending on time of day and time of the year, we use boosted array models (Currie et al. (2006), Bühlmann and Hothorn (2007) and Hothorn et al. (2010)). The suggested array models are highly efficient concerning computing time and memory footprint. Moreover, we include cyclic effects to get smooth cross-day estimates and the traffic density to account for confounding effects.

423: Research Synthesis and Meta Analysis – Network meta-analysis

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1010
Session Chair: Ralf Bender
Session Chair: Peter Schlattmann

423: 1

A graphical tool for locating inconsistency in network meta-analysis

Ulrike Krahn, Jochem König, Harald Binder

Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Johannes Gutenberg University Mainz, Germany

In a network meta-analysis, competing treatments are compared simultaneously by connecting evidence from different randomized trials and allowing for indirect comparisons. While estimation assumes a consistent network of treatment effects, there might be some direct comparisons that lead to incoherence and may have unfortunately side effects on the network estimates of others.

We investigate how such direct comparisons can be identified in the context of fixed-effects meta-analysis with known variances, using the inverse variance weighting method. Analysis can be equivalently performed in two stages: initially summarizing evidence of studies with same treatment arms, and secondly fitting a linear model to the aggregated treatment effects. We propose to explore a set of models that results from successively relaxing the assumption of consistency for studies with same treatment arms by adding new parameters to the model. The change in inconsistency can then be assessed by chi-square statistics. Based on this, we provide a graphical tool to highlight hot spots of inconsistency between specific direct evidence and to render transparent which direct comparison might have introduced it.

We illustrate our approach for published network analyses and discuss the ability to track down inconsistency to only one or at least a few direct comparison. This provides a starting point for more generally identifying conditions under which deviating direct estimates can be identified as a source of inconsistency.

423: 2

Decomposing Cochran's Q for network meta-analysis and visualizing and characterizing the flow of evidence into a mixed treatment comparison

Jochem König, Ulrike Krahn, Harald Binder

Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Johannes Gutenberg University Mainz, Germany

In network meta-analysis, direct evidence of studies comparing two or few treatments is pooled in order to get a broader data basis for decision support. The results are network based effect estimates for all pairs of treatments taking indirect evidence into account. The validity of the results depends on homogeneity and consistency assumptions that are sometimes hard to verify. For routine use we propose Cochran's Q for network meta-analysis to be decomposed into a heterogeneity and an inconsistency Q. To support the evaluation of validity, we propose a display of the flow of evidence into a mixed treatment comparison and introduce new measures to characterize it. Specifically, if the meta-analytic pooling is based on a linear fixed effects model, the mixed comparison of two treatments is a linear combination of direct effect estimates comparing these or other treatments. The linear coefficients can be seen as the generalization of weights known from classical meta-analysis. We summarize properties of these coefficients and display them as a weighted directed acyclic graph, representing the flow of evidence. Furthermore, measures are introduced that quantify the direct evidence proportion, the mean path length, and the minimal parallelism of mixed treatment comparisons. The graphical display and the measures are illustrated for two published network analyses. In these applications, the proposed methods are seen to render transparent the process of data pooling in mixed treatment comparisons. They can be expected to be more generally useful for guiding and facilitating the validity assessment in network meta-analysis.

423: 3

Prüfung der Konsistenzannahme bei Netzwerk Meta-Analysen: Ein Vergleich verschiedener Verfahren

Corinna Kiefer, Sibylle Sturtz, Ralf Bender

Institute for Quality and Efficiency in Healthcare (IQWiG), Köln, Germany

In der letzten Zeit gewinnen Verfahren für indirekte Vergleiche im Rahmen von systematischen Übersichten zunehmend an Bedeutung. Darunter fallen sowohl einfache Verfahren für einen adjustierten indirekten Vergleich, als auch komplexere Verfahren zur Kombination von direkter und indirekter Evidenz. Letztere werden in der Literatur als Mixed Treatment Comparison (MTC) Meta-Analyse, Multiple Treatment Meta-Analyse (MTM) oder Netzwerk Meta-Analyse bezeichnet (Song et al., BMJ 2011). Allen indirekten Vergleichen liegen drei zentrale Annahmen zugrunde: die Ähnlichkeitsannahme, die Homogenitätsannahme und die Konsistenzannahme (Hoaglin et al., Value Health 2011). Besonders die Konsistenzannahme und die damit verbundenen ungelösten methodischen Probleme bei der Prüfung dieser Annahme stehen häufig in der Kritik und führen zu einer allgemeinen Skepsis gegenüber der Anwendung und der Ergebnissicherheit von indirekten Vergleichen. Bei jeder Durchführung von indirekten Vergleichen sollte daher die sorgfältige Prüfung der Konsistenzannahme einen zentralen Aspekt bilden. Hierzu existiert momentan jedoch noch kein allgemein akzeptiertes Verfahren und auch kein etablierter Standard zur

Quantifizierung bedeutsamer Inkonsistenz (Li et al., BMC Medicine 2011). In der Literatur werden zwar verschiedene Verfahren vorgeschlagen (Dias et al., Stat Med 2010; Sturtz & Bender, Res Syn Meth 2012), über die Validität und Vergleichbarkeit der Ergebnisse der Verfahren ist allerdings nur wenig bekannt.

Im Rahmen dieses Vortrags werden verschiedene in der Literatur vorgeschlagene Verfahren zur Untersuchung von Inkonsistenz auf ein komplexes Netzwerk aus einem realen Datenbeispiel angewendet. Es wird gezeigt, dass die unterschiedlichen Verfahren zu differierenden Ergebnissen führen und dass allein aufgrund der unterschiedlichen Ansätze der Verfahren ein Vergleich erschwert wird.

423: 4

The use of two-way linear mixed models in multi-treatment meta-analysis

Hans-Peter Piepho[1], Emlyn Williams[2], Larry Madden[3]

[1]Universität Hohenheim, Germany; [2]Australian National University, Canberra, Australia; [3]Ohio State University, Wooster, USA

Meta-analysis summarizes the results of a series of trials. When more than two treatments are included in the trials and when the set of treatments tested differs between trials, the combination of results across trials requires some care. Several methods have been proposed for this purpose, which feature under different labels, such as network meta-analysis or mixed treatment comparisons. Two types of linear mixed model can be used for meta-analysis. The one expresses the expected outcome of treatments as contrast to a baseline treatment. The other uses a classical two-way linear predictor with main effects for treatment and trial. In this paper we compare both types of model and explore under which conditions they give equivalent results. We illustrate practical advantages of the two-way model using two published datasets. In particular, it is shown that between-trial heterogeneity as well as inconsistency between different types of trial is straightforward to account for.

References:

Piepho, H.P., Williams, E.R., Madden, L.V. (2012): The use of two-way mixed models in multi-treatment meta-analysis. Biometrics (online).

The contribution has been withdrawn.

Ranking von Behandlungsarmen bei multiplen Vergleichen in Meta-Analysen

Lorenz Uhlmann, Katrin Jensen, Meinhard Kieser

University of Heidelberg, Germany

424: Risk Analysis and Risk Prediction

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1098
Session Chair: Axel Benner

424: 1

On criteria for evaluating risk prediction models for public health applications

Ruth Pfeiffer

National Cancer Institute Bethseda, United States of America

We propose and study novel criteria to assess the usefulness of models that predict risk of disease incidence for screening and prevention, or the usefulness of prognostic models for management following disease diagnosis. The proportion of cases followed, PCF(p), is the proportion of individuals who will develop disease who are included in the proportion p of individuals in the population at highest risk. The proportion needed to follow-up, PNF(q), is the proportion of the general population at highest risk that one needs to follow in order that a proportion q of those destined to become cases will be followed. We also propose the integrated PCF, iPCF, and iPNF, the integrated PNF, obtained by integrating PCF and PNF over a range of values of q or p. Under the assumption that the risk model is well calibrated PCF, PNF, iPCF and iPNF can be estimated based on observed risks in a population alone. When the risk models are not well calibrated, PCF, PNF, iPCF and iPNF can be estimated consistently from case control data when the outcome prevalence in the population is known, and from cohort data, with baseline covariates and observed health outcomes. We study the efficiency of the various estimates and propose tests for comparing two risk models, both of which were evaluated in the same validation data.

424: 2

Radiation risk of leukaemia mortality for children and adolescents in the cohort of Japanese A-bomb survivors (1950-2003) from risk models selected for multi-model inference

Jan Christian Kaiser[1], Linda Walsh[2,3]

[1]Helmholtz Zentrum München, Oberschleissheim, Germany; [2]Bundesamt für Strahlenschutz, Oberschleissheim, Germany; [3]Faculty of Medical and Human Sciences, University of Manchester, United Kingdom

To derive estimates of the excess relative risk (ERR) from ionising radiation for leukaemia mortality in Japanese A-bomb survivors the technique of multi-model inference (MMI) is applied. MMI has been recently introduced in radiation epidemiology to improve the leukaemia risk analysis for children and adolescents. This topic is of public concern since it has been suggested that already bone marrow doses of 100 mSv and below enhance the radiation risk notably for children undergoing diagnostic CT scans or children living near nuclear power plants. MMI is based on several plausible models which fit the data about equally well. Risk estimates from single models are combined with weights which are based on the Akaike Information Criterion (AIC). Two approaches of model selection are discussed. Firstly, previously published models are refitted to updated cohort data and secondly, a set of candidate models is submitted to a rigorous selection protocol using likelihood ratio tests. Then the AIC-weighted means of model-specific risk estimates are compared. All previously published models apply a linear-quadratic dose response whereas the selection protocol preferred models with markedly weaker responses. Their AIC-weighted mean ERR is about three times lower at doses around 100 mSv for children and adolescents. The corresponding 95% confidence intervals predict no enhanced risk for doses below 300 mSv. Application of MMI in radiation epidemiology is recommended, since it reduces the bias from relying on a single model of choice for risk assessment and it provides a more comprehensive characterisation of uncertainties.

424: 3

Does calibration of clinical prediction rules assess bias?

Werner Vach

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Calibration is often thought to assess the bias of a clinical prediction rule. In particular, if the rule is based on a linear logistic model it is often assumed that an overestimation of all coefficients results in a calibration slope less than 1, and an underestimation in a slope larger than 1.

We investigated the relation of the bias and the residual variation of clinical prediction rules with the typical behaviour of calibration plots and calibration slopes, using some artificial examples.

We found that calibration is not only sensitive to the bias of the clinical prediction rule, but also to the residual variation. In some circumstances, the effects may cancel out, resulting in a misleading perfect calibration.

We conclude that poor calibration is a clear indication of limited usefulness of a clinical prediction rule. However, a perfect calibration should be interpreted with care, as this may happen even for a biased prediction rule.

425: Computational and Simulation Based Statistics

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1015
Session Chair: Uwe Ligges

425: 1

Implementing False Discovery Rate Procedures For Simulation-Based Tests With Bounded Risk

Georg Hahn, Axel Gandy

Imperial College London, United Kingdom

Consider multiple hypotheses to be tested for statistical significance using a procedure which controls the False Discovery Rate (FDR), e.g. the method by Benjamini-Hochberg. Instead of observing all p-values directly, we consider the case where they can only be computed by simulation. This occurs e.g. for bootstrap or permutation tests.

Naively, one could use an equal number of samples for the estimation of the p-value of each hypothesis and then apply the original FDR procedure. This technique is certainly not the most efficient one, nor does it give any guarantees on how the results relate to the FDR procedure applied to the true p-values.

This talk presents a more sophisticated approach that uses fewer samples for all those hypotheses which can already be classified with sufficient confidence and more samples for all those which are still unidentified. The algorithm is designed to give, with high probability, the same classification as the one based on the exact p-values and performs well in computer simulations.

425: 2

"Performed better on real data sets": an illusion?

Anne-Laure Boulesteix[1], Robert Hable[1,2], Sabine Lauer[1], Manuel Eugster[1,3]

[1]Ludwig-Maximilians-University, Munich, Germany; [2]University of Bayreuth, Germany; [3]University of Helsinki, Finland

In computational sciences, including computational statsitics, machine learning, and bioinformatics, most abstracts of articles presenting new supervised learning methods end with a sentence like "our method performed better than existing methods on real data sets'', e.g. in terms of error rate. However, these claims are often not based on proper statistical tests and, if such tests are performed (as usual in the machine learning literature), the tested hypothesis is not clearly defined and poor attention is devoted to the type I and type II error. In the present paper we aim to fill this gap by providing a proper statistical framework for hypothesis tests comparing the performance of supervised learning methods based on several real data sets with unknown underlying distribution. After giving a statistical interpretation of ad-hoc tests commonly performed by machine learning scientists, we devote special attention to power issues and suggest a simple method to determine the number of data sets to be included in a comparison study to reach an adequate power. Finally, we present simulation results and an exemplary benchmark study based on high-dimensional microarray data sets.

425: 3

Motivation and Collaboration in the R Project

Kurt Hornik

Wirtschaftsuniversität Wien, Austria

We present first results from the "CRAN Motivation Survey", an attempt to answer the question why people participate in or contribute to the R Project for Statistical Computing, and to identify key determinants for doing so. The questionnaire contains scales for work design, values and types of intrinsic and extrinsic motivation, which are used together with socio-demographic variables to model responses reflecting participation.

We also investigate structures of collaboration in R-related software projects hosted on source code repositories including R-Forge and github, analyzing in particular how the concentrations of co-author commit counts develop according to time and underlying technology.

Finally, we discuss possible implications of our findings for the future of R-related software development.

426: Statistical Surveillance

Time: Thursday, 21st Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3042
Session Chair: Sven Knoth

426: 1

Monitoring Medical Outcomes

Stefan Hans Steiner

University of Waterloo, Canada

It is desirable to monitor and compare surgical performance over time. However, surgical performance is subject to change over time due to a variety reasons including patient heterogeneity, learning, deteriorating skills due to age, etc. For instance, due to learning we expect inexperienced surgeons to improve their skills with practice. In this talk I propose a graphical method to monitor surgical performance based on an exponential weighted moving average (EWMA) of previous results that incorporates risk adjustment to account for patient heterogeneity. Using the EWMA allows a down weighting of the performance further in the past thereby focusing on recent outcomes. The monitoring chart is intuitive and clinically interpretable as it plots an estimate of the current failure rate for a “standard” patient and can include a measure of uncertainty. The standard patient is defined by the user. The method also allows a formal and informal comparison of risk-adjusted performance for different surgeons. The proposed method is illustrated with an example from cardiac surgery. We discuss the setup of the monitoring/comparison method and quantify the tradeoff between bias and variability in the estimate of surgical performance when selecting the EWMA smoothing constant.

426: 2

Robust online-surveillance of multivariate time series

Matthias Borowski, Roland Fried

TU Dortmund, Germany

The online- (also real-time or sequential) surveillance of multivariate data stream time series takes place in many fields. The time series are often not stationary but show enduring and suddenly changing trends and level shifts, and the coherences between the univariate components of the data stream may change over time. Due to the high frequency of measurement the data are often corrupted by a changing amount of noise and outliers. Furthermore, single observations as well as long stretches of data can be missing due to technical problems.

The online-surveillance of such complex time series is a challenging task calling for fast and robust (multivariate) methods that are based on weak assumptions and feasible in practice. We present new robust procedures for the online-detection of trend changes and jumps, for the online-estimation of current trends and for the online-surveillance of coherences between univariate series. All procedures assume an unknown locally linear signal carrying the relevant information and are therefore based on robust regression in moving windows. The combination of the new methods results in a multivariate procedure for comprehensive online-extraction and surveillance of relevant information given by the data stream.

426: 3

When to sell a stock? A sequential changepoint problem

Hans Rudolf Lerche, Dominik Stich

Universität Freiburg, Germany

The value of a stock raises up to an unknown timepoint at which it changes its trend. When to sell this stock? We model this problem with geometric Brownian motion and phrases it as an optimal stopping problem. The problem is solved using a filtering approach. Related aspects like expected delay are also dicussed.

430: Plenary session: Risk prediction (Gail, Bauer)

Time: Thursday, 21st Mar 2013: 10:40am - 12:00pm · Location: KG II, Audimax
Session Chair: Martin Schumacher

430: 1

Some Applications of Risk Modeling in Medicine and Public Health

Mitchell H. Gail

National Cancer Institute Bethesda, United States of America

We use the term `risk' to denote absolute risk, namely the probability that a particular event will occur in a defined time interval. Absolute risk depends on the age at the beginning of the time interval, on other covariates that may be present, on the duration of the interval, and on competing risks, that can diminish the absolute risk of the event of interest. Models of the absolute risk of developing a disease have potential applications in disease prevention. Such models can be used to assist in designing disease prevention trials, in assessing the potential population reductions in absolute risk from a prevention program, and in allocating preventive resources to those at highest risk. Absolute risk may also be useful in counseling individuals on prevention strategies, and, more formally, to assist in weighing the risks and benefits of a preventive intervention. Absolute risk is often used to assist in decision-making after a disease is diagnosed, because a patient with a dire prognosis may benefit from intensive and potentially dangerous treatments, whereas a patient with a favorable prognosis may not. We review some of these applications and some methods used to assess the usefulness of risk models.

430: 2

Coherent Modeling of the Risk in Mortality Projections: A Semi-Parametric Approach

Daniel Bauer

Georgia State University, United States of America

Unlike conventional approaches to modeling mortality risk, this paper considers uncertainties in mortality projections rather than uncertainties in realized mortality rates. Specifically, we conduct our analysis on a time series of mortality forecasts generated based on a rolling window of annual mortality data and a fixed forecasting methodology. We find that one to two factors are capable of capturing the great majority of all the variation in the data, and that the shape of these factors is very similar across different forecasting methodologies and populations. Relying on a semi-parametric representation that encompasses all self-consistent models with transitions parameterized by Normal-Distributed random vectors, we identify and estimate suitable models for the dynamic evolution of mortality forecasts. We further propose a model variant which guarantees a non-negative process of the spot force of mortality. Example applications show that resulting error estimates are considerably different from conventional approaches. Hence, the resulting models present coherent, parsimonious, and tractable workhorses in situations where the appraisal of risks within mortality projections is important.

442: Statistics in Practice

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1199
Session Chair: Stephanie Roll

442: 1

Meta-analysis of individual participant data from observational studies (Part 1)

Simon Thompson

Cambridge University, United Kingdom

Meta-analyses of multiple studies, for which individual participant data (IPD) are available, are becoming more common. The aim of this session is to update participants on statistical methods that can be used for such analyses, and the pitfalls to be avoided. The focus will be on observational studies rather than randomised trials. Available software will be discussed. The session will be organised as four 30-minute presentations, each allowing 10 minutes for discussion and questions. Specifically, the presentations will cover the following topics:

1. The basics of meta-analysis

To ensure participants have the relevant background, the session starts with a résumé of conventional statistical methods used in meta-analysis:

* The assumptions behind fixed-effect and random-effects analyses

* The interpretation of overall estimates, confidence and prediction intervals

* Quantifying and investigating heterogeneity

* The advantages of IPD over summary data

* One-step and two-step meta-analysis methods

2. Meta-analysis of observational epidemiological studies

Using examples from the Emerging Risk Factors Collaboration, a consortium of over 100 epidemiological studies that has shared individual data on over 1 million participants, the following issues will be addressed:

* Summarising data in each study in a consistent way

* Meta-analysis of log hazard ratios

* Adjusting for covariates and analysing interactions

* Investigation of heterogeneity: separating within- and between-study information

* Handling confounders that are completely missing in some studies

References:

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Bayesian methods for meta-analysis of causal relationships estimated using genetic instrumental variables. Statistics in Medicine 2010; 29: 1298-1311.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. International Journal of Epidemiology 2011; 40: 755-764.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Methods for meta-analysis of individual participant data from Mendelian randomisation studies with binary outcomes. Statistical Methods in Medical Research, online June 2012.

