MEDDELELSER
Dansk Selskab for Teoretisk Statistik

Kalender 2007

Dato	No.	Aktivitet
11/9	6	Biostatistisk Afdeling, Københavns Universitet. Jason Fine, University of Wisconsin, Madison: *Quantile Inference for Competing Risks Data*
13/9	6	Biostatistisk Afdeling, Københavns Universitet. Jason Fine, University of Wisconsin, Madison: *Nonparametric Association Analysis of Multivariate Competing Risks Data, with Application to Dementia Onset in an Aging Population*
17-19/9, 8-10/10	1	Forskerskolen i Biostatistik, København. *Statistical Analysis of Survival Data for Biostatistical/Statistical PhD students*
19/9	6	Institut for Matematiske Fag, Københavns Universitet. Jiti Gao, University of Western Australia: *A New Test in Parametric Linear Models with Nonparametric Autoregressive Errors*
20/9	6	Thiele Centre, University of Aarhus. Filip Lindskog, Royal Institute of Technology, Stockholm: *Title to be announced*
21-23/9	6	Thiele Centre, University of Aarhus / Sandbjerg Estate, Sønderborg Conference: *Symposium on Future Stereology*
25/10	6	Thiele Centre, University of Aarhus. Ole E. Barndorff-Nielsen, University of Aarhus: *Upsilon Transformations*
5-9/11	3	Department of Mathematical Sciences, University of Copenhagen Ph.D.-course: *Statistical Analysis of Microarray Expression Data with R and Bio-conductor*
20-21/11	6	DSTS, DTU, Lyngby *DSTS 2-day Meeting*

No.: Nummer af meddelelser hvor arrangement er annonceret.

Kontingent og Betalingsservice

Det er nu muligt at betale kontingent for medlemskab af DSTS via Betalingsservice.

Bestyrelsen opfordrer alle medlemmer til at benytte sig af denne mulighed, samt huske at man ved adresse- eller status-ændring skal ændre det på DSTS's hjemmeside (http://www.dsts.dk/da) eller underrette kasseren (kass@dsts.dk) og webmasteren (web@dsts.dk) herom.

Nyt Om Navne

Niels Bindslev er pr. 1. juli 2007 blevet ansat som Senior Biostatistiker i Cyncron Biometrics.

Stine Segel er pr. 1. september 2007 blevet ansat som biostatistiker ved Kardiovaskulært Forskningscenter, Aalborg Sygehus. Stine kommer fra en stilling på Novo Nordisk A/S.

Anne Vingård Olesen er pr. 1. september 2007 blevet ansat som biostatistiker ved Enheden for Psykiatrisk Forskning i Region Nord. Anne kommer fra en stilling på Århus Universitet.

Næste nummer af "MEDDELELSER" udkommer 1. oktober 2007.
Bidrag skal være redaktøren i hænde senest **den 21. september kl. 12.00.**

Deadlines i år 2007

	MEDDELELSER udkommer	Frist for indlevering af bidrag:
7:	1. oktober	21. september
8:	5. november	26. oktober
9:	3. december	23. november

Seminarer i anvendt statistik

Professor Jason Fine
University of Wisconsin, Madison

besøger Biostatistisk Afdeling og holder to seminarer. Seminarerne afholdes kl. 15.15 på det gamle Kommunehospital, Øster Farimagsgade 5, opgang B. Der serveres te i Biostatistisk Afdelings bibliotek (opgang B, 2. sal) en halv time før.

Tirsdag d. 11. september 2007, lokale 5.0.28 (bemærk ugedag).

Quantile Inference for Competing Risks Data

A conceptually simple quantile inference procedure is proposed for cause specific failure probabilities with competing risks data. The quantiles are defined using the cumulative incidence function, which is intuitively meaningful in the competing risks set-up. We establish the uniform consistency and weak convergence of a nonparametric estimator of this quantile function. These results form the theoretical basis for extensions of standard one-sample and two-sample quantile inferences for independently censored data. This includes the construction of confidence intervals and bands for the quantile function, and two sample tests. Simulation studies and a breast cancer example illustrate the practical utility of the methodology.

Torsdag d. 13. september 2007, lokale 5.1.34 (bemærk ugedag).

