

Program

Monday, July 24

08:45-09:00 **Welcome words**

Great hall (Juhlasali)

09:00-10:00 **Plenary session**

PL-1 **Opening lecture**

Plenary, Great hall (Juhlasali)

Chair: Richard Samworth

Martin Wainwright Pairwise ranking and crowd-sourcing: Statistical models
and computational challenges

10:30-12:30 **Invited and contributed sessions**

IS-01 **New trends in survival analysis**

Invited session, Great hall (Juhlasali)

Organizer: Valentin Patilea

Chair: Valentin Patilea

Valentin Patilea A general approach for cure models in survival analysis
Michael R. Kosorok Tree based precision medicine for right censored data
Jane-Ling Wang Semiparametric analysis of transformation models for
doubly-censored clustered survival data

IS-02 **Functional data**

Invited session, Lecture room 1 (Sali 1)

Organizer: Pedro Delicado

Chair: Pedro Delicado

Aldo Goia Evaluating the complexity of functional data
Hans-Georg Müller Mixture inner product spaces and their application to func-
tional data analysis
José Luis Torrecilla Radon-Nikodym derivatives, RKHS and functional classifi-
cation

IS-03 New trends in spatial statistics

Invited session, Lecture room 13 (Sali 13)

Organizer: Marc Hallin

Chair: Peter Robinson

Ngai Hang Chan	Inference for spatial trends
Zudi Lu	On bandwidth choice for spatial data density estimation
Peter Robinson	Inference on trending panel data

TCS-02 Alternate Bayesian models and methods

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Madhuchhanda Bhattacharjee

Chair: Mikko J. Sillanpää

Mikko J. Sillanpää	Estimation of network topology: Genes and individuals
Patrik Waldmann	The use of Bayesian additive regression trees in genomics
Bob O'Hara	Flexible coding of latent factor analyses
Tommi Härkönen	Nonparametric Bayesian intensity model: Exploring time-to-event data on two time scales
Madhuchhanda Bhattacharjee	<i>Canceled</i>

Note: Madhuchhanda Bhattacharjee is replaced by Tommi Härkönen.

TCS-03 Non-standard non-parametric regression

Topic-contributed session, Lecture room 12 (Sali 12)

Organizer: Melanie Birke and Natalie Neumeyer

Chair: Melanie Birke

Christoph Breunig	Regression with selectively missing covariates
Aurore Delaigle	Nonparametric covariate adjusted regression
Gwennaëlle Mabon	Nonparametric testing in boundary models
Ursula U. Müller-Harknett	Estimating the response density in nonlinear and non-parametric regression
Gilles Stupfler	Estimating a frontier function using a high-order moments method
Anne Vanhems	A mollification approach to study the nonparametric instrumental regression model

CS-111 Econometrics

Contributed session, Lecture room 6 (Sali 6)

Chair: Ruey-Ching Hwang

Shih-Feng Huang	Fluctuation reduction of value-at-risk estimation and its applications
Farrukh Javed	Higher moments of the estimated tangency portfolio weights
Bezirgen Veliyev	The realized empirical distribution function of volatility
Ruey-Ching Hwang	A logistic regression point of view toward loss given default distribution estimation
Xu Huang	<i>Canceled</i>

CS-112 Likelihood based inference

Contributed session, Lecture room 10 (Sali 10)

Chair: Artis Luguzis

Angel Angelov	Maximum likelihood estimation for survey data with informative interval censoring
Yuichi Hirose	Profile likelihood approach to a large sample distribution of estimators in joint mixture mode
Manuela Cattelan	(Hybrid) pairwise likelihood inference for Plackett models
Artis Luguzis	Model selection using penalized empirical likelihood
Liga Bethere	<i>Canceled</i>

CS-113 Limit theorems for stochastic processes

Contributed session, Lecture room 7 (Sali 7)

Chair: Lauri Viitasaari

Markus Bibinger	Volatility estimation for stochastic PDEs using high-frequency observations
Carla Taming	Statistical inference for the Wasserstein distance on discrete spaces with applications
Klaus Telkmann	On multivariate kernel density estimation under long memory
Christian Palmes	Nonparametric drift estimation in a Levy driven diffusion model under microstructure noise
Lauri Viitasaari	Parameter estimation for Langevin equation in a general Gaussian framework

CS-114 Time-series analysis: Modeling

Contributed session, Lecture room 8 (Sali 8)

Chair: Carsten Jentsch

Alexander Kreuzer	Heavy tailed spatial autocorrelation models
Nicole Barthel	Vine based modeling of multivariate volatility time-series
Qianqian Zhu	Quantile double autoregression
Arthur Guillaumin	Analysis of nonstationary modulated time series with applications to oceanographic flow measurements
Henri Nyberg	Taking zero lower bound seriously: A structural VAR model containing positive-valued components
Carsten Jentsch	Modeling dynamic networks using high-dimensional binary autoregressive time series processes

14:00-16:00 Invited and contributed sessions

IS-05 Graphical models

Invited session, Great hall (Juhlasali)

Organizer: Piotr Zwiernik

Chair: Piotr Zwiernik

Piotr Zwiernik	The correlation space of Gaussian latent tree models and model selection without fitting
Steffen Lauritzen	Maximum likelihood in multivariate Gaussian models under total positivity
Caroline Uhler	Learning Bayesian networks based on sparsest permutations
Venkat Chandrasekaran	<i>Canceled</i>

IS-06 Developments in Bayesian nonparametrics

Invited session, Lecture room 1 (Sali 1)

Organizer: Sonia Petrone

Chair: Stefano Favaro

Sergio Bacallado	Unbiased estimation and prediction in the hierarchical two-parameter Poisson-Dirichlet process
Tamara Broderick	Edge-exchangeable graphs, sparsity, and paintboxes
Stefano Favaro	Multi-armed bandit for species discovery: A Bayesian non-parametric approach

IS-10 High-dimensional model selection

Invited session, Lecture room 13 (Sali 13)

Organizer: Małgorzata Bogdan

Chair: Małgorzata Bogdan

Małgorzata Bogdan	Sorted L-One Penalized Estimation (SLOPE).
Veronika Rockova	Spike-and-slab LASSO
Stefan Wager	Approximate residual balancing: De-biased inference of average treatment effects in high dimensions
Florian Frommlet	<i>Canceled</i>

TCS-04 Trends in time series analysis

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Konstantinos Fokianos

Chair: Konstantinos Fokianos

Matteo Barigozzi	Sequential testing for structural stability in approximate factor models
Guy Nason	New developments in network time series
Hernando Ombao	Exploratory tools for high dimensional time series with applications to brain signals
Konstantinos Fokianos	The multivariate auto-distance covariance/correlation function
Anne van Delft	<i>Canceled</i>

TCS-14 Inference for dynamic systems

Topic-contributed session, Lecture room 12 (Sali 12)

Organizer: Shota Gugushvili and Itai Dattner

Chair: Shota Gugushvili

Itai Dattner	Statistical learning of dynamic systems
Clemens Kreutz	Inferring dynamical systems using the profile likelihood
Ester Mariucci	Compound Poisson approximation to estimate the Lévy density
Moritz Schauer	Bayesian inference for partially observed diffusion processes
Rami Yaari	A two-stage approach for estimating the parameters of an age-group epidemic model from incidence data

TCS-15 Distance correlation

Topic-contributed session, Lecture room 6 (Sali 6)

Organizer: Carsten Jentsch and Dominic Edelmann

Chair: Carsten Jentsch

Dominic Edelmann	The distance correlation coefficient for right-censored survival data
Phyllis Wan	Applications of distance correlation to time series
Carina Beering	Empirical characteristic function-based estimation and distance correlation for locally stationary processes
Jochen Fiedler	Distance correlation for spatial stochastic processes

CS-121 Advanced topics in regression analysis

Contributed session, Lecture room 10 (Sali 10)

Chair: Toshio Honda

Yue Zhao	Envelopes for censored quantile regression
Frank Werner	Empirical risk minimization as parameter choice rule for general linear regularization methods
Mickaël De Backer	An adapted loss function for censored quantile regression
Toshio Honda	Adaptively weighted group Lasso for semiparametric quantile regression models
Yi-Chen Lin	<i>Canceled</i>
Philip Brown	<i>Canceled</i>

CS-122 Change point analysis

Contributed session, Lecture room 7 (Sali 7)

Chair: Zuzana Praskova

Yubin Tian	Change-point analysis for the coefficients of variation based on empirical likelihood
Andreas Anastasiou	Change-point detection and subset expansion
Ivor Cribben	Methods for estimating spectral clustering change points for multivariate time series
Zuzana Praskova	Sequential monitoring in panel data

CS-123 Robust and nonparametric methods

Contributed session, Lecture room 8 (Sali 8)

Chair: Onur Toka

Martin Kroll	Nonparametric adaptive Poisson regression by model selection
Maria Umlauft	Rank-based permutation approaches for nonparametric factorial designs
Chi Zhang	A new robust regression model: Type II multivariate t distribution with applications
Lubna Amro	Permuting incomplete paired data: A novel exact and asymptotic correct randomization test
Evgenii Pchelintsev	Adaptive improved estimation in heteroscedastic nonparametric regression models
Onur Toka	Robust variable selection in restricted linear regression model

CS-124 Statistical inference for stochastic processes

Contributed session, Auditorium IV

Chair: Lionel Truquet

Andrea Kraus	Estimation in partially observed branching processes under different scenarios
Younes Ommame	On predictive density estimation under S-Hellinger distance and alpha-divergence
Lothar Heinrich	Asymptotic normality for a class of empirical functionals of strongly mixing marked point Processes
Tamar Gadrich	Estimating a Poisson parameter based on incomplete information: Precision vs. simplicity
Lionel Truquet	Local stationarity and time-inhomogeneous Markov chains

Note: The presentation of Samuel Maistre was moved to session CS-413 on Thursday at 10:30-12:30, Lecture room 10.

17:00-18:00 Plenary session

PL-2 Forum lecture I

Plenary, Great hall (Juhlasali)

Chair: Niels Richard Hansen

Mark Girolami	Diffusions and dynamics on statistical manifolds for statistical inference
---------------	--

Tuesday, July 25

09:00-10:00 Plenary session

PL-3 Special invited lecture

Plenary, Great hall (Juhlasali)

Chair: Anne Gégout-Petit

Alison Etheridge Modelling evolution in a spatial continuum

10:30-12:30 Invited and contributed sessions

IS-07 Large sample covariance matrices: Theory and applications

Invited session, Great hall (Juhlasali)

Organizer: Jianfeng Yao

Chair: Jianfeng Yao

Florent Benaych-Georges Kernel spectral clustering of large dimensional data
Johannes Heiny Large sample covariance matrices with heavy tails
Weiming Li On structure testing for component covariance matrices of
a high-dimensional mixture

IS-08 Advances in causality

Invited session, Lecture room 1 (Sali 1)

Organizer: Marloes Maathuis

Chair: Elizabeth L. Ogburn

Antti Hyttinen Exact constraint-based causal discovery
Emilija Perkovic Utilizing background knowledge for the estimation of total
causal effects
Elizabeth L. Ogburn Causal inference for social network data

IS-18 Concentration inequalities

Invited session, Lecture room 13 (Sali 13)

Organizer: Pascal Massart

Chair: Pascal Massart

Devdatt Dubhashi Measure concentration in machine learning algorithms
Ivan Nourdin Stein's method, logarithmic Sobolev and transport inequalities
Daniel Paulin Analysis of optimization based high dimensional inference
based on concentration inequalities

TCS-05 Majorization theory in statistics, probability, econometrics, economics and engineering

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Rustam Ibragimov

Chair: Rustam Ibragimov

Rustam Ibragimov	Majorization, bundling, diversification and voting
Eduard Jorswieck	Majorization results for secure outage capacities in fading wireless channels
Paul Kattuman	Inference ready measures of dispersiveness for small samples
Victor de la Pena	On boundary crossing by stochastic processes

TCS-06 Emerging statistical tools in complex structural data analysis

Topic-contributed session, Lecture room 12 (Sali 12)

Organizer: Linglong Kong

Chair: Dario Gasbarra

Jia Liu	Variational Bayes estimation in constrained kurtosis diffusion imaging under a Rician noise model
Guodong Li	Hybrid quantile regression estimation for time series models with conditional heteroscedasticity
Keren Shen	On a spiked model for large volatility matrix estimation from noisy high-frequency data
Ping-Shou Zhong	<i>Canceled</i>
Sijian Wang	<i>Canceled</i>

Note: New speaker Jia Liu.

TCS-11 Multiplicity control for structured systems

Topic-contributed session, Lecture room 6 (Sali 6)

Organizer: Frank Werner

Chair: Frank Werner

Armin Schwartzman	Confidence regions for spatial excursion sets from repeated random field observations, with an application to climate
Claudia König	Multiscale Poisson scanning: Multivariate limit theory and statistical implications
Guenther Walther	Constructing the essential histogram
James Sharpnack	Adaptive scan statistics with convolutional networks

CS-211 Bayesian inference

Contributed session, Lecture room 10 (Sali 10)

Chair: Mutlu Kaya

Peter Spreij	A non-parametric Bayesian approach to decomposing from high frequency data
Suleyman Ozekici	Bayesian inference on doubly stochastic Markov processes
Alisa Kirichenko	Function estimation on a large graph using Bayesian Laplacian regularization
Shota Gugushvili	Non-parametric Bayesian estimation of a diffusion coefficient
Xavier Loizeau	Optimal aggregation in circular deconvolution by a frequentist Bayesian approach
Mutlu Kaya	Robust Bayesian penalized regression estimation based on Ramsay-Novick distribution

CS-212 Biostatistics

Contributed session, Lecture room 7 (Sali 7)

Chair: Kexin Yu

Hans J. Skaug	Close-Kin Mark-Recapture
Marta Cousido-Rocha	Testing the equality of a large number of populations with applications to genomics
Umberto Picchini	Inference via Bayesian synthetic likelihoods for a mixed-effects SDE model of tumor growth
Cristina Gutiérrez Pérez	Limiting genotype frequencies of Y-linked genes with a mutant allele in a two-sex population
Kexin Yu	Detection of imprinting effects for quantitative traits on X chromosome

CS-213 New approaches for analyzing high dimensional data

Contributed session, Lecture room 8 (Sali 8)

Chair: Tatjana Pavlenko

Maarten Jansen	Optimisation bias correction in sparse structured variable selection
Benedikt Bauer	On deep learning as a remedy for the curse of dimensionality in nonparametric regression
Tom Boot	Confidence intervals in high-dimensional regression based on regularized pseudoinverses
Anne Gégout-Petit	Aggregated methods for covariates selection in high-dimensional data under dependence
Piotr Pokarowski	Improving Lasso for generalized linear model selection
Tatjana Pavlenko	CsCsHM-statistic for detecting rare and weak effects in high-dimensional classification

14:00-16:00 Invited and contributed sessions

IS-11 Bayesian statistics and software

Invited session, Great hall (Juhlasali)

Organizer: Aki Vehtari

Chair: Aki Vehtari

Javier González Hernandez	GPyOpt: Practical Bayesian optimization for model configuration and experimental design
Antti Kangasrääsio	Engine for likelihood free inference (ELFI)
Sebastian Weber	Bayesian aggregation of average data

IS-12 High-dimensional time series with breaks

Invited session, Lecture room 1 (Sali 1)

Organizer: Claudia Kirch

Chair: Claudia Kirch

Haeran Cho	Simultaneous multiple change-point and factor analysis for high-dimensional time series
Moritz Jirak	Optimal change point tests in high dimension
Martin Wendler	Bootstrap and change-point detection in functional time series and random fields

IS-21 Statistical genetics and genomics

Invited session, Lecture room 13 (Sali 13)

Organizer: Stéphane Robin

Chair: Stéphane Robin

Emma Bonnet	Heritability estimation in high-dimensional sparse linear mixed models
John Hickey	Deriving animal breeding benefit from millions of animals with sequence
Mark van de Wiel	Empirical Bayes learning in high-dimensional prediction settings

TCS-01 Advances in statistics on manifolds

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Karthik Bharath

Chair: Karthik Bharath

Anuj Srivastava	A geometric framework for density estimation and its extensions to estimation under shape constraint
Karthik Bharath	Sampling of warp maps for curve alignment
Ian Jermyn	Numerical inversion of SRNF maps for elastic shape analysis of genus-zero surfaces
Yoav Zemel	Fréchet means and Procrustes analysis in Wasserstein space

TCS-08 Causal inference

Topic-contributed session, Lecture room 12 (Sali 12)

Organizer: Fan Li

Chair: Stefan Wager

Federica Licari	Bayesian inference in RDD: An application to students's academic performances
Andrea Mercatanti	Evaluating the European central bank corporate sector purchase program via a regression Discontinuity Design with an ordered forcing variable
Fan Li	Subgroup balancing propensity score
Stefan Wager (Discussant)	

TCS-12 Recent advances in stochastic modeling

Topic-contributed session, Lecture room 6 (Sali 6)

Organizer: Sonia Malefaki and Vasillis Koutras

Chair: Filia Vonta

Agapios Platis	Sojourn time distributions effects on a redundant multi-state deteriorating system with maintenance
Alex Karagrigoriou	On control charts for not necessarily symmetric distributions with simulations
George-Jason Siouris	Improvement of expected shortfall estimations in APARCH and FIGARCH models
Sonia Malefaki	Optimizing the main dependability and performance measures of a diesel engine system

CS-221 Applied statistics

Contributed session, Lecture room 10 (Sali 10)

Chair: Deon Kleinsmith

Ana Kolar	The importance of causal thinking and statistical thinking in causal effect studies
Leena Kalliovirta	Finnish wood purchasing cartel: Damage estimates on pulpwood based on the VECM approach
Han-Ming Wu	graphics.SDA: The R base graphics companion for exploratory symbolic data analysis
Guosheng Yin	Bayesian change-point detection for ordinal data with North Atlantic tropical cyclones
Deon Kleinsmith	Updating the South African sampling frame

CS-222 Depths, outlier detection and classification

Contributed session, Lecture room 7 (Sali 7)

Chair: Samuela Leoni

Pavlo Mozharovskiy	Approximate computation of data depths that satisfy the projection property
Stanislav Nagy	Data depth for measurable random mappings
Klaus Nordhausen	Multivariate outlier detection with ICS
Clémentine Barreyre	Statistical methods for outlier detection in space telemetries
Samuela Leoni	A new class of depth functions based on Bregman divergences
Yuan-Chin Chang	<i>Canceled</i>

CS-223 Generalized linear models and nonlinear regression

Contributed session, Lecture room 8 (Sali 8)

Chair: Vladimir Pastukhov

Takeshi Kurosawa	Monotonic property and asymptotic distribution of regression correlation coefficient
Nicolas Asin	Adaptive functional linear regression in the presence of dependence
Sebastian Letmathe	An iterative plug-in for P-Spline regression
Luigi Augugliaro	A differential-geometric approach to generalized linear models with grouped predictors
Qingfen Liu	Model selection and model averaging for nonlinear regression models
Vladimir Pastukhov	Asymptotic distribution of the isotonic regression estimator over a countable pre-ordered set

17:00-18:00 Plenary session

PL-4 Bernoulli society - European mathematical society lecture

Plenary, Great hall (Juhlasali)

Alexander S. Holevo	Quantum Shannon theory
---------------------	------------------------

Wednesday, July 26

09:00-10:00 Plenary session

PL-5 Special invited lecture

Plenary, Great hall (Juhlasali)
Chair: Ingrid Van Keilegom

Gerda Claeskens Effects of model selection and weight choice on inference

10:30-12:30 Invited and contributed sessions

IS-13 Time series econometrics

Invited session, Great hall (Juhlasali)
Organizer: Pentti Saikkonen
Chair: Pentti Saikkonen

Christian Francq Estimation risk for the VaR of portfolios driven by semi-parametric multivariate models
Dennis Kristensen Nonparametric estimation of time-varying parameters in nonlinear models
Bent Nielsen Tightness of M-estimators for multiple linear regression in time series

IS-14 Computational tools for Bayesian inference

Invited session, Lecture room 1 (Sali 1)
Organizer: Jean-Michel Marin
Chair: Jean-Michel Marin

Jimmy Olsson PaRIS-based Rao-Blackwellization of particle Gibbs
Murray Pollock Exact Bayesian inference for big data: Single- and multi-core approaches
Christian Robert Combining Monte Carlo estimates

IS-17 High-dimensional extremes

Invited session, Lecture room 13 (Sali 13)
Organizer: Holger Rootzen
Chair: Raphaël Huser

Valérie Chavez-Demoulin Non-stationary modeling of tail dependence of two subjects concentration
Clement Dombry Asymptotic properties of likelihood estimators in multivariate extremes
Raphaël Huser Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures

TCS-07 Recent advances in analysis of shapes and directions

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Alfred Kume

Chair: Alfred Kume

Huiling Le	Smooth curve fitting to 3D shape data
Stephen Huckeman	Backward nested subspace inference and applications
Simon Preston	Some novel spherical regression models
Brian Thomas Rooks	Averaging symmetric positive-definite matrices via the minimal scaling-rotation framework

TCS-10 Advances in survival and reliability

Topic-contributed session, Lecture room 12 (Sali 12)

Organizer: Juan Eloy Ruiz-Castro and Mariangela Zenga

Chair: Juan Eloy Ruiz-Castro

Ewa Strzalkowska-Kominiak	Censored functional data
Giuliana Cortese	Regression models for the restricted residual mean time for right-censored and left-truncated data
Rosa Lillo	The Batch Markov Modulated Poisson process, a powerful tool in reliability models
Juan Eloy Ruiz-Castro	A Markovian arrival process with marked transitions to model a complex system with loss of units
Mariangela Zenga	Joint models for time-to-event and bivariate longitudinal data

CS-311 Bootstrapping

Contributed session, Lecture room 6 (Sali 6)

Chair: Mayya Zhilova

Shahab Basiri	Enhanced bootstrap method for statistical inference in the ICA model
Sarah Friedrich	MATS: Inference for potentially singular and general heteroscedastic MANOVA
Justin Chown	Regularization parameter selection by residual-based bootstrap
Anna Dudek	Bootstrap for triangular arrays with growing period
Ali Charkhi	Bootstrap confidence intervals after model selection
Mayya Zhilova	Higher-order Berry-Esseen inequalities and accuracy of the bootstrap

CS-312 Linear regression

Contributed session, Lecture room 10 (Sali 10)

Chair: Jan Picek

Funda Erdugan	A new restricted Liu-type estimator in linear regression model
Ulrike Schneider	On the distribution and model selection properties of the Lasso estimator in low and high dimensions
Nina Senitschnig	Prediction out-of-sample using block shrinkage estimators: Model selection and predictive inference
Zhang Liu	Group variable selection using bootstrap method
Jan Picek	L-moment estimation for linear regression models

CS-313 Model selection

Contributed session, Lecture room 7 (Sali 7)

Chair: Tsung-Shan Tsou

Genevieve Robin	Low-rank interaction contingency tables
Apostolos Batsidis	Model selection tests for count models based on the empirical probability generating function
Natalia Stepanova	Estimating the amount of sparsity in mixture models
Estelle Kuhn	Testing variance components in nonlinear mixed effects models
François Bachoc	Uniformly valid confidence intervals post-model-selection
Tsung-Shan Tsou	Sample size determination using Poisson for comparing two count populations that exhibit over-dispersion

CS-314 Random variables and stochastic processes

Contributed session, Lecture room 8 (Sali 8)

Chair: Jean-Baptiste Aubin

May-Ru Chen	A shoes and doors problem
Wolf-Dieter Richter	Dependence in multivariate distributions
David Stenlund	Time to absorption in the Mabinogion urn model
Jean-Baptiste Aubin	On the correction term in the small ball probability factorisation for Hilbert random elements

Thursday, July 27

09:00-10:00 Plenary session

PL-6 Special invited lecture

Plenary, Great hall (Juhlasali)
Chair: Pauliina Ilmonen

Hannu Oja Scatter matrices and linear dimension reduction

10:30-12:30 Invited and contributed sessions

IS-15 Networks

Invited session, Great hall (Juhlasali)
Organizer: Ernst Wit
Chair: Maria Deijfen

Maria Deijfen Birds of a feather or opposites attract - effects in network modelling
Gesine Reinert Network comparison
Debleena Thacker Nonuniform random geometric graphs with location-dependent radii

IS-16 Stochastic processes

Invited session, Lecture room 1 (Sali 1)
Organizer: Herold Dehling
Chair: Herold Dehling

Rafal Kulik The tail empirical process of regularly varying functions of geometrically ergodic Markov chains
Vladas Pipiras Semi-parametric, parametric and possibly sparse models for multivariate long-range dependence
Jeannette Woerner Inference for the driving Lévy process of continuous-time moving average processes

IS-19 Distributional regression

Invited session, Lecture room 13 (Sali 13)
Organizer: Thomas Kneib
Chair: Thomas Kneib

Giampiero Marra A new approach to fitting generalised additive models for location, scale and shape
Julien Hambuckers A Markov-switching generalized additive model for compound Poisson processes, with applications to operational losses models
Nikolaus Umlauf LASSO penalization in the framework of generalized additive models for location, scale and shape

TCS-09 Inverse problems in econometrics

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Pierre Maréchal and Anne Vanhems

Chair: Anne Vanhems

Jan Johannes	Data-driven estimation by aggregation based on a penalised contrast criterion
Pascal Lavergne	Instrumental regression via spline smoothing
Konstantin Eckle	Tests for qualitative features in the random coefficients model
Pierre Maréchal	A mollification approach to the deconvolution problem

CS-411 Bayesian inference and applications

Contributed session, Lecture room 12 (Sali 12)

Chair: Alberto Pessia

Weichang Yu	Variational diagonal discriminant analysis with model selection
Stéphanie van der Pas	How many needles in the haystack? Adaptive uncertainty quantification for the horseshoe
Nurzhan Nurushev	Uncertainty quantification for biclustering model
Mikko Heikkilä	Differentially private Bayesian learning on distributed data
Alberto Pessia	Bayesian cluster analysis of categorical data with supervised feature selection

CS-412 Modeling high dimensional observations

Contributed session, Lecture room 6 (Sali 6)

Chair: Patrick Tardivel

Dominik Mueller	Selection of sparse vine copulas in high dimensions with the Lasso
Kou Fujimori	The Dantzig selector for some statistical models of stochastic processes
Andreas Artemiou	A first effort to real time sufficient dimension reduction
Ivana Milovic	Conditional means of low-dimensional projections from high-dimensional data: explicit error bounds
Matthias Killiches	Model distances for vine copulas in high dimensions
Patrick Tardivel	Sparsest representation and approximation of a high-dimensional linear system

CS-413 Survival analysis

Contributed session, Lecture room 10 (Sali 10)

Chair: Jelena Valkovska

Sergey Malov	On non precisely observed right censored survival data
Grzegorz Wylupek	A permutation test for the two-sample right-censored model
Anja Rueten-Budde	Interpretation of the discriminative ability of a time-dependent marker in survival models
Lecleire Uriel	Nonparametric goodness-of-t test of the baseline hazard function in the Cox model
Samuel Maistre	Two-step estimation of an additive model for censored data.
Jelena Valkovska	Jackknife empirical likelihood inference for the ROC curve

Note: The presentation of Samuel Maistre was moved from session CS-124.

CS-414 Time-series analysis: Modeling and testing

Contributed session, Lecture room 7 (Sali 7)

Chair: Oliver Stypka

Jacek Leśkow	Fraction of time approach in statistical inference for cyclostationary signals
Ronan Le Guével	Goodness-of-fit test for multistable Lévy processes
Šárka Hudecová	Goodness-of-fit tests for time series of counts
Vanessa Berenguer Rico	Marked and weighted empirical process of residuals
Oliver Stypka	Testing for linear cointegration against the alternative of smooth transition cointegration

14:00-16:00 Invited and contributed sessions

IS-04 Random graphs

Invited session, Great hall (Juhlasali)

Organizer: Maria Deijfen

Chair: Maria Deijfen

Julia Komjathy	Explosion in branching processes and applications to random graph models
Sofia Olhede	Methods of network comparison
Tobias Müller	<i>Canceled</i>

IS-09 Modeling dependence using copulas

Invited session, Lecture room 1 (Sali 1)

Organizer: Marek Omelka

Chair: Marek Omelka

Claudia Czado	D-vine quantile regression
Anastasija Teterova	The realized hierarchical Archimedean copula in risk modelling
Dragan Radulovic	Weak convergence of empirical copula processes indexed by functions

IS-22 New developments in latent components models

Invited session, Lecture room 13 (Sali 13)

Organizer: Alessio Farcomeni

Chair: Fiona Steele

Thomas Kneib	Semiparametric bivariate conditional copula regression with binary and continuous marginals
Brendan Murphy	Latent space stochastic block model for social networks
Fiona Steele	A longitudinal mixed logit model for estimation of push-pull effects in residential location choice

TCS-13 Recent developments in long-range dependence

Topic-contributed session, Lecture room 5 (Sali 5)

Organizer: Herold Dehling

Chair: Herold Dehling

Marie-Christine Dueker	Limit theorems for Hilbert space-valued linear processes under long range dependence
Annika Betken	Estimation of a change point in long-range dependent time series based on the Wilcoxon statistic
Johannes Tewes	Change-point tests and the bootstrap under long-range dependence
Herold Dehling	Testing for structural breaks via ordinal pattern dependence

CS-421 Advanced topics in modeling and estimation

Contributed session, Lecture room 12 (Sali 12)

Chair: Daniel Kraus

Sandra Fortini	Partial exchangeability without stationarity
Xifen Huang	An AD approach for constructing separable minorizing functions in a class of MM algorithms
Derya Karagöz	Robust Shewhart control charts by using ranked set sampling and neoteric ranked set sampling schemes
Tatyana Krivobokova	Simultaneous confidence bands for ratios of quantile functions
Constantinos Petropoulos	Estimating a linear parametric function of a doubly censored exponential distribution
Daniel Kraus	Growing simplified vine copula trees: challenging Dißmanns algorithm

CS-422 Biometry

Contributed session, Lecture room 6 (Sali 6)

Chair: Susanne Ditlevsen

Yuh-Ing Chen	Comparison of two related biomarkers based on the area under ROC curve and Youden index
Ségolen Geffray	Data-driven optimization of marked point processes with applications to object detection in 2D image
Anne Hein	Rigid motion estimation in sparse sequential dynamic imaging
Eugen Pircalabelu	Zooming on edges of L1 penalized graphical models estimated for different coarseness scales
Jessica Stockdale	Bayesian estimation for transmission potential of smallpox
Susanne Ditlevsen	Inferring visual processing in the brain

CS-423 Inference for high dimensional data

Contributed session, Lecture room 10 (Sali 10)

Chair: Stepan Mazur

Tommaso Proietti	A Durbin-Levinson regularized estimator of high dimensional autocovariance matrices
Nestor Parolya	Testing for independence of large dimensional vectors
Paavo Sattler	Inference for high-dimensional split-plot-designs
Sévérien Nkurunziza	On convergence of the sample correlation matrices in high-dimensional data
Stepan Mazur	Central limit theorems for functionals of large dimensional sample covariance matrix and mean vector

CS-424 Markov models and applications

Contributed session, Lecture room 7 (Sali 7)

Chair: Eliane Rodrigues

Birgit Sollie	Parameter estimation for discretely observed infinite-server queues with Markov-modulated input
Paula Bran	MCMC algorithms for misidentification problems
Chrysoula Ganatsiou	On discrete-time birth-death circuit chains in random ergodic environments
Sylvain Le Corff	Optimal scaling of the random walk Metropolis algorithm for L_p mean differentiable distributions
Alexandra Tsimbalyuk	The distribution of a perpetuity
Eliane Rodrigues	A non-homogeneous Markov chain model with seasonal transition probabilities applied to ozone data

CS-425 Time-series analysis: Inference

Contributed session, Lecture room 8 (Sali 8)

Chair: Valeriy Voloshko

Dirk-Philip Brandes	On the sample autocovariance of a Lévy driven moving average process sampled at a renewal sequence
Robert Stelzer	Statistical inference for Lévy-driven CARMA processes based on non-equidistant observations
Imma Valentina Curato	Weak dependence and GMM Estimation for supOU and mixed moving average processes
Timo Teräsvirta	Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model
Valeriy Voloshko	On statistical estimation of parameters for a family of binary autoregressive time series

16:00-17:00 Poster session

2nd floor foyer

Stanislav Anatolyev	Weak unobservable factors in many asset environments: dimensionality asymptotics and new procedures
Idir Arab	On stochastic orders with applications
Antonio Arcos Cebrián	Alternative calibration techniques in panel surveys
Alexander Braumann	A bootstrap procedure in VAR systems with mixed-frequency data
Katarzyna Brzozowska-Rup	Estimating the parameters of the Heston model from stock data using parallel multilevel SMC
Li-Ching Chen	An information matrix goodness-of-fit test of the conditional logistic model
Vivian Yi-Ju Chen	Geographically weighted quantile regression with spatially lagged dependent variable
Zheng Chen	Comparison of two crossing survival curves
Emilio Coelho-Barros	An application of the inverse power Lindley distribution In left censored data
Jordan Franks	Importance sampling type correction of Markov chain Monte Carlo and exact approximations
Mousa Golalizadeh	On application of hierarchical clustering methods for planar shape data
Matias Heikkilä	Multivariate extremes based on a notion of radius
Sami Helander	Pareto depth for functional data
Raúl Hernández-Molinar	Using empirical modeling and extreme value theory in estimating border conditions in quality control
Hsiang-Ling Hsu	Model-robustly D- and A-optimal designs for log-contrast models in mixture experiments
Agnieszka Jach	Testing collinearity of vector time series
Ci-Ren Jiang	Multidimensional functional principal component analysis
Jing Kang	Use of multi-level modelling in the analysis of micro-ct data obtained using micro-ct
Joona Karjalainen	Moment-based parameter estimation in binomial random intersection graph models
Toni Karvonen	Large-scale probabilistic integration
Clemens Kreutz	Inferring dynamical systems using the profile likelihood
Niko Lietzén	Complex valued robust multidimensional SOBI
Pai-Ling Li	Functional principal component analysis with application to air quality data
Kuo-Chin Lin	Order selection for detecting misspecification in the random effects of GLMMs
Frederik Riis Mikkelsen	Degrees of freedom for piecewise Lipschitz estimators
Iliana Peneva	A Bayesian nonparametric model for finding novel pancreatic cancer subtypes
Umberto Picchini	Inference for intractable-likelihoods using a likelihood-free version of SAEM
Jose Antonio Roldán-Nofuentes	Comparison of the predictives values of two diagnostic tests subject to a retrospective Study
Istoni Sant'Ana	Using SA and VNS procedures for estimating the Double-Hubbert diffusion process
Roman Schefzik	Identifying differential distributions for single-cell RNA sequencing data comprising biological replicates
Fabio Sigrist	Recommender systems for purchase data: a novel evaluation approach and an empirical Comparison

Tereza Simkova	L-moment homogeneity test in linear regression
Alina Sinisalo	Automated testing of differences of means to promote the effective use of public data
Alena Skolkova	Many instrument asymptotics: implementation in STATA
Jessica Stockdale	Likelihood approximation methods for stochastic epidemic models
Yaohua Tang	Analyzing high dimensional covariance matrices using convolutional neural networks
Patrick Tardivel	Familywise error rate control with a lasso estimator
Anastasija Tetereva	Sentiment spillover effects for US and European companies
Marko Voutilainen	On the hedging error in mixed models
Samuel Wiqvist	Bayesian inference for modeling of protein folding data
Xu Huang	<i>Canceled</i>
Roberto Souza	<i>Canceled</i>

17:00-18:00 Plenary session

PL-7 Special invited lecture

Plenary, Great hall (Juhlasali)

Chair: Tatyana Krivobokova

John Aston Functional object data analysis

Friday, July 28

09:00-10:00 Plenary session

PL-8 Forum lecture II

Plenary, Great hall (Juhlasali)

Chair: Niels Richard Hansen

Mark Girolami Diffusions and dynamics on statistical manifolds for statistical inference

10:30-12:30 Invited and contributed sessions

IS-20 Recent advances in computationally intensive statistics

Invited session, Great hall (Juhlasali)

Organizer: Nicolai Meinshausen

Chair: Rajen Shah

Gian Thanei The xyz algorithm for fast interaction search
Jonathan Rosenblatt On the optimality of averaging in distributed statistical learning
Rajen Shah Goodness of fit tests for high-dimensional linear models

Note: The order of the talks is changed.

CS-511 Advances in design and modeling

Contributed session, Lecture room 1 (Sali 1)

Chair: José E. Chacón

Adrien Saumard A principled over-penalization of AIC
Yi-Ting Hwang A novel multivariate Bernoulli distribution
Nan-Cheng Su Economic design of two stage control charts with bivariate skew-normal measurements
ChihHao Chang Asymptotic theory for linear mixed-effects model selection
Marek Omelka Testing covariate effects in conditional copulas
José E. Chacón Mixture model modal clustering

Note: The order of the talks is changed.

CS-512 Biostatistics and survival analysis

Contributed session, Lecture room 13 (Sali 13)

Chair: Benjamin Christoffersen

Sanne Willems Optimal scaling for survival analysis with ordinal data
Jorge Alberto Achcar Use of graphical methods in the diagnostic of bivariate life-time distributions with censoring data
Helene Charlotte Rytgaard Variable selection in random survival forests
Pawel Teisseyre An information theory approach to discover hidden associations and interactions in biomedical data
Myriam Vimond A three decision procedure for testing the type of motion of a particle trajectory
Benjamin Christoffersen Dynamic hazard models using state space models

CS-513 Functional data analysis

Contributed session, Lecture room 5 (Sali 5)

Chair: Joni Virta

David Kraus	Inferential procedures for fragmentary functional data
Anirvan Chakraborty	Functional registration and local variations
Paulo Eduardo Oliveira	A weak dependence notion for functional variables
Joni Virta	Independent component analysis of multivariate functional data

CS-514 Modeling and applications

Contributed session, room Lecture room 12 (Sali 12)

Chair: Anna Kiriliouk

Kun-Lin Kuo	Simulating conditionally specified models
Merle Behr	Multiscale Inference for blind demixing with unknown number of sources
Stanislav Anatolyev	Second order asymptotic biases of consistent estimators under many instruments
Anna Kiriliouk	Climate event attribution using the multivariate generalized Pareto distribution

CS-515 Networks and graphs

Contributed session, Lecture room 6 (Sali 6)

Chair: Gino Bertrand Kpogbezan

Jonas Krampe	Time series modeling on dynamic networks
Yuri Pavlov	Limit distributions of maximum vertex degree in a conditional configuration graph
Peter Otto	Coalescent processes and mean minimal spanning trees of irregular graphs
Gino Bertrand Kpogbezan	Network inference using prior knowledge from time-course data
Marina Leri	<i>Canceled</i>

CS-516 New approaches to regression analysis

Contributed session, Lecture room 10 (Sali 10)

Chair: Martin Wagner

Muhammad Naveed Tabassum	Pathwise least angle regression and a significance test for the elastic net
Gulin Tabakan	A new Liu regression estimator in partially linear models
Katharina Proksch	Multiscale scanning in inverse problems
Nil Venet	Gaussian process regression model for distribution inputs
Martin Wagner	Standard fully modified OLS estimation of cointegrating polynomial regressions

CS-517 Principal component analysis and beyond

Contributed session, Lecture room 7 (Sali 7)

Chair: Liang-Ching Lin

Andreas Elsener	Variable selection properties of sparse principal component analysis
Jim Skinner	Structured PCA
Stéphane Robin	Variational inference for probabilistic Poisson PCA
Liang-Ching Lin	Robust principal expectile component analysis

CS-518 Time-series analysis: Regression

Contributed session, Lecture room 8 (Sali 8)

Chair: Yeşim Güney

Markus Matilainen	Sliced average variance estimation for multivariate time series
Britta Steffens	On locally weighted trigonometric regression under long-range dependence
Yetkin Tuğ	Robust variable selection in an autoregressive error term regression model
Marco Singer	Kernel partial least squares for stationary data
Şenay Özdemir	Empirical likelihood estimation for linear regression models with AR(p) error terms
Yeşim Güney	Maximum L _q -likelihood estimation for autoregressive error term regression model

14:00-15:00 Plenary session

PL-9 Closing lecture

Plenary, Great hall (Juhlasali)

Chair: Maarten Jansen

Yann LeCun Deep learning: A statistical puzzle

15:00-15:15 Closing words

Abstracts

Use of graphical methods in the diagnostic of bivariate lifetime distributions with censoring data

Jorge Alberto Achcar
University of São Paulo (Brazil)

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

The choice of an appropriate bivariate parametrical probability distribution for pairs of lifetime data in presence of censored observations usually is not a simple task in many applications. Each existing bivariate lifetime probability distribution proposed in the literature has different dependence structure. Commonly existing classical or Bayesian discrimination methods could be used to discriminate the best among different proposed distributions, but these techniques could be not appropriate to say that we have good fit of some particular model to the data set. In this paper, we explore a recent dependence measure for bivariate data introduced in the literature to propose a graphical and simple criterion to choose an appropriate bivariate lifetime distribution for data in presence of censored data.

Permuting incomplete paired data: A novel exact and asymptotic correct randomization test

Lubna Amro
Ulm University (Germany)

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

Various statistical tests have been developed for testing the equality of means in matched pairs with missing values. However, most existing methods are commonly based on certain distributional assumptions such as normality, 0-symmetry or homoscedasticity of the data. The aim of this work is to develop a statistical test that is robust against deviations from such assumptions and also leads to valid inference in case of heteroscedasticity or skewed distributions. This is achieved by applying a clever randomization approach to handle missing data. The resulting test procedure is not only shown to be asymptotically correct but is also finitely exact if the distribution of the data is invariant with respect to the considered randomization group. Its small sample performance is further studied in an extensive simulation study and compared to existing methods. Finally, an illustrative data example is analyzed.

Monday 24
14:00-16:00
Lecture room 7
(Sali 7)

Change-point detection and subset expansion

Andreas Anastasiou and Piotr Fryzlewicz
London School of Economics (UK)

In this talk, we present a new approach for consistent estimation of the number and locations of multiple change points in time series. We focus on changes of the mean in piecewise-constant signal models. For this purpose, the widely known CUSUM statistic is employed and its values are compared to a threshold. Another approach presented in the talk is concerned with model selection by choosing the candidate model, which optimises a specific information criterion. Our methodology is concerned with offline change-point detection and its main characteristics are firstly an expansion of the subsets of data to be considered and secondly an adaptive localisation, which ensures that we consider subsets that contain at most one previously undetected change-point. This is the key point behind the methods good performance when it comes to both accuracy and speed.

Friday 28
10:30-12:30
Lecture room 12
(Sali 12)

Second order asymptotic biases of consistent estimators under many instruments

Stanislav Anatolyev
CERGE-EI (Czech Republic)

We consider a linear homoskedastic instrumental variables model with many instruments. In an asymptotic framework where their number is proportional to the sample size, we derive the second order asymptotic biases of the jackknife IV, bias corrected 2SLS and LIML estimators, as well as the Fuller-type correction of LIML. The structure of second order asymptotic biases is similar to that of first order asymptotic variances. We compare their expressions whenever possible and elaborate on certain special cases as well as a specific numerical example.

Poster session

Weak unobservable factors in many asset environments: Dimensionality asymptotics and new procedures

Stanislav Anatolyev
CERGE-EI (Czech Republic)

We consider a class of asset pricing models with a sizeable number of assets where observable factors do not fully account for cross-sectional dependence among assets. We show that under such circumstances some popular empirical procedures yield spurious results and cause misleading inferences. We set up a model with comparable cross-sectional and time series dimensions and show that it is possible to correct for influence from omitted factors using an instrumental variables framework, with either strong or weak instruments, depending on the amount of omitted variability. In addition, we propose a test of correct specification and a goodness-of-fit measure within an asymptotic framework with weak omitted factors and many assets.

Maximum likelihood estimation for survey data with informative interval censoring

Angel Angelov and Magnus Ekström
Umeå University (Sweden)

Monday 24
10:30-12:30
Lecture room 10
(Sali 10)

Interval-censored data may arise in questionnaire surveys when, instead of being asked to provide an exact value, respondents are free to answer with any interval without having pre-specified ranges. In this context, the assumption of noninformative censoring is violated, and thus the standard methods for interval-censored data are not appropriate because they can produce biased results. For a given sampling scheme, we suggest a maximum likelihood estimator of the underlying distribution function, assuming that it belongs to a parametric family. The consistency and asymptotic normality of the estimator are established. A bootstrap procedure is presented and its asymptotic validity is discussed. Simulations show a good performance of the proposed methods.

On stochastic orders with applications

Idir Arab and Paulo Eduardo Oliveira
University of Coimbra (Portugal)

Poster session

Stochastic ordering of distributions of random variables may be defined by the relative convexity of the tail functions. This has been extended to higher order stochastic orderings, by iteratively reassigning tail-weights. The actual verification of those stochastic orderings is not simple, as this depends on inverting distribution functions for which there may be no explicit expression. The iterative definition of distributions, of course, contributes to make that verification even harder. We have a look at the stochastic ordering, introducing a method that allows for explicit usage, applying it to the Gamma and Weibull distributions, giving a complete description of the order relations within each of those families.

Alternative calibration techniques in panel surveys

Antonio Arcos Cebrián¹, Maria del Mar Rueda¹ and Manuel Trujillo²
¹Universidad de Granada (Spain), ²IESA-CSIC (Spain)

Poster session

In the last few years we have witnessed a multiplication in the use of probability-based panels that collect data via online or mixed-mode surveys as an answer to the growing concern with the quality of the data obtained with traditional survey modes (Blom Bosnjak, Das, and Lynn 2016; Callegaro et al. 2014). In this paper, we propose two estimation techniques suitable for survey data affected by nonresponse where auxiliary information exists at both the panel level and the population level. The first one consists in a one phase calibration (Deville and Särndal, 1992) at the population level and the second one consists in a calibration in two phases: at population level and at panel level. We illustrate the differences of two methodologies with a real application with data from the Citizen Panel for Social Research in Andalucía (PACIS), a probability-based mixed-mode panel of the Andalusian population.

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

A first effort to real time sufficient dimension reduction

Andreas Artemiou¹, Seung-Jun Shin², Yuexiao Dong³

¹Cardiff University (UK), ²Korea University (Korea), ³Temple University (USA)

In this era where massive datasets are becoming the norm we need of computationally efficient real time algorithms. In this work we propose a new methodology to achieve real time sufficient dimension reduction in regression. This methodology outperforms previous methodology in simple problems, and it is computationally very efficient. We show that in the massive dataset case where new data are constantly collected one can update very fast previous estimation of the algorithm by using the new data only. Moreover, we show how this can be extended to real time sparse sufficient dimension reduction.

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

Adaptive functional linear regression in the presence of dependence

Nicolas Asin

Université catholique de Louvain (Belgium)

We consider the estimation of a slope function in functional linear regression which models a relationship between a response and a random function in presence of dependence in the data generating process. Assuming an independent and identically distributed (iid.) sample it has been shown in Cardot and Johannes (2010) that a least squares estimator based on dimension reduction and thresholding can attain minimax-optimal rates of convergence up to a constant. Comte and Johannes (2012) propose model selection procedures for the orthogonal series estimator introduced first by Cardot and Johannes (2010). As this estimation procedure requires an optimal choice of a dimension parameter with regard amongst others to certain characteristics of the unknown slope function we investigate its fully data-driven choice based on a combination of model selection and Lepskis method inspired by Goldenshluger and Lepski (2011). For the resulting fully data-driven thresholded least squares estimator a non-asymptotic oracle risk bound is derived by considering either an iid sample or by dismissing the independence assumption. In the context of dependent observations the derived risk bounds coincide up to a constant as the independent case assuming sufficiently weak dependence characterized by a fast decay of the mixing coefficients. Employing the risk bounds the minimax optimality up to constant of the estimator is established over a variety of classes of slope functions.

Thursday 27
17:00-18:00
Great hall
(Juhlasali)

Functional object data analysis

John Aston

University of Cambridge (UK)

Functional Data Analysis has grown into a mature area of statistics over the last 20 years or so, but it is still predominantly based on the notion that the data are one dimensional i.i.d. curves belonging to some smooth Euclidean-like space. However, there have been many recent examples arising from the explosion of data being recorded in science and medicine that do not conform to these notions. Functional Object Data Analysis looks at the situation where the objects are functional-like, in that they are well represented in infinite dimensional spaces, but where there are other considerations such as geometry or higher dimensionality. We will examine cases where the data is multidimensional, where it no longer lives in a Euclidean space and where the objects are related such as in space or time. Including the data's intrinsic constraints can profoundly enhance the analysis. Examples from Linguistics, Image Analysis and Forensics will help illustrate the ideas.

On the correction term in the small ball probability factorisation for Hilbert random elements

Jean-Baptiste Aubin¹ and Enea Bongiorno²

¹ICJ INSA Lyon (France), ²Università del Piemonte Orientale (Italy)

Wednesday 26
10:30-12:00
Lecture room 8
(Sali 8)

The small-ball probability of a Hilbert valued process is considered. Recent works have shown that, for a fixed number d and as the radius ε of the ball tends to zero, the small-ball probability is asymptotically proportional to (a) the joint density of the first d principal components (PCs) evaluated at the center of the ball, (b) the volume of the d -dimensional ball with radius ε , and (c) a correction factor weighting the use of a truncated version of the process expansion. Under suitable assumptions on the decay rate of the eigenvalues of the covariance operator of the process, it has been shown that the correction factor in (c) tends to 1 as the dimension increases.

This work in progress studies the properties of the correction factor and introduces an estimator. It turns out that such estimator is consistent and asymptotically normal distributed. Features of such estimator provides insight on the dimensionality of the process.