Emerging Risk Factors Collaboration. Major lipids, apolipoproteins and risk of vascular disease. Journal of the American Medical Association 2009; 302: 1993-2000.

Emerging Risk Factors Collaboration. C-reactive protein, fibrinogen, and cardiovascular disease prediction. New England Journal of Medicine 2012; 367: 1310-1320.

Jackson D, White IR; Fibrinogen Studies Collaboration. Systematically missing confounders in individual participant data meta-analysis of observational cohort studies. Statistics in Medicine 2009; 28: 1218-1237.

Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. Statistics in Medicine 2010; 29: 1282-1297.

Thompson SG, Kaptoge S, White IR, Wood AM, Perry PL, Danesh J; Emerging Risk Factors Collaboration. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. International Journal of Epidemiology 2010; 39: 1345-

1359.

Wood AM, White IR, Thompson SG; Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. Statistics in Medicine 2009; 28: 1067-1092.

443: Statistical Surveillance

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG III, HS 3042
Session Chair: Leonhard Held

443: 1

Surveillance Data Analysis using Varying Coefficient Models

Shireen Assaf

University of Padova, Italy

The analysis of health behaviour and risk surveillance data for describing changing effects of independent variables on health outcomes is not usually performed. The use of varying coefficient models with non-parametric techniques is a useful method which allows coefficients to vary with time using smooth functions. This allows for the study of the changing effects of health risks and behaviours on a health outcome in order to better inform policy interventions. Using the Italian PASSI health and behaviour surveillance data (which has collected monthly data since 2007), a time varying coefficient model is constructed for a smoking status binary outcome variable (current smoker or not). The model includes ten independent variables comprising socio-demographic and health risk and behavior variables. Polynomial spline estimation using B-spline basis functions was used for estimation which allows for different coefficients to have different degrees of smoothness. Selection of knots for the B-spline basis in the polynomial estimation method is conducted for each independent variable using AIC criterion. The final generalized time varying coefficient model was constructed in a stepwise method using chi-squared tests which allows for the exclusion of time varying coefficients if the coefficient is actually constant. The results have shown that the variables income, physical activity and alcohol consumption contain significant time varying coefficients. The Odds Ratio plots of these varying coefficients with time can be used for interpreting the changing characteristics of smokers.

443: 2

New outbreak detection developments in the R package ‘surveillance’

Maëlle Salmon, Michael Höhle

Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany

The R package ‘surveillance’ is free software for the visualization, modeling and monitoring of count data and categorical time series in public health surveillance (Höhle, 2007). It includes methods for the prospective detection of outbreaks in count data time series encouraging practical application in public health institutes, e.g. it is currently used in Finland, Sweden and France. Since the original developments, ‘surveillance’ has been improved on several points. Firstly, time series of routine collected surveillance data are now all represented using S4 objects of class sts allowing for consistent definition of attributes, and their use and transformation by methods. Secondly, surveillance encompasses several extensions of the widely used Farrington method (Farrington et al, 1996): the use of an offset variable in the regression to account for time changing population, and the improvements described in a recent paper (Noufaily et al, 2012) such as the use of the negative binomial distribution and a different reweighting procedure. Finally, surveillance is adapted to different types of monitoring situations: situations with little baseline information, for instance counts of emerging infectious organisms, can be analyzed with specific methods such as the EARS methods C1, C2 and C3 (Fricker et al, 2008) and binomial or categorical data can be studied with the categoricalCUSUM algorithm. Our work provides examples with accompanying R code for application to routine collected data on foodborne pathogens, for which such tools hold a great potential.

443: 3

Estimation of the HIV-prevalence and -incidence in Germany 2012

Matthias an der Heiden, Christian Kollan, Lieselotte Voß, Uli Marcus, Osamah Hamouda

Robert Koch-Institut, Germany

Background: Since the progression time between an HIV-infection and the diagnosis of HIV or AIDS could be quite large, the estimation of the HIV-prevalence and -- incidence is highly non-trivial even with good data about the HIV- and AIDS-diagnoses in the course of the epidemic.

Methods: We newly established a unified back calculation model based on data from the AIDS case register in the time period from 1978 to 1995 -- before an effective HIV treatment was established - and the reported HIV diagnoses from the German mandatory surveillance system since 1993. Missing data were imputed using multiple imputation methods. The unified model enabled us to estimate the complete course of the HIV-epidemic in the main transmission groups -- MSM, male/female IDU, male/female Hetero -- taking into account the age structure. In particular, the number of undiagnosed HIV-infections in these groups can be estimated. Assuming structural similarities in the western federal states including Berlin and in the eastern federal states, allowed us also to estimate the course of the epidemic in each of the federal states of Germany.

Results: For the World AIDS Day 2012 we estimated a total of 78.000 (confidence range 66.000-91.000) persons living with HIV (PLWH) und about 3400 (confidence range 3100-3600) newly infected in 2012. The number of undiagnosed PLWH was estimated as 14.000 (confidence range 13.000-15.000), which corresponds to 18% (confidence range 16%-20%) of all PLWH.

443: 4

Spatial power-law perspectives in modelling infectious disease spread

Sebastian Meyer, Leonhard Held

University of Zurich, Switzerland

Brockmann et al. (2006) recently showed that human travel behaviour can be described by a decreasing power-law function of distance [fn]x[/fn], [fn]f(x) ∝ x^{-d}[/fn], with positive decay parameter [fn]d[/fn]. We propose to incorporate such power-law weights in space-time models for infectious disease surveillance data. To this end, two previously established model classes are extended: the multivariate time series model for aggregated surveillance counts proposed by Held and Paul (2012), and the space-time point process model for individual point-referenced data proposed by Meyer et al. (2012). The decay parameter [fn]d[/fn] is estimated simultaneously with all other unknown parameters of the different formulations. The performance of the new approach is investigated for count data on influenza in Southern Germany (2001-2008) and individual cases of invasive meningococcal disease in Germany (2002-2008). In both applications, the power law substantially improves model fit and predictions. Implementation in the R package 'surveillance' allows to apply the approach in other settings.

References:

* Brockmann, D., Hufnagel, L., and Geisel, T. (2006). The scaling laws of human travel. Nature, 439(7075):462–465.

* Held, L. and Paul, M. (2012). Modeling seasonality in space-time infectious disease surveillance data. Biometrical Journal, 54(6):824–843.

* Meyer, S., Elias, J., and Höhle, M. (2012). A space-time conditional intensity model for invasive meningococcal disease occurrence. Biometrics, 68(2):607–616.

444: Statistics in Agriculture and Ecology -- Modelling spatial-temporal processes

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1098
Session Chair: Hans-Georg Schön

444: 1

Frequency domain approach for scale-dependent design and analysis of agricultural experiments

Ole Wendroth

University of Kentucky, United States of America

In many cases, experiments in agricultural field soils are difficult to analyze because of underlying inherent spatial heterogeneity. If in addition the state-variable of interest is characterized by a considerable local variability, identification of processes and spatial relationships between different variables can become extremely complex or even impossible. In this contribution, two case studies are introduced that are not based on random distribution of treatments and observations but on imposing treatments in cyclic or repetitive patterns. In one case study, the leaching behavior of a bromide tracer applied to the soil surface was investigated, and how the leaching depth is affected by different rainfall characteristics. In another case study, wheat yield response to mineral nitrogen fertilizer was investigated in a farmer's field that is known for topographic and soil textural differences. In both studies, frequency-domain analysis identified the variance components. An additive state-space model was then used to separate small-scale from large-scale variation. The large-scale components could be explained by distinct treatment differences in the first case, and with soil and topographic differences in the second case study using autoregressive state-space models. This approach is a promising way to conduct experiments even in heterogeneous landscapes.

444: 2

Modeling genotype by environment interaction in agricultural trials using latent curves

Sabine K. Schnabel[1,2], Fred A. van Eeuwijk[1,2], Paul H.C. Eilers[1,3]

[1]Biometris, Wageningen University and Research Centre, The Netherlands; [2]Centre for Biosystems Genomics, Wageningen, The Netherlands; [3]Department of Biostatistics, Erasmus Medical Center, Rotterdam, The Netherlands

In plant research, data for a set of genotypes is often collected in repeated field and greenhouse trials representing different environmental conditions. The phenotypic results are summarized in the form of two-way genotype by environment (GxE) tables of means. Models for GxE data should include one or more terms for genotype by environment interaction, with bilinear terms being a popular class of terms to include. Such bilinear terms can be interpreted as genotypes showing differential sensitivity to underlying latent environmental characterizations or gradients. In the case of just one bilinear term, arranging the environments along the latent gradient can be seen as a seriation problem. We propose to arrange the environments based on smooth latent curves. This is a generalization of existing popular approaches to model genotype by environment interaction. The result is an order of the environments and distances along an underlying environmental gradient. For all genotypes, genotype specific curves are produced whose characteristics can be associated with DNA variation in QTL analysis. The model can also be extended to accommodate missing data through a weighting scheme. The method will be illustrated with simulated data as well as empirical data from field trials.

444: 3

Spatio-temporal prediction of tree mortality based on long-term sample plots, climate change scenarios and parametric frailty modeling

Arne Nothdurft

FVA Baden-Württemberg, Germany

An approach is presented to predict the effects climate change may have on mortality of forest trees. Mortality is modeled using long-term observations from the Pan-European Programme for Intensive and Continuous Monitoring of Forest Ecosystems plots, retrospective climate data and frailty models having a parametric baseline hazard function. The linear predictor is modeled by B-spline regression techniques to allow for nonlinear cause-and-effect curves. Spatio-temporal predictions of tree mortality in the German state of Baden-Württemberg were derived in terms of unconditional hazard ratios and based on climate projection data. According to the model, marginal risk of tree death for 100 year old Norway spruce trees will be doubled until 2100.

References:

Nothdurft, A. 2012. Spatio-temporal prediction of tree mortality based on long-term sample plots, climate change scenarios and parametric frailty modeling. Forest Ecology and Management. In press. doi: 10.1016/j.foreco.2012.11.028.

445: Risk Analysis and Risk Prediction

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1010
Session Chair: Ernst Eberlein

445: 1

Dynamics of Correlation Risk

Elena Silyakova, Wolfgang Karl Härdle

Humboldt-Universität zu Berlin, Germany

Equity basket correlation is an important risk factor. It characterizes the strength of linear dependence between assets and thus measures the degree of portfolio diversification. It can be estimated both under the physical measure from return series, and under the risk neutral measure, from option prices. The difference between two estimates lay in the foundation of a so called "dispersion strategy". We study the performance of this strategy on the German market over the recent 2 years and propose several hedging schemes based on implied correlation (IC) forecasts. Modeling IC is a challenging task both in terms of computational burden and estimation error. First the number of correlation coefficients to be estimated would grow with the size of the basket. Second, since the IC is implied from option prices it is not constant over maturities and strikes. Finally, the IC changes over time. The dimensionality of the problem is reduced by an assumption that the correlation between all pairs of equities is constant (equicorrelation). The implied constant correlation is then approximated from implied volatilities of stocks and implied volatility of the basket. In such a way every day one recovers an IC surface (ICS). To analyze this structure and the dynamics of the ICS we employ a dynamic semiparametric factor model (DSFM), which yields a low dimensional representation. The ICS is studied in a functional principal components (FPCA) framework. The resulting factors are analyzed in a time series context.

445: 2

Nonparametric estimation of the value-at-risk of a stationary ergodic time series

Tina Felber

TU Darmstadt, Germany

Let [fn]\left(Z_t\right)_{t\in\Z}[/fn] be a stationary and ergodic sequence of square integrable real-valued random variables satisfying [fn]Z_i=m\left(Z_{-\infty}^{i-1}\right)+\sigma\left(Z_{-\infty}^{i-1}\right) \epsilon_i \quad (i\in\Z)[/fn] for some independent and identically distributed real-valued random variables [fn](\epsilon_i)_{i \in \Z}[/fn] with zero expectation and variance [fn]1[/fn]. Assume that [fn]\epsilon_0[/fn] has a density [fn]f[/fn] with respect to Lebesgue measure. Given [fn]n[/fn] observations of the time series [fn]\left(Z_t\right)_{t\in\Z}[/fn] we use methods from nonparametric regression in order to estimate [fn]m[/fn] and [fn]\sigma[/fn], and based on these estimates we construct an estimate [fn]f_n[/fn] of [fn]f[/fn]. We show that our estimate of [fn]f[/fn] is [fn]L_1[/fn]-consistent whenever the above model holds and we illustrate how our estimates can be used to construct a weakly universally consistent estimate of the value-at-risk of the time series. The finite sample size performance of our estimates is illustrated by simulated data and the estimate of the value-at-risk is compared with some semi-parametric methods.

445: 3

GARCH-extended models: theoretical properties and applications

Giles-Arnaud Nzouankeu Nana[1,2]

[1]Fraunhofer ITWM, Germany; [2]TU Kaiserslautern, Germany

Nowadays the GARCH-models are almost an inescapable instrument when modeling the volatility and especially in discrete time. There are many studies who proved the performances of these models and there are many suggestions to improve them. One type of improvement which is trend are the extension of the GARCH models by including additional (exogenous) variables for a better explanation and comprehension of the volatility movements. These are the so-called CARCHX models. While there are a lot of empirical studies in this direction there are few studies concerning the theoretical properties of these models. The aim of this work is to contribute to the filling of this gap. We establish sufficient conditions for some theoretical properties like stationarity, existence of moments, ergodicity, geometric ergodicity,... of a general class of GARCH-models containing exogenous parameters. For some of these properties we also show that the conditions are also necessary. From these properties we can derive many other classical statistical properties. we also provided some examples and applications to illustrate and give the importances of the properties.

445: 4

Pair copula constructions for mixed outcomes with application to comorbidity among the elderly

Claudia Czado, Jakob Stoeber

Technische Universität München, Germany

Flexible multivariate models for discrete and continuous outcomes are needed in many areas of applications. A pair copula construction for mixed outcomes is developed allowing for an expression of the log likelihood, which can be directly optimized in large data sets. This avoids the need for computational intensive MCMC algorithms. Estimation and model selection algorithms are presented involving the choice of the

factorization, the pair copula families and parameter estimation. Concepts are illustrated using data from the Second Longitudinal Study of Aging (LSOA II) allowing the joint prediction of chronic diseases such as heart disease, stroke, hypertension, diabetes, arthritis and body mass index. Adjustment of gender, age, income, education level and smoking status is facilitated through marginal regression models.

References:

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence Insurance: Mathematics and Economics 44 (2), 182-198.

A. Panagiotelis, C. Czado and H. Joe (2012), Pair copula constructions for multivariate discrete data Journal of the American Statistical Association. 107, 1063-1072.

446: Computational and Simulation Based Statistics

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG I, HS 1015
Session Chair: Kurt Hornik

446: 1

Extensions of SVD-based Kalman filtering and smoothing

Christian Heinze

Bielefeld University, Germany

The literature on linear time series methods is rich in reformulations of the basic Kalman filtering and smoothing equations. The available approaches either yield (conditional) state covariance or information matrices as is or in square root form. Covariance based approaches allow cheap time updates whereas information based approaches excel during the measurement step. Since matrix inverses and square roots are obtained in linear time once a singular value decomposition (SVD) is available, algorithms updating the latter allow to combine the merits of both classes in a natural way. The currently available formulation of SVD-based filtering and smoothing requires non-singular covariances. In addition, the currently available R implementation approaches likelihood evaluation by an additional SVD, an unnecessary but time consuming step. I show how the avoid the latter and dispose with the former. The resulting algorithms are expressible in terms of LAPACK routines shipped with a standard distribution of R. My talk explains both extensions. In addition, I compare the latter with the sequential processing covariance based approach both in terms of speed and accuracy by means of Monte Carlo simulations.

446: 2

Weighted Bootstrap and Clustered Standard Errors

Jan Sebastian Nimczik

University of Mannheim, Germany

In many applications the data are grouped into clusters while observations within clusters are prone to common (unobservable) influences. Hence, in linear regression models standard assumptions for consistency of OLS standard errors are violated. Moreover, in finite samples cluster-robust adjustments of standard errors are biased. This paper deals with (nonparametric) bootstrap methods that aim to improve on approximations from asymptotic theory with only few clusters. Recently, bootstrapping studentized test statistics has become increasingly popular due to the provision of asymptotic refinement for small samples (Cameron, Miller, and Gelbach, 2008). My paper adds to this literature by analyzing the performance of a more generalized weighted bootstrap approach. This method generates resamples by repeatedly attaching random weights to each cluster in the data set (Barbe and Bertail, 1995). This adresses two issues that arise with conventional bootstrap methods. First, the weighted bootstrap allows including cluster-invariant binary variables in the regression model while conventional bootstraps fail to estimate such models with few clusters. Second, small sample performance of standard error bootstraps can be improved through the choice of the weighting scheme. I conduct Monte-Carlo simulations to examine the performance of several bootstrap methods with few clusters using the rejection rate as an outcome criterion. Drawing weights from the uniform distribution the weighted bootstrap offers minimal but persistent improvement on conventional (pairs) bootstrap-se methods. I perform various sensitivity checks and apply the methods to different data generating processes including a design that replicates a potential application concerning the estimation of tuition effects in Germany.

446: 3

Comparison and Analysis of local classification methods

Julia Schiffner, Claus Weihs

TU Dortmund, Germany

Recently there has been a lot of interest in local classification methods in the data mining and machine learning community and thus, by now, several local versions of nearly all standard classification techniques like logistic regression, support vector machines and boosting are available.

However, when proposing a new local approach the authors usually demonstrate superior performance in comparison with the underlying (global) classification method. But in most cases only few data sets and a low number of competitors is considered.

To our knowledge there are no larger studies that compare the performance of different local versions and the underlying base method over a large number of classification problems.

We present an extensive simulation study where we assess the performance of local and global methods in a systematic way. For this purpose we artificially generate data in order to cover a wide range of possible classification problems. Moreover, we consider 6 base classification methods of different complexity (a simple classifier that always predicts the most frequent class, linear and quadratic discriminant analysis, logistic regression, support vector machines and neural networks) and investigate 4 different local versions of each method. As performance measure we use the misclassification rate. In addition, we consider the bias-variance decomposition of the error rate in order to gain deeper insight into the behavior of local learning algorithms.

The obtained misclassification rates as well as bias and variance are related to selected properties of the classification problems as well as the classification methods.

446: 4

The system for testing intelligence based on factor models and self-organizing feature maps

Anastasia Panfilova

Moscow State University of Psychology & Education, Russian Federation

Presented is a new testing system based on using the factor models and self-organizing feature maps as well as the method of filtering undesirable environment influence.

Testing process is described by the factor model with simplex structure, which represents the influences of genetics and environmental factors on the observed parameters -- the answers to the questions of the test subjects in one case and for the time, which is spent on responding to each test question to another. The Monte Carlo method is applied to get sufficient samples for training self-organizing feature maps, which are used to estimate model goodness-of-fit measures and, consequently, ability level. A prototype of the system is implemented using the Raven's Progressive Matrices (Advanced progressive matrices) -- an intelligence test of abstract reasoning. Elimination of environment influence results is performed by comparing the observed and predicted answers to the test tasks using the Kalman's filter, which is adapted to solve the problem. The testing procedure is optimized by reducing the number of tasks using the distribution of measures to belong to different ability levels after performing each test task provided the required level of conclusion reliability is obtained.