Nonparametric Association Analysis of Multivariate Competing Risks Data, with Application to Dementia Onset in an Aging Population

While nonparametric association analyses of bivariate failure times have been widely studied, analogous analyses of bivariate competing risks data have not been investigated. Such analyses are important in familial association studies in genetic epidemiology and demography, where multiple interacting failure types may invalidate nonparametric analyses for independently censored clustered survival data. The scenario is common in population based studies where onset of certain chronic diseases, eg, psychiatric disorders, may be dependently censored by death. I first develop nonparametric estimators for the bivariate cause-specific hazards function and the bivariate cumulative incidence function, which are natural extensions of their univariate counterparts and make no assumptions about the dependence of the risks. The estimators are shown to be uniformly consistent and to converge weakly to Gaussian processes. Time-dependent summary association measures are proposed and yield formal tests of independence in clusters. The practical utility of the methodology is illustrated in an analysis of dementia in the Cache County Aging Study, where dependent censoring by mortality is heavy and the onset associations are strongly time-varying.

Per Kragh Andersen

Institut for Matematiske Fag
Københavns Universitet
Universitetsparken 5, 2100 Kbh. Ø

Seminar i matematisk statistik og sandsynlighedsregning

Seminaret afholdes kl. 15.15 i auditorium 10 på H.C. Ørsted Instituttet.
Efter seminaret serveres der te og chokolade i lokale E325.

Onsdag den 19. september 2007:

Speaker: Jiti Gao, University of Western Australia
Title: A New Test in Parametric Linear Models with Nonparametric Autoregressive Errors

Abstract:

This paper considers a class of parametric models with nonparametric autoregressive errors. A new test is proposed and studied to deal with the parametric specification of the nonparametric autoregressive errors with either stationarity or nonstationarity. Such a test procedure can initially avoid misspecification through the need to parametrically specify the form of the errors. In other words, we propose estimating the form of the errors and testing for stationarity or nonstationarity simultaneously. We establish asymptotic distributions of the proposed test. Both the setting and the results differ from earlier work on testing for unit roots in parametric time series regression. We provide both simulated and real-data examples to show that the proposed nonparametric unit-root test works in practice.

THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Department of Mathematical Sciences
University of Aarhus

Activities at the Thiele Centre

Seminars:

Thursday 20 September 2007, 14:15, building no. 1531, room Koll.D
Filip Lindskog, Royal Institute of Technology, Stockholm:
Title to be announced.

Thursday 25 October 2007, 14:15, building no. 1531, room Koll.D
Ole E. Barndorff-Nielsen, University of Aarhus:
Upsilon Transformations

ABSTRACT: Stochastic integrals of deterministic functions are infinitely divisible and thus, in particular, associate the Lévy measure of the Lévy process to the Lévy measure of the integral. For certain types of integrands the mapping thus established has interesting special properties and is referred to as an Upsilon transformation. More broadly, this term is used for injective regularising mappings on the class of Lévy measures into itself. The talk will survey the properties of such transformations, relating to classical and free infinite divisibility.

Conferences:

21-23 September 2007
Sandbjerg Estate, Sønderborg
Symposium on Future Stereology
http://www.thiele.au.dk/Stereology/

29 January – 1 February 2008
Sandbjerg Estate, Sønderborg
Workshop on Stochastics in Turbulence and Finance

14-18 July 2008
Sandbjerg Estate, Sønderborg
Conference on Efficient Monte Carlo: From Variance Reduction to Combinatorial Optimization. A Conference on the Occasion of R.Y. Rubinstein's 70th Birthday
http://www.thiele.au.dk/Rubinstein/

DSTS 2-day Meeting 20-21 November 2007, DTU, Lyngby.

Programme:

Tuesday 20. November, before lunch:
In line with the previous DSTS 2-day meeting there will be a mini workshop within the topic
"Point process modelling and statistical inference". Everyone interested are welcome to participate.
Detailed programme for this will follow in a later issue of DSTS Meddelelser. The mini workshop
will, just as the 2-day meeting, take place at DTU (same venue) and will finish with a free lunch.

Jesper Møller, Institut for Matematiske Fag, Aalborg Universitet.

Tuesday 20. November

14.00-14.45 **Hans-Peter Piepho,** Professor and leader of the Bioinformatics Unit, University of
            Hohenheim:
            Issues in the design and analysis of plant breeding and variety trials.
14.45-15.15 Coffee
15.15-16.00 **Lars Kai Hansen,** Professor in Intelligent Signal Processing, IMM, DTU:
            Learning to detect signals from labeled and unlabeled data
16.00-16.45 **Per Chr. Hansen,** Professor in Scientific Computing, IMM, DTU:
            Large-Scale Ill-Posed Problems and Applications
16.45-18.15 Beer/water outside the auditorium.