A differential-geometric approach to generalized linear models with grouped predictors

Luigi Augugliaro

University of Palermo (Italy)

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

We propose an extension of the differential-geometric least angle regression method to perform sparse group inference in a generalized linear model. An efficient algorithm is proposed to compute the solution curve. The proposed group differential-geometric least angle regression method has important properties that distinguish it from the group lasso. First, its solution curve is based on the invariance properties of a generalized linear model. Second, it adds groups of variables based on a group equiangularity condition, which is shown to be related to score statistics. An adaptive version, which includes weights based on the Kullback-Leibler divergence, improves its variable selection features and is shown to have oracle properties when the number of predictors is fixed.

Unbiased estimation and prediction in the hierarchical two-parameter Poisson-Dirichlet process

Sergio Bacallado¹, Lorenzo Trippa² and Stefano Favaro³

¹University of Cambridge (UK), ²Harvard University (USA), ³Università degli Studi di Torino (Italy)

Monday 24
14:00-16:00
Lecture room 1
(Sali 1)

Bayesian nonparametric modelling of partially exchangeable sequences requires prior distributions on dependent random measures. A common choice is to let the random measures be conditionally independent given a latent discrete distribution. Such models give rise to natural urn schemes for sampling the prior distribution. On the other hand, posterior inference most often requires Markov chain Monte Carlo techniques. We propose an exact sampler for latent variables which allows exact sampling from the predictive distribution in the hierarchical two-parameter Poisson-Dirichlet process, a flexible family of dependent random measures. We also study the performance of debiasing methods proposed by Rhee and Glynn for the same class of models. It is shown through experiment that the Markov chain samplers considered mix very slowly in certain regions of parameter space, displaying a sharp transition.

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

Uniformly valid confidence intervals post-model-selection

François Bachoc
University Paul Sabatier (France)

We suggest general methods to construct asymptotically uniformly valid confidence intervals post-model-selection. The constructions are based on principles recently proposed by Berk et al. (2013). In particular, the candidate models used can be misspecified, the target of inference is model-specific, and coverage is guaranteed for any data-driven model selection procedure. After developing a general theory we apply our methods to practically important situations where the candidate set of models, from which a working model is selected, consists of fixed design homoskedastic or heteroskedastic linear models, or of binary regression models with general link functions.

Monday 24
14:00-16:00
Lecture room 5
(Sali 5)

Sequential testing for structural stability in approximate factor models

Matteo Barigozzi¹ and Lorenzo Trapani²
¹London School of Economics (UK), ²Cass Business School (UK)

In this paper we propose a class of tests for detecting breaks in the factor structure of large panels of time series. The procedure is based on the sample covariance matrix and its eigenvalues computed on a rolling sample. By testing at each point in time for the number of eigenvalues diverging (as the cross-sectional dimension going to infinity) we can detect breaks in real-time. In an approximate stationary r -factor model the r largest eigenvalues of the covariance matrix of the data diverge as the cross-sectional dimension N diverges while the remaining $N - r$ stay bounded. In this setting, when a break in the factor structure occurs the asymptotic behaviour of the eigenvalues is affected. In particular, when a change in the loadings takes place or a new factor is added the $r + 1$ -th eigenvalue will start diverging too. We therefore propose a new double randomised test statistics where we test for no-break occurrence, that is for $r + 1$ -th eigenvalue to be bounded against the alternative that it diverges with N . The approach proposed allows also to discriminate between breaks in the factor loadings matrix and changes in the number of factors. A numerical study shows the performance of the method in finite samples.

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

Statistical methods for outlier detection in space telemetries

Clémentine Barreyre
Airbus Defence and Space (France)

Satellite monitoring is an important task to prevent the failure of satellites. For this purpose, a large number of time series are analyzed in order to detect anomalies. In this study, we provide a review of such analysis focusing on statistical methods that rely on features extraction. In particular, we set up features based on fixed functional bases and data-based bases. Using statistical tests, we select the features that are the most relevant for the outlier detection purpose. Usual outlier detection methods (One-class SVM, Local Outlier Factor) are compared to a method that takes into account the temporal structure of the data. Those algorithms are applied to simulated and real telemetry data.

Vine based modeling of multivariate volatility time-series

Nicole Barthel¹, Claudia Czado¹ and Yarema Okhrin²

¹Technical University of Munich (Germany), ²University of Augsburg (Germany)

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

Studying realized volatility based on high-frequency data is of particular interest in asset pricing, portfolio management and evaluation of risk. We propose an approach for dynamic modeling and forecasting of realized variance-covariance matrices that allows for model parsimony and automatically guarantees positive definiteness of the forecasts. We use the one-to-one relationship between a correlation matrix and its associated set of partial correlations corresponding to any regular vine specification. Being algebraically independent, the partial correlations of the vine do not need to satisfy any algebraic constraint such as positive definiteness. We present several selection methods to choose, among the large number of vines, a vine structure, which captures the characteristics of the multivariate time-series of the correlation parameters and which has both a clear practical and probabilistic interpretation. The individual partial correlation time-series as well as the series of variances are studied using elaborate HAR models to account for long-memory behavior, non-Gaussianity and heteroscedasticity commonly observed for volatility data. The dependence between assets is flexibly modeled using vine copulas. Forecasts for the correlation matrices, and in combination with the predicted variances forecasts for the variance-covariance matrices, are obtained through the bijection to the forecasts of their associated partial correlation vines. The usefulness of the methodology is demonstrated through loss functions comparing its performance to existing methods which are based on the Cholesky decomposition for parametrization of the realized variance-covariance matrices.

Enhanced bootstrap method for statistical inference in the ICA model

Shahab Basiri, Esa Ollila and Visa Koivunen

Aalto University (Finland)

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

We develop low complexity and stable bootstrap procedures for FastICA estimators. Our bootstrapping techniques allow for performing cost efficient and reliable bootstrap-based statistical inference in the ICA model. Performing statistical inference is needed to quantitatively assess the quality of the estimators and testing hypotheses on mixing coefficients in the ICA model. The developed bootstrap procedures stem from the fast and robust bootstrap (FRB) method, which is applicable for estimators that may be found as solutions to fixed-point (FP) equations. We first establish analytical results on the structure of the weighted covariance matrix involved in the FRB formulation. Then, we exploit our analytical results to compute the FRB replicas at drastically reduced cost. The developed enhanced FRB method (EFRB) for FastICA permits using bootstrap-based statistical inference in a variety of applications (e.g., EEG, fMRI) in which ICA is commonly applied. Such an approach has not been possible earlier due to incurred substantial computational efforts of the conventional bootstrap. Our simulation studies compare the complexity and numerical stability of the proposed methods with the conventional bootstrap method. We also provide an example of utilizing the developed bootstrapping techniques in identifying equipotential lines of the brain dipoles from electroencephalogram (EEG) recordings.

Model selection tests for count models based on the empirical probability generating function

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

Apostolos Batsidis¹ and M. Dolores Jiménez-Gamero²

¹University of Ioannina (Greece), ²Universidad de Sevilla (Spain)

Given a sample from an unknown count population and two possibly misspecified parametric models, which may be separate, overlapping or nested, the problem of model selection consists in testing if the two competing models are equally close to the true population, against the hypothesis that one model is closer than the other. In this talk, measuring the closeness of a model to the truth by means of a distance based on the probability generating functions (PGF), a test approach to the model selection is presented. The motivation for using the probability generating function is that it is usually much simpler than the corresponding probability mass function, fully characterizes the distribution and possesses convenient features not shared by the characteristic or moment generating function. In this context, the unknown parameters are estimated based on a pgf-based divergence. As a prerequisite, some asymptotic properties of these estimators are studied. From these properties, consistent tests for model selection based on PGFs are given for separate, overlapping and nested models. The finite sample performance of the proposed methods is evaluated by simulations.

On deep learning as a remedy for the curse of dimensionality in nonparametric regression

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

Benedikt Bauer

TU Darmstadt (Germany)

In nonparametric regression estimation we have to impose smoothness assumptions on the regression function $m : \mathbb{R}^d \rightarrow \mathbb{R}$ in order to achieve nontrivial rates of convergence. For (p, C) -smooth (roughly speaking p times differentiable) functions, Stone (1982) showed that the optimal rate is $n^{-\frac{2p}{2p+d}}$, which can be extremely bad for high dimensions (so-called curse of dimensionality). Additional assumptions are necessary to circumvent this problem and achieve better rates of convergence. So for our result, we assume that the regression function satisfies a so-called generalized hierarchical interaction model of order $d^* \leq d$, a structure that includes many other types of functions assumed in the literature. Then we can show that least squares estimates based on multilayer feedforward neural networks are able to circumvent the curse of dimensionality and achieve the rate $n^{-\frac{2p}{2p+d^*}}$ (up to some logarithmic factor) after deriving a result concerning the approximation of multivariate smooth functions by multilayer feedforward neural networks with bounded weights and a bounded number of hidden neurons.

Empirical characteristic function-based estimation for locally stationary processes

Monday 24
14:00-16:00
Lecture room 6
(Sali 6)

Carina Beering¹, Anne Leucht¹, Marco Meyer² and Carsten Jentsch³

¹Technische Universität Braunschweig (Germany), ²Helmut Schmidt University Hamburg (Germany)

³Universität Mannheim (Germany)

We propose a kernel-type estimator for the local characteristic function (local CF) of locally stationary processes. Under weak moment conditions, we obtain joint asymptotic normality for local empirical characteristic functions (local ECF). Additionally, we provide a central limit theorem for the local empirical characteristic function process. We apply our asymptotic results to parameter estimation of time-varying α -stable distributions for $\alpha \in (1, 2)$. Finally, the applicability of our local ECF is illustrated by analyzing the pairwise dependence structure over time of log returns of German stock prices.

Multiscale inference for blind demixing with unknown number of sources

Merle Behr¹, Chris Holmes² and Axel Munk¹

¹University of Göttingen (Germany), ²University of Oxford (UK)

Friday 28
10:30-12:30
Lecture room 12
(Sali 12)

We discuss a new methodology for statistical Blind Source Separation (BSS) in a change-point regression setting. More precisely, for a given alphabet one observes the mixture $Y_j = \sum_{i=1}^m \omega_i f^i(x_j) + \epsilon_j$, where m is the unknown number of components, ω_i are unknown mixing weights, f^i unknown step functions taking values in the known alphabet, and ϵ is some normal noise with mean zero. The aim in these models is to identify m , ω_i , and f^i from the observations Y_j . This BSS problem occurs, for example, in digital communications and cancer genetics. Under weak identifiability conditions we obtain exact recovery within an ϵ -neighborhood of the mixture. Based on this we provide uniformly honest lower confidence bounds and estimators with exponential convergence rates for the number of source components. With this at hand, we obtain consistent estimators with almost optimal convergence rates and asymptotically uniform honest confidence statements for the weights and the sources. We explore our procedure with a data example from cancer genetics, where one aims to assign copy-number variations from genetic sequencing data to different tumor-clones and their corresponding proportions in the tumor.

Kernel spectral clustering of large dimensional data

Florent Benaych-Georges¹ and Romain Couillet²

¹Université Paris Descartes (France), ²Centrale-Supélec (France)

Tuesday 25
10:30-12:30
Great hall
(Juhlasali)

In this joint work with Romain Couillet (Central-Supelec), we give an analysis of kernel spectral clustering methods in the regime where the dimension p of the data vectors to be clustered and their number n grow large at the same rate. We demonstrate, under a k -class Gaussian mixture model, that the normalized Laplacian matrix associated with the kernel matrix asymptotically behaves similar to a so-called spiked random matrix. Some of the isolated eigenvalue-eigenvector pairs in this model are shown to carry the clustering information upon a separability condition classical in spiked matrix models. We evaluate precisely the position of these eigenvalues and the content of the eigenvectors, which unveil important properties concerning spectral clustering, in particular in simple toy models. Our results are then compared to the practical clustering of images from the MNIST database, thereby revealing an important match between theory and practice.

Marked and weighted empirical process of residuals

Vanessa Berenguer Rico and Bent Nielsen

University of Oxford (UK)

Thursday 27
10:30-12:30
Lecture room 7
(Sali 7)

In this paper, a new and wide class of marked and weighted empirical processes of residuals is introduced. The framework is general enough to accommodate both stationary and non-stationary regressions as well as a wide class of estimation procedures and it has interesting applications in specification testing and robust statistics. To show the workings and usefulness of the general theory, we apply it in two different scenarios. First, we show that the asymptotic behaviour of linear statistical functionals of residuals, expressed as integrals with respect to their empirical distribution functions, can be easily analyzed given the main theorems of the paper. In our context the integrands can be unbounded provided that the underlying distribution meets a certain moment condition. Second, we consider a normality test for robust regressions. The test is based on the third and fourth moments of robustified (truncated) residuals. The statistical analysis of the test reveals that the rescaling of the moment based statistic is case dependent, i.e., it depends on the estimation method being used. Hence, using the standard least squares normalizing constants in robust regressions will lead to incorrect inferences. However, if appropriate normalizations, which we derive, are used then we show that the test statistic is asymptotically chi-square.

Empirical likelihood ratio confidence interval for the linearly trimmed mean

Monday 24
10:30-12:30
Lecture room 10
(Sali 10)

Liga Bethere and Janis Valeinis
University of Latvia (Latvia)

In our work we obtain empirical likelihood confidence interval for linearly or smoothly trimmed mean (as mentioned in Stigler, 1973). The limiting distribution of the empirical likelihood ratio is a scaled Ξ_1^2 . The comparison to other L-estimates and also standard method of confidence interval estimate for linearly trimmed mean is done using simulations of contaminated distributions.

Reference: Stigler, S.M., 1973. The asymptotic distribution of the trimmed mean. *The Annals of Statistics*, 472–477.

Estimation of a change point in long-range dependent time series based on the Wilcoxon statistic

Thursday 27
14:00-16:00
Lecture room 5
(Sali 5)

Annika Betken
Ruhr-University Bochum (Germany)

We consider an estimator for the location of a shift in the mean of long-range dependent sequences. The estimation is based on the two-sample Wilcoxon statistic. Consistency and the rate of convergence for the estimated change point are established. In the case of a constant shift height, the $1/n$ convergence rate (with n denoting the number of observations), which is typical under the assumption of independent observations, is also achieved for long memory sequences. If the change point height decreases to 0 with a certain rate, the suitably standardized estimator converges in distribution to a functional of a fractional Brownian motion.

References: [1] Betken, A., 2016. Testing for change-points in long-range dependent time series by means of a self-normalized Wilcoxon test. *Journal of Time Series Analysis*, 20, 785–809. [2] Betken, A., 2017. Change-point estimation based on Wilcoxon tests in the presence of long-range dependence. arXiv:1612.0826

Sampling of warp maps for curve alignment

Tuesday 25
14:00-16:00
Lecture room 5
(Sali 5)

Karthik Bharath¹ and Sebastian Kurtek²
¹University of Nottingham (UK), ²Ohio State University (USA)

The problem of pairwise alignment of curve/functional data, when cast within the framework of determining a warp map that best matches the curves, will be discussed. Based on a constructive definition of a distribution on the set of warp maps of an interval and the unit circle, model- and stochastic optimisation-based approaches to the alignment problem will be discussed along with simulated and real-data examples.

Volatility estimation for stochastic PDEs using high-frequency observations

Markus Bibinger¹ and Mathias Trabs²

¹Philipps-Universität Marburg (Germany), ²Universität Hamburg (Germany)

Monday 24
10:30-12:30
Lecture room 7
(Sali 7)

We study the parameter estimation for parabolic, linear, second order, stochastic partial differential equations observing a mild solution on a discrete grid in time and space. The SPDE model covers many interesting applications, including the stochastic heat equation, the cable equation in neurobiology and term structure models. A high-frequency asymptotic regime is considered where the mesh of the grid in the time variable goes to zero. Focusing on volatility estimation, we provide a simple and easy to implement method of moments estimator based on squared increments of the process. The estimator is consistent and admits a central limit theorem. It should be emphasized that the theory considerably differs from the statistics for semi-martingales literature. The performance of the method is illustrated in simulations.

Sorted L-One Penalized Estimation (SLOPE)

Małgorzata Bogdan¹, E. van den Berg, W. Su, C. Sabatti and E. J. Candès

¹University of Wrocław (Poland)

Monday 24
14:00-16:00
Lecture room 13
(Sali 13)

Sorted L-One Penalized Estimator is an extension of LASSO aimed at control of the False Discovery Rate when the true vector of regression coefficients is sparse and regressors are weakly correlated. It also has the property of clustering regression coefficients according to the similarity of true signals and the correlation between different explanatory variables. In this talk we will explain the method and present new theoretical and empirical findings concerning its properties.

Heritability estimation in high-dimensional sparse linear mixed models

Anna Bonnet¹, Elisabeth Gassiat² and Céline Lévy-Leduc¹

¹AgroParisTech/INRA (France), ²Université Paris-Sud (France)

Tuesday 25
14:00-16:00
Lecture room 13
(Sali 13)

The heritability of a biological quantitative feature is defined as the proportion of its variation that can be explained by genetic factors. We propose an estimator for heritability in high dimensional sparse linear mixed models and we study its theoretical properties. We highlight the fact that in the case where the size N of the random effects is too large compared to the number n of observations, a precise estimation for heritability cannot be provided. Since in practice almost all datasets verify the condition $N \gg n$, we perform a variable selection method to reduce the size of the random effects and to improve the accuracy of heritability estimations. However, as shown by our simulations, this kind of approach only works when the number of non-zero components in the random effects (i.e. the genetic variants which have an impact on the phenotypic variations) is small enough. In face of this limitation, we proceeded to define an empirical criterion to determine whether it is possible or not to apply a variable selection approach. As an example of its use, we applied our method to estimate the heritability of the volume of several regions of the human brain

Confidence intervals in high-dimensional regression based on regularized pseudoinverses

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

Tom Boot
University of Groningen (Netherlands)

We introduce a new method to obtain valid confidence intervals for the coefficients in a linear regression model with almost exponentially more variables than observations. The method relies on approximating the inverse covariance matrix by a scaled Moore-Penrose pseudoinverse, and then applying the lasso to perform a bias correction. In addition, we propose random least squares, a new regularization technique which yields narrower confidence intervals with the same theoretical validity. Random least squares estimates the inverse covariance matrix using multiple low-dimensional projections of the data. This is shown to be equivalent to a generalized form of ridge regularization. This form of regularization is compared to standard ridge penalization. The methods are illustrated in Monte Carlo experiments and in an empirical example using quarterly data from the FRED-QD database, where gross domestic product is explained by a large number of macroeconomic and financial indicators.

MCMC algorithms for misidentification problems

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

Paula Bran, Matthew Schofield and Richard Barker
University of Otago (New Zealand)

Uncertainty about the true identities behind observations is known in statistics as a misidentification problem. The observations may be duplicated, wrongly reported or missing which results in error-prone data collection. This error can affect seriously the inferences and conclusions. A wide variety of MCMC algorithms have been developed for simulating the latent identities of individuals in a dataset using Bayesian inference. In this talk, the DIU (Direct Identity Updater) algorithm is introduced. It is a Metropolis-Hastings sampler with an application-specific proposal density. Its performance and efficiency is compared with two other algorithms solving similar problems. The convergence to the correct stationary distribution is discussed by using a toy example where the data is comprised of genotypes which includes uncertainty. As the state space is small, the behaviour of the chains is easily visualized. Interestingly, while they converge to the same stationary distribution, the transition matrices for the different algorithms have little in common.

On the sample autocovariance of a Lévy driven moving average process sampled at a renewal sequence

Thursday 27
14:00-16:00
Lecture room 8
(Sali 8)

Dirk-Philip Brandes and Imma Curato
Ulm University (Germany)

We consider a Lévy driven continuous time moving average process $X = (X_t)_{t \in \mathbb{R}}$ sampled at random times which follow a renewal structure independent of X . More precisely, let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a suitable kernel, $L = (L_t)_{t \in \mathbb{R}}$ a Lévy process,

$$X_t := \mu + \int_{\mathbb{R}} f(t-s) dL_s, \quad t \in \mathbb{R},$$

where $\mu \in \mathbb{R}$, $(T_n)_{n \in \mathbb{N}_0}$ a renewal sequence, and consider the process $Y_n := X_{T_n}$, $n \in \mathbb{N}_0$. We establish asymptotic normality of the sample mean, the sample autocovariance, and the sample autocorrelation of Y under certain conditions. Further, we compare our results to a non-random equidistant sampling and give some applications.

A bootstrap procedure in VAR systems with mixed-frequency data

Poster session

Alexander Braumann and Jens-Peter Kreiss
TU Braunschweig (Germany)

We propose a bootstrap procedure for mixed-frequency vector autoregressive time series, i.e. multivariate data where the univariate time series are observed at different sampling frequencies. The aim of the procedure is to approximate the finite sample distribution of estimators for functions of the parameters of the high-frequency autoregressive system, $y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \nu_t$, $\nu_t \stackrel{i.i.d.}{\sim} (0, \Sigma)$, $t \in \mathbb{Z}$, $\det(I_q - A_1 z - \dots - A_p z^p) \neq 0$, $\forall |z| \leq 1$, for example the autocovariances, autocorrelations, system parameters A_1, \dots, A_p and noise parameters. When mixed-frequency data is observed, the parameters of the autoregressive system can still be identified on an open and dense subset of the parameter space (see Anderson et al. 2016). In contrast to single-frequency data, not all innovations ν_t can be estimated and therefore the classical autoregressive bootstrap procedure can not be applied. Furthermore, it has been shown that the asymptotic distribution of the autocovariances depends on second and fourth moments of the innovations ν_t . Therefore a wild type bootstrap procedure is proposed which samples the innovations in the bootstrap world from a (discrete) distribution whose second and fourth moments are equal to the estimated second and fourth moments of the innovations. Finally the construction of bootstrap prediction intervals is addressed.

Regression with selectively missing covariates

Christoph Breunig
Humboldt-Universität zu Berlin (Germany)

Monday 24
10:30-12:30
Lecture room 12
(Sali 12)

We consider the problem of regression with selectively observed covariates. Identification of the regression function relies on instrumental variables that are independent of selection conditional on potential covariates. We propose a consistent two-step estimation procedure and derive its rate of convergence; also its pointwise asymptotic distribution is established. We demonstrate the usefulness of our method in survey data with non random missingness.

Edge-exchangeable graphs, sparsity, and paintboxes

Tamara Broderick¹, Trevor Campbell¹ and Diana Cai²
¹MIT (USA), ²University of Chicago (USA)

Monday 24
14:00-16:00
Lecture room 1
(Sali 1)

Representation theorems for exchangeable data have proved powerful tools for modeling and inference within Bayesian analysis. For example, a natural assumption for network data is that the distribution over vertices and edges is invariant to reordering the vertices. In this case, a consequence of the Aldous-Hoover representation theorem is that any sequence of such vertex-exchangeable graphs must have a number of edges that grows quadratically in the number of vertices. By contrast, real network-data applications typically exhibit subquadratic growth (i.e. sparsity). Here, the representation theorem can be used to show that all vertex-exchangeable models (e.g. stochastic block model, infinite relational model, infinite latent attribute model, latent feature relational model, random function model) are misspecified. In this work, we develop representations of edge-exchangeable network models, in which the probability of generating a graph sequence is invariant to reordering the arrival of edges. We provide a paintbox representation (akin to the Kingman paintbox representation for clustering) for edge-exchangeable graphs; we call our construction the graph paintbox. We show that it characterizes the distribution of all edge-exchangeable graphs and thereby provides a way to study their properties, such as sparsity and power laws. We also outline the class of graph frequency models, which construct the graph by sequentially sampling edges to instantiate based on latent vertex weights. We provide a characterization for all graph frequency models by defining an exchangeable vertex probability function (EVVPF). Our EVVPF is inspired by the exchangeable partition probability function from clustering, which is useful for developing practical posterior inference procedures.

Estimating the parameters of the Heston model from stock data using parallel multilevel SMC.

Katarzyna Brzozowska-Rup
Kielce University of Technology (Poland)

The project has been funded by the National Centre of Science granted on the basis of decision number DEC-2013/11/D/HS4/04014. We are analysing the question of estimating stochastic volatility from stock data using the Heston model with stochastic interest rates driven by the Cox-Ingersoll-Ross process (CIR). The model is the extension of the Black and Scholes model which takes into account a non-lognormal distribution of the assets returns, leverage effect and the important mean-reverting property of volatility. In addition, it has a semi-closed form solution for European options. The volatility and the model parameters are unobserved variables and they have to be estimated from the discret data (observations), which makes the significant inference and computational challenging. The parameters contained in model are estimated by the Multilevel Monte Carlo (MLMC) Euler discretization of the continuous Heston model and constructing the augmented likelihood function. The difficulty is that a closed form solution of the model is not known, and numerical approximation or discretization methods have to be employed for model evaluation. We propose using Sequential Monte Carlo (SMC) methods. SMC is an important class of online posterior density estimation algorithms. The idea behind this technique is to sample approximately from any sequence of probability distributions using a combination of importance sampling and resampling steps. Since the simulation of particles is computationally expensive, the purpose of the current work is to provide a framework for implementing generic SMC algorithms on parallel hardware. Parallel computing is a form of computation in which large complex problems can be divided into independent smaller ones and can be solved simultaneously. The most relevant benefits of parallel computing consist in improving memory efficiency, which allows us to reduce computation time. The efficacy of the proposed method is demonstrated under simulation and an empirical investigation of the suitability of the model to the S&P 500 Index and WIG Index.

References: [1] Cox, J.C, Ingersoll, J.E, Ross, S.A, 1985. A theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385–407. [2] Heston, S., 1993. A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Rev. Fin. Studies*, 6, 327–343. [3] Hurn, A.S., Lindsay, K.A, McClelland, A.J, 2012. Estimating the Parameters of Stochastic Volatility Models using Option Price Data. *NCER Working Paper Series*. [4] Durham, G., Geweke J., 2011. Massively Parallel Sequential Monte Carlo for Bayesian Inference. Available at SSRN: <https://ssrn.com/abstract=1964731> or <http://dx.doi.org/10.2139/ssrn.1964731> [5] Giles, M.B, 2008. Multilevel Monte Carlo path simulation. *Op. Res*, 56, 607–617. [6] Jasra, A., Kamatani, K., Law, K.J., Zhou, Y. Multilevel particle filter, arXiv preprint arXiv:1510.04977.

(Hybrid) pairwise likelihood inference for Plackett models

Manuela Cattelan¹ and Cristiano Varin²

¹University of Padova (Italy), ²Ca' Foscari University - Venice (Italy)

The specification and estimation of models for dependent discrete data has always attracted great interest. We consider the class of bivariate distributions introduced by Plackett (1965), and propose various parametric models for its dependence parameter. When the observations are binomial, our proposal corresponds to the specification of a parametric model for the cross-ratio parameter. However, in Plackett bivariate distributions, any marginal distribution function can be assumed, so the proposed approach can be used for various type of dependent count data arising in many contexts, including longitudinal and spatial data. In the latter case, we develop parametric models for the lorelogram (Heagerty and Zeger, 1998) adapted from the geostatistical literature. To avoid the need of specifying high-dimensional multivariate distributions, we suggest to employ composite likelihood methods (Varin, Reid and Firth, 2011) for inferential purposes. In particular, we propose to use either the marginal pairwise likelihood or the hybrid pairwise likelihood (Kuk, 2007), since both require only the specification of marginal bivariate distributions. Simulation

studies are performed in order to investigate the finite sample properties of the estimator and the proposed methodology is illustrated through applications to real data sets.

Mixture model modal clustering

José E. Chacón

Universidad de Extremadura (Spain)

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

The two most extended density-based approaches to clustering are surely mixture model clustering and modal clustering. In the mixture model approach, the density is represented as a mixture and clusters are associated to the different mixture components. In modal clustering, clusters are understood as regions of high density separated from each other by zones of lower density, so that they are closely related to certain regions around the density modes. If the true density is indeed in the assumed class of mixture densities, then mixture model clustering allows to scrutinize more subtle situations than modal clustering. However, when mixture modeling is used in a nonparametric way, taking advantage of the denseness of the sieve of mixture densities to approximate any density, then the correspondence between clusters and mixture components may become questionable. In this talk we introduce two methods to adopt a modal clustering point of view after a mixture model fit. Examples are provided to illustrate that mixture modeling can also be used for clustering in a nonparametric sense, as long as clusters are understood as the domains of attraction of the density modes. Finally, a simulation study reveals that the new methods are extremely efficient from a computational point of view, while at the same time they retain a high level of accuracy.

Functional registration and local variations

Anirvan Chakraborty and Victor M. Panaretos

École Polytechnique Fédérale de Lausanne (Switzerland)

Friday 28
10:30-12:30
Lecture room 5
(Sali 5)

We consider the problem of non-parametric registration of functional data that have been subjected to random deformation (warping) of their time scale. The separation of this phase variation ("horizontal" variation) from the amplitude variation ("vertical" variation) is crucial in order to properly conduct further analyses, which otherwise can be severely distorted. We determine precise conditions under which the two forms of variation are identifiable, under minimal assumptions on the form of the warp maps. We show that these conditions are sharp, by means of counterexamples. We then present a non-parametric registration method based on a "local variation measure", which bridges the registration problem of functional data and the problem of optimal transportation. The method is proven to consistently estimate the warp maps from discretely observed data, without requiring any penalization or tuning on the warp maps themselves. This circumvents the problem of over/under-registration often encountered in practice. Similar results hold in the presence of measurement error, with the addition of a pre-processing smoothing step. A detailed theoretical investigation of the strong consistency and the weak convergence properties of the resulting functional estimators is carried out including the rates of convergence. We also give a theoretical study of the impact of deviating from the identifiability conditions, quantifying it in terms of the spectral gap of the amplitude variation. Numerical experiments demonstrate the good finite sample performance of our method, and the methodology is further illustrated by means of a data analysis.

Monday 24
10:30-12:30
Lecture room 13
(Sali 13)

Inference for spatial trends

Ngai Hang Chan¹ and Rongmao Zhang²

¹The Chinese University of Hong Kong (China), ²Zhejiang University (China)

Non-stationary spatial models are widely applicable in diverse disciplines, ranging from biomedical sciences to geophysical studies. In many of these applications, testing for change in the trend and detecting the form of the trend are of crucial importance. By virtue of the m-dependence approximation of a stationary random field, a novel test based on integrated squared errors is proposed in this paper. Furthermore, asymptotic properties of this test are established and the limit distribution is found to be asymptotically normal. The method is illustrated by simulations and data analysis.

Monday 24
14:00-16:00
Great hall
(Juhlasali)

Graphical modeling with latent variables

Venkat Chandrasekaran¹, Armeen Taeb¹, John Reager², Michael Turmon², Pablo Parrilo³, Alan Willsky³

¹Caltech (USA), ²JPL (USA), ³MIT (USA)

Suppose we have a graphical model with sample observations of only a subset of the variables. Can we separate the extra correlations induced due to marginalization over the unobserved, hidden variables from the structure among the observed variables? In other words is it still possible to consistently perform model selection despite the unobserved, latent variables? We demonstrate that the key problem that arises (in Gaussian models) is one of decomposing the precision matrix of the observed variables into a sparse matrix (representing graphical model structure among the observed variables) and a low-rank matrix (representing the effects of marginalization over the hidden variables). Under suitable conditions, such a decomposition can be accomplished by an estimator that is given by a tractable convex program. The geometric aspects of our approach are highlighted, with the algebraic varieties of sparse matrices and of low-rank matrices playing an important role. We also demonstrate the empirical utility of our approach in analyzing the California reservoir system.

Joint work with Armeen Taeb (Caltech); John Reager, Michael Turmon (JPL); Pablo Parrilo, Alan Willsky (MIT)

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

Asymptotic theory for linear mixed-effects model selection

ChihHao Chang

National University of Kaohsiung (Taiwan)

We propose a conditional generalized information criterion (CGIC) for linear mixed-effects model selection, which is extended from the conditional Akaike's information criterion (CAIC) of Vaida and Blanchard (2005). We establish conditions for the consistency and the asymptotic loss efficiency of CGIC in selecting fixed and random effects under some regularity conditions in terms of a squared error loss. In addition, we introduce a notion of second-order risk efficiency to further distinguish among models that satisfy the first-order efficiency. Moreover, the asymptotic result is applied to spatial regression model selection under both the fixed domain and the increasing domain asymptotic frameworks.

On adaptive subject selection in active learning

Yuan-Chin Chang
Academia Sinica (Taiwan)

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

To apply active learning procedures to construct classification rule allows us to select and add new subjects joining an existing training set based on the information obtained with the current training samples. Naturally, from statistical viewpoint, it is a sequential procedure with adaptive sample selection. In machine learning literature, the concepts of the uncertainty sampling and the statistical experimental design are two popularly used ideas in subject selection. In statistical literature, we can also readily find the discussions about the properties of this type of adaptive methods, but there is lack of discussion about subject selection. In this study, we will consider using the idea in robust analysis to subject selection for adaptive learning procedures. Both model-dependent and model-free methods will be discussed.

Bootstrap confidence intervals after model selection

Ali Charkhi
KU Leuven (Belgium)

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

Taking the selection randomness into account for inference is not easy. Different model selection criteria define different types of variability. In this article, we propose a new algorithm which can be used in linear and non-linear models and it is applicable for different model selection criteria. The algorithm is based on conditional m-out of-n bootstrap. The proposed algorithm produces conservative confidence intervals for some parameters and it generates reasonably accurate confidence intervals for other parameters. Simulation results together with real data examples illustrate the working and applicability of the proposed bootstrap algorithm.

Non-stationary modeling of tail dependence of two subjects concentration

Valerie Chavez-Demoulin¹, Kshitij Sharma¹ and Pierre Dillenbourg²

¹University of Lausanne (Switzerland), ²École Polytechnique Fédérale de Lausanne (Switzerland)

Wednesday 26
10:30-12:30
Lecture room 13
(Sali 13)

Collaborative learning research offers many methodological challenges when its focus is on exploring eye-tracking data. A natural approach is to analyze sequences of the observations that specifically represent collaboration. These episodes are, for instance, coincidental times of individual high concentration (defined by low entropy). In this paper, we propose a flexible model that describes the tail dependence structure between two subjects entropy when the pair is collaborating to accomplish a goal. The method is based on the generalized additive models (GAM) for dependence structure developed by Vatter and Chavez-Demoulin (2015) adapted to tail dependence coefficients.

An information matrix goodness-of-fit test of the conditional logistic model

Poster session

Li-Ching Chen

Tamkang University (Taiwan)

The case-control design has been widely applied in clinical and epidemiological studies to investigate the association between risk factors and a given disease. The retrospective design can be easily implemented and is more economical over prospective studies. To adjust effects for confounding factors, methods such as stratification at the design stage and may be adopted. When some major confounding factors are difficult to be quantified, a matching design provides an opportunity for researchers to control the confounding effects. The matching effects can be parameterized by the intercepts of logistic models and the conditional logistic regression analysis is then adopted. This study demonstrates an information-matrix-based goodness-of-fit statistic to test the validity of the logistic regression model for matched case-control data. The asymptotic null distribution of this proposed test statistic is inferred. It needs neither to employ a simulation to evaluate its critical values nor to partition covariate space. The asymptotic power of this test statistic is also derived. The performance of the proposed method is assessed through simulation studies. An example of the real data set is applied to illustrate the implementation of the proposed method as well.

A shoes and doors problem

Wednesday 26
10:30-12:30
Lecture room 8
(Sali 8)

May-Ru Chen

National Sun Yat-sen University (Taiwan)

A man has a house with n doors. Initially he places k pairs of walking shoes at each door. For each walk, he chooses one door at random, and puts on a pair of shoes, returns after the walk to a randomly chosen door and takes off the shoes at the door. Sooner or later he discovers that no shoes are available at the door he has chosen for a further walk (and has to walk barefoot). We derive the asymptotic behavior as n goes to infinity of the distribution and moments of the number of finished walks. We also consider a more general setting where the numbers of pairs of shoes initially placed at the doors are not necessarily equal.

Geographically weighted quantile regression with spatially lagged dependent variable

Poster session

Vivian Yi-Ju Chen¹, Tse-Chuan Yang² and Yin-Jhen Yan¹

¹Tamkang University (Taiwan), ²University at Albany, The State University of New York (USA)

Geographically Weighted Regression (GWR) and Quantile Regression (QR) have been found useful in capturing spatial non-stationarity and modeling the entire distribution of the regressand, respectively. While GWR and QR are popular in the field of geography and statistics, they have never been integrated until recently (Chen et al., 2012). Chen and her colleagues (2012) are among the first to develop the quantile-based spatial analysis tool that permits researchers to investigate the spatial non-stationarity across the whole distribution of a dependent variable. The innovative spatial statistical methodology is called the geographically weighted quantile regression (GWQR). Though the GWQR is a novel approach to thoroughly handling heterogeneity, it is argued that adding the ability to consider spatial dependence to GWQR will make it a more complete spatial analysis methodology. The goal of this study is then to address such perceived gap by developing a new method, called GWQR-SL hereafter, which incorporates the spatial lag (SL) effect into the GWQR. We first formulate the modeling specification, and then develop bootstrap methods to conducting the inference of model parameters. A simulation study is conducted to examine and further validate the performance of the GWQR-SL methodology. Finally, the proposed model is applied to a real dataset as an empirical illustration.

Comparison of two related biomarkers based on the area under ROC curve and Youden index

Yuh-Ing Chen

National Central University (Taiwan)

Thursday 27
14:00-16:00
Lecture room 6
(Sali 6)

The problem under consideration in this paper is to compare two diagnostic markers in a case-control study where the paired marker measurements are possibly correlated. Note that the area under the receiver operation characteristic (ROC) curve, denoted by AUC, is widely employed to evaluate the effectiveness of a diagnostic marker. Moreover, Youden's index J , is important for the marker assessment since it provides an optimal cutoff value of the marker for disease diagnostic. Therefore, to make a comparison of the two markers based on the AUC and J , simultaneously, we construct joint confidence regions for the difference of the two correlated AUCs and that of the correlated J s. The parametric confidence region is derived when paired marker measurements are distributed according to bivariate normal distributions. A nonparametric confidence region is also obtained when bivariate normal distributions are not tenable. A simulation study is further conducted to investigate the coverage probabilities and areas of the proposed confidence regions where data are generated from a variety of bivariate distributions with normal or generalized gamma marginal distributions linked by copula functions. Finally, the proposed confidence regions are illustrated by using a real data set on the diagnosis of Pancreatic cancer.

Comparison of two crossing survival curves

Zheng Chen

Southern Medical University (China)

Poster session

In statistical analysis of cancer and other clinical trial data, we often encounter the problem of comparing the equivalence of survival times with censored observations from two groups. The most commonly used method for such comparisons is the log-rank test, which is the most reliable method for detecting differences between survival curves when the hazard rates of the treatment and control groups are proportional. However, when two survival curves are crossed-over, which is a violation of the proportional hazard assumption, the log-rank test will lose reliability and may become unsuitable because the earlier differences are cancelled out by the emerging later differences in the opposite direction. A number of studies were confronted with analysis of the differences between two crossing survival curves using the log-rank test. In this work, we illustrate a simulated data analysis using several alternative methods and explain how methods should be selected when two survival curves are crossed. This study was supported by the National Natural Science Foundation of China (81673268).

Limit distributions of maximum vertex degree in a conditional configuration graph

Irina Cheplyukova and Yuri Pavlov

Karelian Research Center of Russian Academy of Sciences (Russia)

Friday 28
10:30-12:30
Lecture room 6
(Sali 6)

We consider a configuration graph with N vertices. The degrees of vertices are drawn independently from a power-law distribution with a positive parameter r . These degrees are equal to the number of each vertex's numbered semiedges. The graph is constructed by joining all semiedges pairwise equiprobably to form edges. Such random graphs are widely used for studying the structure and dynamics of various social, telecommunication networks and Internet topology. Let parameter r is a random variable following uniform distribution on the finite interval $[a, b]$, $0 < a < b$. We obtained the limit distributions of the maximum vertex degree under the conditions that number of edges is equal to n and N, n tend to infinity. The study was partially supported by the Russian Foundation for Basic Research, grant 16-01-00005.

Tuesday 25
14:00-16:00
Lecture room 1
(Sali 1)

Simultaneous multiple change-point and factor analysis for high-dimensional time series

Haeran Cho¹, Matteo Barigozzi² and Piotr Fryzlewicz²

¹University of Bristol (UK), ²LSE (UK)

We propose the first comprehensive treatment of high-dimensional time series factor models with multiple change-points in their second-order structure. We operate under the most flexible definition of piecewise stationarity, and estimate the number and locations of change-points consistently as well as identifying whether they originate in the common or idiosyncratic components. Through the use of wavelets, we transform the problem of change-point detection in the second-order structure of a high-dimensional time series, into the (relatively easier) problem of change-point detection in the means of high-dimensional panel data. Our methodology circumvents the difficult issue of the accurate estimation of the true number of factors by adopting a screening procedure. In extensive simulation studies, we show that factor analysis prior to change-point detection improves the detectability of change-points, and identify and describe an interesting spillover effect in which substantial breaks in the idiosyncratic components get, naturally enough, identified as change-points in the common components, which prompts us to regard the corresponding change-points as also acting as a form of factors. Our methodology is implemented in the R package `factorcpt`, available from CRAN.

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

Regularization parameter selection by residual-based bootstrap

Justin Chown, Nicolai Bissantz and Holger Dette

Ruhr Universitat Bochum (Germany)

Residual-based analysis is generally considered a cornerstone of statistical methodology. For a special case of indirect regression, we investigate the residual-based empirical distribution function and provide a uniform expansion of this estimator, which is also shown to be asymptotically most precise. Studying the empirical process of the residuals requires new results concerning the indirect regression estimator, and we link the rate of the strong uniform consistency of the indirect regression function estimator to the already known minimax optimal rate. This investigation naturally leads to a completely data-driven technique for selecting a regularization parameter used in the indirect regression function estimator. The resulting methodology is based on a smooth bootstrap of the model residuals, which exploits smoothness in the error distribution. This work presents a new criteria for selecting an appropriate regularization, and, further, consistency of bootstrap-based quantiles of test statistics based on continuous functionals of the error distribution.

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

Dynamic hazard models using state space models

Benjamin Christoffersen

Copenhagen Business School (Denmark)

State space models provides a computationally efficient way to model time-varying coefficients in survival analysis. This is implemented in the `dynamichazard` package on CRAN. We cover the methods for estimation and models implemented in `dynamichazard` along with examples. Computational pros and cons of the methods will be highlighted. The models and methods are applied to a large data set with hard disk failures. The presented methods can be applied to other topics than survival analysis at the same computational cost for general non-linear filtering problems.

Effects of model selection and weight choice on inference

Gerda Claeskens
KU Leuven (Belgium)

Wednesday 26
9:00-10:00
Great hall
(Juhlasali)

Weights may be introduced in the estimation process in several ways: estimators may be weighted by zero/one weights in a model selection procedure such that only a 'selected' estimator is kept for further consideration; weighted estimators may employ more general weights, which can be optimised in some fashion; or weights can be introduced during the estimation stage, resulting in so-called composite estimators which minimise a weighted loss function. Several such estimation strategies are discussed and compared. In general, the randomness of the weights makes inference challenging. For some special cases, including random 0/1 weights from selection by Akaike's information criterion, it is possible to construct asymptotic confidence regions which are uniformly valid and which incorporate the selection uncertainty.

An application of the inverse power Lindley distribution in left censored data

Emilio Coelho-Barros¹, Jorge Achcar² and Josmar Mazucheli³
¹UTFPR/CP (Brazil), ²USP/FMRP (Brazil), ³UEM (Brazil)

Poster session

In this paper we consider the use of classical and Bayesian inference methods to analyze a data set in presence of left censoring. Several probability distributions have been proposed in the literature, especially with the aim of obtaining models that are more flexible relative to the behaviors of the density and hazard rate functions. Recently, a new generalization of the Lindley distribution was proposed by Ghitany et al. (2013), called power Lindley distribution. Another generalization was proposed by Sharma et al. (2015), known as inverse Lindley distribution. In our approach, we used a distribution obtained from these two generalizations and named as inverse power Lindley distribution. A numerical illustration is presented considering data of thyroglobulin levels present in a group of individuals with differentiated cancer of thyroid.

Regression models for the restricted residual mean time for right-censored and left-truncated data

Giuliana Cortese
University of Padova (Italy)

Wednesday 26
10:30-12:30
Lecture room 12
(Sali 12)

Hazard functions are typically studied via Cox's regression survival models. The resulting hazard ratios estimate the relative risks, and these measures are difficult to interpret and hard to be translated into clinical benefits in terms of increased survival time. The main objective is often to study survival functions, aiming at a global summary over a time period. Therefore, there has been increasing interest in summary measures based on the survival function, easier to interpret than the hazard ratio. For the survival time T , we consider the mean residual time $E(T - t | T \geq t)$, which has recently received increasing interest in literature. This quantity represents a partial area under the survival function and is interpreted as the residual life expectancy of individuals who had survived up to a certain time t . However, due to the presence of right censoring, the tail of the survival distribution is often difficult to be correctly estimated. As a solution, we propose to study the restricted mean residual time $E(\min(T, \tau) - t | T \geq t)$, for any $\tau > 0$. We present regression models for this new measure, based on weighted estimating equations and the inverse probability of censoring weighted estimator to model potential right censoring. We also show how to extend the models to deal with delayed entries. Estimation performance is investigated by simulation studies. Using real data about the Danish Monitoring Cardiovascular Risk Factor Survey, we illustrate an application of these regression models when the link function is identity or exponential.

Testing the equality of a large number of populations with applications to genomics

Tuesday 25
10:30-12:30
Lecture room 7
(Sali 7)

Marta Cousido-Rocha¹, Jacobo de Uña-Álvarez¹ and Jeffrey D. Hart²

¹University of Vigo (Spain), ²Texas A & M University (USA)

The statistical analysis of data from genome-wide studies is non-trivial due to their complexity. In particular, such analysis often involves the comparison of a large number of correlated populations (e.g. gene expression levels) from a small number of individuals. In this work we introduce a goodness-of-fit method to test for the equality of densities in such a small sample, large dimension setting. The method extends the one in Zhan and Hart (2012) in that it allows for correlated outcomes across populations. The asymptotic null distribution of the test as the number of populations grows is derived under mixing conditions. The consistency of the test under certain alternatives is established. Simulation studies are conducted. Applications to genomics are included.

Methods for estimating spectral clustering change points for multivariate time series

Monday 24
14:00-16:00
Lecture room 7
(Sali 7)

Ivor Cribben

University of Alberta (Canada)

Spectral clustering is a computationally feasible and model-free method widely used in the identification of communities in networks. We introduce data-driven methods, including Network Change Points Detection (NCPD), which detect change points in the network structure of a multivariate time series, with each component of the time series represented by a node in the network. Spectral clustering allows us to consider high dimensional time series where the number of time series is greater than the number of time points ($n < p$). The methods allow for estimation of both the time of change in the network structure and the graph between each pair of change points, without prior knowledge of the number or location of the change points. Permutation and bootstrapping methods are used to perform inference on the change points. The methods are applied to various simulated high dimensional data sets as well as to a resting state functional magnetic resonance imaging (fMRI) data set.

Weak dependence and GMM estimation for supOU and mixed moving average processes.

Thursday 27
14:00-16:00
Lecture room 8
(Sali 8)

Imma Valentina Curato and Robert Stelzer

Ulm University (Germany)

We consider a mixed moving average process X driven by a Lévy basis and show that is a weakly dependent process. Using this property, we show that sample mean and autocovariances of X have a limiting normal distribution. As application, we can then apply a Generalized Method of Moments estimation for the supOU process and the supOU stochastic volatility model after choosing a suitable distribution for the mean reverting parameter. For these estimators we analyze the asymptotic behavior in detail.

D-vine quantile regression

Claudia Czado¹, Carole Bernard² and Daniel Kraus¹

¹Technical University of Munich (Germany), ²Grenoble École de Management (France)

Thursday 27
14:00-16:00
Lecture room 1
(Sali 1)

Quantile regression methods have become very popular in applications, where one is interested in assessing high quantiles conditional on other variables. We first show that ordinary quantile regression induces Gaussian dependence, thus might be insufficient for applications requiring tail dependence as is often the case for financial data. We then show how the flexible vine copula class of D-vines can be utilized to allow for non Gaussian dependence. Additionally we derive an algorithm for the selection of important covariates. Further this copula based approach does not have the problem of crossing quantiles for different α levels. The approach is tailored to the estimation of quantiles and is analytical and not simulation based. We illustrate the superiority of the approach through data applications and simulation.

References: [1] Bernard, C., Czado, C., 2015. Conditional quantiles and tail dependence. *Journal of Multivariate Analysis*, 138, 104-126. [2] Kraus, D., Czado, C., 2016. D-vine copula based quantile regression. *Computational Statistics & Data Analysis*. To appear.

Statistical learning of dynamic systems

Itai Dattner

University of Haifa (Israel)

Monday 24
14:00-16:00
Lecture room 12
(Sali 12)

Dynamic systems are ubiquitous in nature and are used to model many processes in biology, chemistry, physics, medicine, and engineering. In particular, systems of (deterministic or stochastic) differential equations as well as discrete models are commonly used for the mathematical modeling of dynamic processes. These systems describe the interrelationships between the variables involved, and depend in a complicated way on unknown quantities (e.g., initial values, constants or time dependent parameters). Learning dynamic systems involves the ‘standard’ statistical problems such as studying the identifiability of a model, estimating model parameters, predicting future states of the system, testing hypotheses, and choosing the ‘best’ model. However, modern dynamic systems are typically very complex: nonlinear, high dimensional and only partly measured. Moreover, data may be sparse and noisy. Thus, statistical learning (inference, prediction) of dynamical systems is not a trivial task in practice. In this talk we will present some recent theoretical results and methodologies concerning identifiability and estimation of dynamic systems. We will also discuss real data examples coming from diverse areas such as infectious diseases and biology.