448: Design of Experiments

Time: Thursday, 21st Mar 2013: 1:00pm - 2:20pm · Location: KG III, HS 3043
Session Chair: Werner g. Müller

448: 1

Design of dose finding studies with multiple end points of different types

Valerii Fedorov

Quintiles, United States of America

Several correlated/mutually depended endpoints are observed in dose finding study. Typically one of them is claimed as a primary end point and the design (dose allocation and sample size) is driven by a single response model. I discuss the design problem with multiple end points (responses) which potentially may be of different nature. For instance, the efficacy end point may be continuous while the toxicity end point may be discrete. I emphasize the necessity to differentiate between the responses and utility functions. The responses are what we observe while the utilities are what should be reported or used in the decision making process. The criteria of optimality are often related to the precision of the estimated utilities.

448: 2

Designs for checking and estimating linear regression models when data are sampled sequentially over time

Wolfgang Bischoff

Katholische Universität Eichstätt-Ingolstadt, Germany

Linear regression models are checked usually by a lack-of-fit (LOF) test to be sure that the model is at least approximatively true. In many practical cases data can be sampled sequentially only. Such a situation can appear in industrial production when goods are produced one after the other. So as time goes by the mean may also depend on time, i.e. the mean is not only a function of the covariates but it may also be a function of time. This dependence on time is difficult to detect by a LOF test. Tests based on the residual partial sum process are more suitable for that. Therefore, in such a situation we suggest to apply a LOF test, for example the F-test, and a test based on the residual partial sum process, for example a test of Kolmogorov type. In case the linear regression model is not rejected by both tests least squares estimation can be used to estimate the parameters of the linear regression model. For the situation just discussed we are interested in a design with which we can efficiently run the two tests and estimate the linear model. Usually, classical optimal designs and LOF-optimal designs have not these properties.

448: 3

Fast algorithms to generate individualized designs for the mixed logit choice model

Martina Vandebroek, Marjolein Crabbe, Deniz Akinc

KU Leuven, Belgium

The mixed logit choice model has become the common standard to analyze transport behavior. Efficient design of the corresponding choice experiments is therefore indispensable to obtain precise knowledge of travelers' preferences. Accounting for the individual-specific coefficients in the model, this research advocates an individualized design approach. Individualized designs are sequentially generated for each person separately, using the answers from previous choice sets to select the next best set in a survey. In this way they are adapted to the specific preferences of an individual and therefore more efficient than an aggregate design approach. In order for individual sequential designs to be practicable, the speed of designing an additional choice set in an experiment is obviously a key issue. This paper introduces three design criteria used in optimal test design, based on Kullback-Leibler information, and compares them with the well-known D-efficiency criterion to obtain individually adapted choice designs for the mixed logit choice model. Being equally efficient to D-efficiency and at the same time much faster, the Kullback-Leibler criteria are well suited for the design of individualized choice experiments.

452: Statistics in Practice

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1199
Session Chair: Stephanie Roll

452: 1

Meta-analysis of individual participant data from observational studies (Part 2)

Simon Thompson

Cambridge University, United Kingdom

Meta-analyses of multiple studies, for which individual participant data (IPD) are available, are becoming more common. The aim of this session is to update participants on statistical methods that can be used for such analyses, and the pitfalls to be avoided. The focus will be on observational studies rather than randomised trials. Available software will be discussed. The session will be organised as four 30-minute presentations, each allowing 10 minutes for discussion and questions. Specifically, the presentations will cover the following topics:

3. Meta-analysis of risk prediction metrics

Whether risk factors aid the prediction of clinical events is often assessed by metrics such as Harrell's C-index, Royston-Sauerbrei's D-statistic, and Pencina's net reclassification index (NRI). Meta-analysis of these metrics across studies presents some problems in practice:

* Deriving C and D in individual studies

* Choice of weighting in a meta-analysis

* The change in C and D, and the NRI, on adding a new risk factor

* Heterogeneity between studies

* Examples from the Emerging Risk Factors Collaboration

4. Meta-analysis of Mendelian randomisation studies

Mendelian randomisation is the application of instrumental variable techniques using genetic variants to estimate the causal effect of a risk factor on an outcome from observational data. Large studies are typically needed to estimate causal effects precisely, and meta-analysis is often required. Topics to be addressed include:

* The principles of Mendelian randomisation, and instrumental variable analysis

* Two-stage analyses and weak instrument bias

* A one-stage analysis that minimizes weak instrument bias

* Dealing with partially missing risk factor data

* Example of C-reactive protein and coronary heart disease

References:

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Bayesian methods for meta-analysis of causal relationships estimated using genetic instrumental variables. Statistics in Medicine 2010; 29: 1298-1311.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. International Journal of Epidemiology 2011; 40: 755-764.

Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Methods for meta-analysis of individual participant data from Mendelian randomisation studies with binary outcomes. Statistical Methods in Medical Research, online June 2012.

Emerging Risk Factors Collaboration. Major lipids, apolipoproteins and risk of vascular disease. Journal of the American Medical Association 2009; 302: 1993-2000.

Emerging Risk Factors Collaboration. C-reactive protein, fibrinogen, and cardiovascular disease prediction. New England Journal of Medicine 2012; 367: 1310-1320.

Jackson D, White IR; Fibrinogen Studies Collaboration. Systematically missing confounders in individual participant data meta-analysis of observational cohort studies. Statistics in Medicine 2009; 28: 1218-1237.

Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. Statistics in Medicine 2010; 29: 1282-1297.

Thompson SG, Kaptoge S, White IR, Wood AM, Perry PL, Danesh J; Emerging Risk Factors Collaboration. Statistical methods for the time-

to-event analysis of individual participant data from multiple epidemiological studies. International Journal of Epidemiology 2010; 39: 1345-1359.

Wood AM, White IR, Thompson SG; Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. Statistics in Medicine 2009; 28: 1067-1092.

453: Statistics in Agriculture and Ecology -- Statistical Genetics

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1098
Session Chair: Hans-Peter Piepho

453: 1

Untersuchung der Korrelationsmatrix von Markereffekten in Kreuzungsversuchen

Sarah Bonk, Friedrich Teuscher, Norbert Reinsch

Leibniz-Institut für Nutztierbiologie, Institut für Genetik und Biometrie, Germany

Im Allgemeinen sind mehrere Loci (Genorte) für die Ausprägung eines quantitativen Merkmals verantwortlich. Dabei kann der Einfluss einzelner Loci sehr schwach sein, wodurch ihre Identifikation erschwert wird. Eine weitere Schwierigkeit bei üblichen Verfahren zur Genkartierung ist die Vernachlässigung der genetischen Kopplung zwischen den Markern. Da mit modernen Genotypisierungsverfahren immer mehr Marker gleichzeitig untersucht werden können (und diese somit immer dichter beieinander liegen), wird die genetische Kopplung wichtiger und sollte berücksichtigt werden.

In diesem Beitrag werden aktuelle Untersuchungen zur Korrelation von Markern in Kreuzungsversuchen vorgestellt, wobei der Fokus auf F2-Inzuchtlinienkreuzungen (mit additiven, dominanten und epistatischen Effekten) und Rückkreuzungen (mit Einzeleffekten und Interaktionen) liegt. Die für diese Situationen berechneten Korrelationsmatrizen können als Block-Diagonalmatrix dargestellt werden, was die Verarbeitung dieser Matrizen stark vereinfacht. Weiterhin wird gezeigt, dass die Inversen der Korrelationsmatrizen extrem spärlich besetzt sind, was die Berechnung noch weiter vereinfacht und z.B. den Speicheraufwand drastisch reduziert.

453: 2

Basis function for modelling marker effects in mapping experiments

Norbert Reinsch, Bonk Sarah, Wittenburg Dörte, Teuscher Friedrich

Leibniz-Institut für Nutztierbiologie, Institut für Genetik und Biometrie, Germany

Putative genetic effects of markers in mapping experiments are spatially correlated. As a result there are nearly linear dependencies between effects to be estimated. For F2- and backcross populations expected correlation matrices have been derived in full detail, including main effects and all their pairwise interactions (as presented in the contribution of Bonk et al.). These kernel matrices implicitely provide basis functions, allowing for a regularized estimation of marker effects. Results can be interpreted as effects of chromosome segments of variable size. Also, there is the possibility for constructing new approximate kernels, thereby modelling the effects of chromosome segments without defining them a priori. This keeps model dimensionality modest, while the number of genetic markers may grow almost unlimited.

453: 3

Systematischer Softwarevergleich für paarweise Alignments anhand von PRRSV-Sequenzen

Katharina Klose[1], Jan Böhmer[2], Cornelia Frömke[1], Ramona Zeimet[1], Katrin Strutzberg- Minder[2], Lothar Kreienbrock[1]

[1]IBEI, Stiftung Tierärztliche Hochschule Hannover, Germany; [2]Gesellschaft für Innovative Veterinärdiagnostik mbH (IVD GmbH), Hannover, Germany

Die Sequenzierung von Genomen ist seit vielen Jahren technisch möglich, diese wurde jedoch wegen des hohen Zeit- und Kostenaufwands im Regelfall nicht in großen Kollektiven angewendet. In den letzten Jahren hat die Anwendung von Sequenzierungstechniken aber stark zugenommen.

Ein häufiges Ziel der biometrischen Analysen ist der Vergleich multipler Sequenzen und die Erstellung phylogenetischer Bäume. Die Basis dafür stellt das paarweise Alignment dar. Für die entsprechende statistische Auswertung steht eine Vielzahl von Programmpaketen zur Verfügung, deren unterschiedliche Funktionalitäten es dem praktischen Anwender nicht ersichtlich machen, welche Software mit welcher Parametrisierung für seine individuelle Fragestellung optimal ist. Im Rahmen einer systematischen Prüfung wurden daher unterschiedliche Softwarepakete hinsichtlich ihrer Algorithmen und Parametrisierung verglichen. Neben der Gegenüberstellung der von den Programmen verwendeten Verfahren, lag der besondere Fokus auf dem Vergleich der Ergebnisse der paarweisen Alignments anhand von Nukleotid- und daraus abgeleiteten Aminosäuresequenzen des Porcinen Reproduktiven und Respiratorischen Syndrom Virus (PRRSV), einem Erreger, der für erhebliche gesundheitliche und wirtschaftliche Schäden in der deutschen Schweineproduktion sorgt.

Dabei werden die Unterschiede und Gemeinsamkeiten bezüglich der Funktion des paarweisen Alignments von frei verfügbaren Programmen (BioEdit, „Biostrings“ von Bioconductor etc.) dargestellt. Anhand von PRRSV-Sequenzen wurde der Einfluss verschiedener Substitutionsmatrizen und Gap-Kosten auf die Alignment Ergebnisse untersucht.

453: 4

Estimation of pooling weight variability in gene expression data

Henrik Rudolf, Gerd Nürnberg, Friedrich Teuscher, Norbert Reinsch

Leibniz Institute for farm animal Biology, Germany

Pooled samples may be used in experiments measuring gene expression data. The possibility to analyze observations with different pool sizes jointly was introduced theoretically in Rudolf et al. (2012). By adequately choosing the experimental design and statistical method, a splitting of data sets or an exclusion of pools with deviating sizes is no longer needed, enabling statistical inferences based on a larger sample size. In order to model the variability of pooled observations not only random biological effects have to be included in the analysis, but additionally a special kind of technical error induced by inaccuracies in blending aliquots of mRNA from different individuals into common pools. By means of simulation we verified the estimation of the pooling weight variability on the log-scale based on an approximation for variance of pools from Zhang et al. (2007) in two different simulation settings. Using a Reml-Log-Likelihood ratio test we found significance of the newly introduced technical error component for between 0.1% and 42% of the genes in 4 experimental data sets from tissues of mice, rats, bees and humans. Sets of genes, which are detected as differentially expressed show occasionally low coincidence between a model with and without the new variance component.

References:

Rudolf, H., Pricop-Jeckstadt, M. and Reinsch, N. (2012): Flexible pooling in gene expression profiles: design and statistical modeling of experiments for unbiased contrasts. (in review)

Zhang, W., Carriquiry, A., Nettleton, D. and Dekkers, J.C. (2007): Pooling mRNA in microarray experiments and its effect on power. Bioinformatics, 23(10), 1217--24.

454: Computational and Simulation Based Statistics

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1015
Session Chair: Achim Zeileis

454: 1

Bayesian generalized additive models for location scale and shape in zero-inflated count data

Nadja Klein[1], Thomas Kneib[1], Stefan Lang[2]

[1]Georg-August-University Göttingen, Germany; [2]University of Innsbruck, Austria

Generalized additive models for location scale and shape (GAMLSS) define a flexible, semi-parametric class of regression models in which the exponential family assumption for the response is relaxed. While ordinary regression analyses only the effects of covariates on the mean of a response, GAMLSS describes more complex parameters of the underlying distribution using semi-parametric additive predictors. In contrast to quantile regression, which avoids distributional assumptions, GAMLSS has the advantage that effects of covariates directly impact the parameters of the underlying distribution, and are therefore easier to interpret. Efficient Markov chain Monte Carlo techniques (MCMC) are an alternative to likelihood-based estimations, whereby the former one has the advantage that additional effects like spatial variations can easily be included. Especially constructing adequate proposal densities, which automatically deliver approximations of the full conditionals, plays a crucial role. Our focus is on zero-inflated Poisson (ZIP) and negative binomial (ZINB) distributions, whereas the classical regression models for count data are extended in the way that excess of zeros can be accounted for. The presented models will be applied to the number of forward citations of patents and to claim frequencies in car insurance.

454: 2

Structured Additive Regression Models: An R Interface to BayesX

Nikolaus Umlauf[1], Daniel Adler[2], Thomas Kneib[2], Stefan Lang[1], Achim Zeileis[1]

[1]Department of Statistics, University of Innsbruck, Austria; [2]Chair of Statistics, Georg-August-University, Germany

Structured additive regression (STAR) models provide a flexible framework for modeling possible nonlinear effects of covariates: They contain the well established frameworks of generalized linear models (GLM) and generalized additive models (GAM) as special cases but also allow a wider class of effects, e.g., for geographical or spatio-temporal data, allowing for specification of complex and realistic models. BayesX is a standalone software package providing software for fitting general classes of STAR models. Based on a comprehensive open-source regression toolbox written in C++, BayesX uses Bayesian inference for estimating STAR models based on either modern Markov chain Monte Carlo (MCMC) simulation techniques, or based on a mixed model representation of STAR models, or with stepwise regression techniques combining penalized least squares estimation with model selection. BayesX not only covers models for responses from univariate exponential families, but also models from less-standard regression situations such as models for multi-categorical responses with either ordered or unordered categories, continuous time survival data, or continuous time multi-state models. We present a new fully interactive R interface to BayesX implemented in the R add-on package R2BayesX. With the new package, STAR models can be conveniently specified using R's formula language (with some extended terms), fitted using the BayesX binary, represented in R with objects of suitable classes, and finally printed/summarized/plotted. This makes BayesX much more accessible to users familiar with R and adds extensive graphics capabilities for visualizing fitted STAR models.

454: 3

Evaluation of Imputation Methods for Hierarchical Datasets

Matthias Speidel

Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit, Germany

Missing values are a common problem in survey datasets. To avoid biases from nonresponse, often imputation methods are applied. However, most imputation methods do not account for hierarchical data structures (e.g. students within schools). Unfortunately there is only a limited number of hierarchical imputation methods and no consensus about an adequate method so far. In this talk I present a simulation study to evaluate the impact of different imputation methods under the assumption that the analysis model of interest is a linear mixed model. Four methods are compared: the available case analysis, an imputation model based on Bayesian linear regression, an imputation model including dummy variables to allow cluster specifc regression lines and a mixed effects imputation model that is congenial to the linear mixed model. In the simulation study missing values in the target variable are generated under the assumption of Missing At Random. To have simulated data as realistic as possible, the simulation parameters are based on an analysis of the Youth Risk Behavior Survey 2011. The results show that the available case analysis has the expected nonresponse bias and the Bayesian linear regression imputation distorts the hierarchical relationships in the data. The dummy variable imputation is unstable as the cluster sizes are too small for its application. The mixed effects imputation shows the best results although some parameters seem to be slightly biased.

454: 4

MCMC-Based Bayesian Model Selection for Regular Vine Copulas

Lutz Gruber

Technische Universität München, Germany

Vine copulas are increasingly popular models to capture complex multivariate dependency patterns. As a pair copula construction, vine copulas factorize the multivariate copula density into the product of (conditional) bivariate copula densities. These so-called pair copula building blocks may come from different copula families, or even mixtures thereof. A set of linked trees -- the ``vine'' -- organizes the decomposition of the densities by the level of conditioning.

Model selection procedures for vine copulas must select suitable pair copulas for each of the factor of the vine factorization along with the vine itself. The lack of any analytical approach to model selection necessitates the use of computational methods. Current model selection procedures proceed tree-by-tree. However, tree-by-tree approaches have been shown to find adequate models only under very restrictive assumptions.

We present a first-of-its-kind algorithm to estimate the vine and its pair copulas jointly over all trees. Our algorithm does not resort to the use of heuristics and is fully likelihood-based. We employ reversible jump Markov chain Monte Carlo to estimate the posterior distribution of regular vine copulas in a Bayesian context or to perform approximate maximum likelihood estimation using simulated annealing. We present simulation studies to show the viability of our approach and illustrates its outperformance over current model selection methods.

455: Design of Experiments

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG III, HS 3043
Session Chair: Valerii Fedorov

455: 1

Designing RNA-Seq Experiments for Multiple Groups

Johanna Mazur, Harald Binder

University Medical Center of the Johannes Gutenberg University Mainz, Germany

With next-generation sequencing techniques, such as RNA-Seq, one is able to obtain genome-wide measurements. For example, genes can be identified that are differentially expressed between groups of samples. To obtain reliable results, a large number of biological replicates in each group are needed, which is very expensive.

To optimize allocation of resources for measurement, we present a three-step approach where a small number of replicates is available initially for multiple groups of interest. Our goal is to perform experimental design such that further experiments are done primarily for those groups, where the most improvement in terms of identifying differentially expressed genes can be obtained. In a first step, we perform hierarchical clustering to obtain different group expression profiles. Thus, we are able to focus our further analysis on genes following a specific expression profile over the groups. In a second step, the groups for further experiments are identified such that the true positive rate is expected to be the largest after addition of the next replicate. Finally, we obtain the number of additional replicates that are needed to obtain a specific power for the remaining groups. This approach is illustrated on a real data RNA-Seq data from a patients with kidney renal clear cell carcinoma. Specifically, the results from a small subsample of patients are compared to the results from the whole dataset.

The results highlight how our approach can help to save costs for RNA-Seq experiments, while still obtaining reliable results on differentially expressed genes.