19.00-   Dinner (Downtown/Frederiksberg area)

Wednesday 21. November

09.00-09.45 **Mogens Bladt,** Guest Professor **& Bo Friis Nielsen,** Assoc. Professor, IMM, DTU:
            Distributions with rational Laplace transform
09.45-10.30 **Thomas Gerds,** Assoc. Professor, Dep. of Biostatistics, Copenhagen University: The
            VAML game [Validation and Assessment of Machine Learning]
10.30-11.00 Coffee
11.00-11.45 **Julie Lyng Forman,** Dep. of Biostatistics, Copenhagen University:
            Inference from Diffusion Driven Models.
11.45-12.30 **Uffe Høgsbro Thygesen,** Senior Researcher, Danish Institute for Fisheries Research,
            DTU:
            Markov models of animals moving in the sea
12.30-13.00 Sandwiches

Venue:
Aud 31, Building 306, DTU Campus, Lyngby (Matematiktorvet, 2800 Kgs. Lyngby)

Registration:
You register by sending an email to Ellen Borup (emb@imm.dtu.dk), IMM no later than 2. Nov.
2007. Please indicate clearly whether you want to participate in the mini workshop. The fee is DKK
225 for students (excl. Ph.D. students) and DKK 450 for all others. The fee is to be paid to: DSTS,
Jyske Bank, Reg. no. 7853, Account no 1117188. Please indicate clearly the person(s) for which the
payment is intended.

Abstracts:

Hans-Peter Piepho: Issues in the design and analysis of plant breeding and variety trials.

Plant breeding and variety testing relies heavily on field experiment conducted in multiple
environments. This talk will cover both the analysis of individual trials and the joint analysis of
series of trials.

Most agricultural field trials involve a layout of plots in rows and columns. Resolvable row-column
designs have proven effective in obtaining efficient estimates of treatment effects. Further
improvement may be possible by post-blocking or by inclusion of spatial model components. This
paper reviews options for augmenting a baseline row-column model for individual trials by the
addition of spatial components having a linear variance structure (Lee and Piepho, 2007; Piepho
and Williams, 2007). The main focus will be on spatial variance-covariance structures that have a
linear form, including separable linear variance models and simple state-space models such as
random walks (Pieho and Ogutu, 2007). Relationship to nearest-neighbour methods based on first
and second differences will be discussed. The usefulness of these options is assessed by analysing a
number of field trials, most of which are uniformity trials.

When integrating results from a series of field experiments, modelling of genotype-by-environment
interaction becomes the prime issue. This paper will review mixed models for this purpose,
focussing on factor-analytic variance-covariance structures (Piepho, 1997), which have a fixed
effects counterpart known as the "Additive Main Effects Multiplicative Interaction" (AMMI)
model. Such models may be used to exploit similarities among different agro-ecological regions
(Piepho and Möhring, 2005), and tey are now being implemented in the German variety testing
system as will be exemplified in the talk. I will also discuss the validity of standard mixed model
analysis as dependent on the missing data pattern (Piepho and Möhring, 2006) as well as the merits
and demerits of a two-stage analysis, where in the first step adjusted genotype means from
individual trials are computed, which are then summarized by a joint analysis across the series in
the second step (Piepho and Möhring, 2007).

References

Piepho, H.P. (1997): Analyzing genotype-environment data by mixed models with multiplicative
     effects. *Biometrics* **53**, 761-766.
Piepho, H.P., Möhring, J. (2005): Best linear unbiased prediction for subdivided target regions.
     Crop Science **45**, 1151-1159.
Piepho, H.P., Möhring, J. (2006): Selection in cultivar trials – is it ignorable? *Crop Science* **146**,
     193-202.
Piepho, H.P., Möhring, J. (2007): On weighting in two-stage analysis of series of experiments.
     Biuletyn Oceny Odmian **32**, 109-121.
Piepho, H.P., Ogutu, J.O. (2007): Simple state-space models in a mixed model framework. *The
     American Statistician* **61**, 224-232.
Piepho, H.P., Williams, E.R. (2007): Two-dimensional linear variance structures for field trials.
     submitted.
Lee, Y., Piepho, H.P. (2007): Ambiguity in random walks and implications for the analysis of field
     trials. *submitted*

Lars Kai Hansen: Learning to detect signals from labeled and unlabeled data

Labeled examples are "expensive" in many real world signal detection applications, hence, it is of interest to formulate algorithms that can learn to classify from mixtures of labeled and unlabeled data. I will present a theoretical analysis of the labeled-unlabeled problem, an algorithm than can learn from unlabeled data, and an application to email classification.