An adapted loss function for censored quantile regression

Mickaël De Backer¹, Anouar El Ghouch¹ and Ingrid Van Keilegom²

¹Université catholique de Louvain (Belgium), ²KU Leuven (Belgium)

Monday 24
14:00-16:00
Lecture room 10
(Sali 10)

Since the pioneering work of Koenker and Bassett (1978), quantile regression has become a pre-eminent substitute to the classical least-squares regression in both theoretical and applied statistics. While mean regression models solely grasp the central behavior of the data, quantile regression allows the analyst to investigate the complete distributional information of the dependence of the response variable on a set of one or more covariates at hand. In that sense, quantile regression provides a more complete view of relationships between the response and the covariates. Existing literature on the estimation of a quantile regression function includes numerous methodologies for fully observed response observations. In practice however, many interesting applications are affected by possible right censoring of the responses due, for instance, to the withdrawal of patients in biomedical studies, or the end of the follow-up period of a clinical trial. When confronted to such incomplete data, the statistical methodologies are to be adapted in order to avoid underestimation of the quantiles of the true response variable. In this perspective, the main rationale of the current

literature has been so far to take censoring into account through the formulation of ad hoc synthetic observations or weighting schemes. In this project however, we discuss a novel methodology for the estimation of quantiles starting from an alternative point of view as the idea is to tackle censoring at the very level of the loss function usually employed for the computation of quantiles, the so-called “check” function. Subsequently, when considering the inclusion of covariates for conditional quantile estimation, by defining a new general loss function the proposed methodology opens the gate to numerous parametric, semiparametric and nonparametric modelling techniques. In order to illustrate this statement, we consider the well-studied linear regression. Consistency of the resulting estimator is obtained under classical assumptions, while numerical examples and a simple application to a real dataset illustrate the adequateness and finite sample performance of the proposed estimator with respect to the literature.

Tuesday 25
10:30-12:30
Lecture room 5
(Sali 5)

On boundary crossing by stochastic processes

Victor de la Pena
Columbia University (USA)

In this talk we introduce an approach for obtaining (sharp) bounds on the expected crossing time of stochastic process. It can be thought of as an extension of the concept of boundary crossing of non-random functions to the case of stochastic processes. It can also be viewed as an extension of Wald’s equations for sums of independent random variables to the case of processes with arbitrary dependence structures.

Thursday 27
14:00-16:00
Lecture room 5
(Sali 5)

Testing for structural breaks via ordinal pattern dependence

Herold Dehling¹ and Alexander Schnurr²
¹Ruhr-University Bochum (Germany), ²University of Siegen (Germany)

We propose new concepts to analyze and model the dependence structure between two time series. Our methods rely exclusively on the order structure of the data points. Hence the methods are stable under monotone transformations of the time series and robust against small perturbations or measurement errors. Ordinal pattern dependence can be characterized by four parameters. We propose estimators for these parameters, and we calculate their asymptotic distribution. Furthermore, we derive a test for structural breaks within the dependence structure.

Reference: [1] Schnurr, A., Dehling, H., 2017. Testing for Structural Breaks via Ordinal Pattern Dependence. *Journal of the American Statistical Association*, DOI:1080/01621459.2016.1164706.

Thursday 27
10:30-12:30
Great hall
(Juhlasali)

Birds of a feather or opposites attract - effects in network modelling

Maria Deijfen¹ and Robert Fitzner²
¹Stockholm University (Sweden), ²Eindhoven University of Technology (Netherlands)

We study properties of some standard network models when the population is split into two types and the connection pattern between the types is varied. The studied models are generalizations of the Erdős-Renyi graph, the configuration model and a preferential attachment graph. For the Erdős-Renyi graph and the configuration model, the focus is on the component structure. We give expressions for the critical parameter, indicating when there is a giant component in the graph, and study the size of the largest component by aid of simulations. For the preferential attachment model, we analyze the degree distributions of the two types and derive explicit expressions for the degree exponents.

Joint work with Robert Fitzner (Eindhoven University of Technology).

Inferring visual processing in the brain

Susanne Ditlevsen¹, Kang Li¹, Claus Bundesen¹, Søren Kyllingsbæk¹, Stefan Treue² and Vladislav Kozyrev³

¹University of Copenhagen (Denmark), ²Goettingen University (Germany), ³Cognitive Neuroscience Laboratory Goettingen (Germany)

Thursday 27
14:00-16:00
Lecture room 6
(Sali 6)

A fundamental question concerning the way the visual world is represented in our brain is how a cortical cell responds when its classical receptive field contains more than a single stimulus object. It is a statistically challenging problem how to infer such behavior and distinguish between different explanatory models from neurobiological data. Particular challenges are that the data is partially observed, highly noisy and autocorrelated. A standard way to deal with noisy data is to average over trials. In this talk I will argue that this might blur or entirely remove essential characteristics and mechanisms, which are fundamental for understanding brain function. Two opposing models have been proposed in the literature. In the response-averaging model [1], the firing rate of the cell to a pair of stimulus objects is a weighted average of the firing rates to the individual objects. By contrast, in the probability-mixing model [2], the cell responds to the pair of objects as if only one of the objects was present in any given trial. Here we compare the abilities of the two models to account for spike trains recorded from single cells in the middle temporal visual area (MT) of rhesus monkeys, using point process techniques. The results support the probability-mixing model [3].

References: [1] Reynolds, J. H., Chelazzi, L., Desimone, R., 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19, 1736–1753. [2] Bundesen, C., Habekost, T., Kyllingsbæk, S., 2005. A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological Review* 112, 291. [3] Li, K., Kozyrev, V., Kyllingsbæk, S., Treue, S., Ditlevsen, S., Bundesen, C., 2016. Neurons in primate visual cortex alternate between responses to multiple stimuli in their receptive field. *Frontiers in Computational Neuroscience*, 10, 141.

Asymptotic properties of likelihood estimators in multivariate extremes

Clement Dombry¹, Sebastian Engelke² and Marco Oesting³

¹Université de Franche-Comté (France), ²École Polytechnique Fédérale de Lausanne (Switzerland), ³University of Siegen (Germany)

Wednesday 26
10:30-12:30
Lecture room 13
(Sali 13)

Max-stable distributions and processes are important models for extreme events and the assessment of tail risks. The full, multivariate likelihood of a parametric max-stable distribution is complicated and only recent advances enable its use in both frequentist and Bayesian statistics. The asymptotic properties of the maximum likelihood estimator and the median of the posterior distribution in multivariate extremes are mostly unknown. We provide natural conditions on the exponent function and the angular measure of the max-stable distribution that ensure asymptotic normality of these likelihood estimators. We show the effectiveness of this result by applying it to popular parametric models in multivariate extreme value statistics. and to the most commonly used spatial max-stable processes.

Measure concentration in machine learning algorithms

Devdatt Dubhashi
Chalmers University (Sweden)

Tuesday 25
10:30-12:30
Lecture room 13
(Sali 13)

We will discuss applications of concentration of measure in machine learning with three examples: large scale first order stochastic methods for optimization, matrix completion and related problems and Thompson sampling. Some challenging open problems in Deep Learning algorithms possibly related to concentration phenomena will be discussed at the end.

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

Bootstrap for triangular arrays with growing period

Anna Dudek
Universite Rennes 2 (France)

In the recent years the number of bootstrap applications for periodic processes has constantly grown. An important class of periodic processes are periodically correlated (PC) processes. They are widely used for modeling in many different settings such as climatology, hydrology, mechanics, vibroacoustics and economics. PC processes can be applied when the period length is constant. In our presentation we focus on a different case, i.e. when the period length is changing over time. A very important example illustrating this phenomena is a chirp signal. It is a signal in which frequency increases or decreases over time. Chirps are commonly met in nature. For example in audio signals (animal communication, echolocation), radar and sonar systems, astrophysics (gravitational waves radiated by coalescing binaries), mechanics and vibrations (e.g. car engines), medicine (EEG data - epileptic seizure) and seismography. In our work we consider a case, when period length is growing in time. This corresponds to the so-called down-chirp signal. To model such phenomena periodic random arrays can be used. We focus on the frequency domain i.e., on the Fourier coefficients of the mean and autocovariance functions. To obtain confidence intervals for parameters of interest we use the circular version of the Generalized Seasonal Block Bootstrap method. We illustrate our results by an example based on simulated data.

Referenses: Dudek, A. E., 2016. First and second order analysis for periodic random arrays using block bootstrap methods. *Electron. J. Statist.*, 10(2), 2561–2583.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655394.

Thursday 27
14:00-16:00
Lecture room 5
(Sali 5)

Limit theorems for Hilbert space-valued linear processes under long range dependence

Marie-Christine Düker
Ruhr University Bochum (Germany)

Let $(X_k)_{k \in \mathbb{Z}}$ be a linear process with values in a separable Hilbert space \mathbb{H} given by $X_k = \sum_{j=0}^{\infty} (j+1)^{-N} \varepsilon_{k-j}$ for each $k \in \mathbb{Z}$, where $N : \mathbb{H} \rightarrow \mathbb{H}$ is a bounded, linear normal operator and $(\varepsilon_k)_{k \in \mathbb{Z}}$ is a sequence of independent, identically distributed \mathbb{H} -valued random variables with $E\varepsilon_0 = 0$ and $E\|\varepsilon_0\|^2 < \infty$. We investigate the central for $(X_k)_{k \in \mathbb{Z}}$ when the series of operator norms $\sum_{j=0}^{\infty} \|(j+1)^{-N}\|_{op}$ diverges.

Thursday 27
10:30-12:30
Lecture room 5
(Sali 5)

Tests for qualitative features in the random coefficients model

Konstantin Eckle
Ruhr University Bochum (Germany)

The random coefficients model can be thought of as linear regression model with random regression coefficients. Given data from this model, the statistical challenge is to recover information about the joint density of the random coefficients which is a multivariate and ill-posed problem. Because of the curse of dimensionality and the ill-posedness, pointwise nonparametric estimation of the joint density is difficult and suffers from slow convergence rates. Larger features, such as an increase of the density along some direction or a well-accentuated mode can, however, be much easier detected from data by means of statistical tests. In this article, we follow this strategy and construct such statistical tests and confidence statements for qualitative features of the joint density of the random coefficients, such as increases, decreases and modes. We propose a multiple testing approach based on aggregating single tests that are designed to extract shape information on fixed scales and directions. Using recent tools for Gaussian approximations of multivariate empirical processes, we derive expressions for the critical value. We apply our method to simulated and real data.

The distance correlation coefficient for right-censored survival data

Dominic Edelmann
German Cancer Research Center (Germany)

Monday 24
14:00-16:00
Lecture room 6
(Sali 6)

Székely, Rizzo, and Bakirov introduced the powerful concept of distance correlation as a measure of dependence between random variables. In contrast to Pearson correlation, which only measures linear dependence, distance correlation can detect any kind of dependence including nonlinear or even nonmonotone associations.

In biomedical applications, one is interested in finding associations between patient characteristics (e.g. the expression of a certain gene) and survival time. This kind of data usually involves right-censoring, implying that the exact survival time cannot be observed for individuals who leave the study before they die. Under this missingness of information, it is not straightforward to derive consistent estimates for the distance correlation.

In this talk, we present a suitable distance correlation coefficient for right-censored survival data accounting for censored observations via inverse probability of censoring weights (IPCW). In particular, we show that the population distance correlation between covariate data and survival times can be approximated by an IPC-weighted U-statistic. Using a simulation study, we demonstrate that this coefficient is helpful to reveal associations which can not be found with standard methods. Finally, we apply this coefficient on variable screening for ultrahigh dimensional DNA methylation data.

Variable selection properties of sparse principal component analysis

Andreas Elsener and Sara van de Geer
ETH Zurich (Switzerland)

Friday 28
10:30-12:30
Lecture room 7
(Sali 7)

Variable selection properties of sparse Principal Component Analysis (PCA) have attracted the interest of the researchers in the past years. Many of the proposed methods and theoretical analyses rely on convex surrogates of the original nonconvex sparse PCA problem. In this work, we consider a rather direct approach. Under several assumptions on the population covariance matrix (e.g. the spiked covariance model and its generalizations) we derive the model selection properties of sparse PCA. We show that the stationary points of the nonconvex problem share many of the properties of the (local) minimum (e.g. the probability for a type I error). Simulations confirm that stationary points of the sparse PCA problem are indeed able to recover the support of the underlying truth.

A new restricted Liu-type estimator in linear regression model

Funda Erdugan¹ and Gulin Tabakan²
¹Kirikkale University (Turkey), ²Aksaray University (Turkey)

Wednesday 26
10:30-12:30
Lecture room 10
(Sali 10)

Kurnaz and Akay (2015) introduced a new general Liu-type estimator (NL) to deal with the well-known multicollinearity problem in linear regression model. In this paper, the restricted Liu-type estimator (RNL) based on the NL is presented. The quadratic bias and mean squared error of the proposed estimator is discussed and compared with the ordinary least squares (OLS), restricted least squares (RLS) and Liu-type (NL) estimators in terms of the mean square error (MSE) criterion. Furthermore, we obtain the estimator of the biasing parameter and give a numerical example to illustrate some of the theoretical results.

Tuesday 25
9:00-10:00
Great hall
(Juhlasali)

Modelling evolution in a spatial continuum

Alison Etheridge
University of Oxford (UK)

Since the pioneering work of Fisher, Haldane and Wright at the beginning of the 20th Century, mathematics has played a central role in theoretical population genetics. In turn, population genetics has provided the motivation both for important classes of probabilistic models, such as coalescent processes, and for deterministic models, such as the celebrated Fisher-KPP equation. Whereas coalescent models capture 'relatedness' between genes, the Fisher-KPP equation captures something of the interaction between natural selection and spatial structure. What has proved to be remarkably difficult is to combine the two, at least in the biologically relevant setting of a two-dimensional spatial continuum.

In this talk we describe some of the challenges of modelling evolution in a spatial continuum, present a model that addresses those challenges, and, as time permits, describe some applications.

Monday 24
14:00-16:00
Lecture room 1
(Sali 1)

Multi-armed bandit for species discovery: a Bayesian nonparametric approach

Stefano Favaro
University of Torino and Collegio Carlo Alberto (Italy)

This work proposes a new methodology for discovering new species, when observations are sampled from different populations. Using a metaphor, we imagine J populations of animals to be available and we can sequentially choose from which of these populations to collect further samples. Both labels and frequencies of these species are unknown a priori. At each time step, the proposed strategy suggests where to collect the next observation in order to maximize the number of total species observed. This strategy is based on a joint use of the hierarchical Pitman-Yor process, to estimate the unknown distributions of animals, and of Thompson Sampling for the sequential allocation problem. Performances of the algorithm are compared to those of other three strategies through simulations.

Monday 24
14:00-16:00
Lecture room 6
(Sali 6)

Distance correlation for spatial stochastic processes

Jochen Fiedler
Fraunhofer ITWM (Germany)

In spatial statistics, the Pearson correlation function is undoubtedly the most frequently used dependence measure. Despite its usefulness, it is not reasonable to use the Pearson correlation with processes admitting non-finite second moments or highly non-linear dependencies, like non-Gaussian stable or certain autoregressive processes, respectively. In such cases, it is possible that the empirical Pearson correlation function fails to detect existing dependencies or that it wrongly implies the existence of linear dependencies.

In this talk, the *distance correlation function* and *distance variogram* as dependence measures for spatial processes are introduced. Both dependence measures are based on the concept of distance correlation between random vectors. The distance correlation has the advantage that it is defined in cases where the Pearson correlation is not, and that it is able to measure non-linear dependencies. It is demonstrated that the distance correlation function is a suitable dependence measure for large classes of processes, in particular stable processes, and that the distance variogram for stable processes with finite first moments is a natural generalization of the usual variogram for Gaussian processes. Furthermore, the distance correlation function and the distance variogram can be consistently estimated. With the help of simulated examples, properties of the empirical versions of the Pearson correlation and the distance correlation function are compared.

The multivariate auto-distance covariance/correlation function

Konstantinos Fokianos
University of Cyprus (Cyprus)

Monday 24
14:00-16:00
Lecture room 5
(Sali 5)

We introduce the notions of multivariate auto-distance covariance and correlation functions for time series analysis. These concepts have been recently discussed in the context of independent and time series data but we extend them in a different direction by putting forward their matrix version. We discuss their interpretation and we give consistent estimators for practical implementation. Additionally, we develop a test for testing pairwise independence of multivariate time series data. The proposed test statistic performs better than the standard multivariate version of Ljung-Box test statistic. Several computational aspects are discussed and some data examples are provided for illustration of the methodology.

Partial exchangeability without stationarity

Sandra Fortini, Sonia Petrone and Polina Sporysheva
Bocconi University (Italy)

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

Exchangeability is a central notion in many areas of probability and related fields and has a fundamental role in Bayesian statistics. However, in many problems forms of competition, selection, and other sources of non stationarity can break exchangeability, although the phenomenon can converge asymptotically to an exchangeable steady state. Thus, weaker notions of stochastic dependence, that do not assume stationarity and yet preserve some main asymptotic properties of exchangeable processes are of theoretical and applied interest. Based on the results of Kallenberg (Ann. Probab., 1988), Berti, Pratelli and Rigo (Ann. Probab., 2004) introduced a notion of conditionally identically distributed (c.i.d.) sequences as a form of stochastic dependence weaker than exchangeability but equivalent to it in the presence of stationarity. A natural question is how the notion of c.i.d. sequences can be extended to a notion of partially c.i.d. processes, in this sense paralleling the extension from exchangeability to (de Finetti) partial exchangeability. We introduce a notion of partially c.i.d. families of sequences, that is shown to be equivalent to partial exchangeability under stationarity. Partially c.i.d. sequences preserve some main limit properties of partially exchangeable sequences. In particular, they are asymptotically partially exchangeable and the joint predictive and empirical distributions converge to the same random limit. Furthermore, we prove a law of large numbers and two central limit theorems for the scaled cumulative forecast errors and the scaled difference between empirical means and predictions, respectively. C.i.d. sequences can be employed to model interacting evolutionary phenomena that, while not being stationary, tend towards an equilibrium state of partial exchangeability. Areas of application include adaptive clinical trials, sequential design, two armed bandit problems and reinforcement learning, and Bayesian inference. The notion of partially c.i.d. sequences extends the spectrum of application to multiple experiments, for example clinical trials in multiple centers. A natural class of partially c.i.d. sequences are interacting randomly reinforced processes satisfying certain conditions on the random reinforcement. There is an increasing interest in interacting randomly reinforced processes in many fields and the notion of partially c.i.d. processes may provide a useful reference framework. Another area of application is Bayesian nonparametric inference, where partially c.i.d. constructions may characterize novel classes of prior distributions for dependent random measures, as the weak limits of the joint predictive distributions.

Keywords: Exchangeability. Reinforced processes. Spreadability. Limit theorems. Prediction. Bayesian nonparametrics. Interacting systems.

Reference: Fortini S., Petrone S., Sporysheva P., 2017. On a notion of partially conditionally identically distributed sequences. <https://arxiv.org/abs/1608.00471>

Estimation risk for the VaR of portfolios driven by semi-parametric multivariate models

Wednesday 26
10:30-12:30
Great hall
(Juhlasali)

Christian Francq and Jean-Michel Zakoian
CREST-ENSAE and University of Lille (France)

Joint estimation of market and estimation risks in portfolios is investigated, when the individual returns follow a semi-parametric multivariate dynamic model and the asset composition is time-varying. Under ellipticity of the conditional distribution, asymptotic theory for the estimation of the conditional Value-at-Risk (VaR) is developed. An alternative method - the Filtered Historical Simulation - which does not rely on ellipticity, is also studied. Asymptotic confidence intervals for the conditional VaR, which allow to simultaneously quantify the market and estimation risks, are derived. The particular case of minimum variance portfolios is analyzed in more detail. Potential usefulness, feasibility and drawbacks of the two approaches are illustrated via Monte-Carlo experiments and an empirical study based on stock returns.

Importance sampling type correction of Markov chain Monte Carlo and exact approximations

Poster session

Jordan Franks, Matti Vihola and Jouni Helske
University of Jyväskylä (Finland)

We use an importance sampling (IS) type correction of approximate Markov chain Monte Carlo (MCMC) output to provide consistent estimators. The IS approach, based on unbiased estimators, provides a natural alternative to delayed acceptance (DA) MCMC, but contrary to DA-MCMC, benefits from straightforward parallelisation. We focus on the case where the MCMC simulation is conducted on an approximate marginal distribution, and the IS type correction leads to consistent inference on the corresponding joint distribution. Such a case allows for importance sampling analogues of pseudo-marginal type MCMC methods, such as the grouped independence Metropolis-Hastings and the particle marginal Metropolis-Hastings. We detail strong consistency of the suggested estimators under mild assumptions, and provide central limit theorems with expressions for asymptotic variances. The methods are illustrated in practice with state space models, where our experimental results are promising and show that IS type approach can be favourable against a computationally comparable DA scheme.

MATS: Inference for potentially singular and general heteroscedastic MANOVA

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

Sarah Friedrich
Ulm University (Germany)

In many experiments in the life sciences several endpoints, potentially measured on different scales, are recorded per subject. Classical MANOVA models assume normally distributed errors and homogeneity of the covariance matrices, two assumptions that are often not met in practice. In the context of high-dimensional one-way layouts, Srivastava and Kubokawa (2013) proposed a test statistic which is invariant under the change of units of measurements. However, it relies on the assumption of a common non-singular covariance matrix. We extend this test statistic to factorial MANOVA designs, incorporating general heteroscedastic models. In particular, we only postulate existence of the group-wise covariance matrices, which may even be singular. Similar to Konietzschke et al. (2015), who resampled the Wald-type statistic, we use bootstrap techniques in order to improve the small sample behaviour of our test statistic. In a large simulation study, we analyse both a parametric and a wild bootstrap approach. The parametric bootstrap yields the best results in various simulation settings. In addition, we derive confidence regions and simultaneous confidence intervals for contrasts of the mean vector based on the bootstrap quantiles. The results are applied to a real data example from cardiology.

References: [1] Srivastava, M. S., and Kubokawa, T., 2013. Tests for multivariate analysis of variance in high dimension under non-normality. *Journal of Multivariate Analysis*, 115, 204–216. [2] Konietzke, F., Bathke, A. C., Harrar, S. W., and Pauly, M., 2015. Parametric and nonparametric bootstrap methods for general MANOVA. *Journal of Multivariate Analysis*, 40, 291–301.

The Dantzig selector for some statistical models of stochastic processes

Kou Fujimori and Yoichi Nishiyama
Waseda University (Japan)

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

The Dantzig selector, which was proposed by Candés and Tao in 2007, is an estimation procedure for regression models in a high-dimensional and sparse setting. The Dantzig selectors for some statistical models of stochastic processes are studied in this paper. We apply this procedure for Cox’s proportional hazards model and some specific models of diffusion processes and prove the l_q consistency of the estimators for every $q \in [1, \infty]$. Based on partial likelihood and quasi-likelihood methods which were studied intensively in low-dimensional settings, we study these statistical models of stochastic processes in high-dimensional and sparse settings, which require some mathematically challenging tasks to prove the consistency of the estimators. The consistency of the estimators depend on some matrix factors and conditions on Hessian matrices of likelihood functions. We use factors called restricted eigenvalue, compatibility factor and weak cone invertibility factor since it is known that the conditions for these factors are weaker than other conditions for statistical models in high-dimensional and sparse settings.

Estimating a Poisson parameter based on incomplete information: Precision vs. simplicity

Tamar Gadrich
ORT Braude College (Israel)

Monday 24
14:00-16:00
Auditorium IV

Consider a sequence of events occurring according to the Poisson process, with an unknown incidence rate λ . Two independent imperfect observers record some of these events, during T time intervals of fixed length (e.g., one day), where the probability that observer i ($i = 1, 2$) records an event is an unknown parameter p_i . Each of the two observers reports the number of observations it made during each time interval. Our estimation problem is to estimate the incidence rate λ together with the observation rates p_1, p_2 . The necessity of more than one imperfect observer for identifiability is clear. In contrast with the well-known problem of Multiple Systems Estimation (also known as Capture-Recapture), here the number of events which are reported by both observers is unknown. We show that the estimation problem under consideration can be cast into a problem of estimating parameters for a Bivariate Poisson distribution, though with a different parameterization than the standard one. We consider two kinds of estimators for the three unknown parameters: moment-type and maximum-likelihood. Both estimators are consistent, that is, converge in probability to the true values of the parameters as the number of observation periods T increases. The performance of these estimators is studied analytically and by means of simulations. The advantage of the moment estimator is its simplicity, while the maximum-likelihood estimator can only be found numerically, as a solution to a maximization problem which cannot be solved for in closed form. We compare the two methods by numerical simulation, and show that the maximum likelihood estimator is (on the average) more precise than the moment estimator. Thus the higher complexity of estimation using the maximum likelihood method is re-paid by better estimates.

On discrete-time birth-death circuit chains in random ergodic environments

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

Chrysoula Ganatsiou

University of Thessaly (Greece)

It is known that in order to investigate representations of Markov processes (with discrete or continuous parameter) having an invariant measure, as decompositions in terms of the well-known circuit passage functions, a systematic work has been developed (Kalpazidou [1], MacQueen [5], Qian Minping and Qian Min [6] and others). The representations are called circuit representations while the corresponding discrete parameter Markov processes generated by directed circuits are called circuit chains. It is also known that the classical birth-death chain is a special case of homogeneous, aperiodic, irreducible Markov chain (discrete-time or continuous-time), where state changes can only happen between neighboring states. This means that the state transitions are of only two types: births which increase the state variable by one and deaths which decrease the state variable by one. The models name comes from a common application, the use of such models to represent the current size of a population where the transitions are literal births and deaths. In general the birth-death processes have many applications in demography, epidemiology or in biology since they may be used to study the evolution of bacteria or the number of people with a disease within a population. By using the circuit representation theory of Markov processes, the present work arises as an attempt to investigate suitable criteria regarding positive/null recurrence and transience for the corresponding adjoint Markov chains describing uniquely the discrete -time birth-death chains by directed circuits and weights in random ergodic environments. (For the study of transience and recurrence for discrete-time birth-death circuit chains in fixed ergodic environments as special cases of random walks we refer the reader to Ch. Ganatsiou [2-4]). In order to investigate suitable criteria for the transience and recurrence of the corresponding Markov chains representing uniquely by directed circuits and weights a discrete-time birth-death chain in a random ergodic environment, we may use the criteria given in the study for fixed ergodic environments restricted to the half-lines $[j, +\infty)$ with reflection in j . Therefore, we have the following: Proposition. The discrete-time birth-death chain in random ergodic environments, is transient, for μ -a.e. environment ω , if and only if its adjoint Markov chain is positive recurrent and reciprocally. Moreover, the adjoint Markov chains are null recurrent simultaneously.

References: [1] Kalpazidou, S., 1995. Cycle representations of Markov Processes, Springer, New York, NY, USA. [2] Ganatsiou, Ch., 2011. On cycle representations of random walks in fixed, random environments, *Proceedings of the 58th World Congress of the International Statistical Institute, Dublin, 21th-26th August 2011*. [3] Ganatsiou, Ch., 2013. On the study of transience and recurrence of the Markov chain defined by directed weighted circuits associated with a random walk in fixed environment, *Journal of Probability and Statistics*. [4] Ganatsiou, Ch., 2013. On cycle representations of discrete-time birth-death processes, *29th European Meeting of Statisticians, Budapest, Hungary, 20-25/7/2013*. [5] MacQueen, J., 1981. Circuit processes. *The Annals of Probability*, 9(4), 604–610. [6] Minping, Q., Min, Q., 1982. Circulation for recurrent Markov chains. *Probability Theory and Related Fields*, 59(2), 203–210.

Data-driven optimization of marked point processes with applications to object detection in 2D image

Thursday 27
14:00-16:00
Lecture room 6
(Sali 6)

Ségolen Geffray

Université de Strasbourg (France)

We focus on the problem of locating, counting and measuring cells in 2D images in an accurate, reliable and robust manner. We adopt the stochastic viewpoint for its ability to give rise to a detection method insensitive both to image local defects such as noise, blur, and to the fact that grey level intensities may vary from an object to another similar object. An image is then considered as an imperfect (noisy, blurred, non-uniformly illuminated,...) observation of a random process in an observation window. When the objects of interest can be modelled as simple parametric shapes such as disks, a suitable class of random processes for this modelling purpose is the class of marked

spatial point processes. Object detection, counting and measuring can then be made by fitting a suitable marked spatial point process to the image. The object detection consists of finding the optimal object configuration which is done by maximizing an objective function $h(\cdot)$. The objective function consists of a penalized version of the quasi-likelihood associated to the image. The quasi-likelihood reflects the suitability of a given point configuration with respect to the image at hand. The penalty term prevents excessive object overlap. The task of locating a global mode of $h(\cdot)$ is a non-convex optimization problem and our objective function is intractable. To circumvent this difficulty, we use the simulated annealing algorithm. The simulated annealing algorithm repeatedly produces a random sample from $h(\cdot)$ for which there is an increasing probability of producing the most probable point configuration. The sampling step of the simulated annealing algorithm motivates the use of a reversible jump Markov chain Monte-Carlo sampler also termed as Metropolis-Hastings-Green sampler. We present a novel data-driven version of both the objective function $h(\cdot)$ and of the optimization algorithm. The algorithm is theoretically grounded by the demonstration of the convergence of the Markov chain. We apply our method on both real and simulated 2D images and illustrates the reliability and the efficiency of our data-driven technique to guide the Markov chain search and get the required location, counting and morphology characterisation.

Aggregated methods for covariates selection in high-dimensional data under dependence

Anne Gégout-Petit^{1,3}, Bérangère Bastien², Aurélie Gueudin^{1,3} and Yaojie Shi¹

¹Inria Nancy Grand Est (France), ²Transgene Strasbourg (France), ³Université de Lorraine (France),

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

We propose a new methodology to select and rank covariates associated to a variable of interest in a context of high-dimensional data under dependence but few observations. The methodology imbricates successively rough selection, clustering of variables, decorrelation of variables using Factor Latent Analysis, selection using aggregation of adapted methods and finally ranking through bootstrap replications. Simulations study shows the interest of the decorrelation inside the different clusters of covariates. The methodology is applied to real data with aim of selecting factors linked to the efficiency of a treatment among more than 50.000 covariates but only two groups of less than 30 subjects.

Diffusions and dynamics on statistical manifolds for statistical inference

Mark Girolami

Imperial College London (UK)

Friday 28
9:00-10:00
Great hall
(Juhlasali)

The use of Differential Geometry in Statistical Science dates back to the early work of C.R.Rao in the 1940s when he sought to assess the natural distance between population distributions. The Fisher-Rao metric tensor defined the Riemannian manifold structure of probability measures and from this local manifold geodesic distances between probability measures could be properly defined. This early work was then taken up by many authors within the statistical sciences with an emphasis on the study of the efficiency of statistical estimators. The area of Information Geometry has developed substantially and has had major impact in areas of applied statistics such as Machine Learning and Statistical Signal Processing.

A different perspective on the Riemannian structure of statistical manifolds can be taken to make breakthroughs in the contemporary statistical modelling problems. Langevin diffusions and Hamiltonian dynamics on the manifold of probability measures are defined to obtain Markov transition kernels for Monte Carlo based inference. This work was motivated by the many challenges presented by contemporary problems of statistical inference, such as for example inference over partial differential equations describing complex physical engineering systems. These lectures aim to provide an accessible introduction to the Langevin and Hamiltonian Monte Carlo methods.

Evaluating the complexity of functional data

Aldo Goia¹, Enea Giuseppe Bongiorno and Philippe Vieu

¹Università del Piemonte Orientale "A. Avogadro" (Italy)

We propose a two step approach in order to study the "complexity" of functional data. In the first step we introduce the log-Volugram, a new graphical tool for assessing to which family of processes (finite-dimensional, fractal, gaussian, ...) the data belong. It is based on a nonparametric $k - NN$ estimation of Small Ball Probability. Once the family is specified, the second step consists in estimating some specific indexes related to the assigned family. The developed methodology is fully free from assumptions on model, distribution as well as dominating measure. Practical performances over finite samples are analyzed by means of numerical studies. An application to financial time series is illustrated.

Poster session **On application of hierarchical clustering methods for planar shape data**

Mousa Golalizadeh

Tarbiat Modares University (Iran)

Determination of the class (group) that each individual from a random sample belong to, is of interest for many disciplines. The job is relatively simple if the data on hand describing abstract features of the individuals, such as weight, length and height. However, one should pay particular attention whenever the features for each sample, simultaneously, representing the geometrical configuration. For instances, consider the case in which the Cartesian coordinates of some landmarks set on the boundary of each object are our sample. Then, one may be interested on clustering the objects in which the full geometry of those objects are the main input of the problem. In other words, one should not be concerned about the changes on the location, scale and rotation of the objects since they need to be filtered out prior to initiate the cluster analysis. Following Kendall (1984), the objects went through such filtering procedure are called the shapes. Mathematically, those shapes lie in non-Euclidean space and so the usual procedures to classify them need to be properly adopted. Obviously, one should utilize proper Riemannian distances defined in the shape analysis contexts where the Procrustes distances play the key role. In this paper, we compare performance of the various versions of Procrustes distances as well as shape distributions to cluster the shape data. Also, simulation studies are employed to evaluate clustering procedures. Moreover, the cross-validation criterion as well as other clustering validations are invoked in order to select a preferred procedure for shape data. We confined ourselves to planar shape data in this paper. As expected, shape clustering is highly related to the way in which one is registering the objects. Unlike the landmark-based procedure followed in our view in this paper, considering the boundary of the objects as the continuous closed curves and proposing general procedure to cluster and classify shapes has been proposed in Srivas et. al. (2005). Taking the length of rays crossed the objects (shape of tumors) and then classifying the children based upon their renal tumors has been studied in Giebel et. al. (2013). Rare studies follow the clustering the objects based upon the landmark-based view to the statistical shape analysis. We can mention the activity reported in Amaral et. al. (2010) and Huang et. al. (2015). While former used the K-means clustering along with adapting the Procrustes distances, latter invoked latent variable models linked to some covariates to improve the clustering the planar shape data. Similar to the former, but confining to three-dimensional objects, was reported in Viu et. al. (2016). Following Dryden & Mardia (1998), we consider a general linear models, adopted for the shape analysis, as

$$X_i = \beta_i(\mu + E_i)\Gamma_i + 1_k\gamma_i^T, \quad \text{vec}(E_i) \sim N(0, \sigma^2 I_{km}), \quad i = 1, \dots, n,$$

where $\beta_i > 0$ is a scale parameter, μ , the mean shape, E_i , complex random error term, Γ_i , a $m \times m$ rotation matrix, 1_q a q -dimensional vector of ones, γ_i a $m \times 1$ location vector and I_q is a $q \times q$ identity matrix. Here, X_i 's, the $k \times 1$, dimension vectors, are describing the configurations of the objects and are our main inputs to investigate the clustering the samples. To derive the similarity matrix, it is common to define the product matrix $S = \sum_{i=1}^n Z_i Z_i^*$ where $Z_i = HX_i / \|HX_i\|, i = 1, \dots, n$

are known as pre-shapes, Z_i^* is the complex conjugate of the transpose of Z and H is $(k1) \times k$ Helmert sub-matrix. After computing the mean shape, proximity matrix is constructed based upon Z_i 's. This is the non-parametric view to cluster the shape data. Full details of our approach have already been reported in Nabil & Golali.

GPyOpt: Practical Bayesian optimization for model configuration and experimental design

Javier González Hernandez
Amazon.com

Tuesday 25
14:00-16:00
Great hall
(Juhlasali)

Bayesian optimization has emerged during the last few years as an effective approach for optimizing black-box functions where direct queries of the objective are expensive. Using a Gaussian process as a surrogate of the objective function, Bayesian optimization has been applied to a wide variety of applications ranging from the design of wet-lab experiments to the automatic tuning of complex models and algorithms. One of the aspects that has triggered this success is the availability of software that has automated the use of these techniques in practice. In this talk, we will review the most elementary aspects of Bayesian optimization and well show the main features of GPyOpt, a general and modular framework for Bayesian optimization that has captured the attention of a wide scientific community. We'll discuss a variety of real examples and show the basic functionalities of the library.

Non-parametric Bayesian estimation of a diffusion coefficient

Shota Gugushvili¹, Frank van der Meulen², Moritz Schauer¹ and Peter Spreij³

¹Leiden University (Netherlands), ²Delft University of Technology (Netherlands), ³University of Amsterdam (Netherlands)

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

Based on discrete time observations on the solution of a stochastic differential equation over the time interval $[0, T]$, we consider a non-parametric Bayesian approach to estimation of its (deterministic) diffusion coefficient. As a prior on the diffusion coefficient, we employ a histogram-type prior with piecewise constant realisations on bins forming a partition of $[0, T]$. On the theoretical side, we justify our approach by demonstrating that the corresponding posterior distribution asymptotically, as the number of observations in the interval $[0, T]$ grows to infinity, concentrates around the true diffusion coefficient and derive the corresponding posterior contraction rate. For a specific choice of the prior based on the use of the inverse gamma distribution, this posterior contraction rate turns out to be optimal for estimation of a Hölder-continuous diffusion coefficient of smoothness $0 < \lambda \leq 1$. On the implementational side, we show that our approach is straightforward to implement and leads to good practical results in a wide range of simulation examples. Finally, we apply our method on the exchange rate data set.

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

Analysis of nonstationary modulated time series with applications to oceanographic flow measurements

Arthur Guillaumin¹, Adam M. Sykulski², Sofia C. Olhede¹, Jeffrey J. Early³ and Jonathan M. Lilly³

¹University College London (UK), ²Lancaster University (UK), ³NorthWest Research Associates (USA)

We propose a new class of univariate nonstationary time series models, using the framework of modulated time series, which is appropriate for the analysis of rapidly-evolving time series as well as time series observations with missing data. We extend our techniques to a class of bivariate time series that are isotropic. Exact inference is often not computationally viable for time series analysis, and so we propose an estimation method based on the Whittle-likelihood, a commonly adopted pseudo-likelihood. Our inference procedure is shown to be consistent under standard assumptions, as well as having considerably lower computational cost than exact likelihood in general. We show the utility of this framework for the analysis of drifting instruments, an analysis that is key to characterising global ocean circulation and therefore also for decadal to century-scale climate understanding.

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

Maximum Lq-Likelihood estimation for autoregressive error term regression model

Yeşim Güney¹, Yetkin Tuğaç¹, Şenay Özdemir² and Olcay Arslan¹

¹Ankara University (Turkey), ²Afyon Kocatepe University (Turkey)

Maximum likelihood estimation method is among the most commonly used estimation procedures for the regression models with autoregressive errors under the normality assumption (Alpuim and El-Shaarawi, 2008). In this study, we consider an alternative estimation method, which is called the maximum Lq-likelihood estimation (MLqE) method proposed by Ferrari & Yang (2010), to estimate the parameters of the regression models with autoregressive errors under the normality assumption. We provide a simulation study and real data examples to illustrate the performance of the MLqE and MLE methods for regression models with AR(p) error terms. We observe from the results of the simulation study and real data examples that the MLq estimators are superior to the ML estimators when the data sets have outliers. Keywords: autoregressive stationary process; maximum lq-likelihood estimation; linear regression.

References: [1] Alpuim, T., El-Shaarawi, A., 2008. On the efficiency of regression analysis with AR(p) errors. *Journal of Applied Statistics*, 35:7, 717–737. [2] Ferrari, D., Yang, Y., 2010. The Maximum Lq-Likelihood Estimation. *The Annals of Statistics* 38, 753–783.

Tuesday 25
10:30-12:30
Lecture room 7
(Sali 7)

Limiting genotype frequencies of Y-linked genes with a mutant allele in a two-sex population

Cristina Gutiérrez Pérez¹, Miguel González¹, Rodrigo Martínez¹ and Gerold Alsmeyer²

¹University of Extremadura (Spain), ²University of Münster (Germany)

In González et al. (2012) a two-type two-sex branching process was introduced to model the evolution of the number of carriers of an allele and its mutations of a Y-linked gene. In such paper, conditions for the extinction and survival of such alleles was studied. The purpose of the present work is to study the growth rates as well as the limiting genotype frequencies for the previous model. Based on two-sex branching processes results, the growth rate of the mutant allele on its set of survival and the extinction of the non-mutant allele has been obtained. Moreover, using the results studied in González et al. (2012) the growth rate of the non-mutant allele on the set of coexistence of both alleles has been found. Also on that set of coexistence, it has been studied the rate of growth of the mutant-allele. In this case, the study turns out to be more difficult due to the dependency of the survival of the mutant-allele on the behavior of the non-mutant allele.

For that reason, to carry out this study, it is necessary to distinguish different cases depending on the probability that an offspring to be female is equal to, less or greater than 0.5 and also on the relation between the probability of mutation and the mean number of offsprings generated by the couples formed by males with the different genotypes (with mutant or not mutant alleles) in the population. Finally, the main results have been illustrated with a simulated study.

References: González M., Gutiérrez C., Martínez R., 2012. Extinction conditions for Y-linked mutant-alleles through two-sex branching processes with blind-mating structure. *Journal of Theoretical Biology*, 307, 104–116.

Nonparametric Bayesian Intensity Model: Exploring Time-to-Event Data on Two Time Scales

Tommi Härkönen¹, Anna But² and Jari Haukka²

¹National Institute for Health and Welfare (Finland), ²University of Helsinki (Finland)

Time-to-event data have been extensively studied in many areas. Although, multiple time scales are often observed, commonly used methods are based on a single time scale. Analyzing time-to-event data on two time scales can offer a more extensive insight into the phenomenon. We introduce a nonparametric Bayesian intensity model to analyze two-dimensional point process on Lexis diagrams. After a simple discretization of the two-dimensional process, we model the intensity by a one-dimensional piecewise constant hazard functions parametrized by the change points and corresponding hazard levels. Our prior distribution incorporates a built-in smoothing feature in two dimensions. We implement posterior simulation using the reversible jump Metropolis-Hastings algorithm. We demonstrate the applicability of the method using both simulated and empirical survival data. Our approach outperforms commonly applied models by borrowing strength in two dimensions.

Monday 24
10:30-12:30
Lecture room 5
(Sali 5)

A Markov-switching generalized additive model for compound Poisson processes, with applications to operational losses models

Julien Hambuckers¹, Thomas Kneib¹, R. Langrock² and A. Sohn¹

¹Georg-August-Universität Göttingen (Germany), ²Bielefeld University (Germany)

This paper is concerned with modeling the behavior of random sums over time. Such models are particularly useful to describe the dynamics of operational losses, and to correctly estimate tail-related risk indicators. However, time-varying dependence structures make it a difficult task. To tackle these issues, we formulate a new Markov-switching generalized additive compound process combining Poisson and Generalized Pareto distributions. This flexible model takes into account two important features: on the one hand, we allow all parameters of the compound loss distribution to depend on economic covariates in a flexible way. On the other hand, we allow this dependence to vary over time, via a hidden state process. A simulation study indicates that, even in the case of a short time series, this model is easily and well estimated with a standard maximum likelihood procedure. Relying on this approach, we analyze a novel dataset of 817 losses resulting from frauds at the Italian bank UniCredit. We show that our model improves the estimation of the total loss distribution over time, compared to standard alternatives. In particular, this model provides estimations of the 99.9% quantile that are never exceeded by the historical total losses, a feature particularly desirable for banking regulators.

Keywords: GAMLSS, distributional regression, generalized Pareto distribution, hidden Markov model, operational losses, compound Poisson process.

Thursday 27
10:30-12:30
Lecture room 13
(Sali 13)

Multivariate extremes based on a notion of radius

Matias Heikkilä¹, Yves Dominicy² and Pauliina Ilmonen¹

¹Aalto University School of Science (Finland), ²Université libre de Bruxelles (Belgium)

We propose an approach to multivariate extreme value theory that is based on the notion of radius: A reduction of a multivariate random variable to one dimension in an affine invariant way that attempts to preserve many features of the tail behavior. The radius depends on population parameters and cannot be observed directly. We derive a convergence result for its empirical counterpart and, as an application, show that the extreme value index of the radius can be estimated in a consistent way. Limiting normality for an estimator is as well obtained in the case of a radius with positive extreme value index. In particular, convergence and limiting normality of the separating Hill estimator is obtained under estimated location and scatter. Asymptotic behavior of the separating Hill estimator was previously known only under known location and scatter.

Thursday 27

10:30-12:30

Lecture room 12

(Sali 12)

Differentially private Bayesian learning on distributed data

Mikko Heikkilä¹, Yusuke Okimoto², Samuel Kaski³, Kana Shimizu² and Antti Honkela¹

¹University of Helsinki (Finland), ²Waseda University (Japan), ³Aalto University (Finland)

Many applications of machine learning, for example in health care, would benefit from methods that can guarantee privacy of data subjects. Differential privacy (DP) has become established as a standard for protecting learning results, but the proposed algorithms require a single trusted party to have access to the entire data, which is a clear weakness. We consider DP Bayesian learning in a distributed setting, where each party only holds a single sample or a few samples of the data. We propose a novel method for DP learning in this distributed setting, based on a secure multi-party sum function for aggregating summaries from the data holders. Each data holder adds their share of Gaussian noise to make the total computation differentially private using the Gaussian mechanism. We prove that the system can be made secure against a desired number of colluding data owners and robust against faulting data owners. The method builds on an asymptotically optimal and practically efficient DP Bayesian inference with rapidly diminishing extra cost.

Thursday 27

14:00-16:00

Lecture room 6

(Sali 6)

Rigid motion estimation in sparse sequential dynamic imaging

Anne Hein

University of Göttingen (Germany)

This talk addresses an issue arising in many modern superresolution fluorescence microscopy techniques, one of them being single marker switching (SMS) microscopy. The challenge is to correctly align long sequences of sparse but spatially and temporally highly resolved images. This is necessary due to rigid motion of the displayed object of interest over time. A purely statistical reconstruction method based on parametric models for the motion functions describing drift, rotation and scaling of the imaged specimen is presented. M-estimators for the parameters are derived and consistency as well as asymptotic normality of the estimators are proven. Furthermore, the results of a simulation study and an application to real SMS microscopy data are presented. The method is shown to be competitive with state of the art calibration techniques which require to incorporate fiducial markers. This technique is also applicable in many other scenarios, where an aggregation of a collection of sparse images is employed to obtain a good reconstruction of the underlying structure, like, for example, in real time magnetic resonance imaging (MRI).

Asymptotic normality for a class of empirical functionals of strongly mixing marked point processes

Lothar Heinrich

University of Augsburg (Germany)

Monday 24
14:00-16:00
Auditorium IV

We study the asymptotic behaviour of empirical functionals of stationary marked point process (short: MPP) observed (location of points and their marks) within an unboundedly increasing sequence of convex observation windows in some Euclidean space. The marks may belong to an arbitrary Polish space. We assume weak dependencies between parts of the MPP defined over suitable classes of two disjoint sets formulated in terms of the strong-(or alpha-)mixing coefficient. Our main result is a CLT for a rather general functional after suitable centering and scaling. The proof relies on a new variant of Bernstein's blocking method adapted to an increasing space-filling sequence of convex sets. Our CLT contains a number of special cases: (i) Different types of marking: unmarked PP, independently MPP, geostatistically MPP, (ii) Different functionals: empirical mark distribution and mark moments, empirical (marked) K-function, empirical (marked) nearest-neighbour distance function, kernel product density estimators, etc. As a by-product we found the apparently best bound of the covariance $\text{cov}(X, Y)$ in terms of the strong-mixing coefficient and certain absolute moments of X and Y .

Large sample covariance matrices with heavy tails

Johannes Heiny¹, Richard Davis² and Thomas Mikosch³

¹Aarhus University (Denmark), ²Columbia University New York (USA), ³University of Copenhagen (Denmark)

Tuesday 25
10:30-12:30
Great hall
(Juhlasali)

We study the joint distributional convergence of the largest eigenvalues of the sample covariance matrix of a p -dimensional heavy-tailed time series when p converges to infinity together with the sample size n . We generalize the growth rates of p existing in the literature. Assuming a regular variation condition with tail index $\alpha < 4$, we employ a large deviations approach to show that the extreme eigenvalues are essentially determined by the extreme order statistics from an array of i.i.d. random variables. The asymptotic behavior of the extreme eigenvalues is then derived routinely from classical extreme value theory. The resulting approximations are strikingly simple considering the high dimension of the problem at hand. We develop a theory for the point process of the normalized eigenvalues of the sample covariance matrix in the case where rows and columns of the data are linearly dependent. Based on the weak convergence of this point process we derive the limit laws of various functionals of the eigenvalues.

Pareto depth for functional data

Sami Helander¹, Germain Van Bever², Sakke Rantala³ and Pauliina Ilmonen¹

¹Aalto University School of Science (Finland), ²Université libre de Bruxelles (Belgium), ³Kemijoki OY (Finland)

Poster session

We introduce a new concept of depth for functional data. It is based on a new multivariate Pareto depth applied after mapping the functional observations to a vector of statistics of interest. These quantities allow to incorporate the inherent features of the distribution, such as shape or roughness. In particular, in contrast to most existing functional depths, the method is not limited to location only. Properties of the depths are explored and the benefits of a flexible choice of features are illustrated on several examples. In particular, its excellent classification capacity is demonstrated on real data example.

Using empirical modeling and extreme value theory in estimating border conditions in quality control

Raúl Hernández-Molinar, Dora Erika Espericueta and Vicente Hernández-García
Universidad Autónoma de San Luis Potosí (México)

The work focuses on the employment of empirical modeling to analyze domains of attraction of critical values that can be used as boundary conditions in control charts. Some results are based on the simulation of order statistics that belong to the tails of a probability distribution function. The data sets were constructed using a random data generation system. The results were validated. The goodness of fit tests were performed, data were validated considering criteria based on the bias and the mean square of the error. Some results show evidence that in the case of data of low voltage electrical current, it is possible to use Normal, Lognormal, Weibull distributions or the distribution of Extreme Value. Data sets are studied considering the 90th, 95th, and 99th percentiles as upper limits. Statistical analysis is performed using concepts related to extreme value theory. Results associated with the 99th percentile are shown. Based on the generated scenarios, it is possible to propose limits that can generate acceptance criterias in contexts with where no previous information exists.