455: 2

Efficient Prediction Designs for Random Fields

Werner G. Müller[1], Luc Pronzato[2], Joao Rendas[2], Helmut Waldl[1]

[1]Johannes Kepler University Linz, Austria; [2]CNRS/Universite de Nice-Sophia Antipolis, France

For estimation and predictions of random fields it is increasingly acknowledged that the kriging variance may be a poor representative of true uncertainty. Experimental designs based on more elaborate criteria that are appropriate for empirical kriging are typically non-space-filling and very costly to determine. In this presentation, we investigate the possibility of using a compound criterion inspired by an equivalence theorem type relation to build designs quasi-optimal for the empirical kriging variance. Two algorithms are proposed, one relying on stochastic optimization to explicitly identify the Pareto front, while the second uses the surrogate criterion as local heuristic to chose the points at which the (costly) true Empirical Kriging variance is effectively computed. We illustrate the performance of the algorithms presented to both a simple simulated example and to a real oceanographic dataset

455: 3

Optimal Design for Count Data with Binary Predictors in Item Response Theory

Heinz Holling[1], Ulrike Graßhoff[2], Rainer Schwabe[2]

[1]University of Münster, Germany; [2]University of Magdeburg, Germany

The Rasch Poisson counts model allows for the analysis of mental speed which represents a basic component of human intelligence. An extended version of the Rasch Poisson counts model, which incorporates covariates in order to explain the difficulty, provides a means for modern rule-based item generation. After a short introduction into the extended Rasch Poisson counts model locally D-optimal calibration designs for this model will be developed. Therefore, the Rasch Poisson counts model is embedded in a particular generalized linear model. Finally, the robustness of the derived designs will be investigated.

456: Official Statistics, Survey Statistics and Demography

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG I, HS 1010
Session Chair: Peter Schmidt

456: 1

Realized and potential wages in Germany: Does gender matter?

René Söllner

Federal Statistical Office, Germany

Despite repeated commitments to foster gender equality, the wage gap between women and men does not vanish. Official statistics of the statistical office of the European Union (Eurostat) show that the German gender pay gap is among the highest in Europe. In order to analyze gender wage differentials, most studies usually apply the decomposition approach proposed by Oaxaca (1973) and Blinder (1973). The decomposition approach involves the estimation of separate wage functions for both gender groups using an ordinary least squares (OLS) model. The OLS approach has been criticized in the past. In order to analyze wage inequality we therefore employ the stochastic frontier approach (SFA) here in this paper. Like the OLS method, the SFA provides information about the determinants of wages, but additionally it allows evaluating the individual gap between realized and potential wages.

Based on the official Structure of Earnings Survey in Germany (Verdienststrukturerhebung), we show that females suffer from a discount in potential wages of about 9% in 2010. In 2006 the gender wage gap in maximum attainable wages was three percentage points larger. Since this wage gap cannot be explained by differences in human capital endowment it is an indicator for gender discrimination in wages. The empirical analysis further demonstrates that the difference between realized and potential wages is nearly the same across gender. Both males and females realize about 91% of their potential wages.

456: 2

Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder – Aktuelle Entwicklungen bei der Bereitstellung von Mikrodaten der amtlichen Statistik

Katharina Cramer[1], Heike Habla[2]

[1]Information und Technik Nordrhein-Westfalen, Germany; [2]Statistisches Bundesamt, Germany

Mit der Gründung der Forschungsdatenzentren (FDZ) der Statistischen Ämter des Bundes und der Länder in den Jahren 2001 und 2002 hat sich die informationelle Infrastruktur für die empirische Sozial- und Wirtschaftswissenschaft in Deutschland nachhaltig verändert. Über die FDZ, den Schnittstellen zwischen der Datenproduktion amtlicher Statistiken und der Wissenschaft, ist seither ein vereinfachter und transparenter Zugang zu diesen Mikrodaten möglich. Verschiedene Zugangswege – Scientific-Use-Files, Kontrollierte Datenfernverarbeitung, Gastwissenschaftsarbeitsplätze – ermöglichen es heute, dass nahezu das gesamte Informationspotential amtlicher Mikrodaten für eine international wettbewerbsfähige empirische Forschung und eine fundierte Politikberatung genutzt werden kann.

Gleichzeitig hat sich durch die Arbeiten in den FDZ der Kenntnisstand im Bereich der Anonymisierung von Mikrodaten deutlich verbessert. Bis vor einigen Jahren galten insbesondere die wirtschaftsstatistischen Mikrodaten als nicht anonymisierbar. Mittlerweile wurden jedoch auch für diese Statistiken Datenmaterialien entwickelt, die außerhalb der geschützten Räume der amtlichen Statistik von der Wissenschaft genutzt werden dürfen.

Insgesamt ist festzustellen, dass der Bedarf an amtlichen Mikrodaten nach wie vor wächst. Nachdem zunächst haushalts- und dann wirtschaftsstatistische Querschnittsdaten nachgefragt wurden, gilt das Hauptinteresse inzwischen den Paneldaten sowie Produkten, die durch die Verknüpfung verschiedener Statistiken entstanden sind. In Zukunft werden die FDZ ihren Nutzerinnen und Nutzern auch Einzeldaten des Zensus 2011 über ihre bestehende Infrastruktur zur Verfügung stellen.

456: 3

Gravity Sampling - Eine Anwendung am Beispiel der nationalen Minderheiten im deutsch-dänischen Grenzland

Christian Hoops[1], Adrian Schaefer-Rolffs[2], Prof. Dr. Kai-Uwe Schnapp[2]

[1]Ipsos Public Affairs; [2]Universität Hamburg

Es existiert eine Vielzahl von Minderheitengruppen auf der Welt. Da die Zugehörigkeit von Behördenseite nicht nachprüfbar ist, ist die Befragung dieser Gruppen sehr problematisch. Bei derartigen Erhebungen mit einer sehr geringen Inzidenz wurde bisher keine Möglichkeit gefunden, auf Basis eines Random-Digit-Dialings zu erheben (vgl. Fernandez et al. 2006: 74). Deshalb werden Zugehörige solcher Bevölkerungsgruppen mit Schneeball-Techniken oder onomastischen Voruntersuchungen ausgewählt, wodurch die Ergebnisse methodenbedingt erheblichen Verzerrungen unterliegen.

Der Vortrag zeigt eine Vorgehensweise, wie gravitationsanalytisch Stichproben gezogen werden können, so dass eine Zufallstelefonbefragung ermöglicht wird und größere Repräsentativität entsteht. Dabei werden die probabilistischen Gravitationsmodelle nach Huff (1964) bzw. Nakanishi & Cooper (1974) angewendet und anhand der Distanz zur deutsch-dänischen Grenze, der Wahlergebnisse bzw. der Dichten nationaltypischer Familiennamen und Organisationen eine Wahrscheinlichkeit berechnet, sich der Minderheit zugehörig zu

fühlen. Eine Projektion auf die gewünschte Stichprobengröße ermittelt die zu erhebenden Fallzahlen pro Ort.

Anhand einer Studie in 32 dänischen Städten wird gezeigt, dass der MCI-Index eine relativ große Korrelation mit der tatsächlichen Verteilung der Minderheit in Dänemark besitzt und sich die Technik im Spezialfall von Geo-Stichproben durchaus anwenden lässt.

Literatur:

Huff, D. L., 1964: Defining and Estimating a Trade Area. Journal of Marketing 28(3): 34-38.

Fernandez, O. S. / N. Rother / M. Braun, 2006: Stichprobenziehung für Migrantenpopulationen in fünf Ländern – Eine Darstellung des methodischen Vorgehens im PIONEUR-Projekt, ZUMA-Nachrichten 59: 72-88.

Nakanishi, M. und L. G. Cooper, 1974: Parameter Estimation for a multiplicative competitive interaction model – least squares approach. Journal of Marketing Research 11(3): 303-311.

456: 4

Die Auswertungsdatenbank zum Zensus 2011

Karsten Lamla

Statistisches Landesamt Baden-Württemberg, Germany

Der Zensus 2011 dient der Ermittlung einer neuen Basis für die Fortschreibung der amtlichen Einwohnerzahlen. Zusätzlich werden bundesweit fachlich detaillierte Informationen über die Bevölkerung und den Gebäude- und Wohnungsbestand gewonnen. Diese Strukturdaten bilden die Grundlage für eine Vielzahl politischer Entscheidungen. Darüber hinaus gibt es für die Zensusdaten ein breites Spektrum weiterer Verwendungsmöglichkeiten. Neben den Entscheidungsträgern aus Politik und Verwaltung gehören deshalb auch interessierte Bürgerinnen und Bürger ebenso wie Vertreter aus Medien, Wissenschaft und Wirtschaft zur Zielgruppe. Das breite Spektrum an Nutzern verlangt ein vielschichtiges Datenangebot. Neben klassischen Printveröffentlichungen wird es deshalb eine Zensus-Auswertungsdatenbank geben, welche frei im Internet zugänglich sein wird. Die Auswertungsdatenbank umfasst zum einen ein Angebot für Datennutzer mit wenig Statistik-Erfahrung oder nur grundlegende Datenbedürfnissen. Der Schwerpunkt liegt hierbei auf dem Angebot von vorgefertigten Tabellen und Grafiken. Darüber hinaus besteht für Nutzer mit einem Interesse an fachlichen Details und mit methodischen Kenntnissen die Möglichkeit, Tabellen selbst zusammenzustellen. Aus dem umfangreichen Themenkatalog des Zensus können Merkmale individuell und flexibel kombiniert, mit Grafiken visualisiert und in verschiedenen Datenformaten abgerufen werden. Dabei ist selbstverständlich sichergestellt, dass die Regeln der statistischen Geheimhaltung berücksichtigt sind und keine Angaben über einzelne Personen an die Öffentlichkeit gelangen können.

457: Industrial Statistics

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG III, HS 3042
Session Chair: Christian Weiß

457: 1

Exponential Smoothing with Covariates

Kristina Lurz, Rainer Göb

University of Würzburg, Germany

For long, exponential smoothing (ES) was considered rather a heuristic forecasting technique without a precise model foundation which guarantees optimality. In 1997, Ord et al. provided the basis for a solid model framework by formulating the single source of error (SSOE) state space scheme, which allowed to demonstrate the optimality of the classical ES predictors. By introducing an additive term depending linearly on exogeneous variables into the observation equation of the SSOE model, Wang (2006) developed the method of exponential smoothing with covariates (ESCov). The present study considers extensions and variants of ESCov in the following respects: i) the way of including covariates, ii) multiple seasonalities, iii) prediction intervals. The models are illustrated in industrial applications, in particular from the area of electrical load forecasting.

457: 2

Clustering of Electricity Networks based on Network Topology and Electrical Stability

Sebastian Krey, Sebastian Brato, Uwe Ligges, Claus Weihs, Jürgen Götze

TU Dortmund, Germany

In modern interconnected powersystems (often comprising whole continents) the control of the power flows is a very complex task. The ongoing replacement of classical power plants with renewable energy provided by highly distributed and relatively small power stations introduces a lot of challenges to maintain the currently in Europe very high quality of the supply with electricity. An central control of large powersystems like the continental European network is not possible. In case of failure of network components or large changes of power flow within the network fast control reactions are necessary to maintain the stability of voltage and frequency. This is only possible form local or regional control stations. Hence a clustering of the network into local areas is necessary. For a sensible clustering different criteria have to be taken into acount. Beside the topology of existing power lines, substations and powerplants also regulatory conditions on quality and redundancy of the network have an impact on the possible clustering results. Maintance and construction work results in a constantly changing of the network topology. In this work we present methods to cluster the network graph under these constraints. Beside clustering based on static information about the electrical characteristics of the network components we also incorporate the results of a stability assessment of the network using dynamic measurements of voltage and frequency into the clustering process. We will compare the results of the different methods and assess their applicability under different conditions.

457: 3

Easy experimental design for complex mixtures

Stefanie Feiler, Philippe Solot

AICOS Technologies AG, Switzerland

Dealing with mixtures, various restrictions may occur in practical applications, such as:

* treating some mixture factors as alternatives (e.g. chemistry / analytics: use of different solvents, pharma: different fillers)

* presence of multiple mixture classes (e.g. preservatives, lubricants)

* restrictions on some mixture class (e.g. not more than 5% of preservatives altogether)

* ratios (e.g. amount of educt 1 at least twice the amount of educt 2)

* having to also include process factors (e.g. temperature) into the model

We present how such problems can be treated in practice, showing as an example a typical situation arising in the pharmaceutical industry (formulation development). On the one hand, we discuss how the underlying statistical model has to be adapted in order to correctly capture the problem setting. But we also demonstrate that, using a modern software tool, statistical design of experiments can be used in this complex situation as effortlessly as in standard ones.

457: 4

The Benefit of Functional Data Analysis in a Thermal Spraying Process

André Rehage, Sonja Kuhnt

TU Dortmund, Germany

Functional data analysis (FDA) has become a growing area of research, which can be credited to Ramsay and Silverman (1997) among others. FDA is a very flexible way to describe data which is measured continuously over time. In industrial applications, this is a typical data situation. We treat a thermal spraying process where due to advanced technology spraying particles can be measured during the whole process (Tillmann et al., 2012). Here common tasks of FDA occur, including modelling a certain process by some kind of regression with functional predictors, functional responses or both.

Drawbacks of functional linear models (FLM) are e.g. the interpretability due to the functional structure of the regression coefficient function (James et al., 2008) or the difficulty in finding suitable basis functions; hence the functional approach is sometimes (almost) not feasible. Furthermore it is often not clear if the FLM benefits from its functional part at all. It is obvious that particularly for only moderately fluctuating data this might be true. Therefore we compare the performance of functional and non-functional linear regression in general and in case of the thermal spraying process.

References:

James, G.M., Wang, J., Zhu, J.: Functional linear regression that's interpretable. Ann. Statist., 37 (5A): 2083-2108, 2009.

Ramsay, J.O., Silverman, B.W.: Functional Data Analysis. Springer, New York, 1997.

Tillmann, W., Kuhnt, S., Hussong, B., Rehage, A., Rudak, N.: Introduction of a day effect estimator in order to improve the prediction of particle properties in an HVOF spraying jet. Thermal Spray Bulletin, 5 (2): 132-139, 2012.

459: Lehrertag

Time: Thursday, 21st Mar 2013: 2:50pm - 4:10pm · Location: KG III, HS 3044

459: 1

Von Daten zur Funktion: Vernetzungen zwischen Stochastik, Geometrie, Algebra und Analysis

Joachim Engel

Pädagogische Hochschule Ludwigsburg, Germany

Der Vortrag diskutiert anhand konkreter Beispiele schultaugliche Konzepte, wie funktionale Zusammenhänge zwischen zwei Variable in realen Sachsituationen moduliert werden können. Dabei werden unterschiedliche Anforderungsniveaus analysiert, inner- wie außermathematische Vernetzungsmöglichkeiten hervorgehoben und die Rolle von Software betrachtet. Kennzeichen des vorgestellten Ansatzes sind 1. Reale Daten als Grundlage für authentische und glaubwürdige Modellierungen 2. Einsatz von Technologie als Werkzeug zum Problemlösen und zur Illustrierung von Konzepten und Zusammenhängen 3. Vernetzung von mathematischen Inhalten: Statistik, elementare Funktionen, Analysis, Geometrie.

4610: Lehrertag -- Daten und Zufall im Übergang von der Sekundarstufe I zur Sekundarstufe II

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: Peterhof, HS 2

4610: 1

Daten und Zufall im Übergang von der Sekundarstufe I zur Sekundarstufe II

Markus Vogel

PH-Heidelberg, Germany

Im Zentrum des Workshops stehen die Erarbeitung und Reflektion unterrichtspraktischer Beispiele im Übergang von der Sekundarstufe I zur Sekundarstufe II. Anhand der Beispiele soll dabei insbesondere die Ideen der beurteilenden Statistik verdeutlicht werden, die später in der Sekundarstufe II systematisiert werden. Der Rechner kann dabei als Hilfsmittel verwendet werden.

4611: Lehrertag -- "Statistisch denken und forschen lernen“ -- explorative Datenanalyse mit der Software TinkerPlots

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: 2114a

4611: 1

"Statistisch denken und forschen lernen" -- explorative Datenanalyse mit der Software TinkerPlots

Daniel Frischemeier

Universität Paderborn, Germany

Die Leitidee „Daten und Zufall“ sieht u.a. das Arbeiten mit realen Datensätzen sowie die Durchführung einer statistischen Erhebung mit anschließender Auswertung und Interpretation der Daten im Mathematikunterricht der Sekundarstufe I vor. Dabei kommt der Auswertung und Interpretation ein besonderer Aspekt zu. Im Sinne der explorativen Datenanalyse ist es wünschenswert, dass die Schülerinnen und Schüler dabei als „Datendetektiv“ agieren und durch „Drehen und Wenden“ der Daten ihren Fragestellungen und Hypothesen nachgehen. Die Software TinkerPlots, eine in den USA entwickelte dynamische Datenanalyse- und Simulationssoftware, vorgesehen für den Stochastikunterricht in den Klassen 3-8, kann die Lernenden im Prozess der Auswertung unterstützen. In diesem Workshop soll das Potential von TinkerPlots als Software zur Unterstützung bei der Auswertung von Daten im Stochastikunterricht diskutiert und vorgestellt werden. In einer Selbstarbeitsphase werden wir dann anschließend in einen realen Datensatz eintauchen und gemeinsam interessanten statistischen Fragestellungen mit TinkerPlots nachgehen.

4612: Lehrertag -- Visualisierungen bei Statistischen Tests mit der Binomialverteilung

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: Peterhof, HS 4

4612: 1

Visualisierungen bei Statistischen Tests mit der Binomialverteilung

Arnold Zitterbart

Regierungspräsidium Freiburg, Germany

Fragestellungen zu Statistischen Tests können in der Unterrichtspraxis sehr einfach mithilfe eines Schemas beantwortet werden. Dabei ist die Gefahr sehr groß, dass wegen der Komplexität des Themas das Schema angewendet wird, ohne den mathematischen Hintergrund verstanden zu haben. Die digitalen Medien bieten inzwischen Visualisierungshilfen, die Schülerinnen und Schülern helfen können, den Gedankengang Statistischer Tests intensiver zu durchdringen. Dabei hat sich GeoGebra sehr erfolgreich bewährt, aber auch die von Schülerinnen und Schülern im Unterricht und in den Klausuren verwendeten Handhelds bieten einige Möglichkeiten, die gegenüber GeoGebra den Vorteil haben, dass problemlos im Unterricht darauf zurückgegriffen werden kann. Diese Möglichkeiten, aber auch die Grenzen sollen in dem Workshop vorgestellt und ihr didaktischer Nutzen diskutiert werden. Für eigene Erfahrungen werden den Teilnehmern TI-Nspire-Handhelds zur Verfügung gestellt.

480: Educational Research -- Competence Modeling in Economics

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1199
Session Chair: Sigbert Klinke

480: 1

Curricula wirtschaftswissenschaftlicher Studiengänge an deutschen Hochschulen und ihre Bedeutung für die Kompetenzmessung

Oliver Lauterbach

HIS-Institute for Research on Higher Education, Germany

Im WiwiKom-Projekt werden internationale Testinstrumente zur Erfassung wirtschaftswissenschaftlicher Fachkompetenz für den Einsatz mit deutschen Studierenden adaptiert. Hierbei sind neben sprachlichen und kulturellen Aspekten auch Unterschiede der Curricula der Studienfächer zu berücksichtigen.

Auf Basis von Analysen der Modulhandbücher von mehr als 80 wirtschaftswissenschaftlichen Bachelor-Studiengängen wird der Frage nachgegangen, welche Studieninhalte über die betrachteten Hochschulen und Studiengänge hinweg als wirtschaftswissenschaftliches Kerncurriculum angesehen werden können. Hierbei zeigt sich unter anderem, dass die einzelnen Studiengänge je nach Hochschultyp im Hinblick auf Pflicht- und Wahlpflichtanteile heterogen strukturiert sind und je nach Studienfach der Betriebswirtschaftslehre, Volkswirtschaftslehre und Wirtschaftswissenschaften unterschiedlich hohe Anteile spezifischer Fachinhalte aufweisen. Weiterhin wird überprüft, inwieweit die inhaltliche Struktur der zu adaptierenden Testinstrumente den angebotenen Studieninhalten entspricht und ob möglicherweise einzelne Bereiche über- oder unterrepräsentiert sind. Hier wird deutlich, dass innerhalb der wirtschaftswissenschaftlichen Fachgebiete weitgehende Übereinstimmung zwischen den Studien- und Testinhalten besteht. Die Ergebnisse werden hinsichtlich ihrer Bedeutung für die weitere Selektion der Testaufgaben und die Gewichtung der einzelnen Inhaltsbereiche im Gesamttest diskutiert.