Per Chr. Hansen: Large-Scale Ill-Posed Problems and Applications

Inverse problems arise in geophysics, tomography, image deblurring and many other areas where the goal is to compute interior or hidden information from exterior data. A common feature of these ill-posed problems is that the solution is extremely sensitive to perturbations, and hence some form of stabilization or regularization is need in order to compute a meaningful solution. In this talk we present some state-of-the-art Krylov subspace method for large-scale inverse problems. The key feature of these methods is that they provide a good "signal subspace" for the desired solution. This, in turn, allows us to compute regularized solutions via a projection onto this lower-dimensional subspace - a feature which is extremely important when we do not know a good a priori value of the regularization parameter.

Mogens Bladt & Bo Friis Nielsen: Distributions with rational Laplace transform

We review what is currently known about one-dimensional distributions on the non-negative reals with rational Laplace transform, also known as matrix-exponential distributions. In particular we discuss a flow interpretation which enables one to mimic certain probabilistically inspired arguments which are known from theory of phase-type distributions. We then move on to present ongoing research for higher dimensions. We discuss a characterization result, some closure properties, and a number of examples. Finally we present open problems and future perspectives.

Thomas Gerds, Anders Albrektstein, Tune Pers, Svend Kreiner, Thorkild Sørensen, Claus Holst: The VAML game [Validation and Assessment of Machine Learning]

In this talk I describe a game where the players build a risk prediction model. Each model predicts the conditional probability distribution of a response variable given a vector of covariates. To build the model, each player chooses a statistical tool, and gets an input matrix with the covariate values of independent subjects and also the true values of the response variable for the same subjects. The predictive performance of the models is evaluated with a strictly proper scoring rule [1] and the 632+ bootstrap method [2]. The game will be illustrated with data from 99 subjects that took part in a NUGENOB trial [3]. The response is the continuous measurement of a respiratory quotient and the covariate vector includes the high-dimensional information from the metabolic platforms 1-H NMR and LC-MS and the subjects age, BMI, blood values, etc. For this game, the players (Anders, Tune, Svend, Thomas) use tools like Random Forests, LASSO, PRIM, Logic Regression, Item Response Theory, Graphical Models, Support Vector Machines and BAMSE. In the talk I will explain how to assess and compare the performances of prediction models obtained with these rather different statistical tools.

[1] Leonard J Savage (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66:783-801.
[2] Bradley Efron and Robert Tibshirani (1997). Improvements on cross-validation: The .632+ bootstrap method. *Journal of the American Statistical Association*, 92:548-560.
[3] http://www.nugenob.com/

Julie Lyng Forman: Inference from Diffusion Driven Models.

Diffusion-type models provide a natural and flexible framework for modeling phenomena that evolve continuously and randomly with time. However their statistical analysis is complicated. Only rarely the functional form of the likelihood function of a diffusion is explicitly known. Inference from diffusion driven models such as summed diffusions, integrated diffusions, and stochastic volatility models of course is no less involved. General estimating equations often present a simple alternative means for fitting these models. I will illustrate the use of estimating equations for diffusion-type models driven by diffusions having linear drift and quadratic squared diffusion coefficient. These I term *the Pearson diffusions* as their invariant distributions belong to the Pearson system. The Pearson diffusions have polynomial eigenfunctions which in turn yield explicit recursive formula for their conditional moments.
Further I will present a new goodness of fit test based on estimating equations for varying sampling frequencies. The basic idea is to compare the parameter estimates based on the full data to those obtained from downsamples. The test can be used to distinguish different kinds of diffusion driven models.

Uffe Høgsbro Thygesen:
Markov models of animals moving in the sea

Markov processes are often used as models for the movements of animals. Examples from the marine environment are bacteria moving to locate food, plankton being advected with a turbulent flow, and fish swimming. In this talk I will give examples of questions we aim to answer using such models, and the analysis of the models we undertake. A common issue in our work is the range of spatial and temporal scales, and therefore I will give and example of diffusion approximations that are useful approximations on a larger scale. I will also discuss deriving optimal behaviour. Finally, I will give an example of geolocation of fish based on date storage tags, i.e. state and parameter estimation in a diffusion model for the movements of a fish.