Tuesday 25
14:00-16:00
Lecture room 13
(Sali 13)

Deriving animal breeding benefit from millions of animals with sequence

John Hickey¹, Mathew Cleveland, William Herring, Alan Mileham, Andreas Kranis, Janez Jenko, Serap Gonen, Mara Battagin, Roger Ros-Freixedes, Adriana Somavilla, Roberto Antolin, John Woolliams and Gregor Gorjanc

¹The Roslin Institute (UK)

Genomic selection has high economic value in breeding programs and this value will result in large data sets of genotyped and phenotyped individuals being generated. With appropriate developments in sequencing methods and strategies, and bioinformatics these large data sets could be sequenced. Such data will help to unravel the genetic and broader biological basis of traits in livestock, including traits associated with reproduction. Unravelling the genetics and biology of traits will lead to more accurate selective breeding which will lead to greater genetic gains. In the present paper we will outline how we are generating this type of data in large pig and chicken populations and we will report on some initial analysis that we are performing.

Monday 24
10:30-12:30
Lecture room 10
(Sali 10)

Profile likelihood approach to a large sample distribution of estimators in joint mixture mode

Yuichi Hirose

Victoria University of Wellington (New Zealand)

We consider a semiparametric joint model that consists of item response and survival components, where these two components are linked through latent variables. We estimate the model parameters through a profile likelihood and the EM algorithm. We propose a method to derive an asymptotic variance of the estimators in this model.

Quantum Shannon theory

Alexander S. Holevo

Steklov Mathematical Institute, Moscow (Russia)

Tuesday 25
17:00-18:00
Great hall
(Juhlasali)

The notions of channel and information capacity are central to the classical Shannon theory. Quantum Shannon theory is a mathematical discipline which uses operator and matrix analysis and various asymptotic techniques to study the laws for information processing in the systems obeying rules of quantum physics. From the mathematical point of view quantum channels are normalized completely positive maps of operator algebras, the analog of Markov maps in the noncommutative probability theory, playing a role of morphisms in the category of quantum systems.

This talk presents basic coding theorems providing analytical expressions for the capacities of quantum channels in terms of various entropic quantities. The remarkable role of specific quantum correlations entanglement as a novel communication resource, is stressed. We report on solution of exciting mathematical problems, such as " <Gaussian optimizers" >, concerning computation of the entropic quantities for both theoretically and practically important class of Bosonic Gaussian channels.

Adaptively weighted group Lasso for semiparametric quantile regression models

Toshio Honda

Hitotsubashi University (Japan)

Monday 24
14:00-16:00
Lecture room 10
(Sali 10)

We propose adaptively weighted group Lasso procedures for simultaneous variable selection and structure identification for varying coefficient quantile regression models and additive quantile regression models. We consider the situation where some initial estimates for calculating weights for weighted Lasso procedures are available. We allow ultra-high dimensional covariates in this paper and we have established the selection consistency of the proposed procedures. This means we can identify partially linear varying coefficient and additive models from larger high-dimensional varying coefficient and additive models, respectively, with probability tending to 1 if we suitably choose the tuning parameter. These semiparametric models are among the most popular semiparametric models. As for tuning parameter selection of the procedures, we have established the consistency of BIC type criteria for selecting partially linear varying coefficient and additive models from larger high-dimensional varying coefficient and additive models, respectively. We have also carried out numerical studies to examine the finite sample properties of the proposed procedures and report the results in this talk.

Model-robustly D- and A-optimal designs for log-contrast models in mixture experiments

Hsiang-Ling Hsu

National University of Kaohsiung (Taiwan)

Poster session

We investigate the issue of the model robust optimal designs in a mixture experiment. A mixture experiment in $(q-1)$ -dimensional probability simplex $S^{(q-1)}$ is an experiment in which the q factors are non-negative and subject to the simplex restriction, i.e., the factors represent relative proportions of q ingredients in the experiment. In this work, we develop the model-robustly D- and A-optimal designs for mixture experiments with consideration of uncertainties between the first-order and second-order models which proposed by Aitchison and Bacon-Shone (1984). The corresponding optimal designs and complete classes for each of the log contrast models are adopted to carry out the structure and the algorithms for obtaining the model-robustly D- and A-optimal designs, verified through the well-known Kiefer-Wolfowitz equivalence theorem. Numerical model-robustly optimal designs are illustrated in our study and we also compare the behaviours of the efficiencies for the linear combination of the D- and A-optimal designs from the individual models.

Monday 24
10:30-12:30
Lecture room 6
(Sali 6)

Fluctuation reduction of value-at-risk estimation and its applications

Shih-Feng Huang

National University of Kaohsiung (Taiwan)

Value-at-Risk (VaR) is a fundamental tool for risk management and is also associated with the capital requirements of banks. Banks need to adjust their capital levels for satisfying the Basel Capital Accord. On the other hand, managements do not like to change the capital levels too often. To achieve a balance, this study proposes an approach to reduce the fluctuation of VaR estimation. The first step is to fit a time series model to the underlying asset returns and obtain the convenient VaR process. A new VaR (NVaR) estimation of the convenient VaR process is then determined by applying change-point detection algorithms and a proposed combination scheme. The capital levels computed from the NVaR process are capable of satisfying the Basel Accord and reducing the fluctuation of capital levels simultaneously. On the basis of the NVaR process, a novel one-step ahead VaR estimation is proposed by using the technique of statistical control charts. The return processes of 30 companies on the list of S&P 500 from 2005 to 2014 are employed for our empirical investigation. Numerical results indicate that the proposed method has a satisfactory performance.

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

An AD approach for constructing separable minorizing functions in a class of MM algorithms

Xifen Huang

The University of Hong Kong (China)

The minorization–maximization (MM) principle provides an important and useful tool for optimization problems and has a broad range of applications in statistics because of its conceptual simplicity, ease of implementation and numerical stability. A key step in developing an MM algorithm is to construct an appropriate minorizing function. This is quite challenging to many practitioners as it has to be done case by case and its success often involves and heavily depends on a clever and specific use of Jensen’s inequality or a similar kind. To address this problem, in this paper, we propose a new assembly and decomposition (AD) approach which successfully constructs separable minorizing functions in a general class of MM algorithms. The AD approach constructs a minorizing function by employing two novel techniques which we refer to as the assembly technique (or A-technique) and the decomposition technique (or D-technique), respectively. The A-technique first introduces the notions of assemblies and complementary assembly, consisting of several families of concave functions that have arisen in numerous applications. The D-technique then cleverly decomposes the high-dimensional objective function into a sum of one-dimensional functions to construct minorizing functions as guided and facilitated by the A-technique. We demonstrate the utility of the proposed approach in diverse applications which result in novel algorithms with theoretical and numerical advantages. Extensive numerical studies are provided to assess its finite-sample performance. Further extensions of the AD techniques are also discussed.

Backward nested subspace inference and applications

Stephan Huckemann
University of Göttingen (Germany)

Wednesday 26
10:30-12:30
Lecture room 5
(Sali 5)

For sequences of random backward nested subspaces as occur, say, in dimension reduction for manifold or stratified space valued data, asymptotic results are derived. In fact, we formulate our results more generally for backward nested families of descriptors (BNFD). Under rather general conditions, asymptotic strong consistency holds. Under additional, still rather general hypotheses, among them existence of a.s. local twice differentiable charts, asymptotic joint normality of a BNDF can be shown. If charts factor suitably, this leads to individual asymptotic normality for the last element, a principal nested mean or a principal nested geodesic, say. It turns out that these results pertain to principal nested spheres (PNS) and principal nested great subspheres (PNGS) analysis by Jung et al. 2010. For illustration of a nested bootstrap two-sample test, derived, simulations are conducted and PNGS is then applied to track early human mesenchymal stem cell differentiation over a coarse time grid and, among others, to locate a change point with direct consequences for the design of further studies.

Goodness-of-fit tests for time series of counts

Šárka Hudecová¹, Marie Hušková¹ and Simos G. Meintanis²

¹Charles University (Czech Republic), ²National and Kapodistrian University of Athens (Greece)

Thursday 27
10:30-12:30
Lecture room 7
(Sali 7)

In recent years, there has been a growing interest in studying time series of counts. Various models have been proposed in the literature, and this leads to a need for goodness-of-fit tests. We propose a testing procedure for the hypothesis that a given set of discrete observations may be formulated as a particular time series of counts with a specific conditional law. Special emphasis is given to the popular models of integer autoregression (INAR) and Poisson autoregression (INARCH). The new test statistics is based on the empirical probability generating function. The asymptotic properties of the test statistics are studied and a bootstrap version of the test is proposed.

Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures

Raphael Huser¹, Thomas Opitz² and Emeric Thibaud³

¹King Abdullah University of Science and Technology (KAUST) (Saudi Arabia), ²INRA (France), ³École Polytechnique Fédérale de Lausanne (Switzerland)

Wednesday 26
10:30-12:30
Lecture room 13
(Sali 13)

Gaussian scale mixtures are constructed as Gaussian processes with a random variance. They have non-Gaussian marginals and can exhibit asymptotic dependence unlike Gaussian processes, which are always asymptotically independent except for perfect dependence. Motivated by the analysis of spatial extremes, we propose a flexible but parsimonious Gaussian scale mixture copula model, which smoothly interpolates from asymptotic dependence to independence. We show how this new model can be fitted to high threshold exceedances using a censored likelihood approach, and we demonstrate that it provides valuable information about tail characteristics. The methodology will then be illustrated with an application to wind speed data in the Pacific Northwest, US, showing that it adequately captures the data's extremal properties.

A logistic regression point of view toward loss given default distribution estimation

Monday 24
10:30-12:30
Lecture room 6
(Sali 6)

Ruey-Ching Hwang

National Dong Hwa University (Taiwan)

We propose a new procedure to estimate the loss given default (LGD) distribution. Owing to the complicated shape of the LGD distribution, using a smooth density function as a driver to estimate it may result in a decline in model fit. To overcome this problem, we first apply the logistic regression to estimate the LGD cumulative distribution function. Then, we convert the resulting result into the LGD distribution estimate. To implement the newly proposed estimation procedure, we collect a sample of 5269 defaulted debts from Moodys Default and Recovery Database. A performance study is performed using 200 pairs of in-sample and out-of-sample datasets with different sizes that are randomly selected from the entire sample. Our results show that the newly proposed procedure has better and more robust performance than its alternatives, in the sense of yielding more accurate in-sample and out-of-sample LGD distribution estimates. Thus, it is useful for studying the LGD distribution.

A novel multivariate Bernoulli distribution

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

Yi-Ting Hwang

National Taipei University (Taiwan)

Many recent studies often observe the response variables repeatedly to understand the influence of certain conditions longitudinally. The general linear model and generalized linear model for longitudinal data are used to make inference of this kind of data. Since the response variable is observed repeatedly, the model settings and estimations would need the multivariate distribution. Many continuous multivariate distributions have been proposed in the literatures. However, owing to the complexity of describing the association among the multivariate discrete random variables, it is lack of the well-known distribution. Based on the construction of the bivariate binomial distribution proposed by Biswas and Hwang (2002) and incorporating the first-order Markov chain to describe the association among the repeat measures, this paper proposes a new multivariate distribution for binary data. The basic properties of this distribution are proved and illustrated numerically. Also, the asymptotic properties of the maximum likelihood estimators (MLEs) of the parameters are discussed. Furthermore, the performance of the MLEs is evaluated through the Monte Carlo simulations.

Exact constraint-based causal discovery

Tuesday 25
10:30-12:30
Lecture room 1
(Sali 1)

Antti Hyttinen

University of Helsinki (Finland)

Causal graphical model structure discovery is a challenging part of causal inference, graphical model research and machine learning. If absence of latent confounders can be assumed, the tool of choice are exact score-based methods, as they provably find the globally optimal equivalence class of structures. Constraint-based methods on the other hand are able to handle more general model spaces with cycles and latent confounders, but offer seriously poorer accuracy. We present research that bridges this gap between constraint-based and score-based methods to some extent. We use a combination of score-based and constraint-based ideas together with Boolean (maximum) satisfiability solving, to obtain an approach that retains exactness and the generality of the model space.

Majorization, bundling, diversification and voting

Rustam Ibragimov¹, Artem Prokhorov² and Johan Walden³

¹Imperial College Business School (United Kingdom), ²University of Sydney Business School (Australia),
³University of California at Berkeley (USA)

Tuesday 25
10:30-12:30
Lecture room 5
(Sali 5)

Using majorization theory and extending earlier works in the field, we develop a framework that allows one to model the optimal bundling problem of a multiproduct monopolist providing interrelated goods with an arbitrary degree of complementarity or substitutability. Characterizations of optimal bundling strategies are derived for the seller in the case of heavy-tailed valuations and tastes for the products. We show, in particular, that if goods provided in a Vickrey auction or any other revenue equivalent auction are substitutes and bidders' tastes for the objects are moderately heavy-tailed, then the monopolist prefers separate provision of the products. However, if the goods are complements and consumers' tastes are extremely thick-tailed, then the seller prefers providing the products on a single auction. We also present results on consumers' preferences over bundled auctions for complements and substitutes in the case when their valuations exhibit heavy-tailedness. In addition, we obtain characterizations of optimal bundling strategies for a monopolist who provides complements or substitutes for profit-maximizing prices to buyers with heavy-tailed tastes. The results and approaches presented in the paper are applicable in a number of other fields in economics, finance, risk managements and related fields, including the analysis of diversification optimality under heavy-tailedness and dependence, comparisons of voting mechanisms, robust statistical and econometric methods and other areas.

Testing collinearity of vector time series

Agnieszka Jach¹ and Tucker McElroy²

¹Hanken School of Economics (Finland), ²US Census Bureau (USA)

Poster session

We investigate the collinearity of vector time series in the frequency domain, by examining the rank of the spectral density matrix at a given frequency of interest. Reduced rank corresponds to collinearity at the given frequency. When the time series data is nonstationary and has been differenced to stationarity, collinearity corresponds to co-integration at a particular frequency. We pursue a full understanding of rank through the Schur complements of the spectral density matrix, and test for rank reduction via assessing the positivity of these Schur complements, which are obtained from a nonparametric estimator of the spectral density. We provide new asymptotic results for the Schur complements, under the fixed bandwidth ratio paradigm. The test statistics are $O_P(1)$ under the alternative, but under the null hypothesis of collinearity the test statistics are $O_P(T^{-1})$, and the limiting distribution is non-standard. Subsampling is used to obtain the limiting null quantiles.

Optimisation bias correction in sparse structured variable selection

Maarten Jansen

Université libre de Bruxelles (Belgium)

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

The selection of sparse variables in high-dimensional problems suffers from the ubiquity of false positives among the many candidate models. More precisely, the false positives have a disturbing effect on the search for an optimal model according to a given information criterion. For a given, fixed model, the information criterion (nearly) unbiasedly estimates a distance between the model and the data generating process. Through the optimisation of the criterion, however, the statistics of the errors in the models under consideration change, leading to bias tending to select false positives. The bias has been described by a mirror reflecting effect with respect to an oracular selection, based on the observation of the noise-free response (see Jansen, 2014, *Biometrika*, 101(1), pages 37-55). The amount of the optimisation bias depends on the dimensionality of the problem, being marginal in low-dimensional problems, but problematic in high-dimensional problems. In

this talk we discuss the presence of the effect in structured model selection, with emphasis on optimal binary subtree selection. A selection algorithm is presented and analysed, leading to the conclusion that although the structured selection reduces the dimensionality, the optimisation bias is still an important effect to take into account.

Monday 24
10:30-12:30
Lecture room 6
(Sali 6)

Higher moments of the estimated tangency portfolio weights

Farrukh Javed and Stepan Mazur
Örebro University (Sweden)

In this paper we consider the estimated tangency portfolio weights. We derive analytical expressions for the higher central and non-central moments of these weights. Main focus has been given to skewness and kurtosis due to the importance of asymmetry and heavy tails of the data. We complement our results with an empirical study where we analyze an international diversified portfolio.

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

Modeling dynamic networks using high-dimensional binary autoregressive time series processes

Carsten Jentsch, Shaikh Tanvir Hossain and Lena Reichmann
Universität Mannheim (Germany)

Suppose a time series of networks is identified by their adjacency matrices A_1, \dots, A_T , where $A_t = (a_{ij;t})_{i,j=1,\dots,N}$ with $a_{ij;t} \in \{0, 1\}$ and $a_{ij;t} = 1$ indicating that there is a directed edge pointing from vertex i to vertex j at time t . In this project, we propose to model the joint dynamics of the edges by multivariate binary time series processes. We adopt the class of Discrete AutoRegressive Moving-Average (DARMA) models introduced by Jacobs and Lewis (1983) to model univariate categorical time series. For our purpose of modeling multivariate binary data, we extend their approach in two aspects. First, we consider vector-valued extensions. Second, we propose a simple modification that allows for negative autocorrelations. The resulting model class is flexible enough to capture very general autocorrelations driving the dynamic network structure. The model class fulfills Yule-Walker-type equations that allow in principle an explicit estimation of the model parameters. However, as the dimension of the adjacency matrices grows quadratically with the number of vertices, we have to make use of Lasso-penalization techniques to estimate sparse models. Here, we adopt the approach of Basu and Michailidis (2015), who established that consistent estimation for high-dimensional vector autoregressive models is possible via ℓ_1 -regularization for a large class of stable processes under sparsity constraints.

References: [1] Basu, S., Michailidis, G., 2015. Regularized Estimation in Sparse High-dimensional Time Series Models. *The Annals of Statistics*, 43(4), 1535–1567. [2] Jacobs, P. A., Lewis, P. A. W., 1983. Stationary Discrete Autoregressive-Moving average Time Series generated by Mixtures. *Journal of Time Series Analysis*, 4(1), 19–36.

Numerical inversion of SRNF maps for elastic shape analysis of genus-zero surfaces

Ian Jermyn¹, Hamid Laga², Qian Xie³ and Anuj Srivastava³

¹Durham University (UK), ²Murdoch University (Australia), ³Florida State University (USA)

Tuesday 25
14:00-16:00
Lecture room 5
(Sali 5)

Recent developments in elastic shape analysis are motivated by the fact that it provides a comprehensive framework for simultaneous registration, deformation, and comparison of shapes. These methods achieve computational efficiency using shape representations that transform invariant elastic metrics on shapes into Euclidean metrics, allowing for the application of standard algorithms and statistical tools. For embeddings of S^2 in R^3 , the relevant representation is called the square-root normal field (SRNF). For the SRNF to be fully exploited, a method is needed for transferring solutions (deformations, averages, modes of variations, etc.) computed in SRNF space, back to the original surface space for visualization and inference, i.e. an inverse of the SRNF map is needed. This raises very difficult questions in differential geometry that have been around for many decades. The difficulty of these questions has prompted the development of an alternative, approximate approach to the problem: an efficient multiresolution algorithm, based on solving an optimization problem in surface space, that estimates a surface corresponding to given SRNF. This solution is found to be effective even for complex shapes that undergo significant deformations including bending and stretching, e.g. human bodies and animals. I will describe this method, and show how it can be used for computing elastic shape deformations, transferring deformations, summarizing shapes, and for finding modes of variability in a given collection of shapes, while simultaneously registering the surfaces. I will demonstrate the efficacy of the proposed algorithms using statistical analysis of human body shapes, classification of generic surfaces, and analysis of brain structures.

Multidimensional functional principal component analysis

Ci-Ren Jiang

Academia Sinica (Taiwan)

Poster session

Functional principal component analysis is one of the most commonly employed approaches in functional and longitudinal data analysis and we extend it to analyze functional/longitudinal data observed on a general d -dimensional domain. The computational issues emerging in the extension are fully addressed with our proposed solutions. The local linear smoothing technique is employed to perform estimation because of its capabilities of performing large-scale smoothing and of handling data with different sampling schemes (possibly on irregular domain) in addition to its nice theoretical properties. Besides taking the fast Fourier transform strategy in smoothing, the modern GPGPU (general-purpose computing on graphics processing units) architecture is applied to perform parallel computation to save computation time. To resolve the out-of-memory issue due to large-scale data, the random projection procedure is applied in the eigendecomposition step. We show that the proposed estimators can achieve the classical nonparametric rates for longitudinal data and the optimal convergence rates for functional data if the number of observations per sample is of the order $(n/\log n)^{d/4}$. Finally, the performance of our approach is demonstrated with simulation studies and the fine particulate matter (PM 2.5) data measured in Taiwan.

Majorization results for secure outage capacities in fading wireless channels

Tuesday 25
10:30-12:30
Lecture room 5
(Sali 5)

Eduard Jorswieck
TU Dresden (Germany)

Recently, the secrecy capacities of fading wiretap channels are derived for different types of channel state information and different system models. The case with slow fading and imperfect information at the transmitter is important for practical secure data transmission. We investigate the impact of spatial correlation of the legitimate and eavesdroppers channel on the outage performance using Majorization theory. The operating regions in which the secrecy outage probability is Schur-convex and Schur-concave are partly characterized. The impact on the secure system design is discussed.

Finnish wood purchasing cartel: Damage estimates on pulpwood based on the VECM approach

Tuesday 25
14:00-16:00
Lecture room 10
(Sali 10)

Leena Kalliovirta
Natural Resources Institute Finland (Finland)

We estimate cartel effects in the highly persistent pulpwood prices in Finland by employing publicly available monthly data and the VEC (vector error correction) model approach. We generalize the before-and-after comparator-based approach into the multivariate nonstationary setting and show that then the dummy variable approach yields a reliable estimate of the cartel damage even if the data is nonstationary. Further, the approach allows us to study the existence of a transition period after the cartel. We estimate a large economic model and its statistically sufficient sub-models to depict the pulpwood markets and compute the but-for prices implied by these VEC models. While operational the purchasing cartel was able to generate a price lowering deterministic trend. The possibility of cartel creating a trend has already been predicted in game theoretic models. During the cartel period, the economic model implies that the monthly prices were decreasing by 0.68% (birch), 0.14% (spruce), and 0.71% (pine) whereas the conservative, statistically adequate sub-models imply slower decreases. Thus, the damage accumulated along the duration of the cartel and the average but-for prices were between 9% to 36% (birch), 2% to 7% (spruce), and 15% to 38% (pine) higher than the observed prices.

Use of multi-level modelling in the analysis of micro-CT data obtained using micro-CT

Poster session

Jing Kang
University of Leeds (UK)

X ray micro-computerised tomography (micro-CT) is an important tool in researching skeletal tissue pathologies, including dental caries (tooth decay) but evaluating interventions based on before and after reconstructed image stacks is complex and the use of over-simplified statistical tools may mask potentially important relationships or detect false positives. Our aim was to compare and contrast outcomes obtained using multi-level modelling (MLM) of data with traditional statistical analyses (e.g. ANOVA or Kruskal-Wallis test) for micro-CT data sets. Datasets were obtained from experiments evaluating the effects of 9 different interventions on the remineralisation of artificial caries-like lesions in human enamel. Two to six lesions were generated in enamel slabs (one lesion per slab) per tooth for 13 teeth and the mineral density of each whole lesion volume quantified using micro-CT before and after the intervention. Data was evaluated either 1) using standard ANOVA or Kruskal Wallis statistical testing or 2) MLM considering hierarchical structuring of data and any baseline coupling of the data. Results showed that traditional statistical analyses such as ANOVA or non-parametric Kruskal-Wallis test with post-hoc comparison may underestimate or overestimate the size of effect of an intervention by assuming independence of

samples and ignoring complex hierarchical relationships. These methods also fail to take in to account any heterogeneity between 1) individual enamel slabs used to generate the lesions and 2) the different teeth from which the slabs were originally taken. MLM was able to consider the clustering effect among data obtained from multiple images from a single lesion and between lesions derived from a single tooth and any heterogeneity in the data. In addition, the impact of baseline lesion mineral values on the final outcome measures could also be accommodated. We conclude that for complex laboratory investigations such as this, MLM be used to take in to account the hierarchical nature of sample populations and heterogeneity existing amongst them. Simple statistical analyses could underestimate or overestimate the effect of a given intervention, potentially faulting significant findings and reducing the potential impact and reach of the research.

Engine for likelihood free inference (ELFI)

Antti Kangasrääsio
Aalto University (Finland)

Tuesday 25
14:00-16:00
Great hall
(Juhlasali)

Engine for Likelihood Free Inference (ELFI) is an open-source Python library for performing likelihood free inference. The library is structured so that the model construction and inference are decoupled. The user has free hands to define the generative model as a directed acyclic graph, which can contain arbitrary program code in the graph nodes. Inference for these user-defined models can be performed with any of the provided approximate Bayesian computation (ABC) algorithms. ELFI is designed to be extensible: there is an API that allows e.g. novel inference methods to be implemented. The library also supports external simulators, such as Matlab code. Simulated data can be stored and reused when needed, and the computation can be parallelized from single computer to cluster environments.

Robust Shewhart control charts by using ranked set sampling and neoteric ranked set sampling schemes

Derya Karagöz and Nursel Koyuncu
Hacettepe University (Turkey)

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

This study proposes to construct Shewhart mean and range control charts by using Ranked Set Sampling and Neoteric Set Sampling schemes with robust estimators under contaminated bivariate normal process. The proposed method is simple method in which sample mean and range estimators are replaced with robust trimmed mean and interquartile range, respectively. The modified method is evaluated in terms of its type I risks and average run lengths. The performance of the simple, ranked set sampling and neoteric ranked sampling schemes is assessed when the Phase I and Phase II data are uncontaminated as well as when the Phase I and Phase II data are contaminated. The modified Shewhart control charts by using ranked set sampling and neoteric sampling schemes have the best performance in the case of non-contamination and contamination.

On control charts for not necessarily symmetric distributions with simulations

Tuesday 25
14:00-16:00
Lecture room 6
(Sali 6)

Alex Karagrigoriou

University of the Aegean (Greece)

Control charts are widely used to analyze and understand process variables, monitor effects of the variables on the difference between target and actual performance and determine if a process is under statistical control. They can also investigate whether the variability is random or it is the result of a shift in the mean and/or the standard deviation. If the reason for the observed deviation is non-random then the process has to be monitored by eliminating the causes that create the problem as well as by ensuring the non-recurrence of them. In non-symmetric distributions where a tail behavior plays a pivotal role in understanding the underlying inequality, the Gini index is preferable than the coefficient of variation since not only it captures the extreme behavior of the underlying distribution but also can provide a satisfactory interpretation of the inequality involved. In addition, the Gini index can be used for the comparison of the underlying distribution across different regions or states and also for indicating the way the quantity of interest dynamically changes over time. Finally, using the Gini index we are able to quantify differences among states. The above special features of the Gini index and its appropriateness for not necessarily symmetric distributions motivates the present work for control charts based on the Gini index. This work is filling up the literature gap by providing the methodology for control charts for the Gini index which is preferable for the analysis of any phenomenon described by any heavy or medium-tailed distributions (e.g. Pareto, Gamma, Exponential, Weibull etc). Note that such situations are frequent in the time to event in reliability theory for the analysis of time to failure (TTF) or time to repair (TTR) in technical, mechanical or technological systems. The goal of this work is to propose a new control chart monitored by the Gini index and investigate its performance in the case of not necessarily symmetric distributions.

Moment-based parameter estimation in binomial random intersection graph models

Poster session

Joona Karjalainen and Lasse Leskelä

Aalto University (Finland)

Binomial random intersection graphs can be used as parsimonious statistical models of large and sparse networks, with one parameter for the average degree and another for transitivity, the tendency of neighbours of a node to be connected. This paper discusses the estimation of these parameters from a single observed instance of the graph, using moment estimators based on observed degrees and frequencies of 2-stars and triangles. The observed data set is assumed to be a subgraph induced by a set of n_0 nodes sampled from the full set of n nodes. We prove the consistency of the proposed estimators by showing that the relative estimation error is small with high probability for $n_0 \gg n^{2/3} \gg 1$. As a byproduct, our analysis confirms that the empirical transitivity coefficient of the graph is with high probability close to the theoretical clustering coefficient of the model.

Large-scale probabilistic integration

Poster session

Toni Karvonen
Aalto University (Finland)

The probabilistic approach to numerical integration and scientific computing in general has attracted significant attention lately. However, application of probabilistic integration methods on integration problems requiring more than at most a few thousand data points has been hindered by the cubic computational cost characteristic to Gaussian process regression that underlies most of these methods. This poster describes a point design scheme that involves no approximations and allows for tremendous reduction in computational complexity. The presented scheme is efficient in high dimensions and for up tens of millions of points.

Inference ready measures of dispersiveness for small samples

Paul Kattuman
University of Cambridge (UK)

Tuesday 25
10:30-12:30
Lecture room 5
(Sali 5)

We present a family of inference-ready measures of dispersiveness, in the general sense that encompasses inequality, concentration, heterogeneity, diversity and so on, for non-negative small samples. Dispersiveness is measured as the probability of majorization relations between the sample of interest and a random sample drawn from a suitable symmetric multi-variate reference distribution chosen to serve as the benchmark. One example of such a reference distribution is the uniform distribution on the standard $(n - 1)$ -simplex, per which all possible samples of size n are equiprobable. The probabilities of majorization so defined are the P-values for the hypothesis that the sample of interest, and a random sample from the reference distribution, are of equal dispersiveness. Unlike other summary indices of dispersion which do not admit inference for small samples, these probability measures directly enable inference, and satisfy properties desirable in any general measure of dispersion.

Robust Bayesian penalized regression estimation based on Ramsay-Novick distribution

Mutlu Kaya¹, Emel Çankaya¹ and Olcay Arslan²
¹Sinop University (Turkey), ²Ankara University (Turkey)

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

Attempts to construct robust estimation procedures have shown the importance of considering more general structures for regression modeling. Bayesian approaches in this respect are inherently robust as they do not condition on a single model but rather seek to accommodate uncertainties within a joint posterior probability distribution. The sampling distribution however needs to be suitably chosen as the likelihood ingredient of an Bayesian analysis is very vulnerable to outlying observations. Heavy-tailed distributions are hypothesized to be "outlier resistant" and representing the error term distribution by them makes the resulting inferences more robust. Ramsay-Novick (RN) distribution within the class of heavy tailed distributions is here proposed to be an intriguing choice as the robustness of this distribution is formed by the modification of an unbounded influence function of a non-robust density. Literature review reveals that practicing with Ramsay- Novick distribution for this purpose has been largely avoided due to the fact that its analytically intractable form complexes the posterior distribution. Representation of previous knowledge for the unknown quantities in the form of a prior density can also be interpreted as a penalty function, which is the Bayesian interpretation for regularization. Laplace prior distribution is here chosen for $L1$ regularization so as to implement a preference for model parameters that are zero. The prior of this form creates Bayesian variable selection procedure famously known as Bayesian Lasso. The aim of this study is therefore to show theoretical evaluation of robust Bayesian estimation of a regression model with RN distributed errors and Laplace prior distribution for the parameters. Complicated forms of full conditional posterior densities were simplified, to some extent, via a

series expansion and Markov Chain Monte Carlo (MCMC) methods were employed for drawing samples from those. A simulation study was conducted to assess the finite sample performance under different scenarios created by including outliers in x -, y - and in both x - and y -directions to the data generation process and also considering variables a number of which are redundant. Resulting Bayesian inferences were found to be promising for the accomplishment of resistance to the outliers in all directions along with the selection of correct variables.

Keywords: Robust Bayesian regularized regression, Ramsay-Novick distribution, Heavy-tailed distributions, Laplace prior

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

Model distances for vine copulas in high dimensions

Matthias Killiches

Technical University of Munich (Germany)

Vine copulas are a flexible class of dependence models consisting of bivariate building blocks and have proven to be particularly useful in high dimensions. Classical model distance measures require multivariate integration and thus suffer from the curse of dimensionality. We provide numerically tractable methods to measure the distance between two vine copulas even in high dimensions. For this purpose, we develop new distance measures based on the Kullback-Leibler distance, using the result that it can be expressed as the sum over expectations of KL distances between univariate conditional densities, which can be easily obtained for vine copulas. To reduce numerical calculations we approximate these expectations only on selected diagonals. In a simulation study we show the validity of our approach and outperform Monte Carlo-integration with respect to computational time. A parametric bootstrap based testing procedure is presented, facilitating the application of our distance measures to highly relevant model selection issues: We investigate if the frequently discussed simplifying assumption for vine copulas is justified for a given data set. Another application is the determination of an optimal truncation level in order to reduce model complexity.

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

Function estimation on a large graph using Bayesian Laplacian regularization

Alisa Kirichenko and Harry van Zanten

University of Amsterdam (Netherlands)

We study a Bayesian approach to estimating a smooth function in the context of regression or classification problems on large graphs. We present a mathematical framework that allows to study the performance of nonparametric function estimation methods on large graphs and we obtain minimax convergence rates for these problems within the framework. We derive theoretical results that show how asymptotically optimal Bayesian regularization can be achieved under an asymptotic shape assumption on the underlying graph and a smoothness condition on the target function, both formulated in terms of the graph Laplacian. The priors we study are randomly scaled Gaussians with precision operators involving the Laplacian of the graph. We also consider the case of missing observations and investigate the generalization performance for various missing mechanisms.

Climate event attribution using the multivariate generalized Pareto distribution

Anna Kiriliouk¹, Philippe Naveau², Holger Rootzén³, Johan Segers¹ and Jennifer Wadsworth⁴

¹Université catholique de Louvain (Belgium), ²Université de Versailles (France), ³Chalmers University (Sweden), ⁴Lancaster University (UK)

Friday 28
10:30-12:30
Lecture room 12
(Sali 12)

Modelling univariate extreme events such as floods and heat waves is often done using the peaks over thresholds method, that is, by fitting a generalized Pareto distribution to all observations that exceed some high threshold. However, many applications in extreme value theory are inherently multivariate, and one is interested in characterizing the degree of tail dependence between the components of a multivariate vector. The multivariate generalized Pareto distribution arises as the limit of a suitably normalized vector conditioned upon at least one component of that vector being extreme. We discuss a construction device which allows us to develop new parametric tail dependence models and illustrate their use in an application aiming at estimating the Fraction of Attributable Risk, a measure often used by climate scientists to compare extreme events under different scenarios in order to illustrate the human influence on climate change.

Updating the South African sampling frame

Deon Kleinsmith

Statistics South Africa (South Africa)

Tuesday 25
14:00-16:00
Lecture room 10
(Sali 10)

Population surveys and censuses involve collecting data from households and traditionally such collections in many countries has occurred using paper format that are generally resource intensive. Sampling frames must meet quality criteria so that the target population are represented without omission and duplication. This paper aims to explain the process the statistical agency has undertaken to update the sampling frame in South Africa using innovative methods and provide results from the updated frame that was used for the countrys large scale Community Survey 2016. The desktop approach using an online capturing tool and satellite imagery with supporting GIS vector data enabled the sampling frame to be updated for downstream processes. This method allowed the field-based listing process to be significantly reduced. A hybrid data access model was implemented to provide improved access and speed to spatial data and satellite imagery at provincial and district offices. The sampling frame was first evaluated to determine the scope of the work to be undertaken and thereafter the frame was updated to capture all the dwelling units. The updated sampling frame was also compared to the South African Master Samples Primary Sampling Units (PSUs) to determine the level of significance between the frames. The results showed that there was not a significant difference between the updated frame and the Master Sample providing further evidence for more innovative approaches in updating sampling frames.

Semiparametric bivariate conditional copula regression with binary and continuous marginals

Thomas Kneib

Georg-August-Universität Göttingen (Germany)

Thursday 27
14:00-16:00
Lecture room 13
(Sali 13)

Copulas are a general and versatile tool for constructing multivariate distributions by combining a specific type of dependence structure specified via the copula with arbitrary marginal as long as these marginals are continuous. This allows for the construction of regression models for bivariate responses within the framework of distributional regression where regression predictors are placed on potentially all distributional parameters including the dependence parameter of the copula. We extend this framework by considering bivariate regression models where at least one of the responses is binary and therefore discrete. Based on the latent utility representation of binary regression models, we can still formulate a copula specification and combine the binary response part with a flexible specification of the dependence structure and the second marginal (unless this

is also binary). We develop both penalized likelihood and Bayes inference and compare the results in an application on adverse birth outcomes where we combine a binary regression model for the presence/absence of low birth weight with a two parameter Gumbel specification for gestational age.

Multiscale Poisson scanning: Multivariate limit theory and statistical implications

Tuesday 25
10:30-12:30
Lecture room 6
(Sali 6)

Claudia Koenig, Axel Munk and Frank Werner
Georg-August-University of Göttingen (Germany)

This talk will address the problem of detecting anomalies in a multivariate Poisson random field $Y_i \sim \text{Pois}(\lambda_i)$, $i \in I_n = \{1, \dots, n\}^d$. Therefore we consider a calibrated scan statistic based on local likelihood ratio tests of homogeneity against heterogeneity, i.e. for $R \subset I_n$

$$\mathcal{H} : \lambda_i = \lambda_0 \text{ for all } i \in R \quad \text{vs.} \quad \mathcal{K} : \lambda_j \neq \lambda_0 \text{ for some } j \in R.$$

Two different settings are considered, namely known and unknown expected intensity λ_0 . In the latter case, it will be estimated by the sample mean. To calibrate a corresponding scan test, we derive limit theorems for the scan statistic over a wide range of scales $\#R$. Using the a priori known limit distribution, we then obtain a procedure to scan the field for anomalies on a prescribed significance level. We present a real data example from SMS microscopy which shows the superior performance of our multiscale method compared to the most commonly used thresholding.

The importance of causal thinking and statistical thinking in causal effect studies

Tuesday 25
14:00-16:00
Lecture room 10
(Sali 10)

Ana Kolar
TaraStats, Statistical Consultancy (Finland)

Causal inference is one of the most challenging statistical topics. Data applications aiming to produce causal claims have to face causal inference challenges in a responsible manner. Statistical thinking and causal thinking are at the heart of reliable causal effect studies. We define causal thinking as a set of thinking processes required to develop an effective causal effect design and analyses procedures, and to provide reliable causal claims. The foundation of causal thinking is grounded in the knowledge of causal inference statistical methodologies. The causal thinking processes are then defined in two parts: (i) as understandings of what is causal conceptually, i.e., physical versus factual cause (Mohr, 1996); and (ii) as understandings of what is causal according to the statistical definition of causality (Kolar and Rubin, 2015). These understandings are crucial for successful employment of causal inference methodologies. Statistical thinking is presented as a skeleton for causal thinking because it enables one to see the big picture. We extend the statistical definition as provided by Box et al. (1978), Britz et al. (1996, 2000), Moore (1990), and Snee (1990) in order to make it applicative also for those who are exposed to the digital world, but professionally do not deal with data analytics. In a nutshell, we define statistical thinking as a conscious thought process that is based on statistical concepts of sampling theory, descriptive and inferential statistics, and which provides one with observational skills enabling to distinguish between different subjective realities and leaning critically towards objectivity. The paper provides an in-depth presentation of the extended statistical thinking definition and how it applies, together with the required causal thinking, to causal effect studies. The importance of both kinds of thinking is illustrated with two main types of causal studies using observational data: (i) studies evaluating effects of known causes; and (ii) studies investigating causes of realised effects. The former are based on counterfactual theories of causality (Imbens & Rubin 2015, Pearl 2000), whereas the later are based on system dynamics approaches (Ford 1999, Forrester 1968, Sterman 2000). The paper concludes with guidelines on how to approach and implement a causal effect study and what to be aware of when providing conclusions about causal claims.

Keywords: statistical thinking, causal thinking, causal effect studies, potential outcomes, system dynamics, observational data

References: [1] Box, G. E. P., Hunter, W. G., and Hunter, J. A., 1978. *Statistics for Experimenters*, New York: John Wiley and Sons. [2] G. Britz et al., Spring 1996. *Statistical Thinking. A special publication of the ASQ Statistics Division Newsletter*. [3] Ford, A., 1999. *Modeling the Environment, An Introduction to System Dynamics Modeling of Environmental Systems*, Island Press. [4] Forrester, J. W., 1968. *Principles of Systems*, Wright-Allen Press. Imbens, G., Rubin, D.B., 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press. [5] Kolar, A., Rubin, D.B., 2015. *Comparative studies with Propensity Score Methods: Theory, Practice and Nomenclature*. Manuscript in preparation. [6] Mohr, L. B., 1996. *The causes of human behavior: Implications for theory and method in the social sciences*. The University of Michigan Press. [7] Moore, D. S., 1990. Uncertainty, in *On the Shoulders of Giants*, ed. L. A. Steen, National Academy Press, 95-173. [8] Pearl, J., 2000. *Causality*, Cambridge University Press. [9] Snee, R., 1990. *Statistical Thinking and its Contribution to Quality. The American Statistician*, 44(2), 116–121. [10] Sterman, J. D., 2000. *Business Dynamics, System Thinking and Modeling for a Complex World*, Irwin McGraw-Hill Education.

Explosion in branching processes and applications to random graph models

Julia Komjathy¹, Enrico Baroni and Remco van der Hofstad

¹Eindhoven University of Technology (Netherlands)

Thursday 27
14:00-16:00
Great hall
(Juhlasali)

In this talk I will investigate the question of explosion of branching processes, i.e., when is it possible that a BP produces infinitely many offspring in finite time. Two important cases in terms of application are age-dependent BPs and BPs arising from epidemic models where individuals are only contagious in a possibly random interval after being infected.

The motivation for studying the explosiveness question is to understand weighted distances in power-law degree random graphs, such as the configuration model, in the regime where the degree distribution is a power-law with exponent between (2,3). Here, the local neighborhood of a vertex and thus the initial stages of the spreading can be approximated by an infinite mean offspring BP. I will explain the recent results on this area. This part is joint work with Enrico Baroni and Remco van der Hofstad.

Tree based precision medicine for right censored data

Michael R. Kosorok

University of North Carolina at Chapel Hill (USA)

Monday 24
10:30-12:30
Great hall
(Juhlasali)

Estimating individualized treatment rules is a central task of personalized or precision medicine. In this presentation, we develop a new, tree based nonparametric method for discovering high dimensional treatment rules when the primary outcome of interest is right censored. The new approach avoids both inverse probability of censoring weighting and semiparametric modeling of either the censoring or failure times. We establish consistency and convergence rates for our proposed estimators. In simulation studies, our estimators demonstrate improved performance compared to existing methods. We also illustrate the proposed method on a phase III clinical trial of non-small cell lung cancer.

Friday 28
10:30-12:30
Lecture room 6
(Sali 6)

Network inference using prior knowledge from time-course data

Gino Bertrand Kpogbezan¹, Aad van der Vaart¹, Mark van de Wiel² and Wessel van Wieringen²
¹Leiden University (Netherlands), ²VUmc and VU Amsterdam (Netherlands)

High-dimensional network reconstruction may benefit substantially from the inclusion of prior knowledge. In the case of gene regulatory networks such prior knowledge may come from data of a different archetypal study. Two types of omics studies can be discerned namely, an observational study and a time-course study. Data from both studies can be used to reconstruct the gene regulatory network. Here, we present a non-model based and a model based approach of borrowing prior information from time-course data into conditional independence graph based gene regulatory network. In a simulation study we show that both the non-model based and the model based approach can greatly reconstruct the underlying gene regulatory network and outperform known competitors if the prior knowledge is relevant. We demonstrate the benefits of the methods in an analysis of gene expression data from GEO.

Friday 28
10:30-12:30
Lecture room 6
(Sali 6)

Time series modeling on dynamic networks

Jonas Krampe and Jens-Peter Kreiss
TU Braunschweig (Germany)

We consider multivariate time series on dynamic networks with a fixed number of vertices. Each component of the time series is assigned to a vertex of the underlying network. The dynamic dependency between the various components of the time series is modeled by means of the edges. We make use of a multivariate doubly stochastic time series framework, that is we assume linear processes for which the coefficient matrices are stochastic processes themselves. We explicitly allow for dependence in the dynamics of the coefficient matrices, including of course an i.i.d. structure as is typically assumed in random coefficients models. In the paper conditions for stationarity will be given and asymptotic normality of simple statistics like the sample mean is investigated. Furthermore, autoregressive moving average models are defined in this framework. Estimators of the parameters are discussed and some interesting features of these processes are shown in simulations.

Monday 24
14:00-16:00
Auditorium IV

Estimation in partially observed branching processes under different scenarios

Andrea Kraus
Masaryk University (Czech Republic)

Originally proposed as population-growth models, branching processes have found applications in various fields, and their popularity is increasing as tools for their statistical analysis are becoming available. While their definition makes them suitable for modelling the growth or the decrease of a phenomenon in time, the implied non-linear dynamics makes them notoriously difficult from the statistical perspective. An additional complication arises when the analysis needs to be based on incomplete observations, which is often the case if one is to be realistic about the collected data. Unless specific information suggesting a suitable offspring distribution is available, non-parametric estimators of crucial characteristics, such as the Malthusian parameter, with desirable asymptotic properties are employed. Several types of asymptotics can be considered: the observation time or the number of ancestors or both may reasonably be assumed large in a specific application. Depending on the type of asymptotics that is of interest and the supercriticality or criticality or subcriticality of the process, different properties of the estimators may be expected and different estimators may be preferred. In our contribution, we examine classical estimators and their extensions for the case of incomplete observation under different scenarios. We study their statistical properties from the theoretical perspective as well as their performance on simulations. We illustrate the methodology on an application to epidemic modelling and an application to modelling a

population of an endangered species. A small number of ancestors and a not-so-short observation time is considered in the first application, while a not-so-small number of ancestors and a short observation period is considered in the second application. Assuming incomplete observation brings the model closer to reality in both cases.

Growing simplified vine copula trees: Challenging Dißmann's algorithm

Daniel Kraus

Technical University of Munich (Germany)

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

Vine copulas are pair-copula constructions enabling multivariate dependence modeling in terms of bivariate building blocks. One of the main tasks of fitting a vine copula is the selection of a suitable tree structure. For this the prevalent method is a heuristic called Dißmann's algorithm. It sequentially constructs the vine's trees by maximizing dependence at each tree level, where dependence is measured in terms of absolute Kendall's τ . However, the algorithm disregards any implications of the tree structure on the simplifying assumption that is usually made for vine copulas to keep inference tractable. We develop two new algorithms that select tree structures focused on producing simplified vine copulas for which the simplifying assumption is violated as little as possible. For this we make use of a recently developed statistical test of the simplifying assumption. In a simulation study we show that our proposed methods outperform the benchmark given by Dißmann's algorithm by a great margin. Several real data applications emphasize their practical relevance.

Inferential procedures for fragmentary functional data

David Kraus

Masaryk University (Czech Republic)

Friday 28
10:30-12:30
Lecture room 5
(Sali 5)

Methods of functional data analysis have been developed for analysing collections of functions, for example curves, surfaces or images. It is usually assumed that all functions are completely (or densely) observed on the same domain. Recently, new applications brought attention to situations where each functional variable may be observed only on a subset of the domain while no information about the function is available on the complement. The principal difficulty for inference in this context is that temporal averaging (e.g., in inner products) is precluded by the incompleteness of the observed functions. Motivated by partially observed medical data, we develop new inferential procedures on the first and second-order properties for functional samples.

Inferring dynamical systems using the profile likelihood

Clemens Kreutz and Jens Timmer

Albert-Ludwigs-University Freiburg (Germany)

Monday 24
10:30-12:30
Lecture room 10
(Sali 10)

Ordinary differential equation models are frequently used to describe dynamical systems. However, due to nonlinearity, limited amount of data and large number of parameters, classical statistical approaches for the calculation of uncertainties for estimated parameters and for model predictions are often hardly feasible. The profile likelihood generalizes classical methods for assessing uncertainties like the Fisher Information for the nonlinear and finite-sample case. In addition, it enables data-based identifiability and observability analyses and offers a model reduction strategy guided by experimental data. In this talk, an overview about the capabilities of the profile likelihood is provided using dynamic models of biochemical reaction networks as they occur in living cells or organisms. Nevertheless, the presented concepts are general and applicable to any kind of deterministic model.

Profile likelihood-based uncertainty analyses

Clemens Kreutz and Jens Timmer
Albert-Ludwigs-University Freiburg (Germany)

Ordinary differential equation models are frequently used to describe dynamical systems. However, due to nonlinearity, limited amount of data and large number of parameters, classical statistical approaches for the calculation of uncertainties for estimated parameters and for model predictions are often hardly feasible. The profile likelihood generalizes classical methods for assessing uncertainties like the Fisher Information for the nonlinear and finite-sample case. In addition, it enables data-based identifiability and observability analyses and offers a model reduction strategy guided by experimental data. In this talk, an overview about the capabilities of the profile likelihood is provided using dynamic models of biochemical reaction networks as they occur in living cells or organisms. Nevertheless, the presented concepts are general and applicable to any kind of deterministic model.

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

Heavy tailed spatial autocorrelation models

Alexander Kreuzer, Thomas Nagler, Tobias Erhardt and Claudia Czado
Technische Universität München (Germany)

Appropriate models are necessary for spatially autocorrelated data. They account for the fact that observations are not independent. A popular model in this context is the simultaneous autoregressive (SAR) model that allows to model the spatial dependency structure of a response variable and the influence of covariates on this variable. This spatial regression model assumes that the error follows a normal distribution. Since this assumption cannot always be met it is necessary to extend this model to other error distributions. We propose the extension to the t-distribution, the tSAR model, which can be used if we observe heavy tails in the fitted residuals of the SAR model. In addition we provide a variance estimate that considers the spatial structure of a variable which helps us to specify inputs for our models. In an extended simulation study we show that the proposed estimators of the tSAR model are performing well. The last part consists of an application to fire danger. Fire danger is measured with many variables. In our modeling approach we use less covariates which allows us to predict fire danger in other locations where less information is available. Furthermore we see in this application that the tSAR model is a notable improvement compared to the SAR model.