480: 2

Calibration of the itempool for assessing competencies in business and economics

Manuel Förster, Sebastian Brückner, Olga Zlatkin-Troitschanskaia

Johannes Gutenberg-University Mainz, Germany

A very important goal of the WiwiKom project is to assess competencies of undergraduate students in business and economics in German higher education. To achieve that goal, two international instruments were translated in cooperation with translation experts at the Johannes Gutenberg-University Mainz, and their content was first validated through expert interviews and through cognitive interviews with. Based on a first pilot stage in the summer term 2012 and other multiple statistical and qualitative content-related criteria 220 of the initial 402 items could be adapted this way and were subsumed in 43 test booklets in the form of several nested Youden square Designs (Frey, Hartig & Rupp, 2009). During the current winter term 2012/2013, these 43 booklets are being deployed in a first major field study with about 3000 students of business and economics at 18 universities. The data from this survey will be analyzed with methods from classical test theory and item response theory in order to determine whether the items are adequate and suited for further empirical modeling of competence structures and levels. The presentation will discuss the current progress and will then offer an insight into the results of the first major field study. Apart from pure descriptive results, a particular focus will be laid on central analyzes concerning the calibration of the item pool. Furthermore first analyzes of the dimensionality of the test instrument and the measurement quality of the items will be presented. Based on the results, implications for further test development will be drawn.

480: 3

An item response model with a generalized logistic link function

Alexander Robitzsch

BIFIE Salzburg, Austria

We introduce a mutidimensional item response model (IRM) with a generalized logistic link function (Stukel, 1988). Deviations from a logistic link function are modelled by introducing two additional link function parameters. Therefore, our model is a generalization of the Rasch model (with employs the ordinary logistic link function) while the link function can be estimated parametrically without sacrificing the property that person ability and item difficulty appear only as main effects in the item response function.

The proposed IRM also handles multiple groups and enables the estimation of latent regression models and drawing plausible values (Adams & Wu, 2007). In addition, the model allows the specification of (positive) local dependencies of items in terms of a marginal modeling approach (Braeken, 2011).

Statistical properties of the proposed IRM with generalized logistic link function are investigated in a simulation study which is based on an implementation of the IRM in the R package sirt (Robitzsch, 2012). For several data sets, the model fit of the proposed IRM is compared with the Rasch model.

480: 4

Analysis of the test items measuring competencies in business and economics in WiwiKom project using IRT framework

Tetyana Sydorenko, Sigbert Klinke

Humboldt-Universität zu Berlin, Germany

Project WiwiKom aims to measure competencies of undergraduate students in business and economics in German higher education. For this purpose test items from already available instruments such as TUCE and EGEL were adapted and after complex expertise adjusted to incorporate the curricular requirements of German universities for the field of business and economics.

One of the main goals of the project is to prove empirically how the verified test is performing in German universities. Based on data gathered from around ten different German universities with business and economics students, who complete the WiWiKom test, we will investigate how far the test can model particular dimensions of competencies in business and economics. For this purpose various Item Response Theory (IRT) models are applied to measure the item characteristics as well as general properties of the test. Moreover, by investigating the differential item functioning based on location, gender, native language we aim to answer the question, whether the test could be applied as an appropriate tool to measure competencies in business and economics in various German universities and across different groups of students.

481: Visualizations and Exploratory Data Analysis

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1098
Session Chair: Göran Kauermann

481: 1

Quantile sheets

Paul H. C. Eilers[1], Sabine K. Schnabel[2]

[1]Erasmus University Medical Centre, The Netherlands; [2]Biometris, The Netherlands

Quantile smoothing is a popular tool for modelling the conditional quantiles of a response as a function of one (or more) independent variables. By definition, quantile curves of different probabilities cannot cross, but in practice we encounter this phenomenon frequently. It appears more often in small data sets, but is not limited to these situations. The crossing of curves may influence further analysis, e.g. when it comes to studying the conditional distribution. In the literature one can find a number of different proposals to prevent these crossings such as restricted regression quantiles, special kernel estimators and additional constraints.

We propose an alternative approach. The basic idea introduces a surface over a two-dimensional domain of the covariate x and the probability p. This surface is called a quantile sheet. Cutting this surface at a fixed probability gives a smooth quantile curve. All quantile curves are estimated simultaneously and so the crossing problem disappears if the sheet is monotonically increasing in probability. The sheet is constructed as a sum of tensor product B-splines with penalties to control the smoothness in both directions. For the estimation iteratively weighted least squares is used in combination with fast array regression methods.

481: 2

Visualization of classification decisions based on geometric predictors

B Balliu, S Boehringer

Leiden University Medical Center, The Netherlands

Classification problems with geometrically interpretable predictors, e.g. graphs, allow for visualization of classifiers. For regression techniques, coefficients associated with predictor variables have to be associated with their corresponding geometric interpretation and used for highlighting.

Often data gets transformed prior to classification, for example by converting graph vertices into pair-wise distances or similar transformations. Such transformations can induce ill-posed problems by creating more predictors than observations which require penalized methods to achieve regularized solutions.

We consider visualization of such penalized predictors which are known to have biased coefficient estimates. A first approach assigns weights to each point in a mean graph corresponding to the importance of that point for the classification process and illustrate the choice of weights for various transformations (distances, angles, areas). A second approach tries to create caricatures, i.e. graphs that overemphasize features of one class with respect to the other. We define penalty functions on graph deformations that are chosen such as to produce caricatures when the penalty is minimized. The penalties account for the transformations and the biased nature of the estimates. We discuss algorithms to compute these transformations.

We illustrate our visualization techniques on a discrimination problem of 2D images where genetic syndromes are discriminated using LASSO regression.

481: 3

Identifying periods of market stress: Multivariate regime switching models

Jakob Stöber[1], Claudia Czado[1], Harry Joe[2]

[1]Technische Universität München, Germany; [2]University of British Columbia, Canada

Misperceptions about extreme dependencies between different financial assets have been an important element of the recent financial crisis, which is why regulating entities do now require financial institutions to account for different behavior under market stress. In particular, European banks have to report a stressed Value at Risk (SVaR) in addition to the standard VaR measure. In this talk, we will demonstrate how appropriate stress periods can be identified using Markov switching time series models.

Our focus will be on the dependence structure where we use flexible R-vine copula models. These are constructed from bivariate copulas only by hierarchical conditioning. To keep them tractable for inference, one usually assumes that copulas of conditional distributions do not depend on the values of variables that are conditioned on. This simplifying assumption has recently come under scrutiny, and we will show for which classes of distributions it is applicable. Further, we will discuss how time-varying or Markov switching parameters generalize simplified copula models.

We will then apply marginal Markov switching models as well as the discussed dependence models to exchange rate and stock return data.

Here, we show that regime switches in marginal time series and the dependence structure are strongly associated, which justifies the SVaR approach.

481: 4

Visualizations for Exploratory PLS Path Modeling

Armin Monecke[1], Marko Sarstedt[2]

[1]Ludwig-Maximilians-Universität München, Germany; [2]Otto-von-Guericke-Universität Magdeburg, Germany

In various disciplines structural equation modeling (SEM) is a well established technique for theory confirmation. But often SEM is used in an exploratory fashion, thus demanding for methods which help the researcher to better understand their data in the context of an a-priori hypothesized cause-effect relationship. Diagnostic plots as known from regression are suitable for exploring the multiple relationships in SEMs when factor or component scores are at hand. As within the PLS-approach to SEM component scores are explicitly calculated it is especially well suited for an exploratory analysis. To present the idea we analyze a customer satisfaction study by means of graphical tools such as density estimates and partial residual plots based on additive models.

482: Official Statistics, Survey Statistics and Demography

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1010
Session Chair: Hans Kiesl

482: 1

A Robust Random Effect Block Bootstrap for Spatially Clustered Data

Timo Schmid[1], Payam Mokhtarian[2], Ray Chambers[2]

[1]Freie Universität Berlin, Germany; [2]University of Wollongong, Australia

Estimates of linear mixed model parameters are sensitive to the presence of outliers in the data. Furthermore, in many environmental applications of linear mixed models, it is reasonable to assume that both group level and individual level random effects are spatially correlated. Extensions to conventional random effects models that account for spatial dependency have been recently proposed in the literature. In this paper we tackle the issue of outlier-robust estimation of the parameters of such a spatial linear mixed model using a re-sampling approach. In particular, a spatial extension of the robust random effects block bootstrap approach to mixed model fitting proposed by Mokhtarian and Chambers (2012) is used to replicate both the hierarchical structure and the spatial structure of the non-outlying data generated by the spatial linear mixed model. Under this approach it is possible to estimate the spatial dependency parameter as well as the variance components in presence of outliers in the sample data. Empirical analyses based on simulation experiments as well as an application to an environmental data set indicate that the method leads to improved model parameter estimates.

482: 2

Estimation of Change over Time in the Presence of Sample Co-ordination

Stefan Zins, Ralf Münnich

Universtiy of Trier, Germany

With Europe 2020, the EU has set up a growth strategy for the coming decade. This strategy includes explicit target values for selected quantitative indicators. Therefore, the evolution of these indicators over time is of high interest to assess the progress towards the agreed EU targets. As an instrument to measure these indicators in a comparable way on an EU wide level, the European Union Statistics on Income and Living Conditions (EU-SILC) has been implemented. EU-SILC is based on annual surveys conducted in the member states that are designed as rotational panels, which allow both for cross-sectional and longitudinal analysis. To carry out a rotational sampling design, a sequence of random samples has to be coordinated over time which introduces stochastic dependencies between selected samples.

In case one is interested in the significance of estimated changes of an indicator value between to two years, one additionally has to estimate the variance of this change. This urges the needs of taking into account the correlation between indicator estimates induced by sample co-ordination.

The aim of the paper is to introduce a framework for sampling co-ordination, which also accommodates changes in the sampling frame due to births and deaths in the studied population. Further, the calculation of the required inclusion probabilities for variance estimation is demonstrated for a co-ordination scheme of simple random samples. Finally, a variance estimator for change in estimates of non-linear statistics, such as the indicators used for Europe 2020, is given.

482: 3

Quantile estimation after multiple imputation

Jörg Drechsler[1], Robin Mitra[2]

[1]Institute for Employment Research, Germany; [2]University of Southampton, United Kingdom

When estimating quantiles of an unknown univariate distribution it is common to use the sample quantile as an unbiased point estimator for the true quantile and estimate its variance using some kind of resampling method, such as the bootstrap or the jack-knife. However, as we illustrate in this talk, using this strategy for a dataset for which missing observations have been multiply imputed will lead to conservative variance estimates based on Rubin's combining rules. The reason is that the sample quantile is not a self-efficient estimator as defined by Meng (1994).

We propose a straightforward maximum likelihood estimator for the quantile using a box-cox transformation that allows valid inferences after multiple imputation. We illustrate through simulation that the estimator is more efficient than the estimator based on the sample quantile and unbiased given that the sample data can be approximated by a normal after the box-cox transformation.

482: 4

Generalized calibration to deal with non-ignorable nonresponse in the German job vacancy survey

Hans Kiesl

Regensburg University of Applied Sciences, Germany

The German job vacancy survey is a quarterly business survey (stratified by sector and size class). Calibration is used to deal with high nonresponse rates; the most important calibration variable being the number of registered job vacancies (business units may inform the Federal Employment Agency of their vacancies; these are called registered vacancies). It turns out that design weighting (ignoring nonresponse) results in biased estimates of the total number of registered vacancies. Since we assume that nonresponse is correlated with the number of any vacancies (registered or not), nonresponse in our survey is non-ignorable and estimators like the GREG will produce biased estimates. We compare two estimation strategies in terms of bias and variance that are more appropriate in our situation by means of a simulation study: a new two-stage GREG estimator and the generalized calibration estimator proposed by Deville (2002) and Chang and Kott (2008).

483: Statistics in Finance, Insurance and Banking

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG I, HS 1015
Session Chair: Ostap Okhrin

483: 1

COMFORT-CCClass: A Common Market Factor Non-Gaussian Returns Model

Marc Paolella

University of Zurich, Switzerland

A new multivariate returns model with various attractive properties is motivated and studied. By extending the CCC model in several ways, it allows for all the primary stylized facts of asset returns, including volatility clustering, non-normality of asset returns (excess kurtosis and asymmetry), and also dynamics in the dependency between assets over time. A fast EM-algorithm is developed for estimation. The predictive conditional distribution is a (possibly special case of a) multivariate generalized hyperbolic, so that sums of marginals (as required for portfolios) are tractable. Each marginal is endowed with a common univariate shock, interpretable as a common market factor, and this stochastic process has a predictable component. This leads to the new model being a hybrid of GARCH and stochastic volatility, but without the estimation problems associated with the latter, and being applicable in the multivariate setting for potentially large numbers of assets. Formulae associated with portfolio optimization, risk measures and option pricing based on the predictive density are developed. In-sample fit and out-of-sample conditional density forecasting exercises using daily returns on the 30 DJIA stocks confirm the superiority of the model to numerous competing ones.

483: 2

On the Sparse Estimation of Copula-Based Vector MEM

Nikolaus Hautsch, Ostap Okhrin, Alexander Ristig

Humboldt-Universität zu Berlin, Germany

This paper provides a multi-stage procedure for estimation of copula-based vector multiplicative error models (VMEM). Applications of VMEM include cross-correlations between discrete positive valued time series, such as modeling of trading volumes, realized volatilities and trading intensities. The model's mean equation is sparsely estimated based on a penalized maximum likelihood approach under the assumption of cross-sectional independent innovations. The presumed copula is then calibrated to the filtered residuals as in the classical copula-GARCH models. Finally, the parameters' efficiency is enhanced by a stepwise re-estimation of the VMEM, based on the full log likelihood and fixed submodels. A simulation study demonstrates the accuracy of the proposed method and further aspects are discussed in an empirical analysis.

483: 3

Estimation of continuous-time process with applications in mathematical finance

Thorsten Schmidt

TU Chemnitz, Germany

The availability of almost continuous-time prices for stocks, bonds or other financial derivatives motivates the application of statistical methods for stochastic processes. We discuss the estimation of Markovian processes, in particular affine and polynomial ones and study their implications to financial applications.

484: Open Topics

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG III, HS 3042
Session Chair: Wolfgang Schmid

484: 1

Simultaneous confidence bands for a percentile line in linear regression

Yang Han, Wei Liu

University of Southampton, United Kingdom

Simultaneous confidence bands have been used to quantify unknown functions in various statistical problems. A common statistical problem is to make inference about a percentile line in linear regression. The main purpose of this research is to construct simultaneous confidence bands for a percentile line over a given covariate interval which can be finite or infinite, and to compare the bands under the average width criterion.

A simultaneous confidence band can quantify the plausible range of the percentile line. Any straight line that lies inside the simultaneous confidence band is deemed, by this band, as a plausible candidate for the true percentile line. It is intuitive that the narrower the band becomes, the better it is. Hence the average width is used as an optimality criterion. We present six symmetric confidence bands and the thorough comparison among these six bands under the average width criterion. An application to a real data set is demonstrated.

484: 2

Confidence set for a maximum point of a univariate polynomial regression function.

Fang Wan, Wei Liu

University of Southampton, United Kingdom

A maximum point of a regression function is defined as one at which the regression function attains its maximum value. The determination of a maximum point of the regression function in a constrained covariate region is often of great interest in polynomial regression analysis. Since the regression function needs to be estimated and its maximum point(s) can only be estimated based on the random observations, the focus of this research is to construct a confidence set for a maximum point of the regression function.

A confidence set for a maximum point provides useful information about the regression function and quantifies where a true maximum point lies. In this paper, we propose an approach to the construction of a confidence set for a maximum point of a univariate polynomial regression function by inverting a family of acceptance sets of hypothesis tests. Examples are given to illustrate our confidence sets in different cases.

484: 3

Simultaneous calibration of ensemble river flow predictions over an entire range of lead-times

Stephan Hemri

University of Heidelberg, Germany

When forecasting water levels and river flow, ensemble weather forecasts are frequently used as input to hydrologic process models. As hydrologic models are imperfect and the input ensembles tend to be biased and underdispersed, the output ensemble forecasts for river runoff typically are biased and underdispersed, too. Thus, statistical post-processing is required in order to achieve sharp and calibrated predictions. In this work Bayesian model averaging (BMA) is applied to statistically post-process ensemble runoff forecasts for a catchment in Switzerland, at lead-times ranging from 1 to 240 hours. First, BMA is applied based on mixtures of univariate normal distributions, subject to the assumption of independence between distinct lead-times. Then, the independence assumption is relaxed in order to estimate multivariate runoff forecasts over the entire range of lead-times simultaneously, based on a BMA version that uses multivariate normal distributions. Since river runoff is a highly skewed variable, Box-Cox transformations are applied in order to achieve approximate normality. Both univariate and multivariate BMA approaches are able to generate well calibrated probabilistic forecasts that are considerably sharper than climatological forecasts.

484: 4

Testing Order-Constrained Hypotheses Concerning Circular Data.

Corine Baayen[1,2]

[1]University of Copenhagen, Denmark; [2]H. Lundbeck A/S, Copenhagen, Denmark

Many fields of research make use of data that can be represented on a circle. Consider directional data, or measurements on circular scales such as the circumplex model of affect, which provides a circular representation of mood using two dimensions, pleasantness and activation. Due to periodicity, the evaluation of these so called circular data requires special statistical methods. This talk will introduce two tests for the analysis of order constrained hypotheses for circular data, allowing for comparisons between groups as well as repeated measurements.

Through these tests, researchers can evaluate their expectations regarding the outcome of an experiment directly by representing their ideas in the form of a hypothesis containing inequality constraints. The resulting data analysis is generally more powerful than one using standard null hypothesis testing. Results from a simulation study show that the tests perform well in terms of type I error and power.

485: Industrial Statistics Statistical Surveillance

Time: Thursday, 21st Mar 2013: 4:40pm - 6:00pm · Location: KG III, HS 3043
Session Chair: Stefan Hans Steiner

485: 1

Residuals-based CUSUM Charts for Poisson INAR(1) Processes

Christian Weiß[1], Murat C. Testik[2]

[1]TU Darmstadt, Department of Mathematics, Darmstadt, Germany; [2]Hacettepe University, Industrial Engineering Department, Beytepe-Ankara, Turkey

To monitor count-type observations that are correlated over time, the Poisson INAR(1) CUSUM chart has been proposed for cases where an INteger-valued AutoRegressive process of order 1 with a Poisson marginal is appropriate. In a recent study, it is shown that the Poisson INAR(1) CUSUM control chart is especially well-suited to detection problems that involve mean shifts in a process. Although special causes of variation that result in mean shifts are the most common out-of-control situation in practice, some special causes of variation may also result in shifts of the autocorrelation coefficient or the process variance.

In this research, we define three types of residuals from the time-series model and develop CUSUM control charts for monitoring these residuals. In particular, three CUSUM control charts, one for each of these residuals, as well as a multivariate CUSUM monitoring of the residuals are investigated. A Phase II comparison of these residuals' monitoring approaches together with a benchmark CUSUM of the count type observations indicate that for different types of process shifts, different residual monitoring approaches may perform better.