Per Bruun Brockhoff, IMM, DTU.

Biostatistician – International Clinical Research

Two positions as biostatisticians are open in the Biostatistics Department, International Clinical Research. The department works primarily within clinical research but contributes to the entire drug development process from discovery to market across the range of therapeutic areas within psychiatry and neurology. The current staffs include 20 biostatisticians, 7 statistical programmers, 1 PhD student, and 3 technical staff members. We offer a challenging job with broad career opportunities in a dynamic and open working atmosphere.

Your job
Part of your responsibility will be to provide statistical input for designing and planning of clinical studies and to participate in statistical analysis and interpretation of clinical studies in all phases of development. You participate in preparing publications, which involves exploratory statistical analyses of a diverse range of clinical study data and, where appropriate, research in new statistical methodologies. Other challenges involve providing statistical input for Clinical Development Plans, safety evaluations, statistical modelling and mathematical simulation based on non-clinical and clinical study data for optimisation of early drug development and introduction of new study designs. You work in close collaboration with clinical researchers and other specialists, exerting your expertise in statistical methodology and keeping abreast of current practices of pharmaceutical R&D.

Your qualifications
Our preferred candidate
- holds an MSc or PhD degree in Statistics or Mathematical Sciences
- has programming experience and familiarity with statistical software such as SAS
- has a strong interest in applying statistical methods to biological problems; work experience from the pharmaceutical industry or consulting experience from an industrial or academic setting is highly desirable
- has an interest in clinical aspects as well as statistical aspects and in working with clinicians
- is goal-oriented, innovative, and flexible, work well under pressure, and possesses the ability to listen, be analytic and proactive
- is fluent in oral and written English
- is a team player and able to interact smoothly with colleagues and collaborators from different functional areas and/or companies

Further information
Please contact Head of Department, Ingrid Sofie Harbo, on +45 3643 2004 or Head of Section, Anna Karina Trap Huusom, on +45 3643 2303. We also recommend you to visit our website www.lundbeck.com.

Your application
Please submit your application electronically at http://www.lundbeck.com/careers/jobs/vacancies/default.asp, where you will find this position in the list of 'Current vacancies'. Applications must be received no later than September 28, 2007. Please state in your application where you have seen this advertisement.

larix

Statistiker søges

Vi har travlt og søger nu en eller flere statistikere. Vi er interesserede i dig både hvis du er nyuddannet og hvis du har lang erfaring.

Larix er et mindre konsulentfirma med fokus på statistik og data management. Vores kunder kommer især fra farma- og biotekbranchen. Det drejer sig både om store firmaer hvor vi arbejder for deres statistik og/eller datamanagement afdelinger, og små firmaer hvor vi dækker deres statistik- og data management behov. Vi er lige nu 2 statistikere, 3 data managere og en programmør.

Vi har brug for at du
* har en matematisk statistisk uddannelse som cand. stat., ingeniør e.l.
* er god til mundtlig og skriftlig kommunikation – også på engelsk
* er god til at samarbejde, er nem at omgås og er serviceminded
* kan bevare overblikket i pressede situationer
* meget gerne relevant erfaring med klinisk statistik og ICH-GCP
* er god til at programmere i SAS

Afhængigt af din baggrund vil vi lave en stilling som passer til din erfaring som seniorstatistiker, eller en uddannelsesstilling hvis du kommer frisk fra et universitet.

Vi tilbyder et spændende og udfordrende job, hvor du vil komme til at arbejde med en række forskellige typer af statistiske opgaver, primært ifm. kliniske studier. Eksempelvis vil du komme til at lave de statistiske analyser og afrapporteringer fra kliniske studier, skrive statistiske analyseplaner, og give input til forsøgsprotokoller inklusive sample size beregninger. Nogle opgaver løses hos Larix, men mange løses som konsulent udstationeret hos kunden. Derudover tilbyder vi en konkurrencedygtig løn med bonus, efteruddannelse, fleksibel planlægning af arbejdet og ikke mindst nogle gode kolleger.

Hvis du er interesseret i at høre mere så check vores hjemmeside www.larix.dk, eller ring til Klaus Juel Olsen, Head of Statistics på 61 61 80 11. Send en ansøgning med CV på e-mail til kjo@larix.dk eller med snail mail til Larix Aps, Bymidten 78, 3500 Værløse.