Wednesday 26
10:30-12:30
Great hall
(Juhlasali)

Nonparametric estimation of time-varying parameters in nonlinear models

Dennis Kristensen and Young Jun Lee
University College London (UK)

We propose nonparametric estimators of time-varying parameters in a general class of nonlinear time series models. Under weak regularity conditions, we show the proposed estimators are consistent and follow a normal distribution in large samples. A key concept in our analysis is local stationarity, for which we provide primitive conditions to hold in the case of Markov processes. To demonstrate the usefulness of our general results, we provide primitive conditions for our theory to hold in a number of examples, including ARCH models and Poisson autoregressions with time-varying parameters.

Simultaneous confidence bands for ratios of quantile functions

Tatyana Krivobokova, Fabian Dunker and Stephan Klasen
University of Göttingen (Germany)

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

Ratios of quantile functions are an important tool to evaluate the distributional pattern of growth processes when repeated cross-sectional data are available. The most popular example are Growth Incidence Curves (GIC) that allow assessments whether income growth in developing countries has been pro-poor. We present a construction of uniform confidence bands for GICs and similar ratios of quantile functions. In contrast to existing point-wise confidence bands that are created via bootstrapping, the bands we propose are valid simultaneously for all points in the domain of GICs. They allow for an assessment of the location and on larger scales of the slope of the curves. Furthermore, the construction does not depend on bootstrapping but on an analysis of the asymptotic distribution of GICs. This allows for significantly faster algorithms. The performance of the confidence band is demonstrated in simulations and in an example using income data from Uganda for 1999-2005.

Nonparametric adaptive Poisson regression by model selection

Martin Kroll
Universität Mannheim (Germany)

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

We consider a nonparametric Poisson regression problem where the integer valued response Y is the realization of a Poisson random variable with parameter $\lambda(X)$. The aim is to estimate the functional parameter λ from independent or weakly dependent observations $(X_1, Y_1), \dots, (X_n, Y_n)$ in a random design framework. In the first part of the talk, we determine a lower bound for the mean integrated squared error and propose a projection estimator in the terms of the trigonometric basis that attains this bound up to a numerical constant. The main part of the talk is devoted to the construction of a fully-data driven estimator of λ by model selection. We proceed in two steps: first, we assume that an upper bound for $\|\lambda\|_\infty$ is known. Under this assumption, we construct an adaptive estimator whose dimension parameter is defined as the minimizer of a penalized contrast criterion. Second, we replace the known upper bound on $\|\lambda\|_\infty$ by an appropriate plug-in estimator of $\|\lambda\|_\infty$. The resulting adaptive estimator is shown to attain the minimax optimal rate up to an additional logarithmic factor. Appropriate concentration inequalities for Poisson point processes turn out to be an important ingredient of the proofs. We end up the talk with an illustration of our theoretical results by a short simulation study.

The tail empirical process of regularly varying functions of geometrically ergodic Markov chains

Rafal Kulik
University of Ottawa (Canada)

Thursday 27
10:30-12:30
Lecture room 1
(Sali 1)

We consider a stationary regularly varying time series which can be expressed as a function of a geometrically ergodic Markov chain. We obtain practical conditions for the weak convergence of weighted versions of the multivariate tail empirical process. These conditions include the so-called geometric drift or Foster-Lyapunov condition and can be easily checked for most usual time series models with a Markovian structure. We illustrate these conditions on several models and statistical applications.

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

Testing variance components in nonlinear mixed effects models

Estelle Kuhn¹, Charlotte Baey² and Paul-Henry Cournède²

¹INRA (France), ²CentraleSupélec, MICS (France)

Mixed effects models are widely used to describe inter and intra individual variabilities in a population. A fundamental question when adjusting such a model to the population consists in identifying the parameters carrying the different types of variabilities, i.e. those that can be considered constant in the population, referred to as fixed effects, and those that vary among individuals, referred to as random effects. In this talk, we propose a test procedure based on the likelihood ratio one for testing if the variances of a subset of the random effects are equal to zero. The standard theoretical results on the asymptotic distribution of the likelihood ratio test can not be applied in our context. Indeed the assumptions required are not fulfilled since the tested parameter values are on the boundary of the parameter space. The issue of variance components testing has been addressed in the context of linear mixed effects models by several authors and in the particular case of testing the variance of one single random effect in nonlinear mixed effects models. We address the case of testing that the variances of a subset of the random effects are equal to zero. We prove that the asymptotic distribution of the test is a chi bar square distribution, indeed a mixture of chi square distributions, and identify the weights of the mixture. We highlight that the limit distribution depends on the presence or not of correlations between the random effects. We present numerical tools to compute the corresponding quantiles. Finally, we illustrate the finite sample size properties of the test procedure through simulation studies and on real data.

Friday 28
10:30-12:30
Lecture room 12
(Sali 12)

Simulating conditionally specified models

Kun-Lin Kuo

National University of Kaohsiung (Taiwan)

Conditionally specified statistical models offer several advantages over joint models; one is that the Gibbs sampling can be used. In multiple imputation, fully conditional specification is gaining popularity because it is flexible and computationally straightforward. However, it is rather restrictive to require that every regression/classification must involve all of the variables. Feature selection often removes some variables from the set of predictors, thus making the regression local. Mixture of full and local conditionals is referred to as partially collapsed Gibbs (PCG) sampler, which can achieve faster convergence due to reduced conditioning. However, its implementation requires choosing a correct scan order. Using an invalid scan order brings out an incorrect transition kernel and leads to a wrong stationary distribution. We prove a necessary and sufficient condition for the PCG sampling to simulate the correct joint distribution, and show that certain conditional models should be excluded from consideration because they can never be simulated correctly. We propose an algorithm that identifies all of the valid scan orders for a given conditional model. Checking compatibility among conditionals of different localities is also discussed.

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

Monotonic property and asymptotic distribution of regression correlation coefficient

Takeshi Kurosawa and Nobuhiro Suzuki

Tokyo University of Science (Japan)

We give an asymptotic distribution for the estimator of *regression correlation coefficient* (RCC) that is one of the measures of predictive power for generalized linear models (GLMs). The RCC, proposed by Zheng and Agresti, is a population value and a generalization of the population value for the coefficient of determination. Recently, Takahashi and the first author in this study provided an explicit form of the RCC for a Poisson regression model and proposed a new RCC estimator. This study discusses the monotonic property of the estimator of the RCC. Furthermore, we show an asymptotic normality of the estimator.

Maximum likelihood in multivariate Gaussian models under total positivity

Steffen Lauritzen

University of Copenhagen (Denmark)

Monday 24
14:00-16:00
Great hall
(Juhlasali)

We analyze the problem of maximum likelihood estimation for Gaussian distributions that are multivariate totally positive of order two (MTP2). By exploiting connections to phylogenetics and single-linkage clustering, we give a simple proof that the maximum likelihood estimator (MLE) for such distributions exists based on at least two observations, irrespective of the underlying dimension. We demonstrate that the MTP2 constraint serves as an implicit regularizer and leads to sparsity in the estimated inverse covariance matrix, determining what we name the ML graph. We show that the maximum weight spanning forest (MWSF) of the empirical correlation matrix is a spanning forest of the ML graph. In addition, we show that we can find an upper bound for the ML graph by adding edges to the MSWF corresponding to correlations in excess of those explained by the forest. We conclude with a brief discussion of signed MTP2 distributions.

Optimal scaling of the random walk Metropolis algorithm for L_p mean differentiable distributions

Sylvain Le Corff

CNRS & Université Paris-Sud (France)

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

In this talk, we consider the optimal scaling of random walk Metropolis algorithms for high-dimensional target densities differentiable in L_p mean. These distributions may be irregular at some points (the Laplace density for instance) or supported on an interval. We first discuss scaling issues when applying MCMC algorithms to high-dimensional Bayesian problems. Then, we present the main result which is the weak convergence of the Markov chain produced by the algorithm (appropriately rescaled in time and space) to a Langevin diffusion process as the dimension goes to infinity. The scaling limit is established under assumptions which are much weaker than the one usually required. This result has important practical implications for the use of random walk Metropolis algorithms in Bayesian frameworks based on sparsity inducing priors and is illustrated using numerical experiments.

Goodness-of-fit test for multistable Lévy processes

Ronan Le Guével¹ and Anne Philippe²

¹Université Rennes 2 (France), ²Université de Nantes (France)

Thursday 27
10:30-12:30
Lecture room 7
(Sali 7)

Multistable processes, that is, processes which are, at each "time", tangent to an α -stable process, but where the index of stability varies along the path, have been introduced as models for phenomena where the intensity of jumps is non constant. The original definition of these models produces locally stable processes that are no longer self-similar, without stationary nor independent increments. We will present how to estimate two functions of interest of these models, and we will explain how we can obtain a statistical test in order to decide if a data set comes from an α -stable process or a multistable one.

Wednesday 26
10:30-12:30
Lecture room 5
(Sali 5)

Smooth curve fitting to 3D shape data

Huiling Le
University of Nottingham (UK)

To carry out the smoothing spline fitting method for shapes of configurations in a 3-dimensional Euclidean space using the technique of unrolling and unwrapping, we link parallel transport along a geodesic on Kendall shape space to the solution of a homogeneous first-order differential equation, some of whose coefficients are implicitly defined functions. This enables us to approximate the procedure of unrolling and unwrapping by simultaneously solving such equations numerically, and so to find numerical solutions for smoothing splines fitted to 3-dimensional shape data. We apply this fitting method to the analysis of some moving peptide data, and a test for the best model among those considered is given.

Friday 28
14:00-15:00
Great hall
(Juhlasali)

Deep learning: A statistical puzzle

Yann LeCun
Facebook AI Research & New York University (USA)

Deep learning is at the root of revolutionary progress in visual and auditory perception by computers, and is pushing the state of the art in natural language understanding, dialog systems and language translation. Deep learning systems are deployed everywhere from self-driving cars to social networks content filtering to search engines ranking and medical image analysis. A deep learning system is typically an "almost" differentiable function, composed of multiple highly non-linear steps, parametrized by a numerical vector with 10^7 to 10^9 dimensions, and whose evaluation of one sample requires 10^9 to 10^{10} numerical operations. Training such a system consists in optimizing a highly non-convex objective averaged over millions of training samples using a stochastic gradient optimization procedure. How can that possibly work? The fact that it does work very well is one of the theoretical puzzles of deep learning. For example, the accuracy and the learning speed both get better as the systems capacity is increased. Why? A more important puzzle is how to train large neural networks under uncertainty. This is key to enabling unsupervised learning and model-based reinforcement learning where the machine is trained from raw natural data, without human-supplied labels. A class of methods called adversarial training is currently our favorite method to approach this problem. This may be the key to building learning system that learn how the world works by observation.

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

A new class of depth functions based on Bregman divergences.

Samuela Leoni and Jean-Baptiste Aubin
ICJ INSA Lyon (France)

In multivariate data analysis a depth function measures how 'deep' a point is located in a given data cloud in Euclidean d -space.

We introduce a new class of depth functions based on Bregman divergences. These dissimilarity measures arise as an error function when approximating a strictly convex function by a tangent hyperplane. They include symmetric dissimilarity measures like the squared Euclidean distance and the Mahalanobis distance as well as asymmetric instances as the Kullback-Leibler divergence and the Itakura-Saito divergence.

The properties of the associated depth functions are investigated. Some numerical results are presented.

Power-law configuration graphs robustness in random environment

Marina Leri

Karelian Research Centre of the Russian Academy of Sciences (Russia)

Friday 28
10:30-12:30
Lecture room 6
(Sali 6)

We consider configuration graphs vertex degrees of which are i.i.d. random variables following the power-law distribution. The distribution parameter τ is a random variable following either the uniform distribution or truncated gamma distribution on a predefined interval $[a, b]$ ($0 < a < b < \infty$). We consider a process of graph destruction which is often referred to as a ‘forest fire model’ and can be interpreted as a fire (or any other destructive influence) spreading over the graph edges. The probability of transferring a destructive influence over a graph edge is either predefined and fixed for all the graph edges or is a random variable following the standard uniform distribution. By computer simulation we study the robustness of considered graphs from a viewpoint of vertex survival in the two cases of starting a destruction process (or ‘fire’): ‘random’ and ‘targeted’. The results of comparative analysis of the two graph models for various destruction processes in terms of ensuring maximum vertex survival after the ‘fire’ are presented.

The study is supported by the Russian Foundation for Basic Research, grant 16-01-00005.

Fraction of time approach in statistical inference for cyclostationary signals

Jacek Leśkow

Cracow University of Technology (Poland)

Thursday 27
10:30-12:30
Lecture room 7
(Sali 7)

The traditional modelling approach to cyclostationary signals $\{X(t); t \in \mathbf{R}\}$ is based on the assumption that the signal is viewed as a realization of a stochastic process. Then, under some asymptotic independence assumptions like φ -mixing or weak dependence the properties of the estimation procedures of relevant characteristics of signals are shown. The most common two examples are estimation of the mean function and the covariance function of a cyclostationary signal. In the talk, another approach will be presented. The signal $\{X(t); t \in \mathbf{R}\}$ will be viewed as a function and the corresponding probability model will be based on cumulative distribution functions based on level crossing of the signal. Under such approach it will be shown how estimating the mean, covariance and other popular characteristics of the signal is possible. We will also show applications to signal processing.

References: [1] Napolitano, A., Dehay, D., 2013. Central Limit Theorem in the Functional Approach. *IEEE Transactions on Signal Processing*, 61(16), 4025 – 4037. [2] Napolitano, A., 2007. Non-relatively measurable functions for secure communications signal design. *Signal Processing*, 87, 2765–2780. [3] Napolitano, A., 2006. Foundations of the functional approach for signal analysis. *Signal Processing*, 86, 3796–3825.

An iterative plug-in for P-Spline regression

Sebastian Letmathe, Yuanhua Feng and Hendrik Schmitz

University of Paderborn (Germany)

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

This paper proposes a new IPI- (iterative plug-in) rule for selecting the smoothing parameter in P-Splines based on the asymptotic results in Wand (1999). In comparison to the DPI- (direct plug-in) approach of Wand (1999) the current algorithm is fully automatic and self-contained. The proposed algorithm is a fixpoint-search procedure and the resulting smoothing parameter is (theoretically) independent of the initial value. Like the DPI, the IPI-idea can also be employed as a refining stage to improve the quality of the smoothing parameter selected by other methods, e.g. the Mallows’ C_p , Cross Validation or the residual maximum likelihood. Some numerical features of P-Splines, the practical performance of the IPI-algorithm and behaviours of the resulting P-Spline estimates are studied in detail through a large simulation. The nice properties of the IPI-algorithm are confirmed by the simulation results. It is shown that the proposal performs very well in practice,

runs quickly and converges in most of the cases after a few iterations. Practical relevance of the proposal is illustrated by different data examples, in particular examples from health economics.

Keywords: P-Splines, selection of the smoothing parameter, iterative plug-in, simulation, health economics.

Tuesday 25
14:00-16:00
Lecture room 12
(Sali 12)

Subgroup balancing propensity score

Fan Li
Duke University (USA)

We investigate the estimation of subgroup treatment effects with observational data. Existing propensity score matching and weighting methods are mostly developed for estimating overall treatment effect. Although the true propensity score should balance covariates for the subgroup populations, the estimated propensity score may not balance covariates for the subgroup samples. We propose the subgroup balancing propensity score (SBPS) method, which selects, for each subgroup, to use either the overall sample or the subgroup sample to estimate propensity scores for units within that subgroup, in order to optimize a criterion accounting for a set of covariate-balancing conditions for both the overall sample and the subgroup samples. We develop a stochastic search algorithm for the estimation of SBPS when the number of subgroups is large. We demonstrate through simulations that the SBPS can improve the performance of propensity score matching in estimating subgroup treatment effects. We then apply the SBPS method to data from the Italy Survey of Household Income and Wealth (SHIW) to estimate the treatment effects of having debit card on household consumption for different income groups. This is a joint work with Jing Dong and Junni Zhang.

Tuesday 25
10:30-12:30
Lecture room 12
(Sali 12)

Hybrid quantile regression estimation for time series models with conditional heteroscedasticity

Guodong Li
The University of Hong Kong (China)

Estimating conditional quantiles of financial time series is essential for risk management and many other applications in finance. It is well-known that financial time series display conditional heteroscedasticity. Among the large number of conditional heteroscedastic models, the generalized autoregressive conditional heteroscedastic (GARCH) process is the most popular and influential one. So far, feasible quantile regression methods for this task have been confined to a variant of the GARCH model, the linear GARCH model, owing to its tractable conditional quantile structure. This paper considers the widely used GARCH model. An easy-to-implement hybrid conditional quantile estimation procedure is developed based on a simple albeit nontrivial transformation. Asymptotic properties of the proposed estimator and statistics are derived, which facilitate corresponding inferences. To approximate the asymptotic distribution of the quantile regression estimator, we introduce a mixed bootstrapping procedure, where a time-consuming optimization is replaced by a sample averaging. Moreover, diagnostic tools based on the residual quantile autocorrelation function are constructed to check the adequacy of the fitted conditional quantiles. Simulation experiments are carried out to assess the finite-sample performance of the proposed approach. The favorable performance of the conditional quantile estimator and the usefulness of the inference tools are further illustrated by an empirical application.

Functional principal component analysis with application to air quality data

Poster session

Pai-Ling Li¹ and Hsin-Yu Chan

¹Tamkang University (Taiwan)

In this study, the daily collected pollutant concentrations obtained from Taiwan air quality monitoring network are viewed as functional data. The structures of various air pollutants in Taiwan are explored by functional principal component (FPC) analysis. In addition, the monitoring stations are clustered through functional data clustering via FPC subspace projection, where the optimal number of clusters for each pollutant is determined by a forward functional testing algorithm. The results show that some pollutants are correlated and the grouping structures are related to climate and topography.

On structure testing for component covariance matrices of a high-dimensional mixture

Weiming Li¹ and Jianfeng Yao²

¹Shanghai University of Finance and Economics (China), ²Hong Kong University (China)

Tuesday 25
10:30-12:30
Great hall
(Juhlasali)

By studying the family of p-dimensional scaled mixtures, this paper shows for the first time a non trivial example where the eigenvalue distribution of the corresponding sample covariance matrix does not converge to the celebrated Marcenko-Pastur law. A different and new limit is found and characterized. The reasons of failure of the Marcenko-Pastur limit in this situation are found to be a strong dependence between the p-coordinates of the mixture. Next, we address the problem of testing whether the mixture has a spherical covariance matrix. It is shown that the traditional Johns test and its recent high-dimensional extensions both fail for high-dimensional mixtures, precisely due to the different spectral limit above. In order to find a remedy, we establish a novel and general CLT for linear statistics of eigenvalues of the sample covariance matrix. A new test using this CLT is constructed afterwards for the sphericity hypothesis.

Bayesian inference in RDD: An application to students's academic performances

Federica Licari, Alessandra Mattei and Fabrizia Mealli

University of Florence (Italy)

Tuesday 25
14:00-16:00
Lecture room 12
(Sali 12)

In Italian Universities, School of Engineering requires students to take a pre-test. Enrollment depends on the results of this pre-test. Specifically students with a test score less or equal a prefixed threshold must comply with additional educational requirements to improve their skills, which we refer to as OFA (Obblighi Formativi Aggiuntivi). The assignment rule underlying the OFA defines a Sharp regression Discontinuity design (SRD). Following Li, Mattei, Mealli (2016), we adopt a probabilistic formulation of the assignment mechanism underling RD designs and we select suitable subpopulations around the threshold on the basis of observed covariates using a Bayesian model-based approach.

Our primary aim is to evaluate causal effects of OFA on students' academic performances, properly accounting for the fact that after the test, some students may decide not to enroll in the School of Engineering. In previous analyses on Italian University students, data on enrollment status was usually neglected focusing on students who actually enrolled. Enrollment status is a post-treatment intermediate variable, and assessing causal effects of OFA on enrollment decisions might be of interest *per se*. Students's performances are not defined for students who do not enroll: they are truncated by death. In order to assess causal effects on students's performances, we use the principal stratification framework defining causal estimands within the Rubin Causal Model. Specifically, we focus on assessing local causal effects on academic performances for the principal stratum of students who would always enroll irrespective of their OFA status, and investigate

characteristics of individuals who belong the other principal strata (students who would enroll only under a treatment/OFA condition and students who would never enroll irrespective of their OFA status) for which causal effects are not defined. We use a Bayesian model-based approach to inference.

References: [1] Li, F., Mattei, A., Mealli, F., 2015. Evaluating the effect of university grants on student dropout: Evidence from a regression discontinuity design using bayesian principal stratifications Analysis. *The Annals of Applied Statistics*, 9(4), 1906–1931.

Poster session

Complex valued robust multidimensional SOBI

Niko Lietzén¹, Klaus Nordhausen^{2,3} and Pauliina Ilmonen¹

¹Aalto University School of Science (Finland), ²University of Turku (Finland), ³University of Tampere (Finland)

Several blind source separation applications require methods that allow deviations from model assumptions and are not sensitive to outliers. In this work, we consider a new algorithm, robust multidimensional eSAM-SOBI, for complex valued blind source separation. The algorithm utilizes multidimensional autocovariance matrices to detect multidimensional dependency structures. In applications like functional magnetic resonance imaging, the interesting dependency structures occur in three dimensional space and time simultaneously and hereby traditional SOBI-like procedures lose information by considering only dependencies in a single direction. We illustrate the better performance of our algorithm in a complex valued simulation study, where outliers are present in some settings.

The Batch Markov Modulated Poisson process, a powerful tool in reliability models.

Rosa Lillo

Universidad Carlos III de Madrid (Spain)

The Batch Markov Modulated Poisson process is a subclass of the versatile Batch Markovian Arrival process (BMAP) [Neuts, 1979] which has been widely used in realibility contexts. The considered class allows for modeling correlated inter-events times as well as correlated simultaneous events. Both theoretical and applied aspects are examined here. On one hand, the identifiability of the process is proven. Also, some findings concerning the correlation structures are provided. Second, some properties related to the count process are derived. And finally, an inference methodology for estimating the model parameters given a samples of inter-event times and batches sizes is presented. Illustrations with real dataset describing the behaviour of a call center are given.

Order selection for detecting misspecification in the random effects of GLMMs

Kuo-Chin Lin¹ and Yi-Ju Chen²

¹Tainan University of Technology (Taiwan), ²Tamkang University (Taiwan)

The normality assumption of the random-effects distribution in generalized linear mixed models is practically assumed, but it may be too restrictive to present the major feature of data. We consider more flexible mixing distributions by the semi-nonparametric density approach based on the truncated Hermite expansions alternative to a normal distribution, and propose a goodness-of-fit test for detecting misspecified random effects in generalized linear mixed models using order selection criteria. The limiting distribution of the proposed test follows a scaled chi-squared distribution with degrees of freedom depending on the truncated order and dimension of random effects. The critical values of the proposed test are approximated by simulation results. A real longitudinal binary data set is demonstrated for the application of the proposed test.

Keywords: Generalized linear mixed models, Goodness-of-fit, Longitudinal binary data, Order selection.

Poster session

Wednesday 26
10:30-12:30
Lecture room 12
(Sali 12)

Robust principal expectile component analysis

Liang-Ching Lin

National Cheng Kung University (Taiwan)

Friday 28
10:30-12:30
Lecture room 7
(Sali 7)

Principal component analysis (PCA) is a widely used dimension reduction technique especially for the high dimension data analysis. These principal components are identified by sequentially maximizing the component score variance for observations centered on the sample mean. However, in practice, one might be more interested in the variation captured by the tail characters instead of the sample mean, for example the analysis of expected shortfall. To properly capture the tail characters, principle expectile component (PEC) analysis was proposed based on an asymmetric L2 norm (Tran, Osipenko, and Hardle, 2014). Although, in order to achieve robustness against outliers, we generalize the PEC by integrating with Huber's norm. The newly proposed method is named as principal Huber-type expectile components (PHEC). A derivative free optimization approach, particle swarm optimization (PSO), is adopted to efficiently identify the components in PHEC. Simulation studies show that the PHEC outperforms PCA and PEC in capturing the tail variation in the case of normal mixture distributions. Finally a real example is analyzed for illustration.

A quantile regression approach to parental time with children

Yi-Chen Lin

Tamkang University (Taiwan)

Monday 24
14:00-16:00
Lecture room 10
(Sali 10)

Parents time on child care has profound impact on childrens human capital formation. Previous studies on parental time with children, which use ordinary least squares and Tobit model as their estimation method, focus on conditional mean of parental time with children. As far as we know, no empirical attention has been paid to parents that spent little time with children. Since wage is likely to be endogenous to time spent with children, this thesis uses censored quantile instrumental variable estimator to analyze the determinants of time spent on child care. Based on micro data from the 2003-2010 American Time Use Survey, we found that the dummy variable for non-hispanic non-black and non-white is statistically significantly negatively associated with time spent with children. This result implies that more resources should be directed to facilitating and promoting parental time with children among non-hispanic non-black and non-white parents.

Variational Bayes estimation in constrained kurtosis diffusion imaging under a Rician noise model

Jia Liu , Dario Gasbarra and Juha Railavo

University of Helsinki (Finland)

Tuesday 25
10:30-12:30
Lecture room 12
(Sali 12)

The analysis of diffusion MR-data is often based on the simplifying assumption that the diffusion of water molecules follows locally a centered Gaussian distribution. However, diffusional non-Gaussianity is a common scenario in biological tissues due to potential barriers and compartments. Diffusion kurtosis imaging (DKI) is an extension of diffusion tensor imaging (DTI), quantifying both Gaussian and non-Gaussian diffusivity by means of the 2nd order tensors and diffusion kurtosis and is considered very useful in the diagnosis of brain disorders. The model contains three constraints coming from the physical relevance of diffusion. Diffusion MR-data is subject to noise and artefacts. A general method for image denoising is image regularization, where the parameter estimators at each voxel depend on the information from neighbours voxels. Often the regularization step is applied at a second stage after estimating the tensor parameters voxelwise. An alternative approach is to perform estimation and regularization steps simultaneously. We propose an original and efficient method for regularized Variational Bayes estimation of DKI parameters under constraints, based on a data augmentation of the Rician likelihood. The method is illustrated using synthetic data and a real dataset from a case-control study on patients with Lévy body dementia.

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

Model selection and model averaging for nonlinear regression models

Qingfen Liu¹, Qingsong Yao²

¹Otaru University of Commerce (Japan), ²Renmin University of China (China)

This paper considers the problem of model selection and model averaging for regression models which can be nonlinear in their parameters and variables. We propose a new information criterion called nonlinear model information criterion (NIC), which is proved to be an asymptotically unbiased estimator of the risk function under nonlinear settings. We also develop a nonlinear model averaging method (NMA) and extend NIC to NICMA criterion, which is the corresponding weight choosing criterion for NMA. By taking account of the complexity of model forms into the penalty term, NIC and NICMA achieve significant gain of performance. The optimality of NMA, convergence of the selected weight and other theoretical properties are proved. Simulation results show that NIC and NMA lead to relatively lower risks compared with alternative model selection and model averaging methods under most situations.

Group variable selection using bootstrap method

Zhang Liu

The University of Hong Kong (China)

We consider a bootstrap group Lasso procedure for grouped variable selection under linear models. To enhance selection accuracy, we run group Lasso on bootstrap samples and select variables by taking intersection of the active variables selected for each bootstrap sample. Based on a sub-gradient solution to the group Lasso problem, we investigate conditions under which our procedure selects correct groups of variables with probability converging to one. The efficiency of the bootstrap procedure for grouped variable selection under low-dimensional cases is illustrated with simulations.

Optimal aggregation in circular deconvolution by a frequentist Bayesian approach

Xavier Loizeau¹, Fabienne Comte² and Jan Johannes³

¹Universität Heidelberg (Germany), ²Université Paris Descartes (France), ³Ruprecht-Karls-Universität Heidelberg (Germany)

In a circular deconvolution problem we consider the estimation of the density of a circular random variable X using noisy observations where the additive circular noise is independent of X and admits a known density. In this context it has been shown in Comte and Taupin [2003] and Johannes and Schwarz [2013] that, considering a family of projection estimators, a fully data-driven choice of the dimension parameter using a model selection approach can lead to minimax-optimal rates of convergence up to a constant. In this presentation we propose a fully data-driven aggregation of those projection estimators depending on a tuning parameter which we interpret later on as a number of iterations. Thereby, we obtain a family of fully data-driven circular deconvolution estimators indexed by the iteration parameter. Interestingly in the limiting case, when the iteration parameter tends to infinite, the aggregated estimator coincides with the projection estimator with fully data-driven choice of the dimension by the model selection approach. However, for any element of the family of aggregated estimators we show up to a constant minimax optimality as well as oracle optimality within the family of projection estimators. The proposed aggregation strategy is inspired by an iterative Bayes estimate in the context of an inverse Gaussian sequence space model, which Johannes and Loizeau [2016] have studied from a frequentist Bayesian point of view. In the same spirit we present an iterative Bayesian model with Bayes estimator coinciding with the aggregated estimator with associated iteration parameter and posterior distribution concentrating at optimal rate as the sample size increases.

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

On bandwidth choice for spatial data density estimation

Zudi Lu¹, Zhenyu Jiang, Nengxiang Ling, Dag Tjøstheim and Qiang Zhang

¹University of Southampton (UK)

Monday 24
10:30-12:30
Lecture room 13
(Sali 13)

Bandwidth selection is crucial in spatial kernel estimation that has become a useful tool in exploring non-Gaussian complex spatial data. In this paper, the choice of both adaptive (i.e., local data dependent) and non-adaptive bandwidths for spatial lattice data density estimation is investigated. We first establish optimality properties of a spatial cross validation (CV) choice for a non-adaptive bandwidth. Although the CV idea has been popular for choosing bandwidth in data-driven kernel smoothing of independent and time series data, its theory and application have not been much investigated for spatial data. Conditions that ensure asymptotic optimality of the spatial CV selected bandwidth in terms of various error measures are derived, actually, also extending time series optimality results. Next, the CV estimate is used in a spatially adaptive bandwidth choice, and oracle properties of the corresponding density estimation are obtained. Numerical simulations demonstrate that finite-sample performance of the proposed spatial adaptive CV bandwidth choice works rather well. It outperforms the existing R-routines such as the rule of thumb and the so-called second-generation Sheather-Jones bandwidths for moderate size of samples. Empirical application to a set of spatial soil data is further illustrated with non-Gaussian features more significantly identified by spatial adaptive density estimation.

Model selection using penalized empirical likelihood

Artis Luguzis¹ and Jānis Valeinis³

University of Latvia (Latvia)

Monday 24
10:30-12:30
Lecture room 10
(Sali 10)

In multivariate analysis selecting the most appropriate subset of all available explanatory variables is a crucial part of modeling process. Penalized empirical likelihood (PEL) simultaneously performs model selection and estimates the parameters of the corresponding model in an asymptotically consistent way. PEL has been applied for generalized linear models [1] and even for more complicated setups, for example, in the models taking into account the correlation structure of longitudinal data [2]. The application of PEL to new statistical models is related to defining appropriate estimating equations that determines a corresponding profile PEL function and an adjusting of computation algorithm if necessary. The computation and maximization of PEL function is a key step of the process and it has to be carried out carefully. This talk is based on ongoing work extending PEL method to the count data models [3] that are able to incorporate overdispersion and excess zeros, which are common features in the count datasets in economics and social sciences.

References: [1] Tang, C. Y., Leng, C., 2010. Penalized High-Dimensional Empirical Likelihood. *Biometrika*, 905–919. [2] Nadarajah, T., 2016. Empirical likelihood based longitudinal studies. *Diss. Memorial University of Newfoundland* [3] Zeileis, A., Christian K., Jackman S., 2008. Regression models for count data in R. *Journal of statistical software*, 27(8), 1–25.

Nonparametric testing in boundary models

Gwennaëlle Mabon, Markus Reiss and Martin Wahl

Humboldt-Universität zu Berlin (Germany)

Monday 24
10:30-12:30
Lecture room 12
(Sali 12)

For nonparametric regression with one-sided errors and a boundary curve model for Poisson point processes we consider estimation of nonlinear functionals of the form $\int \Phi(g(x))dx$ for known weakly differentiable Φ . In view of L^p -norms a primary example is $\Phi(x) = |x|^p$. We propose a simple estimator by maximum likelihood already developed by Reiss and Selk (2015) in the same model. In that case unbiased estimation is feasible and optimal convergence rates can be derived. As an application we discuss L^p -separation rates in nonparametric testing. The proofs rely essentially on martingale stopping arguments for counting processes and the underlying point process geometry. Surprising differences with standard models like Gaussian white noise are discussed.

Monday 24
14:00-16:00
Auditorium IV

Two-step estimation of an additive model for censored data.

Samuel Maistre¹, Anouar El Ghouch² and Ingrid Van Keilegom³

¹Université de Strasbourg (France), ²Université Catholique de Louvain (Belgium), ³KU Leuven (Belgium)

Nonparametric additive models have been studied widely since the seminal work of Hastie and Tibshirani (1986). They provide a good compromise between parametric and fully nonparametric modelings. As each function of interest is one-dimensional, it can be pictured and therefore interpreted easily by practitioners. Different techniques aiming to avoid the curse of dimensionality have been proposed, including penalized regression splines and kernel backfitting. A relatively recent method combines the ideas of these two methods, namely Spline Backfitted Kernel (SBK) estimation. We propose to adapt it when the response variable is right censored. This method includes two steps : firstly, estimate the additive model using splines for which the number of knots leads to undersmoothing ; secondly, use the previous estimates to perform a univariate kernel smoothing to estimate each function of interest. We show that when we use the Kaplan-Meier weights, the asymptotic results of this procedure are similar to those of the uncensored case.

Tuesday 25
14:00-16:00
Lecture room 6
(Sali 6)

Optimizing the main dependability and performance measures of a diesel engine system

Sonia Malefaki and Christos Samlis

University of Patras (Greece)

An important characteristic of our modern times is the design of large scale, complexity and accuracy of mechanical and other systems which are extremely demanding in dependability. Many of our societies modern services own their existence to the above mentioned systems. This kind of systems, during their lifetime, is constantly losing its ability to perform to the same extent as it was designed to do. They appear to function in a certain amount of degradation states before they reach a total failure. In many occasions a total failure or degradation of the system provokes financial and/or social losses. To avoid that kind of situations, periodical inspection and maintenance of these systems are of great importance. A typical example of a system that functions in a certain amount of deterioration states is a diesel engine system and its subsystems that are subject to inspection and maintenance. Great emphasis has been given to the asymptotic behavior of this system, due to the fact that the majority of them are designed to operate for a long period of time. In this work the availability and the total operational cost of the diesel engines subsystems will be studied in order to evaluate the aforementioned measures of a diesel engine's system. Moreover, an optimal inspection and also maintenance policy which maximizes the availability of the system or minimizes the total operational cost is determined.

Thursday 27
10:30-12:30
Lecture room 10
(Sali 10)

On non precisely observed right censored survival data

Sergey Malov

Saint Petersburg State University (Russia)

We consider a model of right censored survival data observed at some fixed or random (observation) times. Let (T, U) be a vector of independent failure and censoring times respectively, and $W_i, i = 1, \dots, k$, be k observation times. The right censored observation consists of the event time $X = \mathbb{1}_{\{T \leq U\}}$ and the indicator $\mathbb{1}_{\{T \leq U\}}$. After k -interval censoring with the intervals $I_1 = [0, W_1]$ and $I_j = (W_{j-1}, W_j], j = 2, \dots, k$, the event time X is not observed exactly any more. Then the observation contains of all the times W_i and the vector $(\kappa, \kappa\delta)$, where $\kappa = \sum_{j=1}^k j \mathbb{1}_{\{X \in I_j\}}$. Note that the indicator δ is observed iff $X \leq W_k$. The right censored, the case k interval censored and the life table survival data are particular cases of the interval censoring of the right censored data. In the case of life table survival data (under fixed observation times) the unknown failure time distribution function is not estimable nonparametrically even in the observation times, but

if the observation times are random there is a way to estimate it. We create a two-step pseudo-likelihood estimator of the failure time distribution based on the sample from the distribution of $(\kappa, \kappa\delta)$ obtained from the right censored data after the interval censoring using an independent sample from the distribution of (W_1, \dots, W_k) , and discuss the optimization problem to get the estimator. Consistency of the pseudo-likelihood estimator under the special case $k = 1$ is obtained. Simulation results display good enough approximation of the true distribution function of failure time by the estimator and its bootstrapped version.

Compound Poisson approximation to estimate the Lévy density

Ester Mariucci¹ and Céline Duval

¹Humboldt-Universität zu Berlin (Germany)

Monday 24
14:00-16:00
Lecture room 12
(Sali 12)

We construct an estimator of the Lévy density, with respect to the Lebesgue measure, of a pure jump Lévy process from high frequency observations: we observe one trajectory of the Lévy process over $[0, T]$ at the sampling rate Δ , where $\Delta \rightarrow 0$ as $T \rightarrow \infty$. The main novelty of our result is that we directly estimate the Lévy density in cases where the process may present infinite activity. Moreover, we study the risk of the estimator with respect to L_p loss functions, $1 \leq p < \infty$, whereas existing results only focus on $p \in \{2, \infty\}$. The main idea behind the estimation procedure that we propose is to use that “*every infinitely divisible distribution is the limit of a sequence of compound Poisson distributions*” (see e.g. Corollary 8.8 in Sato (1999)); we can thus profit from the abundance of techniques developed for the estimation of the Lévy density of a compound Poisson process in the high frequency setting. We consider linear wavelet estimators and the performance of our procedure is studied in term of L_p loss functions, $p \geq 1$, over Besov balls. The resulting rates are minimax-optimal for a large class of Lévy processes, including subordinators. The results are illustrated on several examples.

A new approach to fitting generalised additive models for location, scale and shape

Giampiero Marra¹ and Rosalba Radice²

¹University College London (UK), ²University of London (UK)

Thursday 27
10:30-12:30
Lecture room 13
(Sali 13)

Regression models typically involve a response variable and a set of covariates. Modeling flexibly the effects of covariates on the outcome through the specification of additive predictors for all the parameters of the outcome distribution has received increasing attention over the last 10 years. We will provide an overview of the modelling options of the function `gamlss()` within the R package `SemiParBIVProbit` and discuss the main theoretical and computational underpinnings of the generic inferential approach developed to fit the models. We will illustrate some of the developments using a medical case study and also discuss recent developments made in the direction of modelling survival outcomes.

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

Sliced average variance estimation for multivariate time series

Markus Matilainen, Klaus Nordhausen and Hannu Oja
University of Turku (Finland)

Linear supervised dimension reduction has a long tradition for iid data with a rich literature. The idea in this setup is to find all linear combinations of a predictor vector \mathbf{x} which are needed to model a response y , where the functional relationship between the response and the explaining variables is assumed to be unknown. Multivariate time series occur also more and more often in a regression context, where the goal is to model one time series as a function of several other time series. Supervised dimension reduction in time series context is much more difficult as the dependency between the response series and the explaining series might also be lagging in time. Using supervised iid dimension reduction methods by adding lagged time series as new variables to the data increases the dimension of the problem dramatically and at the same time reduces the sample size. A supervised dimension reduction approach aimed especially for time series has recently been proposed. Under quite weak assumptions authors use for their TSIR approach the approximate joint diagonalization of several supervised matrices, which consider the temporal nature of the data. These matrices were inspired by the iid supervised dimension reduction method sliced inverse regression (SIR) and have similar drawbacks as the standard SIR. Similar as SIR for iid data was improved by SAVE (sliced average variance estimation), we would like to suggest a SAVE for time series version TSAVE here. Also a hybrid of TSIR and TSAVE is introduced. Our methods are first illustrated by examples. Then their predictive performance compared to other methods, including TSIR, is measured with a simulation study.

Thursday 27
14:00-16:00
Lecture room 10
(Sali 10)

Central limit theorems for functionals of large dimensional sample covariance matrix and mean vector

Stepan Mazur¹, Taras Bodnar² and Nestor Parolya³

¹Örebro University School of Business (Sweden), ²Stockholm University (Sweden), ³Leibniz University of Hannover (Germany)

In this talk we consider the asymptotic distributions of functionals of the sample covariance matrix and the sample mean vector obtained under the assumption that the matrix of observations has a matrix variate general skew normal distribution. The central limit theorem is derived for the product of the sample covariance matrix and the sample mean vector. Moreover, we consider the product of an inverse covariance matrix and the mean vector for which the central limit theorem is established as well. All results are obtained under the large dimensional asymptotic regime where the dimension p and sample size n approach to infinity such that $p/n \rightarrow c \in (0, 1)$.

Tuesday 25
14:00-16:00
Lecture room 12
(Sali 12)

Evaluating the European central bank corporate sector purchase program via a regression discontinuity design with an ordered forcing variable

Andrea Mercatanti^{1,2}, Fan Li³, Taneli Mäkinen² and Andrea Silvestrini²

¹Luxembourg Institute of Socio-Economic Research (Luxembourg), ²Bank of Italy (Italy), ³Duke University (USA)

The Corporate Sector Purchase Programme (CSPP), which started in June 2016, is part of the expanded Asset Purchase Programme of the European Central Bank (ECB), which is designed to address the risks of a too prolonged period of low inflation in the euro area. One exercise of monetary policy interest, is in assessing the effects of the CSPP on bond spreads. The eligibility criteria of the CSPP make it suitable to be evaluated via a Regression Discontinuity Design. In fact, under the CSPP, the ECB can buy corporate bonds provided they obtained a rating, from a rating agency, above a threshold (BBB- or equivalent). Therefore, the purchase of bonds can be considered a treatment, while the rating is a categorical ordered forcing variable whose non-standard

cardinality represents the main challenge to face in a RDD analysis. Our proposed estimation strategy is based on identifying a new threshold on a continuous predictor of the categorical ordered forcing variable, for the purposes of which we employ the coupon rate of the bond. To move away from the standard inference *at* the threshold we propose to assume unconfoundedness of the treatment given the continuous predictor. We apply the overlap weights to detect an interval *around* the threshold via a balancing assessment, across treatments, of a pre-programme variable available from the firms' balance sheets, then to estimate the effects of the CSPP in the detected interval via a weighted estimator of the Average Treatment Effect. The provisional results show a significant effects of the programme on the outcome of interest. Our current work is focused on enlarging the set of pre-programme data from firms' balance sheets in order to reinforce the balancing assessment.

Degrees of freedom for piecewise Lipschitz estimators

Poster session

Frederik Riis Mikkelsen and Niels Richard Hansen

University of Copenhagen (Denmark)

A representation of the degrees of freedom akin to Stein's lemma is given for a class of estimators of a mean value parameter in \mathbb{R}^n . Contrary to previous results our representation holds for a range of discontinuous estimators. It shows that even though the discontinuities form a Lebesgue null set, they cannot be ignored when computing degrees of freedom. Estimators with discontinuities arise naturally in regression if data driven variable selection is used. For the Lasso-OLS estimator the general representation leads to an estimate of the degrees of freedom based on the Lasso solution path, which in turn can be used for estimating the risk of Lasso-OLS. A similar estimate is proposed for best subset selection. The usefulness of the risk estimates for selecting the number of variables is demonstrated via simulations with a particular focus on Lasso-OLS.

Conditional means of low-dimensional projections from high-dimensional data: Explicit error bounds.

Ivana Milovic

University of Vienna (Austria)

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

Take a random d -vector Z that has a Lebesgue density and so that $EZ = 0$ and $EZZ' = I_d$. Consider two projections defined by unit-vectors α and β , namely a response $y = \alpha'Z$ and an explanatory variable $x = \beta'Z$. Under regularity conditions, Leeb has shown (2013, AoS) that for most β 's, $E[y|x] \approx$ linear in x , and that $Var[y|x] \approx$ const in x , provided that d is large. These results imply that most simple submodels of a high-dimensional linear model are approximately correct. But Leeb's results are asymptotic, as $d \rightarrow \infty$ and no explicit bounds have been established. We provide explicit, finite- d error bounds for the results regarding the conditional expectation. For a fixed d , let E_d be the set of β 's such that $E[y|x] \approx$ linear in x . We find bounds on the size of E_d (w.r.t. the uniform distribution v on the unit d -sphere) and we show that its size increases very fast. Namely, $v(E_d) \rightarrow 1$, as $d \rightarrow \infty$, at a rate faster than any polynomial rate. Our current research suggests that similar results can be expected for the conditional variances.

Tuesday 25
14:00-16:00
Lecture room 1
(Sali 1)

Optimal change point tests in high dimension

Jirak Moritz
TU Braunschweig (Germany)

Suppose we observe a high-dimensional time series and are interested in testing for change points. We provide tests that are minimax optimal subject to sparse alternatives. The construction is based on consistent bootstrap methods or appropriate limit distributions. The conditions are very flexible and include many popular multivariate linear and nonlinear models from the literature, such as ARMA, GARCH and related models. In the case of general factor models our tests are even adaptive with respect to the spatial dependence structure, and the corresponding minimax rate can be expressed in terms of underlying quantiles.

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

Approximate computation of data depths that satisfy the projection property

Pavlo Mozharovskyi¹ and Rainer Dyckerhoff²
¹CREST, Ensai (France), ²University of Cologne (Germany)

Data depth is a concept in multivariate statistics that measures the centrality of a point with respect to a given data cloud in an Euclidean space. If the depth of a point can be represented as the minimum of the depths with respect to all unidimensional projections of the data, then the depth satisfies the so-called projection property. Such depths form an important class that includes many of the depths that have been proposed in literature. For depths that satisfy the projection property an approximate algorithm can be easily constructed since taking the minimum of the depths with respect to only a finite number of unidimensional projections yields an upper bound for the depth with respect to the multivariate data. Such an algorithm is particularly useful if no exact algorithm exists or if the exact algorithm has a high computational complexity, as this is the case with the halfspace depth or the projection depth. To compute these depths in high dimensions, the use of an approximate algorithm with better complexity is surely preferable. Several strategies for selecting the univariate projections are proposed and the performances of the respective algorithms are compared.

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

Selection of sparse vine copulas in high dimensions with the Lasso

Dominik Mueller and Claudia Czado
Technical University of Munich (Germany)

We propose a novel structure selection method for high dimensional ($d > 100$) sparse vine copulas. Current sequential greedy approaches for structure selection require calculating spanning trees in hundreds of dimensions and fitting the pair copulas and their parameters iteratively throughout the structure selection process. Our method uses a connection between the vine and structural equation models (SEMs). The later can be estimated very fast using the Lasso, also in very high dimensions, to obtain sparse models. Thus, we obtain a structure estimate independently of the chosen pair copulas and parameters. Additionally, we define the novel concept of regularization paths for R-vine matrices. It relates sparsity of the vine copula model in terms of independence copulas to a penalization coefficient in the structural equation models. We illustrate our approach and provide many numerical examples. These include simulations and data applications in high dimensions, showing the superiority of our approach to other existing methods.

Latent space stochastic block model for social networks

Brendan Murphy¹, James Ng, Tyler McCormack, Bailey Fosdick and Ted Westling

¹University College Dublin (Ireland)

Thursday 27
14:00-16:00
Lecture room 13
(Sali 13)

A large number of statistical models have been proposed for social network analysis in recent years. In this paper, we propose a new model, the latent position stochastic block model, which extends and generalises both latent space model (Hoff et al., 2002) and stochastic block model (Nowicki and Snijders, 2001). The probability of an edge between two actors in a network depends on their respective class labels as well as latent positions in an unobserved latent space. The proposed model is capable of representing transitivity, clustering, as well as disassortative mixing. A Bayesian method with Markov chain Monte Carlo sampling is proposed for estimation of model parameters. Model selection is performed WAIC and models of different number of classes or dimensions of latent space can be compared. We apply the network model to social network interactions of Irish politicians and are able to identify highly interpretable classes which assist in understanding the political position of independent politicians who are not affiliated to any political party.

Mixture inner product spaces and their application to functional data analysis

Hans-Georg Müller¹, Zhenhua Lin² and Fang Yao²

¹University of California at Davis (USA), ²University of Toronto (Canada)

Monday 24
10:30-12:30
Lecture room 1
(Sali 1)

We introduce the concept of mixture inner product spaces associated with a given Hilbert space, which feature an infinite-dimensional mixture of finite-dimensional vector spaces and are dense in the underlying Hilbert space. While this concept is of interest for data analysis in any infinite-dimensional Hilbert space, the case of functional data that are random elements in the L^2 space of square integrable functions is of special interest. In this mixture space representation, each realization of the underlying stochastic process falls into one of the component spaces and is represented by a finite number of basis functions. The mixture representation also provides a new perspective for the construction of a probability density in function space under mild regularity conditions and leads to a mixture functional principal component analysis, where individual trajectories possess a trajectory-specific dimension. We establish estimation consistency of the functional mixture density and introduce an algorithm for fitting the functional mixture model based on a modified expectation-maximization algorithm.

Estimating the response density in nonlinear and nonparametric regression

Ursula Müller

Texas A&M University (USA)

Monday 24
10:30-12:30
Lecture room 12
(Sali 12)

We consider nonlinear and nonparametric regression models in which the errors have mean zero and are independent of the covariates. The independence assumption is important: it enables us to construct an estimator for the response density that uses all the observed data, in contrast to the usual local smoothing techniques, and which therefore permits a faster rate of convergence. For invertible (monotonic) regression functions, and a suitably chosen bandwidth, this estimator is consistent and converges with the optimal parametric root-n rate. We also discuss cases when the regression function is not invertible and the root-n rate cannot be achieved. If the regression function is a step function, we provide a response density estimator that has the same bias as the usual estimators based on the responses, but a smaller asymptotic variance.

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

Data depth for measurable random mappings

Stanislav Nagy
KU Leuven (Belgium)

Data depth is a mapping, which to a point in a multivariate vector space $s \in S$ and a probability measure P on S assigns a number $D(s; P)$ describing how central s is with respect to P , in an attempt to generalize quantiles to multivariate data. For S having infinite dimension, depth is typically considered only for S being the Banach space of continuous functions over a compact interval. In the contribution, we explore possibilities of extension of known depth functionals beyond this simplest setting. We discuss definitions, theoretical properties, and consistency/measurability issues connected with a straightforward generalization of commonly used depth functionals towards multidimensional random mappings which may lack continuity, be observed discretely, or be contaminated with additive noise.