485: 2

On outliers and interventions in INGARCH time series

Roland Fried, Tobias Liboschik, Konstantinos Fokianos

TU Dortmund, Germany

We review recent research on the modelling of different types of outlier and intervention effects in time series of counts, arising e.g. in epidemiology. We concentrate on time series models within the framework of generalized linear models, and on outlier effects entering the dynamics of the process in different ways, and affecting more than one observation. Iterative likelihood based procedures for outlier detection and correction can be constructed for outliers of such types, while so-called additive outliers describing e.g. pure measurement artefacts are treated more easily within a Bayesian framework. Some ideas for robust estimation of the model parameters are presented as well.

485: 3

Incorporating parameter uncertainty into the setup of EWMA control charts monitoring normal variance

Sven Knoth

Helmut Schmidt University Hamburg, Germany

Most of the literature concerned with the design of control charts relies on perfect knowledge of the distribution for at least the in-control process. Some papers treated the handling of EWMA charts monitoring normal mean in case of unknown parameters -- refer to Jones et al. (2001). In Jensen et al. (2006): ``Effects of Parameter Estimation on Control Chart Properties: A Literature Review'' a nice overview was given. Additionally, it was mentioned that it would be useful to evaluate and take into account these effects also for variance control charts. Here, we consider EWMA charts for monitoring the normal variance. Given a sequence of batches of size [fn]n[/fn], [fn]X_{ij}[/fn], [fn]i=1,2,...[/fn] and [fn]j=1,2,...,n[/fn] calculate the batch sample variance and apply the EWMA smoothing with given lambda. The upper scheme stops as soon as the EWMA statistic is larger than the threshold [fn]c_u[/fn]. The parameters [fn]0<\lambda\leq1[/fn] and [fn]c_u>0[/fn] are chosen to enable a certain useful detection performance. The most popular performance measure is the so-called Average Run Length (ARL), that is [fn]E_\sigma(L)[/fn] for the true standard deviation sigma. If [fn]\sigma_0[/fn] has to be estimated by sampling data during a pre-run phase, this effects, of course, the behavior of the applied control chart. Typically the ARL is increased. Most of the papers about characterizing the uncertainty impact deal with the changed ARL patterns and possible adjustments. Here, a different way of designing the chart is treated: Setup the chart through specifying a certain false alarm probability such as [fn]P_{\sigma_0}(L \leq1000) \leq \alpha[/fn]. This results in a specific [fn]c_u[/fn]. Here we describe a feasible way to determine this value [fn]c_u[/fn] also in case of unknown parameters for a pre-run series of given size (and structure). A two-sided version of the introduced EWMA scheme is analyzed as well.

485: 4

On the usefulness of EWMA charts for econometric strucutral change tests

Dominik Wullers

Helmut-Schmidt-Universität Hamburg, Germany

We show that EWMA control charts are useful to quickly detect changes in econometric models. Evidence is presented that suggests a higher performance regarding detection speed and better size and power distributions for EWMA procedures when compared to the well-known MOSUM detector. To allow for a reasonable comparison, Quartile Run Lengths are used for calibration. Leisch et al. (2000)'s moving estimates statistic is considered to allow for an econometric application of the EWMA chart.

511: Meth Börse

Time: Friday, 22nd Mar 2013: 8:00am - 8:30am · Location: KG I, HS 1199

511: 1

Quantilsregressionen

Bernd Fitzenberger

Albert-Ludwigs-Universität, Germany

tba

521: Statistics in Clinical and Preclinical Research -- Translational Sciences

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG II, Audimax
Session Chair: Richardus Vonk

521: 1

Statistical Considerations in Translational Medicine

Shein-Chung Chow

Duke University School of Medicine, United States of America

This presentation will focus on statistical considerations for assessment of translation in language (e.g., translation of case report forms in multinational clinical trials), information (e.g., translation of basic discovery to clinic) and technology (e.g., translation of Chinese diagnostic techniques to well established clinical study endpoints) in pharmaceutical/clinical research and development. However, most of the effort will be directed to statistical considerations for translation in information. Pizzo (2006) defines translational medicine as "bench-to-bedside" research wherein a basic laboratory discovery becomes applicable to the diagnosis, treatment or prevention of a specific disease and is brought forth by either a physician-scientist who works at the interface between the research laboratory and patient care or by a team of basic and clinical science investigators. Statistics plays an important role in translational medicine to ensure that the translational process is accurate and reliable with certain statistical assurance. For this purpose, statistical criteria for assessment of one-way and two-way translation are proposed. Under a well established and validated translational model, statistical tests for one-way and two-way translation are derived. Strategies for selection of clinical study endpoints (e.g., absolute changes, relative changes, or responder defined either based on absolute change or relative change, etc) are reviewed. Statistical inference for lost-in-translation and for the applicability of an animal model to a human model are also discussed.

Key Words: Bench-to-bedside; One-way translational process; Two-way translational process; Lost-in-translation; Probability of success.

521: 2

Translational Statistics: Enabling key decisions in early drug development

Michael Branson

Novartis Pharma AG, Switzerland

The early development sector of the pharmaceutical industry has placed extensive focus on “proof of concept” as the basis or gatekeeper for making the key decision to kill or transition new therapies to mid-stage drug development. We will review and consider some key aspects and pre-requisites leading to developing a proof of concept strategy, coupled with a quantitative framework.

521: 3

Modeling and Simulation in Early Phases of Drug Development

Günter Heimann

Novartis Pharma AG, Switzerland

Over the past decade, quantitative methods have gained an increasing role in drug development. This includes more modern study designs and analysis methods, but more importantly also the quantitative integration of information across studies, projects, and other sources of information. Modeling and simulation has been a strong catalyst of these modern approaches.

In this presentation I will present a few examples of how quantitative integration of information can be used in early phases of drug development. The range of examples includes using population pharmacokinetic models to inform realistic scenarios for clinical trial evaluation, using historical data to link biomarkers to clinical endpoints and to predict the clinical efficacy result from early biomarker data, and using literature data to benchmark the competitive profile of a drugs in phase I.

522: Research Synthesis and Meta Analysis -- Special methods in meta-analysis

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1010
Session Chair: Jochem König
Session Chair: Oliver Kuß

522: 1

Parametric meta-analyses of trials with incomplete reporting of competing events may be performed for cumulative event probabilities.

Federico Bonofiglio[1], Michael T Koller[2], Martin Schumacher[1], Jan Beyersmann[1], Guido Schwarzer[1]

[1]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [2]Basel Institute for Clinical Epidemiology, University Hospital Basel, Switzerland

Survival endpoints which are not experienced by every individual are subject to competing risks. Meta-analyses should account for this, because cumulative event probabilities depend on all risk-specific hazards. Since reporting on competing events is typically incomplete, we try to recover information on the competing risk-specific hazards under a simplifying constant hazard assumption. Using the recovered information and a delta method argument allows for meta-analyses on the cumulative event probability scale. We illustrate the approach using a meta-analysis on ten randomized clinical trials of implantable cardioverter defibrillator implantation versus control considering the outcomes sudden cardiac deaths and non cardiovascular deaths. We also discuss the relative merits of meta-analyses on the risk-specific hazards scale or on the probability scale. The latter ties in with current research on inference for the ratio of cumulative event probabilities in the presence of competing risks, but relies on choosing appropriate landmark times during follow-up.

522: 2

Summarizing EC50 estimates from several dose-response experiments: a comparison of a meta-analysis approach to a mixed-effects model approach

Xiaoqi Jiang, Annette Kopp-Schneider

German Cancer Research Center, Heidelberg, Germany

Dose-response studies are performed to investigate the potency of the substance under investigation. Data are typically evaluated using a log-logistic model that includes EC50 as one of the model parameters. Often, more than one experiment is performed for a compound under study and it is then necessary to summarize the estimated EC50 values from several experiments. In this context, mixed-effects models can be designed to estimate the average EC50, considering the variability among and within experiments. However, fitting nonlinear mixed-effects models is more complicated than in a linear situation and convergence problems are often encountered. Especially in the case of a small number of experiments, mixed-effects models tend to be overparametrized. An alternative strategy to summarize EC50 values from several experiments is the application of a meta-analysis approach to combine EC50 estimates obtained from separate log-logistic model fitting. In meta-analysis, estimates obtained from several studies are combined in a single superanalysis, which integrates the results from each study and yields one overall estimate with confidence interval. We compared these two modeling strategies in a simulation study and our analysis showed that that meta-analysis is a simple and robust method to combine EC50 estimates.

References:

[1] C. Ritz, Toward a unified approach to dose-response modeling in ecotoxicology. Environmental Toxicology and Chemistry, 29(1):220-229, 2010.

[2] J.C. Pinheiro and D.M. Bates, Mixed-effects models in S and S-PLUS. Springer Verlag, 2000.

[3] W. Viechtbauer, Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics, 30(3):261-293, 2005.

522: 3

Sequentielle Meta-Analysen zur Planung und Auswertung von Safety-Fragestellungen

Daniel Saure, Katrin Jensen, Meinhard Kieser

Institut für Medizinische Biometrie und Informatik, Universität Heidelberg, Germany

Während Meta-Analysen von Wirksamkeitsnachweisen überwiegend retrospektiv durchgeführt werden, stellt sich gerade in der Arzneimittelentwicklung bei dem Vergleich von Safety-Daten die Frage nach einem prospektiven, sequentiell geplanten Vorgehen über mehrere klinische Studien hinweg. Für sequentielle Meta-Analysen existieren einige Ansätze in der Literatur, darunter die Kombination von p-Werten (Jennison und Turnbull, 2005) und die wiederholte kumulative Meta-Analyse (Whitehead, 1997). Für letztere untersuchen Higgins et al. (2011) den Fehler 1. Art in einem Modell mit zufälligen Effekten. Bei Safety-Daten und damit seltenen Ereignissen sind jedoch exakte statistische Methoden in der Regel asymptotischen Verfahren vorzuziehen. In Vandermeer et al. (2009) findet sich ein Vergleich von exakten und asymptotischen klassischen meta-analytischen Methoden.

In unserem Vortrag stellen wir die oben genannten sequentiellen meta-analytischen Methoden vor und überprüfen ihre Anwendbarkeit auf

Safety-Daten. Der Fokus liegt dabei auf dem Modell mit festen Effekten. Anhand von Simulationen untersuchen wir unter anderem die Einhaltung des Fehlers 1. Art in verschiedenen Szenarien, die typische Situationen in der Arzneimittelsicherheit widerspiegeln.

522: 4

Negative binomial rate and overdispersion estimation based on incomplete information in systematic reviews

Christian Röver[1], Stefan Andreas[1,2], Tim Friede[1]

[1]Universitätsmedizin Göttingen, Germany; [2]Lungenfachklinik Immenhausen, Germany

The analysis of count data is commonly done using Poisson models [1,2]. Negative binomial models are a straightforward and readily motivated generalization for the case of overdispersed data, i.e., when the observed variance is greater than expected under a Poissonian model [3]. The joint estimation of rate and overdispersion parameters however is not trivial in this case. We review the properties of different parameter estimation approaches, including the utilization of partial (censored) information. The methods are motivated and illustrated by a systematic review in chronic obstructive pulmonary disease (COPD). Here diverse sources of information are utilized to infer event rates and overdispersion, and often only summary statistics of the data, like mean and frequencies of sub-categories of outcomes, are given.

References:

[1] R. Nicholas, S. Straube, H. Schmidli, S. Schneider, T. Friede. Trends in annualized relapse rates in relapsing-remitting multiple sclerosis and consequences for clinical trial design. Multiple Sclerosis Journal, 17(10):1211-1217, October 2011.

[2] S. M. Steinvorth, C. Röver, S. Schneider, R. Nicholas, S. Straube, T. Friede. Explaining temporal trends in annualized relapse rates in placebo groups of randomized controlled trials in relapsing multiple sclerosis: systematic review and meta-regression. Submitted for publication.

[3] J. F. Lawless. Negative binomial and mixed Poisson regression. The Canadian Journal of Statistics, 15(3):109–225, September 1987.

523: Computational and Simulation Based Statistics

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1015
Session Chair: Bernd Bischl

523: 1

Effectiveness of population-based skin cancer screening – a microsimulation of melanoma mortality

Nora Eisemann[1], Annika Waldmann[2], Alexander Katalinic[1,2]

[1]Institute of Cancer Epidemiology, University of Lübeck, Germany; [2]Institute of Clinical Epidemiology, University of Lübeck, Germany

Background: A nation-wide skin cancer screening (SCS) was implemented in 2008 in Germany. It aims at improving early detection of melanoma in order to reduce melanoma mortality. While the idea of early detection is compelling, demonstrating the effectiveness of melanoma screening is crucial. The effectiveness of SCS on melanoma mortality depends on factors such as targeted age groups and screening interval. Since it is not feasible to conduct studies to investigate the impact of these factors, we developed a microsimulation model which is able to predict melanoma mortality in Germany under several SCS options.

Materials and Methods: Based on German cancer registry data, population data, and other published data on melanoma progression and screening participation, we developed a stochastic microsimulation model. 10,000 populations of each 100,000 female or male persons and their melanoma-related life histories were simulated and calibrated to observed melanoma incidence and mortality. In a second step, life histories were changed by a SCS.

Results: Compared to a non-screening scenario (age-standardised mortality rate: 1.6 and 2.6 per 100,000 for women and men), a SCS of the 35-85 year old population every second year and a participation rate of 27.2% for women and 10.4% for men (according to experiences of a pilot project) reduced mortality by up to 0.7 death per 100,000 person-years for women and men, corresponding to a relative reduction of up to 47% in women and 27% in men.

Conclusion: The model gave plausible melanoma mortality predictions and allows the comparison of different SCS scenarios.

523: 2

Raytracing in the Lexis diagram

Ralph Brinks, Sandra Landwehr, Giani Guido

German Diabetes Center, Düsseldorf, Germany

Background: Relevant events of the illness-death-model (IDM) described in [1] are diagnosis of a chronic disease and death with or without the disease. In this work a code snippet in R [2] is presented, which simulates persons moving in the IDM. For this, integrals along line segments in the Lexis diagram have to be calculated. Since analytical solutions for those line integrals might not be available, numerical approximations are the most general approach. In numerical integration, the step size at which the integrand is evaluated typically plays an important role for the accuracy of the results.

Materials and Methods: A population of 300000 persons born in 60 consecutive years is simulated using Monte Carlo (MC) techniques. Siddon's raytracer [3] and the iterated trapezoid rule are used to calculate the line integrals. A test case is chosen where analytical solutions of the line integrals exist. Run-time of the simulation and accuracy of the prevalence in the age range 40-90 years (compared to the analytic solution) are examined.

Results: Step sizes of 2/1/0.5/0.25 (years) on a standard desktop PC lead to run-times of 31/59/112/223 minutes. The absolute error in the prevalence is about 2 per mille, independent of the step size.

Conclusions: Run-time and step size are inversely proportional. In the example the accuracy is determined by the MC sampling error, not by the step size.

References:

[1] Keiding, J Roy Stat Soc A 1991:154

[2] R Development Core Team 2012

[3] Siddon, Medical Phys 1985:12(2)

523: 3

Which model to match?

Roxana Halbleib[1], David Veredas[2], Matteo Barigozzi[3]

[1]University of Konstanz, Germany; [2]ECARES, Free University of Bruxelles, Belgium; [3]London School of Economics, United Kingdom

The asymptotic efficiency of indirect estimation methods, such as the efficient method of moments and indirect inference, depends on the choice of the auxiliary model. Up to date, this choice is somehow ad hoc and based on an educated guess. In this article we introduce a class of information criteria that helps the user to optimize the choice among nested and non–nested auxiliary models. They are the indirect analogues of the widely used Akaike–type criteria. A thorough Monte Carlo study based on two simple and illustrative models shows the usefulness of the criteria.

523: 4

Statistical modelling of thermo-physic materials properties using segmented regression

Irina Roslyakova[1], Holger Dette[2]

[1]Department of Scale Bridging Thermodynamic and Kinetic Simulation, ICAMS, Ruhr-Universität Bochum, Germany; [2]Ruhr-Universität Bochum, Fakultät für Mathematik - Mathematik III, Bochum, Germany

Frequently in regression problems, a model is assumed to be a single parametric function throughout the entire domain of analysis. However, in many phenomena is it necessary to consider regression models which have different analytical form in different regions of domain. Such type of regression models used for the parameter estimation is called segmented regression. The proposed segmented regression was developed for modeling of the heat capacity.

Traditionally in applied thermo-physics the temperature dependence of the heat capacity is described by high-order polynomials, with adjustable parameters fitted to experimental data. This approach led to fitting coefficients that lack any physical meaning and it is not valid below 298.15K.

To overcome this difficulties, in this work we used more physical approach that requires the modeling of several contributions (e.g. electronic, vibrational, etc.). Since these contributions appear in different temperature ranges and can be described by different functions, the segmented regression methodology was applied for developing of mathematical model for heat capacity of materials. Several segmented regressions were considered, analyzed and validated by using statistical tests. Corresponding confidence intervals have been calculated using the bootstrap method.

Additionally to this practical issue, the asymptotic theory for the proposed model will be developed. In particular we prove consistency and asymptotic normality of the nonlinear least square estimator under non-standard assumptions.

524: Statistics in Finance, Insurance and Banking

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1098
Session Chair: Ioana Andreea Duca

524: 1

A goodness-of-fit test for regular vine copula models

Ulf Schepsmeier

Technische Universität München, Germany

A first goodness-of-fit test for regular vine (R-vine) copula models is introduced. R-vine copulas are a very flexible class of multivariate copulas based on a pair-copula construction (PCC) (Aas et al., 2009). The test arises from the information matrix equality and specification test proposed by White (1982) and extends the goodness-of-fit test for copulas introduced by Huang and Prokhorov (2011). The corresponding critical value can be approximated using either its asymptotic theory or bootstrapping. In an extensive simulation study we investigated the observed size and power under several different R-vine model specifications. The new test will be applied to exchange rates comparing different vine copula models.

References:

[1] Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula construction of multiple dependence. Insurance: Mathematics and Economics 44, 182-198.

[2] Huang, W. and A. Prokhorov (2011). A goodness-of-fit test for copulas. submitted to Economic Reviews. White, H. (1982). Maximum likelihood estimation of misspecified models. Econometrica 50, 1-26.

524: 2

A Multivariate Volatility Vine Copula Model

Eike Brechmann[1], Moritz Heiden[2], Yarema Okhrin[2]

[1]Technische Universität München, Germany; [2]University of Augsburg, Germany

This paper proposes a dynamic framework for modeling and forecasting of realized covariance matrices using vine copulas to allow for more flexible dependencies between assets. Our model automatically guarantees positive definiteness of the forecast through the use of a Cholesky decomposition of the realized covariance matrix. We explicitly account for long-memory behavior by using ARFIMA and HAR models for the individual elements of the decomposition. Furthermore, our model allows for non-Gaussian innovations and GARCH effects, accounting for volatility clustering and unconditional kurtosis.

The dependence structure between assets is studied using vine copula constructions, which allow for nonlinearity and asymmetry without suffering from an excessive number of parameters or symmetry restrictions as in conventional multivariate copula models.

Beside studying in-sample properties, we assess the usefulness of our method in a one-day ahead forecasting framework, comparing recent types of models for the realized covariance matrix. In a Value-at-Risk (VaR) forecasting framework, vine models require less capital requirements due to smoother and more accurate forecasts. Further, we evaluate the model performance by a model confidence set approach based on point forecasts and different loss functions.

524: 3

Total loss estimation using copula-based regression models

Nicole Krämer, Eike C. Brechmann, Daniel Silvestrini, Claudia Czado

Technische Universität München, Germany

We present a joint copula-based model for insurance claims and sizes. It uses bivariate copulae to accommodate for the dependence between these quantities. We derive the general distribution of the policy loss without the restrictive assumption of independence. We illustrate that this distribution tends to be skewed and multi-modal, and that an independence assumption can lead to substantial bias in the estimation of the policy loss. Further, we extend our framework to regression models by combining marginal generalized linear models with a copula. We show that this approach leads to a flexible class of models, and that the parameters can be estimated efficiently using maximum-likelihood. We propose a test procedure for the selection of the optimal copula family. The usefulness of our approach is illustrated in a simulation study and in an analysis of car insurance policies.

All developed methods are implemented in the R-package "CopulaRegression -- Bivariate Copula-Based Regression Models" that is avalaible on CRAN.

524: 4

Options Implied Stock Return Distributions

Ioana Andreea Duca[1,2,3], Maria Grith[1,2], Wolfgang Karl Härdle[1,2]

[1]Humboldt-Universität Berlin, Germany; [2]Center for Applied Statistics and Economics, Berlin, Germany; [3]Academy of Economic Studies Bucharest, Romania

A pricing kernel specification reflecting market preferences in a model with state variable and reference point formation is implemented to adjust the risk neutral density of the stock return dynamics. This yields the density of the returns under the physical measure, which captures market expectations. The novelty of the method is that it allows for the non-monotonicity of the pricing kernel which might lead to improved forecasts of the physical density. Based on the DAX Index and option data at EUREX we investigate the existence of the empirical pricing kernel paradox, i.e. a locally increasing pricing kernel, and estimate the physical density of the DAX Index returns. The results show that our method provides better forecasts of the physical densities across different maturities as compared to methods using strictly decreasing pricing kernel specification.

525: Spatial Statistics

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3043
Session Chair: Thomas Kneib

525: 1

Low-Rank Spatial and Spatio-Temporal Models For Large Datasets

Matthias Katzfuß

Universität Heidelberg, Germany

With the proliferation of modern high-resolution measuring instruments mounted on satellites, planes, ground-based vehicles, and monitoring stations, a need has arisen for statistical methods suitable for the analysis of large spatial datasets observed on large, heterogeneous spatial domains.

Most statistical approaches to this problem rely on low-rank models, for which the process of interest is modeled as a linear combination of spatial basis functions plus a fine-scale-variation term. I describe how low-rank models can be used for global data, spatio-temporal data, and data fusion.

For the full-scale approximation, a type of low-rank model that uses a parent covariance and a set of knots to parameterize the model components, I will discuss two extensions: First, I will describe how to make Bayesian inference on the set of knots. Second, I will argue that it is often advantageous to use a nonstationary parent covariance, and I propose a generalization of the Matern covariance to the sphere that can be used for global data.

The models are applied to satellite [fn]CO_2[/fn] measurements.

525: 2

Spatially adaptive probabilistic weather forecasting using SPDE-INLA and Gaussian copulas

Annette Möller[1], Alex Lenkoski[2], Thordis Thorarinsdottir[2]

[1]University of Heidelberg, Germany; [2]Norwegian Computing Center, Oslo, Norway

Numerical Weather Prediction (NWP) is a method of using physical models to derive weather predictions of future weather conditions. Statistical postprocessing methods that construct predictive distributions from ensembles of NWP outputs are used to correct the NWP forecasts for biases and imperfect representation of the forecast uncertainty. Most of the established postprocessing methods focus on a single weather quantity at a given location, but do not take into account spatial structures of the forecast errors. Model parameters are not estimated in a spatially adaptive manner and forecast errors can only be computed at observation locations, not on the whole model grid. Recently, two different classes of methods have emerged that account for these disadvantages. The first one estimates spatially adaptive model parameters at observation locations and interpolates the forecast errors to the grid. The second one dresses probabilistic forecasts of a spatially non-adaptive model with simulated spatially correlated error fields. The method proposed here combines the advantages of both approaches: It provides spatially adaptive model parameters by assuming a spatial Gaussian field (GF) on each bias-correction parameter in the postprocessing model and employs a Gaussian copula to introduce spatial correlation in the predictive samples. A new method, based on the fact that GFs with Matérn covariance function are solutions to a certain stochastic partial differential equation (SPDE), is employed to benefit computationally from obtaining a GMRF approximation of the GF.

525: 3

Spatially adaptive post-processing of ensemble forecasts for temperature

Michael Scheuerer, Luca Büermann, Gottlieb König

Ruprecht-Karls-Universität Heidelberg, Germany

The introduction of ensemble prediction systems has changed the practice of numerical weather prediction. In order to represent forecast uncertainty, ensemble prediction systems generate several different forecasts of the same weather variable by perturbing initial conditions and model parameters. These forecasts are then interpreted as a sample of a predictive distribution. While such a forecast ensemble offers valuable probabilistic information, it often turns out to be uncalibrated, i.e. it suffers from biases and typically underestimates the prediction uncertainty. Methods for statistical post-processing have therefore been proposed to re-calibrate the ensemble and turn it into a full predictive probability distribution.

Weather variables like temperature depend on factors that are quite variable in space which suggests that post-processing should be done at each site individually. At locations where no measurements for calibration are available, post-processing parameters from neighboring stations must be interpolated if a predictive distribution at this location is desired. In this talk we propose an extension of the non-homogeneous Gaussian regression (NGR) model for temperature post-processing that uses an intrinsically stationary Gaussian random field model for spatial interpolation. This model is able to capture large scale fluctuations of temperature, while additional covariates are integrated into the random field model to account for altitude-related and other local effects.

Based on the dynamical forecasts by the COSMO-DE-EPS, an ensemble prediction system operated by Deutscher Wetterdienst, and

observational data over Germany we evaluate the performance of our method. Our method yields locally calibrated and sharp probabilistic forecasts and compares favorably with other approaches.

525: 4

A new hierarchical copula construction with an application to air pollution modeling

Eike Brechmann

Technische Universität München, Germany

While there is substantial need for dependence models in higher dimensions, most existing models are rather restrictive and barely balance parsimony and flexibility. In this talk, the class of hierarchical Kendall copulas is proposed as a new approach to tackle these problems. By aggregating dependence information of groups of variables in different hierarchical levels using the Kendall distribution function, which is the multivariate analog of the probability integral transform, hierarchical Kendall copulas are able to model complex dependence patterns without severe restrictions.

The talk explicitly discusses properties as well as inference techniques for hierarchical Kendall copulas. A closed-form sampling algorithm is derived for Archimedean copulas, while for general copulas an approximative method is proposed. For estimation, a sequential and a joint approach are discussed.

Due to its negative health effects air pollution today is a major concern. To better understand the joint behavior of pollutants we therefore propose a hierarchical Kendall copula model which explicitly allows for spatial effects between sites and accounts for meteorological effects on air pollution by marginal regression models. As an illustration we analyze [fn]PM_{10}[/fn], [fn]NO_2[/fn] and [fn]CO[/fn] time series from seven cities in south-western California.

References:

Eike C. Brechmann (2012). Hierarchical Kendall copulas: Properties and inference. Preprint, http://arxiv.org/abs/1202.1998.

526: Industrial Statistics

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG III, HS 3042
Session Chair: Sonja Kuhnt

526: 1

Measurement Uncertainty of Immunoassay Concentration Estimates

Katy Klauenberg, Bernd Ebert, Clemens Elster

Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

We developed a Bayesian framework to coherently quantify the uncertainty of immunoassay concentration estimates. This framework encompasses a heteroscedastic non-linear regression and subsequent prediction of an unknown concentration from indirect measurements. Substantive prior knowledge (such as non-negativity of concentrations) usually exists for immunoassays.

Applied to each of 42 fluorescent sandwich ELISAs conducted within the international comparability study CCQM-P58, the Bayesian method (introduced in [1]) results in uncertainties that are considerably larger than previously reported uncertainties in [2]. This leads to different conclusions about the comparability study, e.g. for the certainty, consistency and repeatability of these tests.

An immunoassay is a type of bio-analytical test that exploits the highly specific binding between antibodies and antigens. Immunoassays are able to measure very small concentrations of substance in solutions and have an immense range of applications, e.g. to detect the presence of an infection, of hormones or drugs.

Development of the radioimmunoassay detecting the hormone insulin was crowned by the Nobel Prize in Medicine. Nowadays, immunoassays such as the pregnancy test are produced on an industrial scale. Improving immunoassay procedures in the laboratory or as reference standard (including statistical methods for their evaluation) will ultimately have economic consequences for the manufacturers.

References:

[1] Klauenberg, Ebert, Voigt, Walzel, Noble, Knight and Elster. Bayesian analysis of an international ELISA comparability study. Clin Chem Lab Med 2011;49(9):1459-68.

[2] Noble, Wang, Cerasoli, Knight, Porter, Gray, et al. An international comparability study to determine the sources of uncertainty associated with a non-competitive sandwich fluorescent ELISA. Clin Chem Lab Med 2008;46:1033-45.

526: 2

A variability study of industrial production data

Thomas Mühlenstädt

W. L. Gore & Associates, Germany

A variability study of industrial production monitoring data is presented. Here many different products are rated by the same measurement and the data are jointly modeled. The considered model is the following:

[fn]Y_{ijk} = \beta_i + e_j + e_k[/fn]

where [fn]\beta_i[/fn] is a main effect for product [fn]i[/fn], [fn]e_j[/fn] is a random effect for production run [fn]j[/fn] of product [fn]i[/fn] and [fn]ek[/fn] is a random effect for variability within production run [fn]j[/fn]. A standard model would assume, that [fn]var(e_j) = \sigma^2_{PR} [/fn] and [fn]var(ek) = \sigma^2_e[/fn] and both random effects independent. Since there is a strong heteroscedasticity in the data these model assumptions are not adequate. As it is of special interest to learn about the variability of different products, just applying a log transformation is not adequate here. A more attractive model for the data set at hand assumes [fn]var(e_j) = g(\beta_i)[/fn] and [fn]var(e_k) = h(\beta_i)[/fn], both functions possibly depending on further parameters to describe their shape. Two approaches for this model situation are presented, the first being a two step procedure using the R-library lmer while the second approach directly estimates all parameters via Maximum Likelihood.

526: 3

Optimization of sheet metal forming processes by computer experiments

Momchil Ivanov

Technische Universität Dortmund, Germany

In the automobile industry sheet metal forming is used for the production of automobile body parts. Unfortunately, defects such as, tearing, wrinkling or spring back frequently occur in the formed parts. With increasing computing capabilities it has become customary to perform simulation studies beforehand, where the sheet-metal characteristics can be cheaply optimized. However, simulations generally take a long time to run, ranging from hours to days, making it impossible to perform direct optimization on the computer code. Instead, the simulator can be considered as a black-box function and an approximation model which is cheaper to evaluate is used to interpolate the simulation.

The Efficient Global Optimization algorithm (EGO) has become a standard tool for optimizing black-box functions in simulation experiments. One of the reasons for the popularity of EGO is its close relation with a well-known metamodel, called Kriging. This connection is also possibly a limitation of the method. It is plausible to assume that in certain situations alternative metamodels exist which perform better. A small example of such a scenario is presented in this talk, alongside with a Kriging-alternative metamodel called kernel interpolation.

Although EGO may theoretically be implemented with any metamodel, the generalization is not obvious since the classical method depends on Kriging-specific assumptions. The algorithm also relies on the model-uncertainty predictor. Kernel interpolation has this important feature -- a built in uncertainty measure. The aim of this talk is to introduce a generalization of EGO with kernel interpolation and to compare the latter with the classical EGO algorithm.

526: 4

Automatic Assessment of Hearing Aid Algorithms for Music Signals

Klaus Friedrichs, Claus Weihs

TU Dortmund, Germany

Instead of being dependent on listening tests, a method for automatic assessment of hearing aid algorithms is very valuable to facilitate their optimization process. Therefore, we propose a concept for evaluating signal transformations for music signals with respect to individual hearing deficits by using a computational model of the human auditory periphery. In this model hearing deficits can be simulated by changing specific model parameters. Our idea is using classification and regression methods to estimate the musical attributes rhythm, pitch, loudness and timbre and thereby providing a basis for comparing the outputs of the models with and without hearing loss. Results are presented for the original model and three modified models, each simulating a specific hearing loss.

527: Open Topics - p-values and reporting

Time: Friday, 22nd Mar 2013: 8:50am - 10:10am · Location: KG I, HS 1199
Session Chair: Helmut Küchenhoff

527: 1

Reverse-Bayes analysis of two common misinterpretations of significance tests

Leonhard Held

University of Zurich, Switzerland

Misunderstanding of significance tests and P values is widespread in medical research and elsewhere. I present a reverse-Bayes analysis (Greenland, 2006, Held, 2012) to assess the implications of two common mistakes in the interpretation of statistical significance tests. The first one is the misinterpretation of the type I error rate as the expected proportion of false positive results among all those called significant, also known as the false positive report probability (FPRP). The second is the misinterpretation of a P value as (posterior) probability of the null hypothesis. The analysis shows that the two misinterpretations imply strong and often unrealistic assumptions on the prior proportion or probability of truly effective treatments. The reverse-Bayes approach turns out to be useful to link and contrast frequentist and Bayesian concepts to statistical inference.

References:

[1] Greenland, S. (2006). Bayesian perspectives for epidemiological research: I. Foundations and basic methods. International Journal of Epidemiology 2006;35:765–775

[2] Held, L. (2012) Reverse-Bayes analysis of two common misinterpretations of significance tests. Clinical Trials (to appear).

527: 2

Decision theory as an alternative to testing many hypotheses

Nicholas Tibor Longford

SNTL, Spain

Testing a large number of hypotheses is a key problem in the analysis of microarray experiments and in other studies in which many elementary experiments are conducted, and the exceptions among them, for which a particular hypothesis does not hold, have to be identified. A class of approaches to this problem is based on controlling the false discovery rate, even though failure to discover should also be considered. We develop a decision-theoretical approach in which errors (misclassifications) of the two kinds are associated with uneven losses, and the total expected loss in the collection of the classifications (decisions made) is minimised. This paper is an output of a research programme to supplant hypothesis testing in mainstream statistics. Hypothesis testing is seen as deficient because it is oblivious to the consequences of the two kinds of errors that can be made.

527: 3

Deriving prediction rules in omics data requires collaboration between different disciplines – on weaknesses in reporting and statistical analysis in a large study

Sam Doerken, Anika Buchholz, Willi Sauerbrei

Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany

Hundreds of prediction models based on omics data have been developed and hundreds of methodological papers are published on issues around deriving, validating and applying methods for prediction. Surprisingly, several key issues known from classical statistics seem to be often ignored in the "high dimensional community".

Using clustering methods, the METABRIC group has developed and validated in independent data novel subgroups for patients with breast cancer [1]. However, presenting results concerning the prognostic value of the derived subgroups weaknesses become obvious. The group has made their omics and clinical data publicly available, allowing re-analysis. Severe weaknesses in analysis (eg. Cox proportional hazards model used despite of strong violations of the PH assumption) and reporting (eg. it is unclear which specific model was used for adjustment of the estimated effects in the main model) will be illustrated. Issues of reporting and more suitable analyses will be presented.

References:

[1] Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486:346-352.

The contribution has been withdrawn.

Bayesian Inference about Odds Ratio Structure in Ordinal Contingency Tables

Maria Kateri[1], Alan Agresti[2]

[1]RWTH Aachen University, Germany; [2]University of Florida, USA

530: Invited Session: Meta Analysis (Copas, Stanley)

Time: Friday, 22nd Mar 2013: 10:40am - 12:00pm · Location: KG II, Audimax
Session Chair: Gerta Rücker
Session Chair: Katja Rost

530: 1

A joint sensitivity analysis for outcome reporting bias and publication bias in meta analysis

John Copas

University of Warwick, United Kingdom

Most systematic reviews concentrate on a single outcome of interest (y) and only consider published studies. But most research studies consider multiple outcomes, and not all studies are published. By only reviewing published studies which report y, meta analysis risks two important sources of bias: outcome reporting bias (for example when studies only report significant outcomes), and publication bias (studies reporting significant outcomes are more likely to be accepted for publication). Ignoring these biases can lead to misleading estimates and exaggerated treatment effects. My talk will extend some recent work in Kirkham et al (British Medical Journal 2010) and Copas (Applied Statistics 2012) and suggest a new statistical model for sensitivity analysis covering both sources of bias. Some theory and examples will be dicussed.

530: 2

Can meta-analysis falsify theory?

Tom Stanley

Hendrix College, United States of America

The question of whether theory is genuinely testable has always been contentious with deep and unavoidable philosophical undertones. If anything can falsify an economic theory, then two logically linked meta-analyses of mature areas of research can. Meta-analyses of two core economic theories — the theory of the firm and the natural rate (of unemployment) hypothesis — along with meta-analyses of their ‘falsifying hypotheses — the efficiency wage hypothesis and unemployment hysteresis—are used to illustrate the issues at stake. It is claimed that the natural rate hypothesis has been falsified in a sophisticated Popperian sense.

531: Statistical Teaching and Consulting

Time: Friday, 22nd Mar 2013: 10:40am - 12:00pm · Location: KG III, HS 3043
Session Chair: Geraldine Rauch

531: 1

EMOS – The European Master in Official Statistics a Vision Infrastructure Project

Markus Zwick

Eurostat, Luxembourg

EMOS is an infrastructure project aiming at developing a program for Training and Education in Official Statistics. The idea is to set up conditions for labelling courses and university programmes, which would lead to a qualification in European Official Statistics. EMOS is a joint project of different stakeholders (NSI, Eurostat, Universities and National Schools in statistics) with the aim of reaching a higher level of knowledge in various ways:

1.) Firstly, statistical producers could benefit from young and well-qualified researchers in official statistics.

2.) Secondly, other organisations with a link to statistics (ministries, central banks, research institutes, consulters etc.) could acquire better qualified staff in statistics on the labour market.

3.) A third point is that NSI and Universities stand to learn a great deal from each other through having this project in common.

In spring 2012, Eurostat launched the Call for Tender for the feasibility study 'Towards a European Master in Official Statistics'. The purpose of this study is to contribute to the creation of a European Master in Official Statistics and to create a network of programmes dealing with Masters in Official Statistics at European level.

The feasibility study should start in autumn 2012 and should draw to a close twelve months later with a final technical report. Based on the feasibility study and assuming that the universities and relevant stakeholders remain interested and the systems of education across Europe are not too different, it is expected that the first courses of EMOS will start in the winter semester of 2014.

531: 2

Garanten für den Lernerfolg?!

Aline Naumann

Eberhard Karls Universität Tübingen, Germany

Es fällt nicht immer leicht, fachfremden Zuhörern statistische Inhalte und Denkweisen zu vermitteln. Für den Lernerfolg des Zuhörers sind nicht nur die Inhalte, sondern auch die Art und Weise der Präsentation von enormer Bedeutung. Daher versuchen wir uns stets zu verbessern. Aber was ist das Ziel? Auf welcher Lernebene (Hören, Lesen, Sehen, Handeln) erreichen wir unsere Zuhörer?

Befragt werden sowohl Studierende und Mitarbeiter der Universität Tübingen, als auch Angestellte in der freien Marktwirtschaft. Ein Fragebogen gibt den Teilnehmern die Möglichkeit, im statistischen Kontext ihre bevorzugte Lernebene und den bevorzugten Aufbau einer Lehrveranstaltung mitzuteilen. Zudem soll unter anderem die Wichtigkeit von Formeln, Graphiken und Beispielen beim Erlernen statistischer Inhalte beurteilt werden.