Monday 24
14:00-16:00
Lecture room 5
(Sali 5)

New developments in network time series

Guy Nason
University of Bristol (UK)

We explain some new developments in the modelling and analysis of time series observed over time-varying networks. Such networks can change their edge structure over time as well as nodes (customers) appearing/disappearing in the model and reappearing as apparent new nodes. We consider the interplay between trend modelling and Network-ARIMA models in this exciting context and show how our models can be used for forecasting and anomaly detection in several current data-rich examples.

Wednesday 26
10:30-12:30
Great hall
(Juhlasali)

Tightness of M-estimators for multiple linear regression in time series

Bent Nielsen¹ and Søren Johansen²
¹University of Oxford (UK), ²University of Copenhagen (Denmark)

We show tightness of a general M-estimator for multiple linear regression in time series. The positive criterion function for the M-estimator is assumed lower semi-continuous and sufficiently large for large argument. Particular cases are the Huber-skip and quantile regression. Tightness requires an assumption on the frequency of small regressors. We show that this is satisfied for a variety of deterministic and stochastic regressors, including stationary and random walks regressors. The results are obtained using a detailed analysis of the condition on the regressors combined with some recent martingale results.

Thursday 27
14:00-16:00
Lecture room 10
(Sali 10)

On convergence of the sample correlation matrices in high-dimensional data

Sévérien Nkurunziza and Yueleng Wang
University of Windsor (Canada)

In this paper, we consider an estimation problem concerning the matrix of correlation coefficients in context of high dimensional data settings. In particular, we revisit some results in Li and Rolsalsky [Li, D. and Rolsalsky, A. (2006). Some strong limit theorems for the largest entries of sample correlation matrices, *The Annals of Applied Probability*, 16, 1, 423-447]. Four of the main theorems of Li and Rolsalsky (2006) are established in their full generalities and we simplify substantially some proofs of the quoted paper. Further, we generalize a theorem which is useful in deriving the existence of the p th moment as well as in studying the convergence rates in law of large numbers.

Multivariate outlier detection with ICS

Klaus Nordhausen¹, Aurore Archimbaud² and Anne Ruiz-Gazen²
¹University of Turku (Finland), ²University of Toulouse 1 Capitole (France)

Tuesday 25
14:00-16:00
Lecture room 7
(Sali 7)

Invariant Coordinate Selection (ICS) jointly diagonalizes two scatter matrices to find structures in data. In this talk we will demonstrate how to use ICS for multivariate outlier detection when assuming that the proportion of outliers is small as it is typical in many industrial quality control applications. For this purpose simulations results and real data examples are shown as well as the how the package implementing the methods works.

Stein's method, logarithmic Sobolev and transport inequalities

Ivan Nourdin¹, Michel Ledoux² and Giovanni Peccati¹
¹Luxembourg University (Luxembourg), ²Université de Toulouse (France)

Tuesday 25
10:30-12:30
Lecture room 13
(Sali 13)

I will review recent connections between Stein's approximation method, logarithmic Sobolev and transport inequalities.

Uncertainty quantification for biclustering model

Nurzhan Nurushev and Eduard Belitser
VU Amsterdam (Netherlands)

Thursday 27
10:30-12:30
Lecture room 12
(Sali 12)

We study the problem of inference on the unknown parameter in the biclustering model by using the penalization method which originates from the empirical Bayes approach. The underlying biclustering structure is that the high-dimensional parameter consists of a few blocks of equal coordinates. The main inference problem is the uncertainty quantification (i.e., construction of a confidence set for the unknown parameter), but on the way we solve the estimation and posterior contraction problems as well. We pursue a novel local approach in that the procedure quality is characterized by a local quantity, the oracle rate, which is the best trade-off between the approximation error by a biclustering structure and the best performance for that approximating biclustering structure. The approach is also robust in that the additive errors in the model are not assumed to be independent with some known distribution, but (in general dependent) only satisfying certain mild exchangeable exponential moment conditions. We introduce the excessive bias restriction under which we establish the local (oracle) confidence optimality of the proposed confidence ball. Adaptive minimax results (for the graphon estimation and posterior contraction problems) follow from our local results. The results for the stochastic block model follow, with implications for network modeling.

Taking zero lower bound seriously: A structural VAR model containing positive-valued components

Henri Nyberg
University of Turku & University of Helsinki (Finland)

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

In the conventional structural vector autoregressive (SVAR) analysis, the variables included in the SVAR model are treated as real-valued components. This is not, however, always the case. In this study, a SVAR model containing at least one strictly positive-valued component is developed. Despite the nonlinearity originating from the inclusion of positive-valued components, the generalized impulse response functions and forecast error variance decompositions can be expressed with explicit analytical formulae without a need to resort simulation-based methods. In the classic three-variable U.S. monetary policy VAR system, containing the positive nominal short-term interest rate and thus the recent zero lower bound period, the empirical results point out important differences to the linear SVAR model.

Tuesday 25
10:30-12:30
Lecture room 1
(Sali 1)

Causal inference for social network data

Elizabeth L. Ogburn¹, Oleg Sofrygin, Ivan Diaz and Mark J. van der Laan

¹Johns Hopkins University (USA)

Interest in and availability of social network data has led to increasing attempts to make causal and statistical inferences using data collected from subjects linked by social network ties. But inference about all kinds of estimands, from simple sample means to complicated causal peer effects, is difficult when only a single network of non-independent observations is available. Estimation of causal effects is complicated not only by dependence, but also by the fact that one subject's exposure or treatment may affect not only his/her own outcome but also the outcomes of his/her social contacts. This phenomenon, known as interference, poses challenges for nonparametric identification of causal effects. There is a dearth of principled methods for identifying causal effects using observational data of this kind and for statistical inference that takes into account the dependence that such observations can manifest. We extend recent work by van der Laan (2014) on causal inference for causally connected units to more general social network settings. Our asymptotic results allow for dependence of each observation on a growing number of other units as sample size increases. We are not aware of any previous methods for inference about network members in observational settings that allow the number of ties per node to increase as the network grows. While previous methods have generally implicitly focused on one of two possible sources of dependence among social network observations, we allow for both dependence due to contagion, or transmission of information across network ties, and for dependence due to latent similarities among nodes sharing ties. We describe estimation and inference for causal effects that are specifically of interest in social network settings.

Thursday 27
9:00-10:00
Great hall
(Juhlasali)

Scatter matrices and linear dimension reduction

Hannu Oja¹, Klaus Nordhausen¹, David E. Tyler² and Joni Virta¹

¹University of Turku (Finland), ²The State University of New Jersey (USA)

Most linear dimension reduction methods proposed in the literature can be formulated using a relevant pair of scatter matrices, see e.g. Tyler et al. (2009), Bura and Yang (2011) and Liski et al. (2014). The eigenvalues and eigenvectors of one scatter matrix with respect to another one can be used to determine the dimension of the signal subspace as well as the projection to this subspace. In this talk, three classical dimension reduction methods, namely principal component analysis (PCA), fourth order blind identification (FOBI) and sliced inverse regression (SIR) are considered in detail. The first two moments of subsets of the eigenvalues are used to test for the dimension of the signal space. The limiting null distributions of the test statistics are given and bootstrap strategies are suggested for small sample sizes. The theory is illustrated with simulations and real data examples. The talk is in part based on Nordhausen et al. (2017).

Keywords: Independent component analysis; Invariant coordinate selection; Principal component analysis; Sliced inverse regression.

References: [1] Bura, E., Yang, J., 2011. Dimension estimation in sufficient dimension reduction: a unifying approach. *Journal of Multivariate Analysis*, 102, 130–142. [2] Liski, E., Nordhausen, K., Oja, H., 2013. Supervised invariant coordinate selection. *Statistics: A Journal of Theoretical and Applied Statistics*, 48, 711–731. [3] Nordhausen, K., Oja, H., Tyler, D.E., 2017. Asymptotic and bootstrap tests for subspace dimension. Submitted. [4] Tyler, D., Critchley, F., Dümbgen, L., Oja, H., 2009. Invariant coordinate selection. *Journal of Royal Statistical Society B*, 71, 549–592.

Methods of network comparison

Sofia Olhede

University College London (UK)

Thursday 27
14:00-16:00
Great hall
(Juhlasali)

The topology of any complex system is key to understanding its structure and function. Fundamentally, algebraic topology guarantees that any system represented by a network can be understood through its closed paths. The length of each path provides a notion of scale, which is vitally important in characterizing dominant modes of system behavior. Here, by combining topology with scale, we prove the existence of universal features which reveal the dominant scales of any network. We use these features to compare several canonical network types in the context of a social media discussion which evolves through the sharing of rumors, leaks and other news. Our analysis enables for the first time a universal understanding of the balance between loops and tree-like structure across network scales, and an assessment of how this balance interacts with the spreading of information online. Crucially, our results allow networks to be quantified and compared in a purely model-free way that is theoretically sound, fully automated, and inherently scalable.

A weak dependence notion for functional variables

Paulo Eduardo Oliveira and Idir Arab

CMUC (Portugal)

Friday 28
10:30-12:30
Lecture room 5
(Sali 5)

Dependence notions have always tried to describe some behaviour about the influence between past and future. The mixing conditions achieved this looking directly at the probabilities, while more recent literature proposed some control through conditions of covariances of suitable transformations of the past and future. The popular positive or negative dependence variations do just this, with controls defined by the covariances of the random variables. Weak dependence notions were introduced replacing these covariances by suitable coefficients. We propose a new definition in this spirit considering the family of Lipschitz transformations together with suitable coefficients. This new notion is strictly larger than the classical positive or negative dependence notions, as it allows for non monotone transformations of the variables. We prove a Bernstein type inequality with a sufficiently fast decay rate to be able to derive almost sure convergence rates. In fact, for real random variables, we may extend the classical asymptotic results, including convergence in distribution of empirical processes. The fact that this dependence depends on the family of Lipschitz transformation and that there are no covariances of the variables involved allows the extension to operatorial random variables, which is a drawback of the positive or negative dependence notions. Hence, we may prove some almost sure asymptotic results under dependence for operatorial variables.

PaRIS-based Rao-Blackwellization of particle Gibbs

Jimmy Olsson¹ and Eric Moulines²

¹KTH Royal Institute of Technology (Sweden), ²École Polytechnique (France)

Wednesday 26
10:30-12:30
Lecture room 1
(Sali 1)

Given a set of observations from a general state-space hidden Markov model, we consider the problem of sampling from the joint smoothing distribution, i.e. the joint posterior of the unobserved states generating these observations. In recent years, particle MCMC methods, especially in the form of particle Gibbs, have become popular for solving this complicated task. A drawback with standard implementations of particle Gibbs is that each MCMC iteration, generating just a single smoothed state trajectory, requires the observation record to be processed by a full particle filter, leading to significant wastage of computational work. In this talk, we discuss how a recent algorithm, namely the particle-based, rapid incremental smoother (PaRIS), tackles - at least in the case where the Monte Carlo objective function of interest is additive - this shortcoming by means of efficient Rao-Blackwellization.

Spectral causality: Exploring dependence between oscillatory activities in multivariate signals

Monday 24
14:00-16:00
Lecture room 5
(Sali 5)

Hernando Ombao¹, Yuxiao Wang², Xu Gao², Ellen Wann² and Ron Frostig²

¹King Abdullah University of Science and Technology (KAUST) (Saudi Arabia), ²University of California at Irvine (USA)

This work is motivated by the problem of characterizing multi-scale changes in brain signals prior to and after a stroke. A model of occlusive stroke in the rat brain was developed at the Frostig laboratory at UC Irvine. Local field potentials (LFPs), which are physiological signals with good spatial and temporal resolution, were recorded in the experimental rats for about one hour prior to occlusion through many hours post-occlusion. Preliminary analyses show that there are both short-term (immediate) responses to stroke and long term as affected neuronal populations undergo a reorganization in response to an injury. Spectral analysis will be used to study dependence between neuronal populations. Of prime interest here is the notion of “spectral causality which is broadly characterized as the extent to which an oscillatory activity in a population of neurons can predict various oscillatory activities in another region at a future time point. Our approach is to extract different oscillatory components via linear filtering and then examine cross-dependence between the filtered signals. The proposed spectral causality approach overcomes the limitations of classical measures such as coherence and partial coherence since these do not indicate directionality. In addition, the proposed approach is superior to partial directed coherence because it is able to precisely capture the time lag in the between oscillatory activity at different regions. Interesting results from exploratory analyses, showing the immediate changes and long-term brain response, will be reported.

Testing covariate effects in conditional copulas

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

Marek Omelka

Charles University (Czech Republic)

We consider copula modeling of the dependence between two or more random variables in the presence of a multivariate covariate. The dependence parameter of the conditional copula possibly depends on the value of the covariate vector. In this talk we present non-parametric and semiparametric methods for testing the null hypothesis that the covariate does not influence the conditional copula. While the nonparametric test is based on the conditional Kendall’s tau, the semiparametric approach is based on a Rao-type score statistic and it requires choice of a working model for the conditional copula. Nevertheless even this test statistics is constructed so that it holds the level of the test asymptotically even if the working model is misspecified. Moreover, the semiparametric approach opens plenty of new possibilities for testing how the conditional copula depends on the multivariate covariate and also for variable selection in copula model building. Last but not least we compare the suggested methods with the other methods available in the literature.

On predictive density estimation under S-Hellinger distance and alpha-divergence

Younes Ommame

Cadi Ayyad University (Morocco)

Monday 24
14:00-16:00
Auditorium IV

In this paper we consider the investigation of the frequentist risk performance of some predictive density estimators under S-Hellinger distance and α -Divergence as loss functions. Our results relates to the multivariate (d-variate) spherically symmetric observed variable $X \sim p(x|\theta)$, where the mean vector θ is unknown, where $Y \sim q(y|\theta)$ to be estimated based on X . We study the Bayes estimation, Minimum Risk Equivariant (MRE) estimation and minimax estimation. We establish a link between the Alpha-Divergence risk of plug-in type estimators and the risk under reflected normal loss, for point estimation. Furthermore, then we exploit this duality to get improvements upon the benchmark MRE Estimator. We prove the inadmissibility of the MRE estimator when $d \geq 3$, then we suggest dominating estimators with or without the presence of restrictions on the unknown mean parameter.

Coalescent processes and mean minimal spanning trees of irregular graphs

Peter Otto

Willamette University (USA)

Friday 28
10:30-12:30
Lecture room 6
(Sali 6)

It is known that in many cases the cluster dynamics of a random graph process can be replicated with the corresponding coalescent process. The cluster dynamics of a coalescent process (without merger history) is reflected in a auxiliary process called the Marcus-Lushnikov process. The merger dynamics of the Marcus-Lushnikov processes will correspond to a greedy algorithm for finding the minimal spanning tree in the respective random graph process. This observation allows one to express the limiting mean length of a minimal spanning tree in terms of the solutions of the Smoluchowski coagulation equations that represent the hydrodynamic limit of the Marcus-Lushnikov process corresponding to the random graph process. In this talk, I will present our work on breaching the gap between the Smoluchowski coagulation equations for Marcus-Lushnikov processes and the theory of random graphs, concentrating on the case of irregular bipartite graphs and deriving the limiting mean length of minimal spanning trees with random edge lengths for this sequence of graphs.

Empirical likelihood estimation for linear regression models with AR(p) error terms

Şenay Özdemir¹, Yeşim Güney², Yetkin Tuğç² and Olcay Arslan²

¹Afyon Kocatepe University (Turkey), ²Ankara University (Turkey)

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

Linear regression models are widely used statistical methods for analyzing data sets in applications. In linear regression models it is usually assumed that the error terms are uncorrelated with zero mean and constant variance. However, in practice the errors may not be uncorrelated. To deal with the correlated error terms the regression model with AR(p) error terms are proposed and studied (Alpuim and El-Shaarawi, 2008). In both case regression parameters are usually estimated either using the Maximum Likelihood (ML) estimation method or the least square (LS) estimation method. The ML method can only be used under the distribution assumption on the error terms. But, in practice, specifying the distribution of error terms may not be easy for the data sets. Therefore alternative methods should be used. One of these methods is the Empirical Likelihood (EL) method proposed by Owen (1991). In this study, we introduce the Empirical Likelihood (EL) estimation method to estimate the regression and the other parameters of a linear regression model with AR(p) error terms. Some numerical studies will be provided to assess the performance of the proposed estimator.

Bayesian inference on doubly stochastic Markov processes

Suleyman Ozekici¹, Refik Soyer² and Joshua Landon²

¹Koc University (Turkey), ²The George Washington University (USA)

Markov processes are used in many fields of sciences, engineering and management to represent the stochastic evolution of some random phenomena. The probability law of Markov processes is described by its generator or the transition rate matrix. In this paper, we suppose that the process is doubly stochastic in the sense that the generator is also stochastic. In our model, we suppose that the entries in the generator change with respect to the changing states of yet another Markov process which represents the random environment that the stochastic model operates in. This setup also yields a Markov modulated Markov process which can be modelled as a bivariate Markov process that can be analyzed probabilistically using Markovian analysis. In this setting, however, we are interested in Bayesian inference on model parameters. In particular, we focus on Bayesian analysis when the states of the environmental or modulating process are unobserved based on observed data on the modulated Markov process. We present a computationally tractable approach using Gibbs sampling and demonstrate it by a numerical illustration. We also discuss cases that involve complete and partial data sets on both processes.

Nonparametric drift estimation in a Lévy driven diffusion model under microstructure noise

Christian Palmes

TU Dortmund (Germany)

In this talk, a pointwise nonparametric kernel based estimator for the drift function in a Lévy driven jump diffusion model is proposed. Under ergodicity and stationarity of the underlying process X , we derive asymptotic properties as consistency and asymptotic normality of the estimator. In addition, we propose a consistent estimator of the asymptotic variance. Moreover, we show that this approach is robust under microstructure noise by using the pre-averaging approach proposed in Podolskij and Vetter (2006). Finally, a simulation study is given that satisfactorily validates our theoretical results from a numerical finite sample point of view.

Testing for independence of large dimensional vectors

Nestor Parolya¹, Taras Bodnar² and Holger Dette³

¹Leibniz University Hannover (Germany), ²Stockholm University (Sweden), ³Ruhr-University Bochum (Germany)

In this paper new tests for the independence of two high dimensional vectors are investigated. We consider the case where the dimension of the vectors increases with the sample size and propose multivariate analysis of variance-type statistics for the hypothesis of a block diagonal covariance matrix. The asymptotic properties of the new test statistics are studied under the null hypothesis and the alternative using random matrix theory. For this purpose we study the asymptotic properties of linear spectral statistics of central and non-central Fisher matrices. In particular we derive a central limit theorem for linear spectral statistics of large dimension non-central Fisher matrices, which can be used to analyze the power of the tests under the alternative. The theoretical results are illustrated by means of a simulation study, where we also compare the new tests with the commonly used likelihood ratio test. In particular it is demonstrated that the latter test does not keep its nominal level, if the dimension of one vector is relatively small compared to the dimension of the other vector. On the other hand the tests proposed in this paper provide a reasonable approximation of the nominal level in such situations.

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

Monday 24
10:30-12:30
Lecture room 7
(Sali 7)

Thursday 27
14:00-16:00
Lecture room 10
(Sali 10)

Asymptotic distribution of the isotonic regression estimator over a countable pre-ordered set

Vladimir Pastukhov and Dragi Anevski
Lund University (Sweden)

Tuesday 25
14:00-16:00
Lecture room 8
(Sali 8)

We derive the asymptotic distribution of the isotonic regression estimator over a countable pre-ordered set and analyse its limiting distribution. We consider separately the case of functions on a finite set and the case of square summable functions on a countable pre-ordered set. We show that under some general conditions on the original sequence of functions the limiting distribution of the sequence of isotonic regressions of the functions is given by the concatenation of the isotonized restrictions of the distributional limit of the original sequence to the comparable level sets. As an example we consider the problems of estimation of a bimonotone probability mass function and of a bimonotone regression function.

A general approach for cure models in survival analysis

Valentin Patilea¹ and Ingrid Van Keilegom²
¹Ensai (France), ²KU Leuven (Belgium)

Monday 24
10:30-12:30
Great hall
(Juhlasali)

Cure models are a special topic in the time-to-event analysis in the presence of random censoring, and concern situations where a proportion of subjects will never experience the event under study. There is quite a large biostatistical literature that considered this type of model. The same models appear in other fields, as for instance in economics and econometrics where are used under the name of split-population models. The usual way that cure models have been constructed so far in the literature is to start from a lifetime of interest T and a censoring variable C . The lifetime T could be infinity with non-zero probability and this probability is modeled by some convenient model, for instance parametric. We propose a different approach where we explicitly exploit the map that links the distribution of the latent variables T and C and the latent cure proportion, to the law of the observations. In this way we provide general purpose inverse mappings that allow expressing the latent quantities as closed form functions of the distributions of the observables. As a consequence, using a convenient estimator of the (conditional) law of the observables, we could derive a likelihood criterion to estimate the conditional cure probability and the (conditional) law of the non-cured event times. Our framework provides more insight on some existing approaches in the literature, namely on the underlying constraints they implicitly suppose. We provide the asymptotic analysis of our likelihood estimator and investigate its finite sample properties using simulated and real data.

Analysis of optimization based high dimensional inference based on concentration inequalities

Daniel Paulin¹, Ajay Jasra¹, Dan Crisan² and Alexandros Beskos³
¹National University of Singapore (Singapore), ²Imperial College London (UK), ³University College London (UK)

Tuesday 25
10:30-12:30
Lecture room 13
(Sali 13)

In this talk, we consider the problem of inferring the current or the initial position of high dimensional chaotic dynamical systems, based on partial observations with i.i.d. noise. This is an important problem in weather forecasting. We propose a simple inference method which starts from an appropriately chosen initial position, and proceeds towards the maximum-a-posteriori of the likelihood function based on Newton's method. We show that concentration inequalities for empirical processes are useful for analysing such methods, because they can precisely control the likelihood function and its derivatives in the whole state space.

CsCsHM-statistic for detecting rare and weak effects in high-dimensional classification.

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

Tatjana Pavlenko¹, Natalia Stepanova² and Oskar Stattin¹

¹KTH Royal Institute of Technology (Sweden), ²Carleton University (Canada)

Recently, a new class of goodness-of-fit test statistics, called for CsCsHM-statistic, based on sup-functionals of weighted empirical processes is proposed and studied (see [1]), where weight functions used are Erdős-Feller-Kolmogorov-Petrovski upper class functions of a Brownian bridge. A number of optimal theoretical properties of CsCsHM-statistic is established in a rare and weak asymptotic framework. Specifically, the proposed test is shown to achieve the detection boundary found by Ingster [2] and, when distinguishing between the null and alternative hypotheses, perform optimally adaptively to unknown sparsity and size of the non-null effects. This talk explores properties of CsCsHM-statistic for the problem of detecting sparse heterogeneous mixtures in high-dimensional settings where the dimensionality greatly exceeds the sample size. Examples of such type of problem includes feature selection in a high-dimensional classification models when informative features are rare and weak. We derive the feature selection procedure by CsCsHM-based thresholding and show that it achieves the optimal detection boundary in a variety of high-dimensional, rare and weak classification settings. Performance accuracy of the classifier with the proposed CsCsHM-based thresholding is demonstrated for a number of simulated scenarios and real data sets. This talk is based on the joint work with Natalia Stepanova and Oskar Stattin.

References: [1] Stepanova, N., Pavlenko, T., 2017. Goodness-of-fit tests based on sup-functionals of weighted empirical processes. *SIAM Journal on Theory of Probability and Its Applications (TVP)*, *accepted*. <https://arxiv.org/abs/1406.0526>. [2] Ingster, Y.I., 1997. Some problems of hypothesis testing leading to infinitely divisible distribution. *Math. Methods Statist.*, 6, 47–69.

Adaptive improved estimation in heteroscedastic nonparametric regression models

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

Evgenii Pchelintsev

Tomsk State University (Russia)

We consider the problem of estimating an unknown function in heteroscedastic regression with a conditionally-Gaussian noise. An adaptive model selection procedure, based on the improved weighted least square estimates, is proposed. It is established that the procedure has higher mean square accuracy in comparison with the procedure based on the classical weighted least squares estimates. Under some conditions on the noise distribution, sharp non-asymptotic oracle inequalities for the quadratic risks are proved and the efficiency of the model selection procedure is shown. Results of numerical simulations are given.

A Bayesian nonparametric model for finding novel pancreatic cancer subtypes

Poster session

Iliana Peneva¹, Richard S. Savage¹, Keith Roberts² and Felicity Evison²

¹University of Warwick (UK), ²University Hospitals Birmingham (UK)

Pancreatic cancer is one of the most common causes of cancer-related death in the developed world and is projected to become the second most common in the Western societies by 2030. In this work, we develop a novel Bayesian nonparametric model for clustering of mixed data types called BayesCluster which aims to identify novel subtypes of pancreatic cancer. It uses a latent variable model that models the observed mixed data as the realisation of latent variables which follow a mixture of Gaussian distributions. We incorporate feature learning in the model because of the large number of predictors, and use a Dirichlet Process mixture modelling approach and MCMC sampler to produce the patient subgroups. We apply the model to the datasets provided from Queen Elizabeth Hospital Birmingham. The proposed method has several extensions: it can

be applied to richer biological structures by using sparsity priors and the kernel trick. It can also be extended to integrate and use the information from multiple different datasets. By combining all available information we are able to better identify the different subtypes of pancreatic cancer, opening the way for new personalised treatments.

Utilizing background knowledge for the estimation of total causal effects

Emilija Perkovic, Markus Kalisch and Marloes Maathuis
ETH Zurich (Switzerland)

Tuesday 25
10:30-12:30
Lecture room 1
(Sali 1)

Background knowledge can be used to improve the output of causal discovery algorithms. However, most current causal methods cannot be directly applied to the resulting graphical output. We develop methodology for estimating total causal effects from maximally oriented PDAGs, which are the graphs that arise in the presence of background knowledge when assuming causal sufficiency. Specifically, we adapt the covariate adjustment, as well as the IDA and joint-IDA frameworks.

Bayesian cluster analysis of categorical data with supervised feature selection

Alberto Pessia¹ and Jukka Corander²
¹University of Helsinki (Finland), ²University of Oslo (Norway)

Thursday 27
10:30-12:30
Lecture room 12
(Sali 12)

Cluster analysis is a common statistical technique for partitioning the observed data into disjoint homogeneous groups. In the presence of multivariate data, it is often useful to identify which features are best predictors of cluster association. The problem is formalized as a bidirectional Bayesian cluster analysis, both in the units space and the features space. The aim is obviously to perform a clustering of the observed sample, but also to classify the variables according to pre-specified levels of discrimination power. Split-merge and Gibbs sampler type MCMC algorithms are employed to simultaneously traverse the posterior of partitions of samples and variables. We show how the model can be successfully utilized for clustering genetic data and highlighting sites under selective pressure. Software implementation for clustering categorical data matrices is freely available at <https://github.com/albertopessia/Kpax3.jl>

Estimating a linear parametric function of a doubly censored exponential distribution

Constantinos Petropoulos¹, Yogesh Mani Tripathi², Farha Sultana² and Manoj Kumar Rastogi³
¹University of Patras (Greece), ²Indian Institute of Technology (India), ³National Institute of Pharmaceutical Education and Research (India)

Thursday 27
14:00-16:00
Lecture room 12
(Sali 12)

For an arbitrary strictly convex loss function, we study the problem of estimating a linear parametric function $\mu + \kappa\sigma$, κ is a known constant, when a doubly censored sample is available from a two-parameter exponential $E(\mu, \sigma)$ population. We establish the inadmissibility of the best affine equivariant estimator by deriving an improved estimator. We provide various implications for quadratic and linex loss functions in details. Improvements are obtained for the absolute value loss function as well. Further a new class of estimators improving upon the best affine equivariant estimator is derived using the Kubokawa method. This class is shown to include some benchmark estimators from the literature. Finally, a conclusion is presented.

Keywords: Brewster-Zidek Estimator, Censored Samples, Inadmissibility, Equivariant Estimator.

Inference for intractable-likelihoods using a likelihood-free version of SAEM

Umberto Picchini
Lund University (Sweden)

We present an approximate maximum likelihood methodology for the parameters of incomplete data models. A likelihood-free version of the stochastic approximation expectation-maximization (SAEM) algorithm is constructed to maximize the likelihood function of model parameters. While SAEM is best suited for models having a tractable "complete likelihood" function, its application to moderately complex models is difficult, and results impossible for models having so-called intractable likelihoods. The latter are typically treated using approximate Bayesian computation (ABC) algorithms or synthetic likelihoods, where information from the data is carried by a set of summary statistics. While ABC is considered the state-of-art methodology for intractable likelihoods, its algorithms are often difficult to tune. On the other hand, synthetic likelihoods (SL) is a more recent methodology which is less general than ABC, it requires stronger assumptions but also less tuning. By exploiting the Gaussian assumption set by SL on data summaries, we can construct a likelihood-free version of SAEM. Our method is completely plug-and-play and available for both static and dynamic models, the ability to simulate realizations from the model being the only requirement. We present simulation studies from our on-going work.

Tuesday 25
10:30-12:30
Lecture room 7
(Sali 7)

Inference via Bayesian synthetic likelihoods for a mixed-effects SDE model of tumor growth

Umberto Picchini
Lund University (Sweden)

We consider parameter estimation for a state-space model fitted on repeated measurements of tumour data from several mice (hereafter "subjects"). We construct a nonlinear mixed-effects model where individual dynamics are expressed by stochastic differential equations (SDE). Non-linear SDE state-space models are notoriously difficult to fit: however, advancements in parameter estimation using sequential Monte Carlo methods have made (exact) Bayesian inference for SDE models approachable, if not always straightforward. Instead, here we use summary statistics to encode information pertaining between-subjects variation, as well as individual variation, and resort to a synthetic likelihood approximation for parameter estimation. In particular, we take advantage of a Bayesian (pseudo-marginal) synthetic likelihood approach. We consider longitudinal data from measurements of tumour volumes in mice. Mice are divided in three groups and administered three different treatments. The dataset is challenging because of sparse observations. Synthetic likelihoods based on summary statistics are able to reproduce the results from exact Bayesian inference and, when data are available from a non-negligible number of subjects, are able to identify a specific treatment to be more effective in reducing tumour growth.

Wednesday 26
10:30-12:30
Lecture room 10
(Sali 10)

L-moment estimation for linear regression models

Jan Picek and Martin Schindler
Technical University of Liberec (Czech Republic)

The L-moments are analogues of the conventional moments and have similar interpretations. They are calculated using linear combinations of the expectation of ordered data. It was recently shown that L-moment is a special case of L-estimator. We propose a generalization of L-moments in the linear regression model based on the averaged regression quantiles as special L-estimator. The properties of extended L-moment are illustrated on simulated data.

Semi-parametric, parametric and possibly sparse models for multivariate long-range dependence

Vladas Pipiras¹, Stefanos Kechagias² and Changryong Baek³

¹University of North Carolina (USA), ²SAS Institute (USA), ³Sungkyunkwan University (Korea)

The focus of this talk is on multivariate (vector-valued) time series that exhibit long-range dependence (LRD) and, more specifically, on (semi-)parametric models that account for general phase parameters in the cross spectra of the series at the zero frequency. Several new multivariate LRD time series models are introduced and their estimation is discussed, possibly assuming sparsity of model parameters. Applications to several real time series are also presented.

Thursday 27
10:30-12:30
Lecture room 1
(Sali 1)

Zooming on edges of ℓ_1 penalized graphical models estimated for different coarseness scales

Eugen Pircalabelu

KU Leuven (Belgium)

We estimate graphical models from data obtained at K different coarseness scales. Starting from a predefined scale $k^* \leq K$ the method zooms in or out over scales on particular edges, thus estimating graphs with similar structures, but different levels of sparsity. The graphs are jointly estimated at all coarseness scales and evaluate the evolution of the graphs from the coarsest to the finest scale or vice-versa. We select an optimal coarseness scale to be used for further analyses. The method is motivated by fMRI datasets that do not all contain measurements on the same set of brain regions. For certain datasets some of the regions have been split in smaller subregions and this gives rise to the framework of mixed scale measurements where the purpose is to estimate sparse graphical models. We accomplish this by pooling information from all subjects in order to estimate a common undirected and directed graph at each coarseness scale, accounting for time dependencies and multiple coarseness scales and by jointly estimating the graphs at all coarseness scales. The applicability of the method goes beyond fMRI data, to other areas where data on different scales are observed and where the joint estimation of graphs is desired.

Thursday 27
14:00-16:00
Lecture room 6
(Sali 6)

Sojourn time distributions effects on a redundant multi-state deteriorating system with maintenance

Agapios Platis

University of the Aegean (Greece)

Most of the contemporary technological systems are operating under multiple deterioration stages and have complicated structures mainly due to demanding industrial applications. The performance and the availability of these systems are of prior importance since their deterioration and/or failure may lead to important losses. In order to improve the operation of such systems and increase their availability, redundancy is usually implemented. For further improvement, maintenance actions can be adopted as well. However, although preventive maintenance improves the performance of a system, it incurs cost. Thus, a lot of research effort has been paid in finding an appropriate maintenance policy that manages to reduce the total operational cost, as a performance measure, and improve the availability of such a system system. In this paper, our work focuses on examining the effects of the inspection, maintenance and repair time distributions on the asymptotic behavior, in terms of availability and performance, and on the optimal maintenance policy for a two-unit multi-state deteriorating system with preventive maintenance. The system consists of two identical units where one unit is operational and the other one is in cold standby mode. The system is regularly inspected and depending on its condition, either no action takes place or maintenance is carried out, either minimal or major. The proposed model takes also into account the scenario of imperfect and failed maintenance.

Tuesday 25
14:00-16:00
Lecture room 6
(Sali 6)

Tuesday 25
10:30-12:30
Lecture room 8
(Sali 8)

Improving Lasso for generalized linear model selection

Piotr Pokarowski¹, Agnieszka Prochenka¹, Michał Frej¹ and Jan Mielniczuk²

¹University of Warsaw (Poland), ²Institute of Computer Science, Polish Academy of Sciences and
Warsaw University of Technology (Poland)

The Lasso, that is l_1 -penalized loss estimator is a very popular tool for fitting sparse models to high-dimensional data. However, theoretical studies and simulations established that the model selected by the Lasso is usually too large. The concave regularizations (SCAD, MCP or capped- l_1) are closer to l_0 -penalized loss that is to the Generalized Information Criterion (GIC) than the Lasso and correct its intrinsic estimation bias. That methods use the Lasso as a starting set of models and try to improve it using local optimization. We propose an alternative method of improving the Lasso for Generalized Linear Models which is a generalization of our SOS algorithm for linear models (Pokarowski and Mielniczuk, J. Mach. Learn. Res, 2015). The method, for a given penalty, orders the absolute values of the Lasso non-zero coordinates and then selects the model from a small family by GIC. We derive an upper bound on the selection error of the algorithm and show in numerical experiments on synthetic and real-world data sets that an implementation our algorithm is more accurate than implementations of concave regularizations (sparsenet, cvplogistic).

Wednesday 26
10:30-12:30
Lecture room 1
(Sali 1)

Exact Bayesian inference for big data: Single- and multi-core approaches

Murray Pollock¹, Hongsheng Dai², Paul Fearnhead³, Adam Johansen¹, Divakar Kumar¹ and
Gareth Roberts¹

¹University of Warwick (UK), ²University of Essex (UK), ³University of Lancaster (UK)

This talk will introduce novel methodologies for exploring posterior distributions by modifying methodology for exactly (without error) simulating diffusion sample paths.

The methodologies discussed have found particular applicability to “Big Data” problems. We begin by presenting the Scalable Langevin Exact Algorithm (ScaLE) and recent methodological extensions (including Re-ScaLE, which avoids the need for particle approximation in ScaLE), which has remarkably good scalability properties as the size of the data set increases (it has sub-linear cost, and potentially no cost as a function of data size). ScaLE has particular applicability in the “single-core” big data setting - in which inference is conducted on a single computer. In the second half of the talk we will present methodology to exactly recombine inferences on separate data sets computed on separate cores - an exact version of “divide and conquer”. As such this approach has particular applicability in the “multi-core” big data setting.

We conclude by commenting on future work on the confluence of these approaches.

Monday 24
14:00-16:00
Lecture room 7
(Sali 7)

Sequential monitoring in panel data

Zuzana Praskova

Charles University (Czech Republic)

A sequential monitoring scheme is proposed to detect instability of parameters in a dynamic panel data model when a training sample of stable observations is available. A detector statistic is based on cumulative processes of weighted LSE residuals computed separately in each panel that are assumed to be mutually independent (though the errors in separate panels can be dependent), and on a high-dimensional aggregation of such cumulative processes. The asymptotic distribution of the detector statistic is established under both the null hypothesis of no change and the alternative that a change occurs in case that both the number of panels and the number of training observations are sufficiently large.

Some novel spherical regression models

Simon Preston¹, Phillip Paine, Michail Tsagris and Andy Wood

¹University of Nottingham (UK)

Wednesday 26
10:30-12:30
Lecture room 5
(Sali 5)

Existing parametric models for spherical regression, i.e. where the response variable lies on the unit sphere S^2 , are quite restrictive, typically assuming some simple structure for the covariate (e.g. the covariate is in R^1 or S^2) and/or that the errors are isotropic. Such assumptions are not reasonable for many applications. I will discuss some new and more flexible parametric models in which the covariates can have quite general structure (for example, they may be in S^p , R^p , categorical, or some combination of these) and in which the error distribution is anisotropic. We consider two anisotropic error distributions - the Kent distribution and the elliptically symmetric angular Gaussian distribution - and two parameterisations of each which enable distinct ways to model how the response depends on the covariates. Within this set-up, it is particularly simple to test the significance of particular covariates, or of the anisotropy of the errors. I will show some applications for which such tests indicate strong evidence to favour these novel models over simpler existing ones.

A Durbin-Levinson regularized estimator of high dimensional autocovariance matrices

Tommaso Proietti and Alessandro Giovannelli

University of Rome Tor Vergata (Italy)

Thursday 27
14:00-16:00
Lecture room 10
(Sali 10)

We consider the problem of estimating the high-dimensional autocovariance matrix of a stationary random process, with the purpose of sample prediction and feature extraction. This problem has received several solutions. In the nonparametric framework, the literature has concentrated on banding and tapering the sample autocovariance matrix. We propose and evaluate an alternative approach based on regularizing the sample partial autocorrelation function, via a modified Durbin-Levinson algorithm that receives as input the banded and tapered partial autocorrelations and returns a sample autocovariance sequence which is positive definite. We show that the regularized estimator of the autocovariance matrix is consistent and its convergence rate is established. We then focus on constructing the optimal linear predictor and we assess its properties. The computational complexity of the estimator is of the order of the square of the banding parameter, which renders its feasible for high-dimensional time series. We evaluate the performance of the autocovariance estimator and the corresponding linear predictor by simulation and empirical applications.

Keywords: Toeplitz systems; Optimal linear prediction; Partial autocorrelation function.

Multiscale scanning in inverse problems

Katharina Proksch

University of Göttingen (Germany)

Friday 28
10:30-12:30
Lecture room 10
(Sali 10)

In this talk we propose a multiscale scanning method to determine active components of a quantity f w.r.t. a dictionary \mathcal{U} from observations Y in an inverse regression model $Y = Tf + \xi$ with operator T and general random error ξ . To this end, we provide uniform confidence statements for the coefficients $\langle \varphi, f \rangle$, $\varphi \in \mathcal{U}$, under the assumption that $(T^*)^{-1}(\mathcal{U})$ is of wavelet-type. Based on this we obtain a decision rule that allows to identify the active components of \mathcal{U} , i.e. $\langle f, \varphi \rangle \neq 0$, $\varphi \in \mathcal{U}$, at controlled, family-wise error rate. Our results rely on a Gaussian approximation of the underlying multiscale statistic with a novel scale penalty adapted to the ill-posedness of the problem. The important special case of deconvolution is discussed in detail. Further, the pure regression case, when $T = \text{id}$ and the dictionary consists of moving windows of various sizes (scales), is included, generalizing previous results for this setting. Simulations support our theory and we illustrate the potential of the method as an inferential tool for imaging. As a particular application we discuss super-resolution microscopy and analyze experimental STED data to locate single DNA origami.

Thursday 27
14:00-16:00
Lecture room 1
(Sali 1)

Weak convergence of empirical copula processes indexed by functions

Dragan Radulovic
Florida Atlantic University (USA)

Weak convergence of the empirical copula process indexed by a class of functions is established. Two scenarios are considered in which either some smoothness of these functions or smoothness of the underlying copula function is required. A novel integration by parts formula for multivariate, right continuous functions of bounded variation, which is perhaps of independent interest, is proved. It is a key ingredient in proving weak convergence of a general empirical process indexed by functions of bounded variation.

Thursday 27
10:30-12:30
Great hall
(Juhlasali)

Network comparison

Gesine Reinert¹, C. Deane, R. Gaunt, L. Ospina, A. Wegner and Xiaochuan Xu
¹University of Oxford (UK)

Many complex systems can be represented as networks, and the problem of network comparison is becoming increasingly relevant. There are many such techniques that vary from simply comparing network summary statistics to computationally costly alignment-based approaches. The challenge remains to correctly cluster networks that are of a different size and density, yet hypothesized to be structurally similar.

In this talk we review existing methods for network comparison and introduce a new network comparison methodology that is aimed at identifying common organizational principles in networks; these networks could be undirected or directed. The methodology is simple and intuitive and outperforms existing methods in a variety of settings ranging from classification of chemical compounds to tracking the evolution of networks representing the topology of the internet.

Wednesday 26
10:30-12:30
Lecture room 8
(Sali 8)

Dependence in multivariate distributions

Wolf-Dieter Richter
University of Rostock (Germany)

One of the nowadays most popular method to model dependence in multivariate probability laws is using copulas. On the other hand, it is well known that the components of a continuous axes-aligned elliptically contoured distributed random vector are independent if and only if its density generating function (dgf) is that of the Gaussian law. Dependence among the vector's components is thus due to deviations of the dgf from the Gaussian one. More generally, a continuous random vector following the axes-aligned multivariate p -generalized elliptically contoured distribution law has dependent components if the dgf is any but that of the p -power exponential law. It is further well known from p -generalized elliptically contoured distribution theory that correlation type dependence is caused by rotating axes-aligned heteroskedastic vectors. In this talk, relations between various approaches to modeling dependence are discussed.

References: [1] Dietrich, T., Kalke, S., Richter, W.-D., 2013. Stochastic representations and geometric parametrization of the two-dimensional Gaussian law. *Chilean Journal of Statistics*, 4 (2), 27–59. [2] Müller, K., Richter, W.-D., 2016. Extreme value distributions for dependent jointly $l_{n,p}$ -symmetrically distributed random variables. *Depend. Model.* 4, 30–62. [3] Richter, W.-D., 2014. Geometric disintegration and star-shaped distributions. *Journal of Statistical Distributions and Applications*, 1:20.

Combining Monte Carlo estimates

Christian Robert
Université Paris-Dauphine (France)

Wednesday 26
10:30-12:30
Lecture room 1
(Sali 1)

When considering a Monte Carlo approximation to an unknown integral, there usually exist several representations of the integral that can be exploited for several or even a single simulation output. While some combinations are based on standard variances considerations, like the AMIS algorithm of Cornuet et al. (2012), more non-parametric and machine-learning alternatives can produce additional evaluations that include an assessment of the uncertainty about the estimates.

Low-rank interaction contingency tables

Genevieve Robin¹, Julie Josse¹, Éric Moulines¹ and Sylvain Sardy²
¹École Polytechnique (France), ²Université de Genève (Switzerland)

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

Log-linear models are popular tools to analyze contingency tables, particularly to model row and column effects as well as row-column interactions in two-way tables. In this paper, we introduce a regularized log-linear model designed for denoising and visualizing count data, which can incorporate side information such as row and column features. The estimation is performed through a convex optimization problem where we minimize a negative Poisson log-likelihood penalized by the nuclear norm of the interaction matrix. We derive an upper bound on the Frobenius estimation error, which improves previous rates for Poisson matrix recovery, and an algorithm based on the alternating direction method of multipliers to compute our estimator. To propose a complete methodology to users, we also address automatic selection of the regularization parameter. A Monte Carlo simulation reveals that our estimator is particularly well suited to estimate the rank of the interaction in low signal to noise ratio regimes. We illustrate with two data analyses that the results can be easily interpreted through biplot visualization. The method is available as an R code.

Variational inference for probabilistic Poisson PCA

Stéphane Robin¹, Julien Chiquet¹ and Mahendra Mariadassou²
¹INRA / AgroParisTech (France), ²INRA / Jouy-en-Josas (France)

Friday 28
10:30-12:30
Lecture room 7
(Sali 7)

Many application domains such as ecology or genomics have to deal with multivariate non Gaussian observations. A typical example is the joint observation of the respective abundances of a set of species in a series of sites, aiming to understand the co-variations between these species. The Gaussian setting provides a canonical way to model such dependencies, but does not apply in general. We consider here the multivariate exponential family framework for which we introduce a generic model with multivariate Gaussian latent variables. We show that approximate maximum likelihood inference can be achieved via a variational algorithm for which gradient descent easily applies. We show that this setting enables us to account for covariates and offsets. We then focus on the case of the Poisson-lognormal model in the context of community ecology.

Reference: Chiquet J., Mariadassou, M., Robin, S., 2017. Variational inference for probabilistic Poisson PCA. arXiv:1703.06633

Monday 24
10:30-12:30
Lecture room 13
(Sali 13)

Inference on trending panel data

Peter M. Robinson¹ and Carlos Velasco

¹London School of Economics (UK)

Semiparametric panel data modelling and statistical inference with fractional stochastic trends, nonparametrically time-trending individual effects, and general spatial or cross-sectional correlation and heteroscedasticity in innovations are developed. The fractional stochastic trends allow for a wide range of nonstationarity, indexed by a memory parameter, nesting the familiar I(1) case and allowing for parametric short-memory. The individual effects can nonparametrically vary simultaneously across time and across units. The spatial or cross-sectional covariance matrix is also nonparametric. The main focus is on estimation of the time series parameters. Two methods are considered, both of which entail an only approximate differencing out of the individual effects, leaving an error which has to be taken account of in our theory. In both cases we obtain standard asymptotics, with a central limit theorem, over a wide range of possible parameter values, unlike the nonstandard asymptotics for autoregressive parameter estimates at a unit root. For statistical inference, consistent estimation of the limiting covariance matrix of the parameter estimates requires consistent estimation of a functional of the cross-sectional covariance matrix. We examine efficiency loss due to spatial correlation. A Monte Carlo study of finite-sample performance is included.

Monday 24
14:00-16:00
Lecture room 13
(Sali 13)

Spike-and-Slab LASSO

Veronika Rockova

University of Chicago (USA)

Despite the wide adoption of spike-and-slab methodology for Bayesian variable selection, its potential for penalized likelihood estimation has largely been overlooked. In this paper, we bridge this gap by cross-fertilizing these two paradigms with the Spike-and-Slab LASSO procedure for variable selection and parameter estimation in linear regression. We introduce a new class of self-adaptive penalty functions that arise from a fully Bayes spike-and-slab formulation, ultimately moving beyond the separable penalty framework. A virtue of these non-separable penalties is their ability to borrow strength across coordinates, adapt to ensemble sparsity information and exert multiplicity adjustment. The Spike-and-Slab LASSO procedure harvests efficient coordinate-wise implementations with a path-following scheme for dynamic posterior exploration. We show on simulated data that the fully Bayes penalty mimics oracle performance, providing a viable alternative to cross-validation. We develop theory for the separable and non-separable variants of the penalty, showing rate-optimality of the global mode as well as optimal posterior concentration when $p > n$.

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

A non-homogeneous Markov chain model with seasonal transition probabilities applied to ozone data

Eliane Rodrigues¹, Taramoto Mario² and Tzintzun Guadalupe³

¹Universidad Nacional Autonoma de Mexico (Mexico), ²Universidade Estadual Paulista Julio de Mesquita Filho (Brazil), ³Instituto Nacional de Ecologia y Cambio Climatico (Mexico)

It is a well know fact that inhabitants of large cities around the world suffer with high levels of air pollution. In the present work we consider the problem of estimating the probability that a population is exposed to high levels of ozone in a given day into the future based on the present and past information about the pollutant's level. In order to do that we assume that the sequence recording whether or not an exceedance of an environmental threshold has occurred in a given day is ruled by a non-homogeneous Markov chain of order one. In order to account for the many cycles that might appear in the empirical transition probabilities, a parametric form incorporating seasonal components is considered for them. The parameters are estimated using the Bayesian point of view via Markov chain Monte Carlo (MCMC) method. The model is applied to the ozone data obtained from the monitoring network of Mexico City.

Comparison of the predictives values of two diagnostic tests subject to a retrospective study

Poster session

Jose Antonio Roldán-Nofuentes
University of Granada (Spain)

Diagnostic tests are fundamental in modern Medicine. A diagnostic test is one which is applied to an individual to determine the presence or absence of a specific disease. When the result of a diagnostic test is positive or negative, this is called a binary test, and its accuracy is measured in terms of two probabilities: sensitivity and specificity. Other fundamental parameters of a binary diagnostic test are the predictive values. The (positive and negative) predictive values represent the clinical accuracy of the test, and they depend on the sensitivity and specificity of the diagnostic test and on the disease prevalence. The comparison of the predictive values of two binary diagnostic tests is a topic that has been the subject of different studies in the field of Statistics. In this article, we propose a hypothesis test to compare the predictive values of two binary diagnostic tests subject to retrospective sampling, assuming for this purpose that there is an estimation of the disease prevalence. The hypothesis test is based on the chi-squared distribution. Simulation experiments were carried out to study the type I error and the power of the hypothesis test studied. The method proposed was applied to a real example.