Nun gilt es diese Ergebnisse zu präsentieren und zu diskutieren. Was wird als wichtig und was als unwichtig angesehen? Sollen wir in Zukunft auf diese „unwichtigen“ Aspekte verzichten oder ist genau das unser Ziel: Unsere didaktischen Fähigkeiten so zu verbessern, dass auch diese Aspekte zum Lernerfolg beitragen?

531: 3

Neue Wege um Biometrie zu vermitteln -- ein Elearning-Angebot zu biometrischen Themen

Marius Wirths, Geraldine Rauch

Institut für Medizinische Biometrie und Informatik, Universität Heidelberg, Germany

Biometrie wird innerhalb verschiedener Studiengänge der Medizinischen Fakultät Heidelberg gelehrt. Innerhalb dieser Veranstaltungen werden Übungen derzeit in Form von schriftlichen Aufgaben durchgeführt, entweder direkt während der Vorlesung oder als Hausaufgaben. In den Evaluationen der verschiedenen Kurse wurde von den Studierenden immer wieder der Wunsch nach mehr Übungsaufgaben und zugehörigen Lösungen geäußert, um auch nach Ende der Vorlesung eigenständig weiterarbeiten und um sich auf Klausuren vorbereiten zu können. Aus diesem Grund wird am Institut für Medizinische Biometrie und Informatik an einem Elearning-Tool gearbeitet, welches interaktive Übungsaufgaben anbietet. Die Studierenden haben so die Möglichkeit, den Lernstoff dem eigenen Lerntempo angepasst zu wiederholen und ihr Wissen abzuprüfen. Für den Dozenten entfällt das mühsame Korrigieren von Übungsaufgaben.

Im Vortrag wird das entwicklete Elearning-Tool anhand von Beispielaufgaben vorgestellt und die technische Umsetzung erläutert.

531: 4

Die Notwendigkeit und das Handwerk des Weglassens -- Statistiklehre für Nichtstatistiker

Hubert Merkel

HAWK Hildesheim/Holzminden/Göttingen, Germany

Statistiklehre für Nichtstatistiker ist zeitlich immer so begrenzt, dass eine exakte mathematische Grundlegung nicht möglich ist. So wünschenswert dies grundsätzlich wäre, kann man aber feststellen, dass bei einer angemessenen Lernzieldefinition diese Notwendigkeit (zunächst) gar nicht besteht.

Schliesslich müssen Nichtstatistiker nicht die Verfahren selbst, sondern "nur" ihre korrekte Anwendung verantworten. Sie müssen keine neuen Verfahren entwickeln, aber die Möglichkeiten und Grenzen kennen, die die Statistik bietet, um ihre Probleme zu lösen.

Den resultierenden didaktischen Überlegungen liegen die Theorie des Begriffslernens und das Bild einer kognitiven Struktur als Netzwerk zu Grunde.

Begriffe lassen sich nicht nur von ihren Grundlagen her systematisch aufbauen, sondern auch aus einer realen Problemsituation heraus ausdifferenzieren.

Wenn zu lernende Begriffe so aufgebaut sind, dass sie logisch und nachvollziehbar die relevanten Zusammenhänge abbilden, dann ist auch ohne die Kenntnis der "letzten Gründe" ein Grundverständnis und auch die sinnvolle Anwendung eines Sachverhalts möglich. Dies ist mit wenig, oft (zunächst) sogar ohne sichtbare Mathematik möglich.

Die Darstellung und Ordnung dieser Begriffswelt muss aber so organisiert sein, dass sie schliesslich immer auch mit Blick auf die Mathematik erweiterbar ist.

Dies ist grundsätzlich, vor allem aber vor dem Hintergrund einer Studienstruktur aus Bachelor- und Masterstudiengängen und der damit möglichen Umorientierung zu einem Grundlagenfach nach dem angewandten Erststudium, notwendig.

Und es ist möglich.

Beispiele werden vorgestellt.

532: Statistics in Finance, Insurance and Banking

Time: Friday, 22nd Mar 2013: 10:40am - 12:00pm · Location: KG I, HS 1098
Session Chair: Edo Schinzinger

532: 1

An new test for Benford's distribution

Dieter William Joenssen

University of Technology Ilmenau, Germany

The usage of Benford's law assists auditors in revealing data manipulation in a wide range of areas, from tax audits and corporate accounting to election fraud. Benford's law dictates that the frequency distribution of the first digits, from certain numerical data, follows a discrete distribution known as the Benford distribution. Operationally, goodness-of-fit tests are used to assess deviations, in the data's first digit's distribution, from the Benford distribution, implying that when the null hypothesis is rejected, data manipulation could be present and further investigation is warranted. Goodness-of-fit tests, like all tests of statistical significance, are prone to the type I error, which is limited by the chosen level of significance, and the type II error, which decreases not only with sample size but is also inherently lower for some testing procedures than for others. Thus, using a test with higher inherent power, in this framework, directly translates to a greater ability to detect data manipulation. Current state of the art, when assessing whether or not data conforms to Benford's law, is the usage of the $\chi 2$ and Kolmogorov–Smirnov tests, which are widely known due to their intuitiveness and simplicity but also lack of power in comparison to specialized tests. A new method of testing for Benford's law, similar to the Shapiro-Francia test for normality, is introduced together with critical values needed for statistical testing. Additionally, a simulation study shows that the new method is competitive in terms of power for a range of theoretical distributions.

532: 2

Limits of Bayesian decision related quantities of nonhomogeneous binomial financial processes

Wolfgang Stummer

University of Erlangen-Nürnberg, Germany

We study Bayesian decision making based on observations ([fn]X(t): t=0[/fn], [fn]T/n[/fn], [fn]2T/n[/fn], ... , [fn]nT/n[/fn]) of the discrete-time dynamics of an arbitrary financial quantity of interest, when the hypothesis a nonhomogeneous [fn]n[/fn]-period binomial model and the alternative is a different [fn]n[/fn]-period binomial model. As the observation gaps tend to zero (i.e. [fn]n[/fn] tends to infinity), we obtain the limits of the corresponding Bayes risk (minimal mean decision loss) and some related notions of model risk.

532: 3

Modeling Real Estate Data using Semiparametric Quantile Regression

Alexander Razen, Stefan Lang

University of Innsbruck, Austria

The Basel II framework strictly defines the conditions under which financial institutions are authorized to accept real estate as collateral in order to decrease their credit risk. A widely used concept for its valuation is the hedonic approach. It assumes, that a property can be characterized by a bundle of attributes that involves both structural covariates (the floor space area, the age of the building, etc.) and locational covariates (the proximity to places of work, the buying power index in the respective region, etc.). Each of these attributes can be assigned an implicit price, summing up to the value of the entire property.

With respect to value-at-risk concepts financial institutions are often not only interested in the expected value but in different quantiles of the distribution of real estate prices. For this reason we use quantile regression and develop a hierarchical semiparametric model that involves linear, nonlinear and spatial effects. Nonlinear effects are modeled with P-splines, spatial effects are represented by Gaussian Markov random fields. Due to the high complexity of the model the final estimations are based on a Bayesian approach. The results are analyzed in detail and then compared with those of a classical mean regression approach under the normal distribution assumption. The remarkable differences for some covariates suggest the use of a model selection criterion that finally shows the quantile regression to be superior.

532: 4

Combining Generalized Linear Models with Credibility

Edo Schinzinger

Ulm University, Germany

In actuarial science, generalized linear models (GLM) and Bühlmann's credibility models are both useful tools for rate making.

The GLM allows the first moment to be related to a linear model via a link function [fn]g[/fn], i.e. [fn]g(E[X])=Y b[/fn].

On the other hand, Bühlmann's credibility model considers a portfolio of I risks whose claim sizes [fn]X_i[/fn] are characterized by iid drawn risk profiles [fn]T_i[/fn]. Then, the premiums [fn]E[X_i T_i][/fn] are estimated.

In the application of stochastic mortality modeling, the rates of Poisson distributed death counts are studied. The GLM provides a decomposition of the log-rates into age effects [fn]Y[/fn] and period effects [fn]b[/fn]. If [fn]T_i[/fn] is assumed to be the country specific risk profile, a credibility approach for GLM will give

[fn]g(E[X_i T_i]) = Y b(T_i)[/fn]

We focus on Poisson GLM with the log-link function. It turns out that, asymptotically for a large number of observations, the least square estimator for [fn]b(T_i)[/fn] is a weighted sum of the individual posterior and the collective prior estimators. For the estimation of the weight matrix; we will make use of the asymptotic behavior of maximum likelihood estimators of the classical GLM.

Finally, first results for West European data are presented.

533: Spatial Statistics

Time: Friday, 22nd Mar 2013: 10:40am - 12:00pm · Location: KG I, HS 1199
Session Chair: Thomas Kneib

533: 1

Determining high-risk zones for unexploded bombs by using spatial point process methodology – concepts, application and extensions

Monia Mahling, Michael Höhle, Helmut Küchenhoff

Ludwig-Maximilians-Universität München, Germany

To prevent accidents in Germany caused by unexploded bombs from the Second World War, high-risk zones need to be determined. If suitable aerial pictures of the area in question exist, statistical methods can be used to determine such zones by considering patterns of exploded bombs as realizations of spatial point processes.

Two methods to determine such zones are presented. The first method is based on the intensity of the point process, the second method on its nearest-neighbour distance. Both methods are applied to patterns of bomb craters derived from aerial pictures taken by the allies during and after World War II. In addition, a risk assessment is performed by estimating the probability that there are unexploded bombs outside the high-risk zone. The consequences of clustering are investigated.

Finally, extensions such as the development of high-risk zones for incompletely observed bomb crater patterns or for patterns with a spatially varying probability of non-explosion are discussed.

References:

[1] Mahling, M., Höhle, M. and Küchenhoff, H. (2013). Determining high-risk zones for unexploded World War II bombs by using point process methodology, Journal of the Royal Statistical Society, Series C, published online 27th July 2012. DOI: 10.1111/j.1467-9876.2012.01055.x

533: 2

EEG enhanced fMRI activation detection

Stefanie Kalus, Ludwig Fahrmeir

Ludwig-Maximilians-Universität München, Germany

In human brain mapping there is a strong interest in combining functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) recordings to fuse complementary information about the location of neuronal activity. We propose an advanced statistical model for enhancing fMRI activation detection by the use of EEG-based spatial prior information in stimulus based experimental paradigms, which is based on a spatial Bayesian variable selection in voxelwise linear regressions. We link selection resp. activation probabilities to a predictor via a latent probit model. Mere fMRI activation detection is achieved by using a predictor with a spatially-varying intercept only. For EEG-enhanced schemes, an EEG effect is added, which is either chosen to be spatially-varying or constant. Spatially-varying effects are regularized by either of two Gaussian Markov Random Field (GMRF) prior types. Statistical inference in resulting high-dimensional hierarchical models is based on Markov Chain Monte Carlo approaches, providing posterior estimates of activation probabilities and enhancing formation of activation clusters. Three basic algorithms are proposed depending on GMRF type and update scheme. An application to an active acoustic oddball experiment and a simulation study show that mere fMRI activation detection models achieve a substantial increase in sensitivity compared to benchmark approaches. Carefully selected EEG-prior information additionally increases sensitivity in activation regions that have been distorted by a low signal-to-noise ratio.

533: 3

Doubly stochastic point processes for the analysis of 3D spatial distribution of marked proteins in the cell nucleus

Volker J. Schmid

Ludwig-Maximilans-Universität München, Germany

Recently developed high-resolution fluorescence microscopy (3D-SIM) allows to spot marked proteins in the cell nucleus (Schermelleh et al. 2008). As the architecture of nuclei and the function of certain proteins is still not fully understood, it is of interest to analyze the spatial distribution of these proteins, the interaction between proteins and the interaction of proteins with chromatin (DNA fibres).

To this end, we will present classic approaches as well as two different doubly stochastic point processes in order to model the observed point patterns. For the latter approach we adopt the idea of constructed covariates modelling the interaction between observed points (Illian et al. 2013).

Although the theory of spatial point processes can easily be extended from usual 2D problem into 3D space, in practice a lot of technical and computational challenges arise. We will present computationally feasible methods to analyse the non-stationary marked point processes, along with methods to actually identify protein locations from the 3D-SIM images in a pre-processing step.

References

Schermelleh L, Carlton PM, Haase S, Shao L, Winoto L, Kner P, Burke,B, Cardoso MC, Agard DA, Gustafsson MG, Leonhardt H and Sedat JW (2008). Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. Science, 320, 1332-6.

Illian J, Soerbye S, Rue H (2013). A Toolbox for Fitting Complex Spatial Point Process Models Using Integrated Nested Laplace Approximation (INLA). Annals of Applied Statistics, in print.

533: 4

Spatial modelling of Ras protein structures on the cellular membrane

Martin Schäfer[1], Peter Verveer[2], Katja Ickstadt[1]

[1]Technische Universität Dortmund, Germany; [2]Max Planck Institute of Molecular Physiology, Dortmund, Germany

In cellular signal transduction, it is assumed that spatial effects, e.g., gradients, spatial trends or clustering, play a pivotal role. An example are clusters of GTPase Ras, a small protein adherent to the plasma membrane, which occur in small regions measuring up to approximately 20nm in diameter and contain 4 to 10 proteins. Stochastic effects are believed to be of relevance. The cluster size influences the cellular signal transmission, e.g., suppressing the Ras clustering leads to an inhibition of the signal transduction. As mutated Ras proteins have been linked to different human tumors, the cluster structure is of interest for biomedical research.

In this work, spatial effects and stochastic phenomena are modelled using Bayesian concepts. In particular, Binomial/Beta and Dirichlet process mixture models are combined for application to the cellular signal transduction. Specifically, clustered and single proteins are distinguished by a Binomial/Beta mixture, while parameters of the cluster structure such as cluster size and radius are inferred by additional model aspects. To complete a comprehensive hierarchical model framework, a nonparametric Dirichlet process mixture of multivariate normals eventually makes use of these estimates to infer the intensity surface of the underlying spatial point process.

Results are useful for comparing cells across different experimental conditions, e.g., healthy and tumorous cells.

534: Open Topics

Time: Friday, 22nd Mar 2013: 10:40am - 12:00pm · Location: KG III, HS 3042
Session Chair: Katja Ickstadt

534: 1

From Spearman's general factor to scree plots and beyond

Nanny Wermuth

Chalmers Technical University, Gothenburg, Sweden

When are associations between a set of items well captured by a single hidden variable? And when is a simple sum score a good summary measure for an individual? A survey of answers and concepts is given that have been developed separately but turn out to be closely related. These include tetrad conditions, M-matrices, scree-plots Cronbach's alpha, the Rasch model and traceable regressions for graphs representing complete independence of items given a hidden variable.

References:

[1] Wermuth, N. (2012). Traceable regressions. Intern Statist Review, 80, 415-438.

534: 2

Calibration Tests for Count Data

Wei Wei, Leonhard Held

Institute of Social and Preventive Medicine, University of Zurich, Switzerland

Calibration, the statistical consistency of forecast distribution and observations, is a central requirement for probabilistic predictions. Calibration of continuous forecasts has been widely discussed, and significance tests have proven to be a powerful tool to detect whether a model is in lack of calibration. However, a significance test for discrete forecasts is rarely seen, especially for distributions with unlimited support. In this talk, we propose three types of significance tests for count data: an unconditional test and a regression test based on scoring rules, and a test based on conditional exceedance probabilities. In particular, we will focus on three popular scoring rules: the ranked probability score, the logarithmic score and the Dawid-Sebastiani score. The simulation studies show that all three types of tests are robust with good control of type I error and sufficient power under miscalibration. As an illustration, tests are implemented for two regression models: the Poisson and the negative binomial regression in patent data about research spending and number of registered patents. The results show that the test approach is powerful in detecting miscalibrated forecasts.

534: 3

Comparing Generalized linear mixed-effects models with adaptive Gauss Hermite and integrated Laplace approximation

Rafael Sauter, Leonhard Held

University of Zurich, Switzerland

Clustered observations, such as longitudinal data are preferably analyzed with mixed effects models. Non-normal responses, belonging to the exponential family, are modeled in analogy with the generalized linear model by generalized linear mixed effects models (GLMM). Maximum likelihood estimation (MLE) of a GLMM is usually based on the marginal likelihood and requires numerical integration. A widely used software package for R, which allows for MLE of GLMM's, is the lme4 package by Bates et al. (2011). Although attractive in its vast modeling possibilities the performance and numerical stability of lme4 for GLMM's is sometimes a cause of critique (Zhang et al., 2011). Bayesian estimation of GLMM's can be done by the R package INLA. This package implements the integrated nested Laplace approximation (INLA) method introduced by Rue et al. (2009), which is a fast and accurate approximation method and an alternative to MCMC. We apply lme4 and INLA to a binary response GLMM for toenail infection data resulting from a clinical trial. Numerical problems and variation in estimates arise for lme4, depending on the choice of the number of Gauss Hermite quadrature points. The simplified Laplace approximation strategy in INLA returns stable estimates in short time. We found comparable results for random intercept models with different prior specifications but increasing differences with more complex random effects structures. Fixing the random effects hyperparameters at the MLE's results in decreased differences between INLA and glmer.

534: 4

Multiple Point Hypotheses Testing Problems

Jens Stange

Humboldt University Berlin, Germany

Multiple testing problems arise in many applications, e.g. in genetics. A typical task is the control of an appropriate type I error measure, analogously to single hypothesis testing. A well established measure for the type I error is the Familywise Error Rate (FWER). That is the probability, that at least one true hypothesis is falsely rejected.

Control of the FWER means that this probability is bounded by a given significance level under every possible constitution of true and false hypotheses. Weak control means that this probability is bounded by the significance level under the assumption that all hypotheses were true (the global hypothesis).

Here a special class of multiple testing problems is considered, where each of the hypotheses is a point hypothesis, i.e. each marginal test has a two-sided alternative. Under specific assumptions, e.g. the subset pivotality condition, introduced by Westfall and Young, it is shown that the global hypothesis becomes the least favourable configuration with respect to the FWER. This yields the convenience that control of the FWER is already implied by weak control.

Some examples of such testing problems will be discussed.

540: Closing session (Invited Speaker: Meng)

Time: Friday, 22nd Mar 2013: 12:10pm - 1:10pm · Location: KG II, Audimax
Session Chair: Göran Kauermann

540: 1

I got more data, my model is more refined, but my estimator is getting worse! Am I just dumb?

Xiao-Li Meng

Havard University, United States of America

Possibly, but more likely you are merely a victim of conventional wisdom. More data or better models by no means guarantee better estimators (e.g., with smaller mean squared error), when you are not following probabilistically principled methods such as MLE (for large samples) or Bayesian approaches. Estimating equations are particularly dangerous in this regard, almost a necessary price for their robustness. These points will be demonstrated via common tasks of estimating regression parameters and correlations, under simple models such as bivariate normal and ARCH(1). Some general strategies for detecting and avoiding such pitfalls are suggested, including checking for self-efficiency (Meng, 1994, Statistical Science) and adopting a guiding working model.

Of course, Bayesians are not automatically immune either to being a victim of conventional wisdom. A simple example is given in the context of a stationary AR(1) model where the so-called “non-informative” Jeffreys prior can get arbitrarily close to a point mass at a unit root, hardly non-informative by any measure.

This talk is based on Meng and Xie (2013, Econometric Reviews).