Averaging symmetric positive-definite matrices via the minimal scaling-rotation framework

Brian Thomas Rooks

Wednesday 26
10:30-12:30
Lecture room 5
(Sali 5)

In this talk we present a new method for averaging in the space of symmetric positive-definite (SPD) matrices that utilizes the minimal scaling-rotation (MSR) geometric framework introduced in Jung *et al.* (2015). The MSR sample mean set is defined as the set of sample Frchet means with respect to MSR distance. We will outline an algorithm for computing candidates for the MSR sample mean, and discuss conditions guaranteeing uniqueness of solutions, strong consistency to a population mean set, and a Central Limit Theorem type result, respectively. The talk will conclude with applications of this averaging framework in the analysis of diffusion tensor imaging (DTI) data.

On the optimality of averaging in distributed statistical learning

Jonathan Rosenblatt¹ and Nadler Boaz²

¹Ben Gurion University of the Negev (Israel), ²Weizmann Institute of Science (Israel)

Friday 28
10:30-12:30
Great hall
(Juhlasali)

A common approach to statistical learning on big data is to randomly split it among m machines and calculate the parameter of interest by averaging their m individual estimates. Focusing on empirical risk minimization, or equivalently M-estimation, we study the statistical error incurred by this strategy. We consider two asymptotic settings: one where the number of samples per machine $n \rightarrow \infty$ but the number of parameters p is fixed, and a second high-dimensional regime where both $p, n \rightarrow \infty$ with $p/n \rightarrow \kappa$. Most previous works provided only moment bounds on the error incurred by splitting the data in the fixed p setting. In contrast, we present for both regimes asymptotically exact distributions for this estimation error. In the fixed- p setting, under suitable assumptions, we thus prove that to leading order, averaging is as accurate as the centralized solution. In the high-dimensional setting, we show a qualitatively different behavior: data splitting does incur a first order accuracy loss, which we quantify precisely. In addition, our asymptotic distributions allow the construction of confidence intervals and hypothesis testing on the estimated parameters. Our main conclusion is that in both regimes, averaging parallelized estimates is an attractive way to speed up computations and save on memory, while incurring a quantifiable and typically moderate excess error.

Interpretation of the discriminative ability of a time-dependent marker in survival models

Anja Rueten-Budde¹, Hein Putter² and Marta Fiocco²

¹Leiden University (Netherlands), ²Leiden University Medical Center (Netherlands)

Harrells C-index (Harrell et al., 1996) is a commonly used measure to assess the discriminative ability of a survival model with one end point. It is defined as the proportion of pairs who experience the event of interest in the order predicted by the prognostic model. The C-index can also be defined based on a weighted average of the area under the receiver operating characteristic curve (AUC) over time (Van Houwelingen and Putter, 2012). Saha and Heagerty (2010) extend the idea of time-dependent AUC to competing risks and models with time-dependent covariates, based on the concept of cumulative/dynamic and incident/dynamic cases and controls. Time-dependent covariates are introduced in the incident/dynamic case, which predicts the accuracy of the marker to distinguish between the subject who experiences the event of interest at time t and those who do not experience an event by that time. To investigate the effect of a binary time-dependent covariate on the AUC over time we simulate from a Markov multi-state illness death model with Weibull distributed hazard functions. Different scenarios illustrate the effect the transition hazards have on the AUC over time; the estimated AUC are compared to the theoretical values. Results show that higher transitions hazards for the illness-death transition reflect back in the AUC, where low values are associated with low AUC values and high values with higher AUC values. The simulation study shows that the estimated AUC agree with the theoretically derived AUC.

References: [1] Harrell, F. E., Lee, K. L., Mark, D. B., 1996. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* 15, 361–387. [2] Saha, P., Heagerty, P. J., 2010. Time-dependent predictive accuracy in the presence of competing risks. *Biometrics*, 66(4), 999–1011. [3] Van Houwelingen, H., Putter, H., 2012. Measuring the predictive value of a cox model. *Dynamic Prediction in Clinical Survival Analysis*, 35–56. Boca Raton, Florida: CRC Press/Chapman & Hall.

A Markovian arrival process with marked transitions to model a complex system with loss of units

Juan Eloy Ruiz-Castro and Mohammed Dawabsha

University of Granada (Spain)

A cold standby multi-state system subject to different types of failures and preventive maintenance is modeled by considering a Markovian arrival process with marked transitions. The system is composed of a general number of multi-states units. Several internal and external degradation states are considered which are observed when a random inspection occurs. If a major degradation level is observed by inspection, the unit goes to preventive maintenance. Also, the main unit is subject to internal repairable failure and external shocks. If one internal repairable failure occurs, the unit goes to the repair facility for corrective repair, and when one external shock happens, this one may produce an aggravation of the internal degradation level, cumulative external damage or external extreme failure (non-repairable failure). When a non-repairable failure occurs the unit is removed and the system continues working. When the last unit undergoes a non-repairable failure then a full new system is included. The system is modeled through matrix blocks and matrix-geometric methods are used to work out the stationary distribution. Relevant performance measures are obtained in both regimes, transient and stationary. All results are expressed in algorithmic and computational form and they have been implemented computationally with Matlab.

Acknowledgements. This paper is partially supported by the Junta de Andalucía, Spain, under the grant FQM-307 and by the Ministerio de Economía y Competitividad, España, under Grant MTM2013-47929-P.

Variable selection in random survival forests

Helene Charlotte Rytgaard and Thomas Alexander Gerds
University of Copenhagen (Denmark)

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

A random survival forest is a machine learning approach which is suitable for right censored time-to-event outcome. One of the main target parameters is variable importance. In this talk I will discuss how variations of the splitting rule which is used to grow the trees of a random survival forest affects the results of variable importance measures. One interesting aspect occurs in applications of survival analysis with competing risks where variables can have an effect on the prediction of event risk also indirectly through the hazard function of the other events. Another aspect of the discussion is about the weights given to events occurring at different time points in the follow-up period. The log-rank test is the standard splitting rule for survival trees, having highest power to detect proportional hazard effects. By variation across Flaming-Harrington's (G^p) family of test statistics, variables with short term effects on survival can be distinguished from variables with late or long term effects. I will illustrate these ideas in a Danish national registry wide association study. Here, we search through thousands of drugs which may prevent onset of depression.

Using SA and VNS procedures for estimating the Double-Hubbart diffusion process

Istoni Sant'Ana^{1,2}, Francisco Torres-Ruiz² and Patricia Román-Román²

¹University of Puerto Rico (Puerto Rico), ²University of Granada (Spain)

Poster session

A problem of great current interest is how to accurately chart the progress of oil production. It is well known that oil exploration is cyclical and that after the oil production reaches its peak in a specific system, a fatal decline will begin. In this context, M.H. Hubbert developed his peak theory in 1956 based on a bell-shaped curve that bears his name. In this work, we consider a stochastic model, based on the theory of diffusion processes, associated with the Hubbert curve. The problem of the maximum likelihood estimation of the parameters for this process is also considered. Since a complex system of equation appears, whose solution cannot be guaranteed via the classical numerical procedures, we suggest the use of metaheuristic optimization algorithms such as Simulated Annealing and Variable Neighborhood Search. Some strategies are suggested for bounding the space of solutions, and a description is provided for the application of the algorithms selected. In the case of the Variable Neighborhood Search algorithm, a hybrid method is proposed in which it is combined with Simulated Annealing. In order to validate the theory developed here, we carry out some studies based on simulated data and consider some real scenarios from crude oil production data.

Inference for high-dimensional split-plot-designs

Paavo Sattler and Markus Pauly
University Ulm (Germany)

Thursday 27
14:00-16:00
Lecture room 10
(Sali 10)

Statisticians increasingly face the problem to reconsider the adaptability of classical inference techniques. In particular, divers types of high-dimensional data structures are observed in various research areas; disclosing the boundaries of conventional multivariate data analysis. Such *large d and small n problems* are, e.g., met in life sciences whenever it is easier or cheaper to repeatedly generate a large number d of observations per subject than recruiting many, say n , subjects. In this talk, we discuss inference procedures for such situations in repeated measures and heteroscedastic split plot designs. These will, e.g., be able to answer questions about the occurrence of certain time, group and interactions effects or about particular profiles.

The test procedures are based on standardized quadratic forms involving unbiased and dimension-stable estimators of different traces of the underlying unrestricted covariance structures. In a

general asymptotic framework ($\min(n, d) \rightarrow \infty$) its limit distributions are analyzed and additional small sample approximations are proposed. Beneath discussing the theoretical properties, its small sample performance is investigated in simulations. Moreover, the procedure is illustrated by a real data application.

References: [1] Bai, Z., Saranadasa, H., 1996. Effect of High Dimension: by an Example of a Two Sample Problem. *Statistica Sinica*, 6, 311–329. [2] Chen, S. X., Qin, Y.-L., 2010. A Two-Sample Test for High-Dimensional Data with Applications to Gene-Set Testing. *The Annals of Statistics*, 38, 808–835. [3] Gregory, K. B., Carroll, R. J., Baladandayuthapani, V., Lahiri, S. N., 2015. A two-sample test for equality of means in high dimension. *Journal of the American Statistical Association*, 110, 837–849. [4] Jordan, W., Tumani, H., Cohrs, S., Eggert, S., Rodenbeck, A., Brunner, E., R  ther, E., Hajak, G., 2004. Prostaglandin D Synthase (β -trace) in Healthy Human Sleep. *Sleep*, 27, 867–874. [5] Pauly, M., Ellenberger, D., Brunner, E., 2015. Analysis of High-Dimensional One Group Repeated Measures Designs. *Statistics*, 9, 1243–1261. [6] Srivastava, M.S., 2009. A test for the mean vector with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis*, 100, 518–532.

A principled over-penalization of AIC

Adrien Saumard¹ and Fabien Navarro

¹ENSAI (France)

Stabilization by over-penalization is a well-known phenomenon for specialists of model selection procedures. Indeed, it has been remarked for a long time that adding a small amount to classical penalized criteria such as AIC lead in good cases to an improvement of prediction performances, especially for moderate and small sample sizes. In particular, overfitting tends to be avoided. We propose here the first principled and general over-penalization strategy and apply it to AIC. Very good results are observed in simulations.

Identifying differential distributions for single-cell RNA sequencing data comprising biological replicates

Roman Schefzik

German Cancer Research Center (DKFZ) (Germany)

Major advances in technology in the current decade allow to sequence information from individual biological cells and thus offer a high resolution of cellular differences. Single cell sequencing facilitates fundamental insights into biology and was chosen as the method of the year 2013 by Nature Publishing Group. In particular, performing high-throughput ribonucleic acid (RNA) sequencing at the single-cell level (scRNA-seq) has enabled the quantification of cellular heterogeneity. However, it remains a challenge to account for the natural biological variation for the detection of differential gene expression, cellular heterogeneity and clustering in the scRNA-seq data. Thus, the design of appropriate statistical analysis methods for scRNA-seq data currently is of great interest and importance.

Recently developed tools are able to detect differences in gene expression distributions across conditions which are more involved than a shift in the mean. For instance, a current approach allows for a classification of genes with a differential distribution (DD) into categories that represent distinct DD patterns such as differential modality or differential proportion of cells within each component.

Our aim is to develop a statistical method that identifies DDs in the scenario in which the scRNA-seq data consist of read counts obtained by different biological replicates for each physiological condition. In such a case, the gene expression for each condition can first be represented by a mixture distribution, where each mixture component corresponds to a kernel density estimate fitted to the expression data of an individual biological replicate. Due to this strategy, the information given by the corresponding multiple biological replicates is aggregated for each condition

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

Poster session

separately. Then, the condition-specific distributions so obtained are checked for significant differences applying the L2 Wasserstein distance and a permutation test. Our procedure is illustrated and evaluated using data from a recent scRNA-seq experiment.

On the distribution and model selection properties of the Lasso estimator in low and high dimensions

Ulrike Schneider

Vienna University of Technology (Austria)

Wednesday 26
10:30-12:30
Lecture room 10
(Sali 10)

We derive expressions for the finite-sample distribution of the Lasso estimator in the context of a linear regression model with normally distributed errors in low as well as in high dimensions by exploiting the structure of the optimization problem defining the estimator. In low dimensions, we assume full rank of the regressor matrix and present an expression for the cdf as well as formulas for the densities of the absolutely continuous parts of the estimator. Additionally, we establish a direct relationship between the Lasso and the LS estimator by explicitly characterizing the transformation mapping the LS to the Lasso estimator. We derive analogous results in a less explicit form for the distribution in high dimensions where we make no assumptions on the regressor matrix at all. In this setting, we also investigate the model selection properties of the Lasso and illustrate that what types of models may potentially be selected by the estimator might be completely independent of the observed data vector.

Confidence regions for spatial excursion sets from repeated random field observations, with an application to climate

Armin Schwartzman

University of California at San Diego (USA)

Tuesday 25
10:30-12:30
Lecture room 6
(Sali 6)

The goal of this work is to give simultaneous confidence regions for the excursion set of a spatial function above a given threshold from repeated noisy observations on a fine grid of fixed locations. Given an asymptotically Gaussian estimator of the target function, a pair of data-dependent nested excursion sets are constructed that are sub- and super-sets of the true excursion set, respectively, with a desired confidence. Asymptotic coverage probabilities are determined via a multiplier bootstrap method, not requiring Gaussianity of the original data nor stationarity or smoothness of the limiting Gaussian field. The method is used to determine regions in North America where the mean summer and winter temperatures are expected to increase by mid 21st century by more than 2 degrees Celsius.

Prediction out-of-sample using block shrinkage estimators: model selection and predictive inference

Nina Senitschnig and Hannes Leeb

University of Vienna (Austria)

Wednesday 26
10:30-12:30
Lecture room 10
(Sali 10)

In a linear regression model with random design, we consider a family of candidate models from which we want to select a 'good' model for prediction out-of-sample. We fit the models using block shrinkage estimators, and we focus on the challenging situation where the number of explanatory variables can be of the same order as sample size and where the number of candidate models can be much larger than sample size. We develop an estimator for the out-of-sample predictive performance, and we show that the empirically best model is asymptotically as good as the truly best model. Using the estimator corresponding to the empirically best model, we construct a prediction interval that is approximately valid and short with high probability, i.e., we show that the actual coverage probability is close to the nominal one and that the length of this prediction

interval is close to the length of the shortest but infeasible prediction interval. All results hold uniformly over a large class of data-generating processes. These findings extend results of Leeb (2009, Ann. Stat. 37:2838-2876), where the models are fit using least-squares estimators, and of Leeb and Huber (working paper), where the models are fit using shrinkage estimators without block structure.

Friday 28
10:30-12:30
Great hall
(Juhlasali)

Goodness of fit tests for high-dimensional linear models

Rajen Shah¹ and Peter Bühlmann²

¹University of Cambridge (UK), ²ETH Zurich (Switzerland)

In this talk I will introduce a framework for constructing goodness of fit tests in both low and high-dimensional linear models. The idea involves applying regression methods to the scaled residuals following either an ordinary least squares or Lasso fit to the data, and using some proxy for prediction error as the final test statistic. We call this family Residual Prediction (RP) tests. We show that simulation can be used to obtain the critical values for such tests in the low-dimensional setting, and demonstrate that some form of the parametric bootstrap can do the same when the high-dimensional linear model is under consideration. We show that RP tests can be used to test for significance of groups or individual variables as special cases, and here they compare favourably with state of the art methods, but we also argue that they can be designed to test for as diverse model misspecifications as heteroscedasticity and different types of nonlinearity.

Tuesday 25
10:30-12:30
Lecture room 6
(Sali 6)

Adaptive scan statistics with convolutional networks

James Sharpnack

University of California at Davis (USA)

We propose a framework for fast adaptive scan statistics for shapes in images based on convolutional neural networks. We demonstrate that with a convolutional feed-forward network, one can compute a scan statistic for a single image that adapts to unknown shape and scale. This can be shown to provide optimal guarantees for asymptotic distinguishability, and asymptotically correct p-values. From these insights, we introduce a framework that learns unknown shapes and signals, from a database of images. We discuss current theoretical and algorithmic guarantees for this general framework.

Tuesday 25
10:30-12:30
Lecture room 12
(Sali 12)

On a spiked model for large volatility matrix estimation from noisy high-frequency data

Keren Shen, Jianfeng Yao and Wai Keung Li

The University of Hong Kong (China)

Recently, inference about high-dimensional integrated covariance matrices (ICVs) based on noisy high-frequency data has emerged as a challenging problem. In the literature, a pre-averaging estimator (PA-RCov) is proposed to deal with the microstructure noise. Using the large-dimensional random matrix theory, it has been established that the eigenvalue distribution of the PA-RCov matrix is intimately linked to that of the ICV through the Marčenko-Pastur equation. Consequently, the spectrum of the ICV can be inferred from that of the PA-RCov. However, extensive data analyses demonstrate that the spectrum of the PA-RCov is *spiked*, that is, a few large eigenvalues (spikes) stay away from the others which form a rather continuous distribution with a density function (bulk). Therefore, any inference on the ICVs must take into account this spiked structure. As a methodological contribution, we propose a spiked model for the ICVs where spikes can be inferred from those of the available PA-RCov matrices. The consistency of the inference procedure is established and is checked by extensive simulation studies. In addition, we apply our method to the real data from the US and Hong Kong markets. It is found that our model clearly outperforms the existing one in predicting the existence of spikes and in mimicking the empirical PA-RCov matrices.

Recommender systems for purchase data: a novel evaluation approach and an empirical comparison

Poster session

Fabio Sigrist

Lucerne University of Applied Sciences (Switzerland)

Automatically generated product recommendations are more and more used not just by digital-only companies but also by a wide range of companies operating in traditional industry sectors such as banking. In most cases, companies only have implicit rating data such as unary purchase or click data. We present a novel strategy for evaluating recommender systems for unary data. This approach explicitly takes into account the fact that non-purchased or non-clicked items are not necessarily non-relevant to a user. The proposed evaluation strategy is applied to the task of generating recommendations for bank customers. We compare several collaborative filtering recommendation systems and find that a classification-based system performs best. Furthermore, a classification-based approach has the advantage that additional user content data can be easily included, thus, alleviating the so-called cold-start problem and resulting in increased predictive performance.

Estimation of network topology: Genes and individuals

Mikko J. Sillanpää

University of Oulu (Finland)

Monday 24
10:30-12:30
Lecture room 5
(Sali 5)

Estimation of non-zero elements of the precision matrix is equal to estimation of the undirected network structure. The talk will cover some of the new methods developed for sparse and non-sparse precision matrix estimation under penalized likelihood setting including the method called ROPE (Kuismin et al. 2017). The talk will also cover some issues important in estimating networks involving genes and individuals.

References: Kuismin, M. O., Kempainen, J. T., and M. J. Sillanpää (2017) Precision matrix estimation with ROPE. *Journal of Computational and Graphical Statistics* (accepted). Available online.

L-moment homogeneity test in linear regression

Poster session

Tereza Simkova

Technical University of Liberec (Czech Republic)

An important task of mathematical statistics is to describe the dependence between variables. Koenker and Bassett (1978) introduced regression quantiles as a generalization of traditional quantiles to obtain comprehensive information about behavior of a dependent variable. Recently, Picek and Schindler (in preparation) have proposed a generalization of L-moments in the linear regression model using the averaged regression quantiles introduced by Jurekova and Picek (2014). They have shown the proposed sample averaged regression L-moments may be used to estimate parameters of a distribution of errors via the usual L-moments method. Here, we propose a homogeneity test in linear regression model using sample averaged regression L-moments. Construction of the homogeneity test for the distribution of errors is based on the idea of Masselot, Chebana and Ouarda (2016) by adopting the permutation test. The usefulness of the proposed test is illustrated on simulated and meteorological data.

References: [1] Hosking, J. R. M., Wallis, J. R., 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, UK. [2] Jurečková, J., Picek, J., 2014. Averaged regression quantiles, In: Lahiri, S., et al. (eds.). *Contemporary Developments in Statistical Theory*, 68, 203–216. [3] Koenker, R., Bassett, G., 1978. Regression quantiles. *Econometrica*, 46, 33–50. [4] Picek, J., Schindler, M. L-moments in linear regression model, in preparation.

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

Kernel partial least squares for stationary data

Marco Singer, Tatyana Krivobokova and Axel Munk
Georg-August-Universität Göttingen (Germany)

In a non-parametric regression setting where the input variables form a stationary time series we prove probabilistic convergence rates of the kernel partial least squares algorithm. Kernel partial least squares is an iterative regularized regression technique where the regularization is achieved by early stopping of the algorithm. These rates depend on three parameters: the regularity of the target function measured by a Hölder-type source condition, the effective dimensionality of the data mapped into the reproducing kernel Hilbert space and the range of the dependence in the input time series. If the effective dimensionality decays polynomially fast and the data are independent and identically distributed quasi-optimal convergence rates are obtained that also hold for kernel ridge regression or kernel conjugate gradient. We study the impact of the dependence of the time series on the convergence rates in a Gaussian setting. For short range dependent data we obtain the same rates as in the independent and identically distributed case, whereas long range dependence yields slower rates.

Poster session

Automated testing of differences of means to promote the effective use of public data

Alina Sinisalo, Anne-Mari Sepponen, and Arto Latukka
Natural Resources Institute Finland (Luke) (Finland)

Economy Doctor is web-based reporting service (www.luke.fi/economydoctor) for publishing time series of business activities and income of Finnish agricultural holdings. The results are presented at average level by region, economic size and type of holding based on a sample survey of microeconomic data from 900-1000 voluntary holdings. Our objective was to promote the effective use of the sample survey results by implementing automated calculation and statistical test routines. Currently, since various descriptive statistics are reported at average level, users are not able to distinguish, if the differences of average results between classes are statistically significant and which classes deviate from others. This is due to the fact that variables having very different distributions may have approximately the same calculated average. Due to limitations set by data privacy decrees detailed and individual farm level results cannot be published. Therefore we developed the reporting platform by adding a feature to indicate the differences of means of accounting data and performance figures within a year or between years. The statistical significance, pairwise comparison and the used test criteria are printed out in real time online based on classes defined by user. The aim is to develop Economy Doctor reporting service to produce more concrete help for researchers, decision makers and the public audience.

Tuesday 25
14:00-16:00
Lecture room 6
(Sali 6)

Improvement of expected shortfall estimations in APARCH and FIGARCH models

George-Jason Siouris
University of the Aegean (Greece)

In finance, one of the main goals is the estimation of volatility, since it is crucial in risk analysis and management. For that propose, state of the art quantitative techniques have been mobilized, techniques that use finance, statistics and computer programming at the same time. To forecast the asset volatility several different models have been developed. Each one of them, offers an answer to a specific aspect of the problem at hand. The most recent stops in this never-ending quest for better estimations were:

1. Asymmetric and power GARCH models (APARCH)

2. Fractionally Integrated GARCH models (FIGARCH)

The need that driven us to these models was the fact that after extensive research on the statistical properties of financial returns, three properties have shown to be present in most, if not all, financial returns. Their existence has been the source of the problems in the estimation of the underlying risk of the assets. These are often called the three stylized facts of financial returns and are:

1. Volatility clusters
2. Fat tails
3. Nonlinear dependence

In this work, we analyze the sources of error in these models and explore ways for improving our estimations.

Close-kin Mark-Recapture

Hans J. Skaug¹, Mark V. Bravington² and Eric C. Anderson³

¹University of Bergen (Norway), ²CSIRO Marine Lab, Hobart (Australia), ³Southwest Fisheries Science Center, Santa Cruz (USA)

Tuesday 25
10:30-12:30
Lecture room 7
(Sali 7)

DNA tags have become popular for some types of mark recapture studies. DNA profiles also contain information about kinship among individuals in the sample. This forms the basis for a novel type of methods called Close-kin Mark-Recapture (CKMR), which like ordinary MR aims at estimating parameters such as abundance and mortality. The probability distribution of the full genealogy of the individuals in the sample is intractable and hence we resort to a pairwise pseudo-likelihood. This creates a need to calculate the probability that two sampled individuals are kin of a particular type. We will focus on parent-offspring probabilities, and show how these can be calculated in an age-structured population.

Structured PCA

Jim Skinner and Richard S. Savage
Warwick University (UK)

Friday 28
10:30-12:30
Lecture room 7
(Sali 7)

Inference can be difficult when dealing with datasets with more dimensions than samples ($p > n$). Even the Principal Components estimated from classical PCA become orthogonal to the true Principal Components if $p \gg n$. However, if we have sufficient prior knowledge this can be used to combat this curse of dimensionality. We introduce *Structured PCA* (SPCA). SPCA allows specification of prior knowledge on the *covariance structure* of data; that is, how strongly or weakly each dimension is expected to correlate. This is obtained as a simple extension to the well known model Probabilistic PCA (PPCA); a zero mean Gaussian prior is added to the loadings matrix mapping the latent to the observed data space. The result is a computational tool similar to PCA allowing Principal Components to be estimated and a linear latent space to be inferred. The only additional input required is a prior covariance matrix which can improve performance when uncertainty is high. Unlike PPCA, SPCA does not admit the closed form computation of the Maximum-a-posteriori model parameters. An Expectation-Maximization algorithm has been derived to allow this computation. Hyperparameters specifying the user-provided prior covariance matrix may also be tuned to the data in an Empirical Bayes manner. Simulations show good results on synthetic data, even when the prior covariance matrix is known to be quite different from the real covariance structure. An R package is in development.

Many instrument asymptotics: Implementation in STATA

Alena Skolkova

CERGE-EI (Czech Republic)

We develop a module for the STATA statistical package that implements recent advancements from the literature on dimension asymptotics when the dimensionality of an instrument in an instrumental variables regression is comparable to a sample size. This contributes to the availability to practitioners without strong theoretical background of most recent findings in the theory of estimation and inference under many instruments. Our add-on STATA commands concentrate on consistent estimation such as bias corrected 2SLS, jackknife instrumental variables and limited information maximum likelihood. While some of these estimators are already available as STATA commands, they yield misleading inferences as the built-in inference procedures rely on conventional asymptotic tools. We program the inference procedures so that they are robust to many possibly weak instruments and work under error non-normality and both under homoscedasticity and heteroscedasticity.

Parameter estimation for discretely observed infinite-server queues with Markov-modulated input.

Birgit Sollie¹, Mathisca de Gunst¹, Bartek Knapik¹ and Michel Mandjes²¹VU Amsterdam (Netherlands), ²University of Amsterdam (Netherlands)

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

The Markov-modulated infinite-server queue is a queueing system with infinitely many servers, where the arrivals follow a Markov-modulated Poisson process (MMPP), i.e. a Poisson process with rate modulating between several values. The modulation is driven by an underlying and unobserved continuous time Markov chain $\{X_t\}_{t \geq 0}$. The inhomogeneous rate of the Poisson process, $\lambda(t)$, stochastically alternates between d different rates, $\lambda_1, \dots, \lambda_d$, in such a way that $\lambda(t) = \lambda_i$ if $X_t = i$, $i = 1, \dots, d$. We are interested in estimating the parameters of the arrival process for this queueing system based on observations of the queue length at discrete times only. We assume exponentially distributed service times with rate μ , where μ is time-independent and known. Estimation of the parameters of the arrival process has not yet been studied for this particular queueing system. Two types of missing data are intrinsic to the model, which complicates the estimation problem. First, the underlying continuous time Markov chain in the Markov-modulated arrival process is not observed. Second, the queue length is only observed at a finite number of discrete time points. As a result, it is not possible to distinguish the number of arrivals and the number of departures between two consecutive observations. In this talk we show how we derive an explicit algorithm to find maximum likelihood estimates of the parameters of the arrival process, making use of the EM algorithm. Our approach extends the one used in Okamura et al. (2009), where the parameters of a Markov-modulated Poisson process are estimated based on observations of the process at discrete times. However, in contrast to our setting, Okamura et al. (2009) do not consider departures and therefore do not deal with the second type of missing data. We illustrate the proposed estimation method with a simulation study.

Reference: Okamura H., Dohi T., Trivedi K.S., 2009. Markovian Arrival Process Parameter Estimation With Group Data. *IEEE/ACM Transactions on Networking*. Vol. 17, No. 4, 1326–1339.

A non-parametric Bayesian approach to decomposing from high frequency data

Peter Spreij¹, Shota Gugushvili² and Frank van der Meulen³

¹University of Amsterdam (Netherlands), ²Leiden University (Netherlands), ³Delft University (Netherlands)

Tuesday 25
10:30-12:30
Lecture room 10
(Sali 10)

Shota Gugushvili Frank van der Meulen Peter Spreij Abstract Given a sample from a discretely observed compound Poisson process, we consider non-parametric estimation of the density f_0 of its jump sizes, as well as of its intensity λ_0 . We take a Bayesian approach to the problem and specify the prior on f_0 as the Dirichlet location mixture of normal densities. An independent prior for λ_0 is assumed to be compactly supported and to possess a positive density with respect to the Lebesgue measure. We show that under suitable assumptions the posterior contracts around the pair (λ_0, f_0) at essentially (up to a logarithmic factor) the $\sqrt{n\Delta}$ -rate, where n is the number of observations and Δ is the mesh size at which the process is sampled. The emphasis is on high frequency data, $\Delta \rightarrow 0$, but the obtained results are also valid for fixed Δ . In either case we assume that $n\Delta \rightarrow \infty$. Our main result implies existence of Bayesian point estimates converging (in the frequentist sense, in probability) to (λ_0, f_0) at the same rate. We also discuss a practical implementation of our approach. The computational problem is dealt with by inclusion of auxiliary variables and we develop a Markov chain Monte Carlo algorithm that samples from the joint distribution of the unknown parameters in the mixture density and the introduced auxiliary variables. Numerical examples illustrate the feasibility of this approach.

A geometric framework for density estimation and its extensions to estimation under shape constraint

Anuj Srivastava

Florida State University (USA)

Tuesday 25
14:00-16:00
Lecture room 5
(Sali 5)

The problem of estimating probability density functions (pdfs), from univariate, conditional, or multivariate data, is a fundamental problem in statistics. While numerous solutions have been proposed over the years, the geometry of the space of pdfs has seldom been utilized in these solutions. We describe a geometric approach where we use an action of the warping group to explore the space of pdfs. The estimation process has two steps – initialize the estimate using any fast method (parametric, kernel-based, etc.), and then use warping to improve this solution by maximizing the solution. These ideas are applicable to univariate (conditional and unconditional) and bivariate densities. We demonstrate these ideas using both simulated and real datasets used commonly in the literature. This framework is shown to either match or outperform the state of the art methods in estimation performance and computational costs. We also discuss extension of this basic framework to estimation of densities under shape (unimodal, bimodal or, generally k -modal) constraints. While unimodal density estimation has often been studied in the literature, estimation of a pdf with k modes, for a given $k > 1$, is rarely discussed. Taking a different warping action, one can explore the space of all pdfs while fixing the number of modes; only the locations and shapes of modes are changed. Once again, we demonstrate these ideas using simulated and real data sets.

A longitudinal mixed logit model for estimation of push-pull effects in residential location choice

Thursday 27
14:00-16:00
Lecture room 13
(Sali 13)

Fiona Steele¹, Elizabeth Washbrook², Christopher Charlton² and William Browne²

¹London School of Economics (UK), ²University of Bristol (UK)

We develop a random effects discrete choice model for the analysis of households' choice of neighbourhood over time. The model is parameterised in a way that exploits longitudinal data to separate the influence of neighbourhood characteristics on the decision to move out of the current area ("push" effects) and on the choice of one destination over another ("pull" effects). Random effects are included to allow for unobserved heterogeneity between households in their propensity to move, and in the importance placed on area characteristics. The model also includes area-level random effects. The combination of a large choice set, large sample size and repeated observations mean that existing estimation approaches are often infeasible. We therefore propose an efficient MCMC algorithm for the analysis of large-scale datasets. The model is applied in an analysis of residential choice in England using data from the British Household Panel Survey linked to neighbourhood-level census data. We consider how effects of area deprivation and distance from the current area depend on household characteristics and life course transitions in the previous year.

On locally weighted trigonometric regression under long-range dependence

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

Britta Steffens

University of Konstanz (Germany)

We consider a seasonal time series model under long memory. Unified formulas for the integrated mean squared error (IMSE) and the asymptotic optimal bandwidth of the locally weighted trigonometric regression estimator of the seasonal component are derived. The IMSE turns out to be of the same order as under independence or short-range dependence. Simulations and data examples illustrate the asymptotic results.

Statistical inference for Lévy-driven CARMA processes based on non-equidistant observations

Thursday 27
14:00-16:00
Lecture room 8
(Sali 8)

Robert Stelzer¹, Frank Bosserhoff¹ and Żywilla Fechner²

¹Ulm University (Germany), ²University of Silesia in Katowice (Poland)

Continuous time autoregressive moving averages driven by Lévy processes are a fundamental time series model defined in continuous time. Such models are in particular relevant in the context of high frequency data or irregularly sampled data. In this talk we focus on the question of estimating the parameters of the process based on observations made at non-equidistant time points. We present and discuss an approach for observations made at independent Poissonian times where it can be shown that the identification of the parameters is actually less problematic than in the case of equidistant observations. In this set-up a Whittle-type estimator for the parameters can be used which is consistent and asymptotically normal under certain conditions. As an alternative approach we consider path-wise Fourier transforms of CARMA processes and high-frequency limits along non-equidistant deterministic grids as a first step to obtaining other estimators.

Time to absorption in the Mabinogion urn model

David Stenlund

Åbo Akademi University (Finland)

Wednesday 26
10:30-12:30
Lecture room 8
(Sali 8)

We consider the Mabinogion urn model, which is a stochastic urn model with the following characterization. An urn contains balls of two different colors (black and white). We draw one ball at random from the urn, observe the color and replace it. If the drawn ball was white, one black ball is removed and replaced by a white ball, and vice versa if the drawn ball was black. This is repeated until all balls are of the same color, in which case the process stops. This urn model has two absorbing states, and here we give an expression for the expected time to absorption as a function of the initial state of the urn. In the symmetric case, when you start with an equal number of black and white balls, the expression reduces to a particularly neat formula. For the Mabinogion urn model there is also an optimal control problem called "the Mabinogion sheep problem", described by D. Williams in *Probability with Martingales*. We briefly discuss the problem and the optimal strategy presented by Williams and thereafter give a formula for the expected time to absorption when using said strategy. We furthermore give expressions for the asymptotic behavior of the expected time to absorption when the number of balls in the urn goes towards infinity. This is done both for the standard Mabinogion process and for the optimal control problem, when starting from a symmetric initial state. We show that in both cases the leading-order terms are the same, namely of the order of magnitude $\frac{1}{2}n \log(n)$, where n is the initial number of both colors in the urn.

Estimating the amount of sparsity in mixture models

Natalia Stepanova and Yibo Wang

Carleton University (Canada), University of Alberta (Canada)

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

This work is largely motivated by a variable selection problem in sparse normal mixtures. For this problem, the sharp selection boundaries, that is, the necessary and sufficient conditions for the possibility of successful variable selection in the exact and almost full regimes are available. The existing selection boundaries, as well as the procedure that provides almost full selection, depend on the fraction of non-zero means, which is generally unknown. In this talk, we present a new estimator for the fraction of non-zero means in normal mixture models with relatively few non-zero means that are only moderately large. We show that, in the region where variable selection is possible, the new estimator dominates (in terms of the minimax rate of convergence) the existing estimators proposed earlier in similar contexts; the same conclusion continues to hold for the region where signal detection is possible. Moreover, our estimator nearly attains the optimal rate of convergence. The obtained analytical results are illustrated numerically.

Bayesian estimation for transmission potential of smallpox

Jessica Stockdale, Theodore Kypraios and Philip O'Neill

University of Nottingham (UK)

Thursday 27
14:00-16:00
Lecture room 6
(Sali 6)

Highly contagious and with a 30% fatality rate, smallpox is considered a serious threat to public health. Although officially eradicated in 1959, there is concern about the synthesis of smallpox and its potential use as a bioterrorist weapon. This motivates the development of robust quantitative measures of transmission in order to plan effective mitigation strategies. We consider an outbreak of smallpox which occurred in the Nigerian town of Abakaliki in 1967. The World Health Organisation recorded the outbreak in detail, with information not only on the time series of case detection times but also individuals place of dwelling, vaccination status and membership of a religious group whose members refused this vaccination. This data set has appeared numerous times in the epidemic modelling literature, but in almost all cases the majority of the data are ignored. The only previous analysis of the full data relies upon approximation methods to derive

a likelihood. We present here the first full Bayesian analysis of the Abakaliki data, using Markov Chain Monte Carlo methods and data augmentation to avoid approximation as well as present a wider range of analyses. We also perform model assessment with simulation-based methods, and results include estimates of transmission rates, vaccine efficacy and reproduction numbers.

Poster session

Likelihood approximation methods for stochastic epidemic models

Jessica Stockdale, Theodore Kypraios and Philip O'Neill
University of Nottingham (UK)

A key issue in stochastic epidemic modelling is the development of methods for dealing with partially observed data. For example data regarding infection processes are rarely available, making standard statistical techniques difficult to apply. Methods such as Markov Chain Monte Carlo are widely used for inference of missing data, but can become computationally demanding when population sizes are large. This motivates the development of likelihood approximation methods which avoid the use of data augmentation. We explore several approximation methods which make a series of independence assumptions in order to form likelihood expressions which are reasonably accurate as well as offer increased computational speed.

Wednesday 26
10:30-12:30
Lecture room 12
(Sali 12)

Censored functional data

Ewa Strzalkowska-Kominiak¹ and Juan Romo²

¹Cracow University of Technology (Poland), ²Universidad Carlos III de Madrid (Spain)

In medical research it is natural that patients are followed over time. Consequently, the measurements of particular biomarker for every patient are often registered as curves. Hence it is of interest to estimate the mean function for patients under certain condition as an average of the observed functional data over a given period of time. This, however, often results difficult since in the follow-up studies it is common that some of those patients drop-out of the study before the end. In this case, only a part of the functional observation is available. We propose an estimator for the functional mean, when the functions may be only partly observed and so censored from the right. Additionally, we propose a bootstrap based confidence band for the mean function and show, in an simulation study, that our new method outperforms the only two existing approaches and apply the estimator to the real data example.

References: [1] Cuevas, A., Febrero, M., Fraiman, R., 2006. On the use of the bootstrap for estimating functions with functional data. *Computational Statistics & Data Analysis*, 51, 1063–1074. [2] Delaigle, A., Hall, P., 2013. Classification using censored functional data. *Journal of the American Statistical Association*, 108(504), 1269–1283.

Monday 24
10:30-12:30
Lecture room 12
(Sali 12)

Estimating a frontier function using a high-order moments method

Gilles Stupfler¹, Stéphane Girard² and Armelle Guillou³

¹University of Nottingham (UK), ²INRIA Rhône-Alpes (France), ³University of Strasbourg (France)

Frontier function estimation can be applied in several problems, such as the estimation of the maximal output of a company given a quantity of input, or the estimation of the maximal temperature at a given point on the surface of the Earth. We present here a method to estimate the frontier function of a finite-dimensional sample. The estimator is based on a kernel regression on high order moments, where we assume that the order of the moments goes to infinity while the bandwidth of the kernel goes to zero. We discuss the asymptotics of our estimator and in particular its uniform consistency and pointwise asymptotic normality, when the conditional distribution function decreases at a polynomial rate to zero in a neighbourhood of the frontier. The finite-sample performance of the estimator is illustrated on a simulation study.

Testing for linear cointegration against the alternative of smooth transition cointegration

Oliver Stypka and Martin Wagner
TU Dortmund (Germany)

Thursday 27
10:30-12:30
Lecture room 7
(Sali 7)

We develop tests for the null hypothesis of linear cointegration against the alternative of smooth transition cointegration, with the alternative given by, i. e.,

$$y_t = Z_t' \theta_L + Z_t' \theta_{NL} \times G(s_t, \theta_G) + u_t,$$

where y_t is a scalar time series, $Z_t := [D_t', X_t']'$, D_t a vector of deterministic regressors, X_t a non-cointegrated I(1) vector and $G(s_t, \theta_G)$ a sufficiently smooth and bounded *transition function*. The *transition variable* s_t is either an integrated variable or a time trend. Testing linear cointegration against the alternative of smooth transition cointegration corresponds to $H_0 : \theta_{NL} = 0$ against $H_1 : \theta_{NL} \neq 0$. It is well-known in the literature that in this setting “some parameters are only present under the alternative” (see in particular Davies, 1977, 1987) or, in other words, that the parameters of the model are not identified under the null hypothesis. This problem is often tackled in the smooth transition regression literature by using Taylor approximations of the transition function and LM-type tests (see from a long list of contributions, e. g., Choi and Saikkonen, 2004; Luukkonen *et al.*, 1988). Using a Taylor approximation of order m leads to

$$y_t = Z_t' \beta_0 + \sum_{j=1}^m (Z_t \otimes s_t^j)' \beta_j + u_t^*,$$

with the precise relationship between the parameters θ_L, θ_{NL} and θ_G on the one hand and β_0, \dots, β_m depending upon the transition function chosen. The test statistics are based on fully modified (compare Phillips and Hansen, 1990) or integrated modified OLS (compare Vogelsang and Wagner, 2014) parameter estimators for the above Taylor approximations to the smooth transition regression function. This necessitates the adaptation of the above estimation approaches to models including cross-products of integrated regressors. We stress the implications of the choice of the transition variable for fully modified and integrated modified OLS estimation and the differences in the asymptotic limiting distributions. Moreover, we discuss consistency of the tests against a variety of alternatives. For the integrated modified OLS based test we additionally discuss also fixed- b inference. Due to the form of the Taylor approximation a prerequisite for fixed- b inference, called *full design*, always prevails, such that there are no additional restrictions for fixed- b inference compared to the usual asymptotic chi-squared standard inference. The properties of the tests are evaluated with a simulation study and compared to the test proposed by Choi and Saikkonen (2004). Finally, we apply our results to test linear cointegration against smooth transition cointegration of money demand for nine countries: Australia, Canada, Denmark, Great Britain, Germany, Norway, South Korea, Switzerland and the USA.

References: [1] Choi, I., Saikkonen, P., 2004. Testing Linearity in Smooth Transition Cointegrating Regressions. *Econometrics Journal*, 7, 341–365. [2] Davies, R.B., 1977. Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative. *Biometrika*, 64, 247–254. [3] Davies, R.B., 1987. Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative. *Biometrika*, 74, 33–43 [4] Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988. Testing Linearity against Smooth Transition Autoregressive Models. *Biometrika*, 75, 491–499. [5] Phillips, P.C.B., Hansen, B.E., 1990. Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies*, 57, 99–125. [6] Vogelsang, T.J., Wagner, M., 2014. Integrated Modified OLS Estimation and Fixed- b Inference for Cointegrating Regressions. *Journal of Econometrics*, 714–760.

Economic design of two stage control charts with bivariate skew-normal measurements

Friday 28
10:30-12:30
Lecture room 1
(Sali 1)

Nan-Cheng Su
National Taipei University (Taiwan)

In many instances, the cost is high to monitor primary quality characteristics called performance variable, but it could be economical to monitor its surrogate. To cover asymmetric processes for two-stage charting methods using both performance and surrogate variables, bivariate skew normal distribution is considered as the underlying distribution of process variables. When the correlation relationship between the performance variable and its surrogate is specified, a charting procedure to monitor either the performance variable or its surrogate in an alternating fashion rather than monitoring the performance variable alone is proposed. The proposed two-stage control charts are constructed under an economic design using Markov chain approach. Its advantages over the existing methods will be presented.

A new Liu regression estimator in partially linear models

Friday 28
10:30-12:30
Lecture room 10
(Sali 10)

Gulin Tabakan
Aksaray University (Turkey)

In this study, we propose a new Liu estimator based on prior information for the vector of parameters in a partially linear model (PLM) as an alternative to the classical difference-based regression estimator in the existence of multicollinearity problem. The theoretical properties of the proposed estimator and its relationship with some existing difference-based biased methods designed for PLM are established. Finally, a numerical example is presented to explain the performance of the new estimator.

Pathwise least angle regression and a significance test for the elastic net

Friday 28
10:30-12:30
Lecture room 10
(Sali 10)

Muhammad Naveed Tabassum and Esa Ollila
Aalto University (Finland)

Least angle regression (LARS) by Efron et al. (2004) is a novel method for constructing the piece-wise linear path of Lasso solutions. For several years, it remained also as the de facto method for computing the Lasso solution before more sophisticated optimization algorithms preceded it. LARS method has recently again increased its popularity due to its ability to find the values of the penalty parameters, called *knots*, at which a new parameter enters the active set of non-zero coefficients. Significance test for the Lasso by Lockhart et al. (2014), for example, requires solving the knots via the LARS algorithm. Elastic net (EN), on the other hand, is a highly popular extension of Lasso that uses a linear combination of Lasso and ridge regression penalties. In this paper, we propose a new novel algorithm, called pathwise (PW-)LARS-EN, that is able to compute the EN knots over a grid of EN tuning parameter α values. The developed PW-LARS-EN algorithm decreases the EN tuning parameter and exploits the previously found knot values and the original LARS algorithm. A covariance test statistic for the Lasso is then generalized to the EN for testing the significance of the predictors. Our simulation studies validate the fact that the test statistic has an asymptotic $\text{Exp}(1)$ distribution.

Statistical inference for the Wasserstein distance on discrete spaces - with applications

Carla Taming, Axel Munk and Max Sommerfeld
University Göttingen (Germany)

Monday 24
10:30-12:30
Lecture room 7
(Sali 7)

We derive distributional limits for empirical Wasserstein distances of probability measures supported on discrete sets. Our approach is based on sensitivity analysis of optimal values of infinite dimensional mathematical programs and Hadamard directional differentiability. A careful calibration of the norm is needed in order to combine differentiability and weak convergence of the empirical process. We give an explicit form of the limiting distribution for ultra-metric spaces. Based on this we illustrate how Wasserstein based inference can be used in large scale problems. An application from nanoscale microscopy is given.

Analyzing high dimensional covariance matrices using convolutional neural networks

Yaohua Tang and Philip L.H. Yu
The University of Hong Kong (China)

Poster session

Modeling and forecasting covariance matrices of asset returns play a crucial role in finance. The availability of high frequency intraday data enables the modeling of the realized covariance matrix directly. However, most models in the literature depend on strong structure assumptions and they also suffer from the curse of dimensionality. To solve the problem, we propose a Fully Convolutional Networks which treats each realized covariance matrix as an image. The networks structure is designed with simplicity in mind, and yet provides superior accuracy compared with several advanced statistical methods. The model could handle both low-dimension and high-dimension realized covariance matrices without additional dimension operations.

Sparsest representation and approximation of a high-dimensional linear system

Patrick Tardivel
Toxalim, Université de Toulouse, INRA, ENVT (France)

Thursday 27
10:30-12:30
Lecture room 6
(Sali 6)

We consider a vector $y \in R^n$ and family of vectors $\mathcal{D} = \{d_1, \dots, d_p\}$ spanning R^n . An ϵ -approximation of y in \mathcal{D} is a vector $x = (x_1, \dots, x_p)$ such that $\|y - (x_1 d_1 + \dots + x_p d_p)\|_{l^2}^2 \leq \epsilon$. Our aim is to find at least one of the sparsest ϵ -approximations of y in high dimension when $p > n$. These sparsest ϵ -approximations are defined as the solutions of the problem

$$S_0^\epsilon := \operatorname{argmin} \|x\|_{l^0} \text{ subject to } \|y - Dx\|_{l^2}^2 \leq \epsilon. \quad (1)$$

In the previous expression $\|x\|_{l^0} := \operatorname{Card}\{i \in \llbracket 1, p \rrbracket \mid x_i \neq 0\}$ is the l^0 "norm" of x and $D := (d_1 | \dots | d_p)$ is the $n \times p$ matrix whose columns are $(d_j)_{1 \leq j \leq p}$. In the particular case where $\epsilon = 0$, the constraint becomes $y = Dx$ and S_0^0 is the set of the sparsest representations of y in \mathcal{D} . The problem (1) is intractable and to our knowledge, there is currently no method to solve it. When $\epsilon = 0$ it has been shown that under some conditions, one can find some elements of S_0^0 by substituting the l^0 "norm" by the l^1 norm. This suggests that when $\epsilon \geq 0$, the problem

$$\operatorname{argmin} \|x\|_{l^1} \text{ subject to } \|y - Dx\|_{l^2}^2 \leq \epsilon$$

could give some elements of S_0^ϵ . This is a convex optimisation problem whose solutions are easily computable. When $\epsilon = 0$, this problem is known as the basis pursuit Donoho(2003). Whereas, when $\epsilon > 0$, one can rewrite this problem as a lasso Tibshirani(1996)

$$\operatorname{argmin} \|y - Dx\|_{l^2}^2 + \lambda \|x\|_{l^1} \text{ with } \lambda > 0, \quad (2)$$

with a one to one relation between $\lambda > 0$ and $\epsilon > 0$. The null space property is a necessary and "almost sufficient" condition insuring that the basis pursuit recovers one of the sparsest representations. For the lasso, to our knowledge, there is no theoretical condition insuring that this problem recover one of the sparsest ϵ -approximations. Instead, the irrepresentable condition or the compatibility condition insure the convergence of $x(\lambda)$, solution of (2), toward a solution $x^* \in S_0^0$ when λ converges to 0. In practice, all these conditions are not checkable because they depend on the set $\{i \in \llbracket 1, p \rrbracket \mid x_i^* \neq 0\}$, where x^* is an element of S_0^0 , which is precisely what we are looking for. Furthermore, when these conditions do not hold the solution obtained with the basis pursuit or with the lasso can be very far to the set we wish to recover S_0^ϵ . In this talk, we propose a new tractable problem which allows to recover a sparsest representation or a sparsest ϵ -approximation. The problem we are going to solve writes

$$S_\alpha^\epsilon := \operatorname{argmin} \sum_{i=1}^p |x_i|^\alpha \text{ subject to } \|y - Dx\|_{l^2}^2 \leq \epsilon.$$

Without any condition, we have

- When $\epsilon = 0$, there exists α_0 so that for all $\alpha \leq \alpha_0$, $S_\alpha^\epsilon \subset S_0^0$.
- When $\epsilon > 0$ the set S_α^ϵ becomes arbitrary close to the set S_0^ϵ when α converges to 0.

A Maximisation Minimisation (MM) method allows to solve numerically this problem. Numerical experiments show that this approach gives better results to recover the sparsest representation than both the basis pursuit problem and the reweighted l^1 minimization problem Candès(2007).

This work is a part of the project GMO90+ supported by the Ministry of Ecology, Sustainable Development and Energy in the national research program Risk'OGM.

Poster session

Familywise error rate control with a Lasso estimator

Patrick Tardivel

Toxalim, Université de Toulouse, INRA, ENVT (France)

Let us consider the linear Gaussian model $Y = X\beta^* + \varepsilon$, where X is a $n \times p$ full-rank design matrix, ε is a centered Gaussian vector with an invertible variance matrix Γ and β^* is an unknown parameter. We want to estimate the so-called active set $\mathcal{A} = \{i \in \llbracket 1, p \rrbracket \mid \beta_i^* \neq 0\}$ of relevant variables. A natural way to recover \mathcal{A} is to test the hypotheses $\mathcal{H}_i : \beta_i^* = 0$, with $1 \leq i \leq p$. Several type I errors can be controlled in such multiple hypotheses tests. In this talk, we focus on the Familywise Error Rate (FWER) defined as the probability to reject wrongly at least one hypothesis \mathcal{H}_i . When a sparse estimator $\hat{\beta}$ of β^* is available, a very simple way to test the hypothesis \mathcal{H}_i is to reject it when $\hat{\beta}_i \neq 0$. The lasso estimator Tibshirani(1996) is probably the most popular sparse estimator. It defined by

$$\hat{\beta}(\lambda) = \operatorname{argmin} \left\{ \frac{1}{2} \|Y - X\beta\|_{l^2}^2 + \lambda \|\beta\|_{l^1} \right\}.$$

Obviously, the test's performances depend on the estimator $\hat{\beta}(\lambda)$ and, by consequence, on the choice of λ . When X has orthogonal columns (i.e $X'X$ is diagonal), the lasso estimator has an explicit expression which allows to chose λ_0 to control the FWER. When X has no more orthogonal columns, we propose to apply a linear transformation U such that UX has orthogonal columns. This orthogonalization allows to bring back in the orthogonal columns case. In this talk, among linear transformations U that orthogonalize X , we determine the best one U^* which provide a powerful multiple testing procedure.

In this talk, we provide an explicit tuning parameter λ_0 that controls the FWER at a level α . We compared the performances of our multiple comparisons procedure with existing ones through the control of the FWER and through their "power" here defined as the proportion of good rejection:

$$\frac{1}{\operatorname{Card}(\mathcal{A})} \sum_{i \in \mathcal{A}} 1_{\hat{\beta}_i(\lambda_0) \neq 0}.$$

Numerical experiments showed that our multiple testing procedure have better performances than the lasso type procedure Janson et al.(2016) and than the state-of-the-art procedures : Holm-Bonferroni method Holm(1979) or generic step-down Lehmann et al.(2005). An application to the detection of metabolites in metabolomics is provided.

This work is a part of the project GMO90+ supported by the Ministry of Ecology, Sustainable Development and Energy in the national research program Risk'OGM.

An information theory approach to discover hidden associations and interactions in biomedical data

Paweł Teisseyre

Polish Academy of Sciences (Poland)

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

In biostatistics, there is often a need to analyze associations between pairs of variables. Over the years, various dependence measures have been investigated. However less attention has been paid to the problem of finding hidden associations, i.e. those which are revealed only when conditioning on some additional covariate. It is an important issue, as in some cases the dependence between two variables is weak, but at the same time the conditional dependence (given the additional covariate) may be strong. In this work we use an Interaction Information measure which quantifies how much adding the additional covariate strengthen the association between the original variables. The problem is closely related to finding interactions between variables. The main contribution is a novel test for the positiveness of Interaction Information which in turn allows to find hidden associations and interactions. The method combines permutation and chi-squared tests. The experiments indicate that the proposed approach gives very promising results. The method is applied to find interesting associations and interactions among variables in real datasets.

On multivariate kernel density estimation under long memory

Klaus Telkmann

University of Constance (Switzerland)

Monday 24
10:30-12:30
Lecture room 7
(Sali 7)

We consider kernel estimation of the probability density function and its derivatives for multivariate linear processes with long memory. A functional limit theorem for the empirical processes is derived, and the asymptotic distribution of kernel density estimators is obtained under general assumptions.

Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model

Timo Teräsvirta¹ and Annastiina Silvennoinen²

¹Aarhus University (Denmark), ²QUT Business School (Australia)

Thursday 27
14:00-16:00
Lecture room 8
(Sali 8)

A new multivariate volatility model that belongs to the family of conditional correlation GARCH models is introduced. The GARCH equations of this model contain a multiplicative deterministic component to describe long-run movements in volatility and, in addition, the correlations are deterministically time-varying. Parameters of the model are estimated jointly using maximum likelihood. Consistency and asymptotic normality of maximum likelihood estimators is proved. Numerical aspects of the estimation algorithm are discussed. A bivariate empirical example is provided.

Thursday 27
14:00-16:00
Lecture room 1
(Sali 1)

The realized hierarchical Archimedean copula in risk modelling

Anastasija Teterova¹ and Ostap Okhrin²

¹University of St Gallen (Switzerland), ²Dresden University of Technology (Germany)

This paper introduces the concept of the realized hierarchical Archimedean copula (rHAC). The proposed approach inherits the ability of the copula to capture the dependencies among financial time series, and combines it with additional information contained in high-frequency data. The considered model does not suffer from the curse of dimensionality, and is able to accurately predict high-dimensional distributions. This flexibility is obtained by using a hierarchical structure in the copula. The time variability of the model is provided by daily forecasts of the realized correlation matrix, which is used to estimate the structure and the parameters of the rHAC. Extensive simulation studies show the validity of the estimator based on this realized correlation matrix, and its performance, in comparison to the benchmark models. The application of the estimator to one-day-ahead Value at Risk (VaR) prediction using high-frequency data exhibits good forecasting properties for a multivariate portfolio.

Thursday 27
14:00-16:00
Lecture room 1
(Sali 1)

Sentiment spillover effects for US and European companies

Anastasija Teterova and Francesco Audrino

University of St Gallen (Switzerland)

The fast-growing literature on the news and social media analysis provide empirical evidence that the financial markets are often driven by information rather than facts. However, the direct effects of sentiments on the returns are of main interest. In this paper, we propose to study the cross-industry influence of the news for a set of US and European stocks. The graphical Granger causality of the news sentiments - excess return networks is estimated by applying the adaptive Lasso procedure. We introduce two characteristics to measure the influence of the news coming from each sector and analyze their dynamics for a period of 10 years ranging from 2005 to 2014. The results obtained provide insight into the news spillover effects among the industries and the importance of sentiments related to certain sectors during periods of financial instability.

Thursday 27
14:00-16:00
Lecture room 5
(Sali 5)

Change-point tests and the bootstrap under long-range dependence

Johannes Tewes

Ruhr-University Bochum (Germany)

We consider the change-point problem for the marginal distribution of sub-ordinated Gaussian processes that exhibit long-range dependence. The asymptotic distributions of Kolmogorov-Smirnov- and Cramér-von Mises type statistics are investigated under local alternatives. By doing so we are able to compute the asymptotic relative efficiency (ARE) of the mentioned tests and the CUSUM test. In the special case of a mean-shift in Gaussian data it is always 1. Moreover our theory covers the scenario where the Hermite rank of the underlying process changes. As the asymptotic behavior depends on a nuisance parameter, we also consider the block bootstrap. It is known that this bootstrap might fail under long memory. We show that even in this case it can be used to estimate the unknown parameter.

References: [1] Tewes, J., 2015. Change-point tests under local alternatives for long-range dependent processes. arXiv:1506.07296 [2] Tewes, J., 2016. Block bootstrap for the empirical process of long-range dependent data. arXiv:1601.01122

Nonuniform random geometric graphs with location-dependent radii

Debleena Thacker¹ and Srikanth K. Iyer

¹Lund University (Sweden)

22 Nov
2:15pm
R104

Inspired by wireless networks, like the mobile ad-hoc networks and sensor networks, we suggest a distribution free approach to the study of random geometric graphs. The distribution of vertices follows a Poisson point process with intensity function $nf(\cdot)$, where $n \in \mathbb{N}$, and f is a probability density function on \mathbb{R}^d . It has been observed that for non-uniform distribution of nodes, it is less efficient to use the same cut-off radius of connection for all nodes from the viewpoint of maximising spatial reuse. This consideration lead us to study location-dependent radii. A vertex located at x connects via directed edges to other vertices that are within a *cut-off* distance $r_n(x)$. We prove strong law results for (i) the critical cut-off function so that almost surely, the graph does not contain any node with out-degree zero for sufficiently large n and (ii) the maximum and minimum vertex degrees. We also provide a characterization of the cut-off function for which the number of nodes with out-degree zero converges in distribution to a Poisson random variable. We illustrate this result for a class of densities with compact support that have at most polynomial rates of decay to zero. Finally, we state a sufficient condition for an enhanced version of the above graph to be almost surely connected eventually.

The xyz algorithm for fast interaction search

Gian-Andrea Thanei¹, Nicolai Meinshausen¹ and Rajen Shah²

¹ETH Zürich (Switzerland), ²Cambridge University (UK)

Friday 28
10:30-12:30
Great hall
(Juhlasali)

When performing regression on a dataset with p variables, it is often of interest to go beyond using main linear effects and include interactions as products between individual variables. For small-scale problems, these interactions can be computed explicitly but this leads to a computational complexity of at least $O(p^2)$ if done naively. This cost can be prohibitive if p is very large. We introduce a new randomised algorithm that is able to discover interactions with high probability and under mild conditions has a runtime that is subquadratic in p . We show that strong interactions can be discovered in almost linear time, whilst finding weaker interactions requires $O(p^\alpha)$ operations for $1 < \alpha < 2$ depending on their strength. The underlying idea is to transform interaction search into a closestpair problem which can be solved efficiently in subquadratic time. The algorithm is called xyz and is implemented in the language R. We demonstrate its efficiency for application to genome-wide association studies, where more than 10^{11} interactions can be screened in under 280 seconds with a single-core 1.2 GHz CPU

Change-point analysis for the coefficients of variation based on empirical likelihood

Yubin Tian and Xia Cai

Beijing Institute of Technology (China)

Monday 24
14:00-16:00
Lecture room 7
(Sali 7)

The coefficient of variation is a reliability characteristic which is used to describe the failure mechanism in accelerated life tests. A novel procedure based on the empirical likelihood is proposed to detect the change point in the coefficients of variation for the multi-sample data. The empirical likelihood test statistic is derived, and the asymptotic null distribution is proven to be extreme value distribution. The performance of the test is studied through simulations. The proposed method is applied to the data of MOS transistors in the power distribution system of aircrafts of Chinese Tiangong Series to illustrate the detection procedure.

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

Robust variable selection in restricted linear regression model

Onur Toka¹, Yetkin Tuac², Meral Çetin¹ and Olcay Arslan²

¹Hacettepe University (Turkey), ²Ankara University (Turkey)

Using prior information in the linear regression is very handy method to provide more efficient estimators of regression coefficients. Restricted regression estimation is a useful method when prior information is available. However, restricted linear regression estimation is spoiled when there are outliers. In this study, we will consider robust parameter estimation methods and the variable selection simultaneously in non-stochastic restricted linear regression model, using least absolute shrinkage and selection operator (LASSO) method. In the application part, there is a small simulation study to illustrate the performance of the proposed method for dealing with the variable selection and the robust parameter estimation in restricted linear regression models. In addition, the proposed method is also applied to most-known variable selection data.

Monday 24
10:30-12:30
Lecture room 1
(Sali 1)

Radon-Nikodym derivatives, RKHS and functional classification

José Luis Torrecilla¹, José R. Berrendero² and Antonio Cuevas²

¹Universidad Carlos III de Madrid (Spain), ²Universidad Autónoma de Madrid (Spain)

Supervised classification is the classical problem of identifying the population to which a new observation belongs. In the functional setting, these populations consist of trajectories generated by different stochastic processes. In some relevant cases, the problem is to decide between two different Gaussian processes with absolutely continuous measures. In that cases, the optimal classification rule can be explicitly calculated by using the corresponding Radon-Nikodym derivatives. The expressions of these derivatives can be obtained from some classical results under both homoscedasticity and heteroscedasticity. Here, we provide some of these expressions and study some problems when estimating the rules and connections with variable selection.

Monday 24
14:00-16:00
Auditorium IV

Local stationarity and time-inhomogeneous Markov chains

Lionel Truquet

ENSAI (France)

In this paper, we revisit a notion of local stationarity for fitting time-inhomogeneous Markov chains models. Our approach, which is not restricted to continuous state spaces, allows to consider a variety of models, including most of the locally stationary autoregressive time series studied in the literature. We consider triangular arrays of time-inhomogeneous Markov chains, defined by some families of contracting and slowly-varying Markov kernels. Using the Dobrushin's contraction coefficients for adapted probability metrics, we show that the distribution of such Markov chains can be approximated locally with the distribution of ergodic Markov chains. Mixing properties of these triangular arrays are also discussed. As a consequence of our results, some classical geometrically ergodic homogeneous Markov chains models have a locally stationary version. In particular, we consider new statistical models such as finite state space Markov chains with a time-varying transition matrix or some integer-valued autoregressive processes with time-varying coefficients. For these two examples, nonparametric kernel estimation of the transition matrix is studied.

The distribution of a perpetuity

Alexandra Tsimbalyuk and Kostas Kardaras
London School of Economics (UK)

Thursday 27
14:00-16:00
Lecture room 7
(Sali 7)

We consider the problem of estimating the joint distribution of a perpetuity, e.g. a cumulative discounted loss process in the ruin problem, and the underlying factors driving the economy in an ergodic Markov model. One can identify the distribution in two manners: first, as the explosion probability of a certain locally elliptic (explosive) diffusion; second, using results regarding time reversal of diffusions, as the invariant measure associated to a certain (different) ergodic diffusion. These identifications enable efficient estimation of the distribution through both simulation and numerical solutions of partial differential equations. When the process representing the value of an economic factor is one-dimensional, or more generally reversing, the invariant distribution is given in an explicit form with respect to the model parameters. In continuous time and a general multi-dimensional setup, the lack of knowledge of the invariant distribution could pose an issue. In my talk I will show how one can amend the situation in the discrete-time case.

Sample size determination using Poisson for comparing two count populations that exhibit over-disper

Tsung-Shan Tsou
National Central University (Taiwan)

Wednesday 26
10:30-12:30
Lecture room 7
(Sali 7)

We propose Poisson as a more convenient and useful model for sample size calculation for comparing two populations of count that exhibit over-dispersion. One can expediently fulfill the mission of sample size determination for power analysis without dealing with the dispersion parameter of the negative binomial. We use simulations to make contrast between the new approach with a recently proposed method whose application is quite limited by the assumptions that are less practical.

Robust variable selection in an autoregressive error term regression model

Yetkin Tuuç¹, Yeşim Güney¹, Şenay Özdemir² and Olcay Arslan¹
¹Ankara University (Turkey), ²Afyon Kocatepe University (Turkey)

Friday 28
10:30-12:30
Lecture room 8
(Sali 8)

In regression models, one of the challenging problems occurs when the error terms are dependent. In this case the estimators for the regression parameters obtained without taking care of the depended structure will have incorrect variance estimates. To deal with the dependent errors the autoregressive structure is commonly used. Another challenging problem in a regression analysis is the variable selection, which can be carried out simultaneously with the parameter estimation using LASSO type methods. In a regression model with autoregressive error terms the variable selection is combined with the parameter estimation using the LASSO and the Bridge method given in the paper by Yoon J.Y., et.al (2012). However, in their paper they used the ordinary least square (OLS) estimation method to obtain the estimators and to select the variables. It is well known that the OLS estimation method is very sensitive to the outliers the results will be poor in terms of estimation and variable selection. In order to combat with the outliers and the variable selection in a regression model with autoregressive errors we propose to combine the MM regression estimation method with the Bridge regression estimation method. To illustrate the performance of the proposed method a small simulation study and real data examples are provided. Key Words: model selection; parameter estimation; regression; robustness.

Reference: Yoon, J.Y., et.al, 2012. Penalized Regression Models with Autoregressive Error Terms. *Journal of Statistical Computation and Simulation*, 83:9, 1756-1772.

Monday 24
14:00-16:00
Great hall
(Juhlasali)

Learning Bayesian networks based on sparsest permutations

Caroline Uhler
MIT (USA)

We consider the problem of learning a Bayesian network or directed acyclic graph (DAG) model from observational data. We propose the sparsest permutation algorithm, a nonparametric approach based on finding the ordering of the variables that yields the sparsest DAG. We prove consistency of this method under strictly weaker conditions than usually required. We discuss how to find the sparsest ordering by introducing the DAG associahedron and a simplex-type algorithm on this convex polytope. We end with discussing some applications to estimating gene regulatory networks.

Thursday 27
10:30-12:30
Lecture room 13
(Sali 13)

Lasso penalization in the framework of generalized additive models for location, scale and shape

Nikolaus Umlauf¹, Andreas Groll² and Thomas Kneib²
¹University of Innsbruck (Austria), ²University of Goettingen (Germany)

We propose a regularization approach for high dimensional data setups in the generalized additive model for location, scale and shape (GAMLSS) framework. It is designed for linear covariate effects and is based on L^1 -type penalties. The following three penalization options are provided: the conventional least absolute shrinkage and selection operator (Lasso) for metric covariates, and both group and fused Lasso for categorical predictors. An implementation is provided in the R package `bamlss` (<https://cran.r-project.org/package=bamlss>).

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

Rank-based permutation approaches for nonparametric factorial designs

Maria Umlauf¹, Frank Konietzschke² and Markus Pauly¹
¹Ulm University (Germany), ²University of Texas at Dallas (USA)

In our work, we develop an exact rank-based permutation test for factorial designs. Typical restrictions of standard methods for analysing such data are equal variances across groups, balanced sample sizes or normally distributed error terms. However, these assumptions are often not met in real data. One alternative is given by the well-known Wald-type statistic (WTS) in semiparametric models, which also yields to asymptotically exact results in cases of heteroscedasticity and no particular error distribution. Recently, a modified permutation test based on the WTS has been developed to improve its small sample properties, see Pauly et al. (2015). However, if ordinal or ordered categorical data are present, the classical parametric and semiparametric models show their limits, since calculated means are neither meaningful nor suitable. Consequently, the aim of our work is to extend the Wald-type permutation statistic (WTPS) to a rank-based WTPS and therefore to nonparametric factorial designs allowing for all kind of observations such as discrete, ordinal or count data. Despite the difficulty of dependencies it is shown that this rank-based approach is also approximately exact and consistent. The small sample behaviour of the nonparametric Wald-type permutation statistic is compared to other well-known techniques in an extensive simulation study.

References: Pauly, M., Brunner, E., Konietzschke, F., 2015. Asymptotic permutation tests in general factorial designs. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 77(2), 461–473.

Nonparametric goodness-of-fit test of the baseline hazard function in the Cox model

Lecleire Uriel¹, Cristina Butucea², Thierry Jeantheau³, Émilie Gerardin⁴

¹UPEM (France), ²CREST, ENSAE, Université Paris-Saclay (France), ³LAMA (UMR 8050), UPEM, UPEC, CNRS, ⁴Safran Aircraft Engines (France)

Thursday 27
10:30-12:30
Lecture room 10
(Sali 10)

We observe for n independent individuals: the right-censored times X_i , their failure indicators δ_i , and a vector of p covariates Z^i . The hazard rate for the uncensored times is $h(t|Z) = h(t) \exp(g(Z))$, with h a non specified baseline hazard function. We consider the goodness-of-fit test that has a simple null hypothesis $h \equiv h_0$ for some given function h_0 . The alternative is expressed in a nonparametric setup and separated from h_0 in a weighted \mathbb{L}_2 norm. We use a projection on a orthonormal basis of this weighted \mathbb{L}_2 space to estimate the distance and build a statistic for the goodness-of-fit test. As the weight heavily depends on the distribution of the censored times and of the covariates, as well as on the function g in our model, we study the influence of plugging-in estimators of these functions on the error probabilities of our test procedure. We show that our rates for testing are faster than the rates for estimation in the same model. Thus, we provide a nonparametric test procedure for the baseline hazard function in the Cox model.

Jackknife empirical likelihood inference for the ROC curve

Jelena Valkovska and Janis Valeinis
University of Latvia (Latvia)

Thursday 27
10:30-12:30
Lecture room 10
(Sali 10)

The receiver operating characteristic (ROC) curves is a popular tool for diagnostic tests accuracy, it is broadly used in medicine, biometrics etc., it is increasingly used in machine learning and data mining researches. In many situations the diagnostic decisions are not limited with a binary choice, so the notions of ROC curves and area under the curve (AUC) were extended for the tests with three diagnostics categories - the ROC surface and the volume under the surface (VUS). Nowadays popular methods for constructing confidence intervals for ROC curves are: empirical likelihood (EL), jackknife empirical likelihood (JEL) and smoothed jackknife empirical likelihood (SJEL). Many of them are already implemented for ROC curves and ROC surfaces, whole and partial area under the ROC curve and their differences. Some new results will be presented regarding SJEL method for partial AUC, SJEL for the volume under the surface (VUS) and all the methods for partial VUS.

Empirical Bayes learning in high-dimensional prediction settings

Mark van de Wiel
VU University medical center (Netherlands)

Tuesday 25
14:00-16:00
Lecture room 13
(Sali 13)

Empirical Bayes enables 'learning from a lot' in two ways: first, from a large number of variables and second, from a potentially large amount of prior information on the variables, termed 'co-data'. We discuss empirical Bayes methods for estimating hyperparameters (e.g. penalties) in regression-based prediction models, such as spike-and-slab, the elastic net and ridge. For the latter, some theoretical results are presented for the quality of an empirical Bayes estimator in terms of n and p . Empirical Bayes is particularly useful to estimate multiple hyper-parameters that model the information in the co-data. Some examples of co-data are: p-values from an external study, additional molecular measurements or genomic annotation. The systematic use of co-data can considerably improve predictions and variable selection, which we demonstrate on a cancer genomics application.

Monday 24
14:00-16:00
Lecture room 5
(Sali 5)

Testing for stationarity of functional time series in the frequency domain

Anne van Delft¹ and Alexander Aue²

¹Université catholique de Louvain (UCL) (Belgium), ²University of California at Davis (USA)

Interest in functional time series has spiked in the recent past with papers covering both methodology and applications being published at a much increased pace. We contribute to the research in this area by proposing stationarity tests for functional time series based on frequency domain methods. Setting up the tests requires a delicate understanding of periodogram- and spectral density operators that are the functional counterparts of periodogram- and spectral density matrices in the multivariate world. Properties of the proposed statistics are derived both under the null hypothesis of stationary functional time series and under the smooth alternative of locally stationary functional time series. The methodology is theoretically justified through asymptotic results. Evidence from simulation studies and an application to annual temperature curves suggests that the tests work well in finite samples.

Thursday 27
10:30-12:30
Lecture room 12
(Sali 12)

How many needles in the haystack? Adaptive uncertainty quantification for the horseshoe

Stéphanie van der Pas, Botond Szabó and Aad van der Vaart
Leiden University (Netherlands)

We investigate the credible sets and marginal credible intervals resulting from the horseshoe prior in the sparse multivariate normal means model. We do so in an adaptive setting without assuming knowledge of the sparsity level (number of signals). We consider both the hierarchical Bayes method of putting a prior on the unknown sparsity level and the empirical Bayes method with the sparsity level estimated by maximum marginal likelihood. We show that credible balls and marginal credible intervals have good frequentist coverage and optimal size if the sparsity level of the prior is set correctly. By general theory honest confidence sets cannot adapt in size to an unknown sparsity level. Accordingly the hierarchical and empirical Bayes credible sets based on the horseshoe prior are not honest over the full parameter space. We show that this is due to over-shrinkage for certain parameters and characterise the set of parameters for which credible balls and marginal credible intervals do give correct uncertainty quantification. In particular we show that the fraction of false discoveries by the marginal Bayesian procedure is controlled by a correct choice of cut-off.

Monday 24
10:30-12:30
Lecture room 12
(Sali 12)

A mollification approach to study the nonparametric instrumental regression model

Anne Vanhems
Toulouse Business School (France)

The overall purpose of this research is to provide new tools for the analysis of structural econometric models. Our focus will be on nonparametric procedures that permit estimation of causal effects without strong parametric assumptions and more specifically our research proposal is motivated by the analysis of structural causal models with endogeneity. Endogeneity may be due to omitted variables, measurement errors, or simultaneity and nonparametric regression with endogeneity induces an ill-posed inverse problem to solve. Our contribution is to solve this ill-posed inverse problem using a mollification approach. Mollification has been investigated for linear ill-posed equations, specifically for deconvolution in signal and image processing by Lannes, Roques, and Casanove (1987), and has been recently extended by Maréchal and Bonnefond (2009) to more general situations. We investigate this technique in an econometric context and compare the performance of our estimator with respect to the literature.

The realized empirical distribution function of volatility

Bezirgen Veliyev
Aarhus University (Denmark)

Monday 24
10:30-12:30
Lecture room 6
(Sali 6)

We construct a noise robust estimator of the distribution function of the (stationary) marginal distribution of the latent spot volatility process of an asset price. We call this estimator the realized empirical distribution function. As a first step, we show that over fixed time span our estimator is a consistent estimator of the empirical distribution function. This step relies on the asset price being observed over a grid with mesh going to zero. Noise robust estimators typically converge at a slower rate, however, a Monte Carlo study suggests that this cost is greatly outweighed by the benefits from being able to use more observations. In the second step, we let the time span tend to infinity while still letting the distance between observations tend to zero. We show that in this setting the realized empirical distribution function is a consistent estimator of the (stationary) marginal distribution function of the latent volatility process. Furthermore, we use our estimator to construct a Kolmogorov-Smirnov type test for the marginal distribution. Finally, we apply the newly developed methods to a set of ultra high-frequency equity tick data.

Gaussian process regression model for distribution inputs

Nil Venet¹, François Bachoc², Fabrice Gamboa² and Jean-Michel Loubes²

¹CEA (France), ²Institut de Mathématiques de Toulouse (France)

Friday 28
10:30-12:30
Lecture room 10
(Sali 10)

Monge-Kantorovich distances, otherwise known as Wasserstein distances, have received a growing attention in statistics and machine learning as a powerful discrepancy measure for probability distributions. In this work, we focus on forecasting a Gaussian process indexed by probability distributions. For this, we provide a family of positive definite kernels built using transportation based distances. We provide a probabilistic understanding of these kernels and characterize the corresponding stochastic processes. We prove that the Gaussian processes indexed by distributions corresponding to these kernels can be efficiently forecast, opening new perspectives in Gaussian process modelling.

Parameter estimation for Langevin equation in a general Gaussian framework

Lauri Viitasaari
Aalto University School of Science (Finland)

Monday 24
10:30-12:30
Lecture room 7
(Sali 7)

Stationary Ornstein–Uhlenbeck process X_t^θ with a parameter $\theta > 0$ can be viewed as the solution to the stochastic Langevin differential equation $dX_t^\theta = -\theta X_t^\theta dt + dW_t$, where W is a standard Brownian motion. Furthermore, it has been recently proved that each continuous Gaussian process can be viewed as a solution to the Langevin equation with certain Gaussian noise G that has stationary increments. In this talk we discuss the problem of parameter estimation θ of the Langevin equation in a general Gaussian framework. We build an estimator that is strongly consistent and that the suitably scaled error converge towards a standard normal distribution under general conditions on the noise G .

A three decision procedure for testing the type of motion of a particle trajectory

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

Myriam Vimond
Ensaï Crest (France)

State-of-the-art techniques in passive particle-tracking microscopy provide high-resolution path trajectories of diverse foreign particles in biological fluids. Then measuring confined diffusion (diffusion in a closed domain or in an open but crowned area), free diffusion (or Brownian motion) and directed diffusion (active transport) is central in cell biology since they represent the main modes of mobility of molecules in living cells. We propose a three decision statistical test procedure in order to determine the type of motion of the observed according the above classification. This method is an alternative to Mean Square Displacement analysis. We assess our procedure on both simulations and real cases.

Independent component analysis of multivariate functional data

Friday 28
10:30-12:30
Lecture room 5
(Sali 5)

Joni Virta¹ and Bing Li²

¹University of Turku (Finland), ²Pennsylvania State University (USA)

We extend a classic method of independent component analysis, the fourth order blind identification (FOBI), to vector-valued functional data. The use of multivariate instead of univariate functions allows for natural definitions for both the marginals of a random function and their mutual independence. Our model assumes that the observed functions are mixtures of latent independent functions residing in suitable Hilbert spaces \mathcal{H}_j , mixed with a bounded linear operator from the product space $\times_j \mathcal{H}_j$ to itself. To enable the inversion of the covariance operator we make the assumption that the dependency between the mixed component functions lies in a finite-dimensional subspace. In this subspace we define fourth cross-cumulant operators and use them to construct a novel Fisher consistent method for solving the independent component problem for vector-valued functions. Finally, both simulations and an application on hand gesture data set are used to demonstrate the advantages of the proposed method over its closest competitors.

On statistical estimation of parameters for a family of binary autoregressive time series

Thursday 27
14:00-16:00
Lecture room 8
(Sali 8)

Valeriy Voloshko and Yuriy Kharin
Belarusian State University (Belarus)

We consider the problem of statistical parameter estimation for binary GLM-based autoregression $x_t \in \{0, 1\}$ of order s :

$$G(\mathbf{P}\{x_t = 1 | \mathcal{F}_{t-1}\}) = a_1 \psi_1(X_{t-s}^{t-1}) + \dots + a_m \psi_m(X_{t-s}^{t-1}),$$

where $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_m) \in R^m$ is an m -dimensional parameter. The link function $G(\cdot)$ is assumed to be of rather general form, and the base functions $\{\psi_i(\cdot)\}_{i=1}^m$ are nonlinear w.r.t. the s lagged variables $X_{t-s}^{t-1} = (x_{t-1}, \dots, x_{t-s})$. We propose the new frequencies-based estimator (FBE) for the parameter \mathbf{a} :

$$\hat{\mathbf{a}} = \hat{\mathbf{a}}(X_1^T) = \langle \Psi \rangle_H^{-1} \Psi H \mathbf{G}(\hat{\Theta}),$$

where H is some real-valued non-negative definite symmetric weight $(2^s \times 2^s)$ -matrix, $\langle \Psi \rangle_H$ is the nonsingular Gram $(m \times m)$ -matrix of the base functions $\{\psi_i(\cdot)\}_{i=1}^m$ w.r.t. the kernel H , $\hat{\Theta} = (\hat{\theta}_q)_{q \in \{0,1\}^s}$ is the 2^s -vector of the frequencies

$$\hat{\theta}_q = \frac{\sum_{t=s+1}^T \mathbf{1}\{X_{t-s}^{t-1} = q, x_t = 1\}}{\sum_{t=s+1}^T \mathbf{1}\{X_{t-s}^{t-1} = q\}},$$

where $\mathbf{G}(\hat{\Theta}) = (G(\hat{\theta}_q))_{q \in \{0,1\}^s}$, $\mathbf{1}\{\cdot\}$ is the indicator function. In comparison with the maximum likelihood estimator (MLE) the FBE is explicitly computed (without iterative numerical methods and related difficulties), supports effective recursive re-estimation under the model extension, and has less restrictive sufficient conditions of uniqueness (does not need log-concavity of the inverse link). We show that in addition to the above advantages, the FBE may be highly effective (i.e. may have the variance close to the Cramer-Rao bound) when using the data-adaptive weight matrix $H = H(X_1^T)$. The computer experiments are performed on simulated and real (genetic) data.

On estimation and model calibration of arbitrary stationary processes

Poster session

Marko Voutilainen, Lauri Viitasaari, Pauliina Ilmonen

Aalto University School of Science (Finland)

Stationary processes have been extensively studied in the literature. Their applications include modelling and forecasting numerous real life phenomena such as natural disasters, sales and market movements.

In [1], we present a new way of fitting a model to a data that is assumed to be a realization from a stationary process. The approach taken is based on the characterization of stationary processes given in [2] yielding an AR(1) representation for an arbitrary stationary process. Although the representation in general is not unique, we show that, in most cases, only one value of the autocovariance function of the noise is enough to provide uniqueness.

The approach we adopt has several advantages when compared, for example, to popular ARMA modelling. When modelling existing data with processes of the ARMA family, the first step is to fix the orders of the model. After that, one can estimate the related parameters by using, for example, the maximum likelihood (ML) or least squares (LS) estimators. The final step is to conduct various diagnostic tests to determine whether the estimated model is sufficiently good or not. ARMA fitting assumes that the error terms are iid or white noise. That is not the case with AR(1) representation. When founding estimation on the AR(1) representation, one does not have to choose between different models. In addition, there is only one parameter to estimate. Yet another advantage over the conventional ARMA estimation is that we obtain a closed form for the estimator, whereas in general, ML and LS estimators do not admit closed form representations.

The new way to model arbitrary stationary processes led us to propose a simple estimator of the parameter of the (unique) AR(1) representation. In [1], we prove consistency and asymptotic normality of this estimator and we provide a simulation study to illustrate its excellent finite sample properties.

References: [1] Voutilainen, M., Viitasaari, L., Ilmonen, P. On estimation and model calibration of arbitrary stationary processes. Manuscript in preparation. [2] Viitasaari, L., 2016. Representation of stationary and stationary increment processes via Langevin equation and self-similar processes. *Statistics & Probability Letters*, 115, 45–53.

Approximate residual balancing: De-biased inference of average treatment effects in high dimensions

Stefan Wager¹, Susan Athey and Guido W. Imbens

¹Stanford University (USA)

Monday 24
14:00-16:00
Lecture room 13
(Sali 13)

There are many settings where researchers are interested in estimating average treatment effects and are willing to rely on the unconfoundedness assumption, which requires that the treatment assignment be as good as random conditional on pre-treatment variables. The unconfoundedness assumption is often more plausible if a large number of pre-treatment variables are included in the analysis, but this can worsen the performance of standard approaches to treatment effect estimation. In this paper, we develop a method for de-biasing penalized regression adjustments to allow sparse regression methods like the lasso to be used for \sqrt{n} -consistent inference of average treatment effects. Our method works under substantially weaker assumptions than other methods

considered in the literature: Unlike high-dimensional doubly robust methods recently developed in econometrics, we do not need to assume that the treatment propensities are estimable, and unlike existing de-biasing techniques from the statistics literature, our method is not limited to considering sparse contrasts of the parameter vector. Instead, in addition standard assumptions used to make lasso regression on the outcome model consistent under 1-norm error, we only require overlap, i.e., that the propensity score be uniformly bounded away from 0 and 1. Procedurally, our method combines balancing weights with a regularized regression adjustment.

”Standard” fully modified OLS estimation of cointegrating polynomial regressions

Martin Wagner, Oliver Stypka, Peter Grabarczyk and Rafael Kawka
Technical University Dortmund (Germany)

Friday 28
10:30-12:30
Lecture room 10
(Sali 10)

In the empirical environmental Kuznets curve (EKC) literature, which investigates a potentially inverted U-shaped relationship between measures of economic development and pollution respectively emissions, it is common practice to use *standard* cointegration methods (see, e.g., Wagner, 2015). To exemplify the issues, consider a typical parametric (quadratic) formulation of an EKC-regression given by:

$$y_t = c + \delta t + \beta_1 x_t + \beta_2 x_t^2 + u_t, \quad (1)$$

$$= Z_t' \theta + u_t,$$

$$x_t = x_{t-1} + v_t, \quad (2)$$

with y_t denoting the measure of (per capita) pollution or emissions (in logarithms), x_t denoting, typically, the logarithm of per capita GDP, $Z_t = [1, t, x_t, x_t^2]'$ and θ the stacked parameter vector. In the above setting $\eta_t := [u_t, v_t]'$ is considered to be a jointly stationary (not over-differenced) process, with the assumptions detailed in the paper. Thus, the regressor x_t is an integrated process of order 1, in short an I(1) process. Therefore, the above equation contains as regressors, in addition to the deterministic components, an I(1) regressor and its square. The latter is not an integrated process of any order (see, e.g., Wagner, 2012). In the empirical literature the above-mentioned fact that powers of integrated processes are not themselves integrated processes has been, by and large, neglected and “linear” cointegration methods are typically applied, in other words, both x_t and x_t^2 are considered as I(1) processes and estimators designed for asymptotically valid inference like the Fully Modified OLS (FM-OLS) estimator of Phillips and Hansen (1990) are commonly applied as if there simply were two integrated regressors included in the regression. Thus, the following two-step estimation is performed: First, define $\hat{\eta}_t^L := [\hat{u}_t, w_t'] := [\hat{u}_t, \Delta x_t, \Delta x_t^2]'$, with \hat{u}_t denoting the OLS residuals from estimating (1) and Δ the first-difference operator. For the vector process $\hat{\eta}_t^L$ define (formal – since $\hat{\eta}_t^L$ is not stationary) long-run and half long-run variance estimators as $\hat{\Omega}^L := T^{-1} \sum_{i=1}^T \sum_{j=1}^T k\left(\frac{|i-j|}{M}\right) \hat{\eta}_i^L \hat{\eta}_j^{L'}$ and $\hat{\Delta}^L := T^{-1} \sum_{i=1}^T \sum_{j=i}^T k\left(\frac{|i-j|}{M}\right) \hat{\eta}_i^L \hat{\eta}_j^{L'}$, for chosen kernel function $k(\cdot)$ and bandwidth M . Next define $y_t^{L,+} := y_t - w_t' \left(\hat{\Omega}_{ww}^L\right)^{-1} \hat{\Omega}_{wu}^L$, with the subscripts corresponding to the partitioning of $\hat{\eta}_t^L$. The “linear” FM-OLS estimator of θ is then defined as:

$$\hat{\theta}^{L,+} := \left(\sum_{t=1}^T Z_t Z_t' \right)^{-1} \left(\sum_{t=1}^T Z_t y_t^{L,+} - \begin{pmatrix} \mathbf{0}_{2 \times 1} \\ T \hat{\Delta}_{wu}^{L,+} \end{pmatrix} \right), \quad (3)$$

with $\hat{\Delta}_{wu}^{L,+} := \hat{\Delta}_{wu}^L - \hat{\Delta}_{ww}^L \left(\hat{\Omega}_{ww}^L\right)^{-1} \hat{\Omega}_{wu}^L$. Note that both (the last component of w_t) as well as $\hat{\Omega}^L$ and $\hat{\Delta}^L$ diverge (under usual assumptions on kernel and bandwidth). A *proper* FM-OLS estimator for the considered quadratic equation has been considered in Wagner and Hong (2016) by taking – which is the difference – into account that x_t^2 is a nonlinear function of x_t and thus v_t . Define $\hat{\eta}_t = [\hat{u}_t, v_t]'$ and similar as above $\hat{\Omega}$ and $\hat{\Delta}$, which are in the example considered here given by the upper 2-2 blocks of $\hat{\Omega}^L$ and $\hat{\Delta}^L$. Furthermore define $y_t^+ := y_t - v_t \hat{\Omega}_{vv}^{-1} \hat{\Omega}_{vu}$, with

the subscripts indicating the partitioning of $\hat{\Omega}$, and $\hat{\Delta}_{vu}^+ := \hat{\Delta}_{vu} - \hat{\Delta}_{vv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}$. This leads to the FM-OLS estimator of θ defined as:

$$\hat{\theta}^+ := \left(\sum_{t=1}^T Z_t Z_t' \right)^{-1} \left(\sum_{t=1}^T Z_t y_t^+ - \hat{\Delta}_{vu}^+ \begin{pmatrix} \mathbf{0}_{2 \times 1} \\ T \\ 2 \sum_{t=1}^T x_t \end{pmatrix} \right). \quad (4)$$

We show in this paper that (under appropriate assumptions on kernel and bandwidth) the “linear” estimator defined in (3) has, surprisingly, the same limiting distribution as the (proper) FM-OLS estimator defined in (4). We also show by means of simulations that inference based on (4) outperforms inference based on (3) in terms of both lower size distortions and higher power. The theoretical results hold, of course, not only for the illustrative quadratic specification but for general cointegrating polynomial regressions.

Pairwise ranking and crowd-sourcing: Statistical models and computational challenges

Martin Wainwright¹, Nihar Shah, Sivaraman Balakrishnan and Aditya Guntuboyina

¹University of California at Berkeley (USA)

Monday 24
9:15-10:15
Great hall
(Juhlasali)

Many modern data sets take the form of pairwise comparisons, in which binary judgements are made about pairs of items. Some examples include the outcomes of matches between tennis players, ratings of the relevance of search queries, and the outputs of crowd-sourcing engines. We discuss some statistical models for modeling data of this type, along with the computational challenges that arise in performing estimation and rank aggregation with such models.

The use of Bayesian additive regression trees in genomics

Patrik Waldmann

The Swedish University of Agricultural Sciences, SLU (Sweden)

Monday 24
10:30-12:30
Lecture room 5
(Sali 5)

Bayesian Additive Regression Trees (BART) is a sum-of-trees method where the size and effect of each regression tree is regularized to be small through three different priors. A regression tree consists of three components: a tree structure with internal nodes, decision rules and a set of terminal nodes. The regression trees partition the predictor space into an unknown function which automatically will model nonlinearities within (dominance) and interactions between (epistasis) genomic markers (i.e. SNPs). Hence, BART is a non-parametric and non-linear method that has the important feature that it can handle all forms of genetic effects of SNPs in a very sparse way. Bayesian back-fitting together with a simple MCMC algorithm makes BART computationally efficient. The aim here is to show that BART not only has excellent prediction properties, but also can identify important SNPs and their genetic mode through variable importance and partial dependence measures.

Constructing the essential histogram

Guenther Walther¹, H. Li, A. Munk, and H. Sieling

¹Stanford University

Tuesday 25
10:30-12:30
Lecture room 6
(Sali 6)

The aim of a histogram is to provide a simple visualization of the empirical distribution. One way to formalize this task is as follows: Construct the simplest histogram (i.e. the histogram with the fewest jumps) that provides a good estimate of the empirical distribution (i.e. probabilities are estimated optimally over all intervals). We present results towards the construction of such an ‘essential histogram’. The procedure is based on distribution-free likelihood ratios defined on intervals from a certain collection of intervals. This collection is constructed to allow efficient computation as well as statistically optimal inference.

Monday 24
14:00-16:00
Lecture room 6
(Sali 6)

Applications of distance correlation to time series

Phyllis Wan
Columbia University (USA)

The use of empirical characteristic functions for inference problems, including estimation in some special parametric settings and testing for goodness of fit, has a long history dating back to the 70s (see for example, Feuerverger and Mureika (1977), and Csörgő (1981)). More recently, there has been renewed interest in using empirical characteristic functions in other inference settings. The distance covariance and correlation, developed by Székely and Rizzo (2007) for measuring dependence and testing independence between two random vectors, are perhaps the best known illustrations of this. We apply these ideas to stationary univariate and multivariate time series to measure lagged auto- and cross-dependence in a time series. Assuming strong mixing, we establish the relevant asymptotic theory for the empirical auto- and cross-distance correlation functions. We also apply the auto-distance correlation function (ADCF) to the residuals of an autoregressive processes as a test of goodness of fit. Under the null that an autoregressive model is true, the limit distribution of the empirical adcf can differ markedly from the corresponding one for an i.i.d. sequence.

Monday 24
10:30-12:30
Great hall
(Juhlasali)

Semiparametric analysis of transformation models for doubly-censored clustered survival data

Jane-Ling Wang¹, Cong Xu² and Yu-Ru Su³

¹University of California at Davis (USA), ²Stanford University (USA), ³Fred Hutchinson Cancer Research Center (USA)

Doubly censored data, which occurs when survival times are subject to both left- and right-censoring, occurs frequently in lifetime studies if subjects were not recruited at the onset time of the study. The approaches to handle double-censoring mechanism are very different from those for right-censoring only, and poses great challenges to their statistical inference. A further challenge occurs when data are obtained from familial-type studies, where data from several or all members of a family are collected and observations from the same family are potentially correlated; thus results in clustered survival data. In this paper, we consider a class of linear transformation models, which includes the proportional hazards and the proportional odds models as special cases, on clustered survival data subject to double-censoring. While the transformation models expand the horizon of survival models, they trigger considerable computational challenges to likelihood approaches, especially in the presence of double-censorship and correlation. An effective expectation-maximization (EM) algorithm is developed to overcome the computational difficulties and yields stable nonparametric maximum likelihood estimates (NPMLEs). We investigate the asymptotic properties of the NPMLEs, and demonstrate the semiparametric efficiency of the finite-dimensional parameters. Besides, a computationally efficient method is proposed to estimate the standard errors (SEs) of the NPMLEs. Simulation studies demonstrate that the proposed EM algorithm and SE estimates perform well. Both are then applied to a dataset from a Hepatitis B clinical study in Taiwan.

Bayesian aggregation of average data

Sebastian Weber¹, Andrew Gelman², Daniel Lee², Michael Betancourt², Aki Vehtari³ and Amy Racine-Poon¹

¹Novartis Pharma AG (Switzerland), ²Columbia University (USA), ³Aalto University (Finland)

How can we combine datasets collected under different conditions? This is the well-known problem of meta-analysis, for which Bayesian methods have long been used to achieve partial pooling. Here we consider the challenge when one dataset is given as raw data while the second dataset is given as averages only. In such a situation, common meta-analytic methods can only be applied when the model is sufficiently simple for analytic approaches. When the model is too complex, for example nonlinear, an analytic approach is not possible. The need for meta-analytic methods applied to complex models arises frequently in clinical drug development. Throughout the different phases of a drug development program, randomized trials are used to establish in stages the tolerability, safety, and efficacy of a candidate drug. At each stage one aims to optimize the design of future studies by extrapolation from the available evidence at the time. This includes collected trial data and relevant external data. However, relevant external data are typically available as averages only, for example from trials on alternative treatments reported in the literature. Moreover, realistic models suitable for the desired extrapolation are often complex (longitudinal and nonlinear). We provide a Bayesian solution by using simulation to approximately reconstruct the likelihood of the external summary and allowing the parameters in the model to vary under the different conditions. We first evaluate our approach using fake-data simulations and then demonstrate its application to the problem in drug development that motivated this research, a hierarchical nonlinear model in pharmacometrics, implementing the computation in Stan.

Tuesday 25
14:00-16:00
Great hall
(Juhlasali)

Bootstrap and change-point detection in functional time series and random fields

Martin Wendler, Beatrice Bucchia, Olimjon Sh. Sharipov and Johannes Tewes
Universität Greifswald (Germany)

Tuesday 25
14:00-16:00
Lecture room 1
(Sali 1)

We consider sequences or fields of random variables taking their value in a Hilbert space. The random variables are assumed to be weakly dependent, meaning that they fulfill some mixing conditions. New test for the hypothesis of stationary are proposed. The alternative is at most one change-point in the case of a time series, or an rectangular change-set in the case of random fields. The asymptotic distribution of the test statistics is obtained with the continuous mapping theorem from new functional central limit theorems for the partial sum process in a Hilbert space. Because the limit distribution is difficult to evaluate and depends on a high-dimensional, difficult to estimate variance parameter, we propose to use bootstrap methods. In the case of time series, we study the nonoverlapping block bootstrap, in the case of random fields the dependent wild bootstrap and show the validity of these methods. In a simulation study, we will show that our test outperforms previous proposals based on dimension reduction. Our new test can also be used to detect arbitrary change in the distribution function of real valued observations.

Empirical risk minimization as parameter choice rule for general linear regularization methods

Monday 24
14:00-16:00
Lecture room 10
(Sali 10)

Frank Werner¹ and Housen Li²

¹Max Planck Institute for Biophysical Chemistry (Germany), ²National University of Defense Technology, Changsha (China)

We consider a-posteriori parameter choice rules for filter based linear regularization methods in the statistical inverse problem setting. In particular, we investigate the choice of the regularization parameter by minimizing an unbiased estimate of the predictive risk. This parameter choice rule and its usage are well-known in the literature, but oracle inequalities and optimality results in this general setting are unknown. We prove a (generalized) oracle inequality, which relates the direct risk with the minimal prediction risk. From this oracle inequality, we are then able to conclude that the filter based regularization methods with the investigated parameter choice rule achieve minimax convergence rates with respect to the mean integrated squared error. Finally, we also present numerical simulations, which support the order optimality of the method and the quality of the parameter choice in finite sample situations.

Optimal scaling for survival analysis with ordinal data

Friday 28
10:30-12:30
Lecture room 13
(Sali 13)

Sanne Willems^{1,2}, Marta Fiocco³ and Jacqueline Meulman¹

¹Leiden University (Netherlands), ²Leiden University Medical Center (Netherlands)

Studies often involve measurement and analysis of categorical data with nominal or ordinal category levels. Nominal categories have no ordering property, e.g. gender. Ordinal category levels, however, have an ordering, e.g. education levels classified as low, medium, and high. When analysing survival data, currently two methods can be chosen to include ordinal covariates in the Cox proportional hazard model. One can use dummy covariates to indicate category memberships. Estimated parameters for each category indicate the increase or decrease in risk of experiencing the event of interest compared to the reference category. Since these parameters are estimated independently from each other, the ordering property of the categories is lost in the process. To keep the ordinal property, one can give integer values to the covariate's categories (e.g. low = 0, medium = 1, high = 2), and include it in the model as a numeric covariate. However, now the ordinal data are interpreted as numeric data, so the property of equal distances between consecutive categories is introduced. This assumption is too strict for this data type; distances between consecutive categories do not necessarily have to be equal. We developed a method to include ordinal data in the Cox model without losing the ordering of category levels, or introducing equal distances between them. A simulation study was carried out to compare the performance of this method with the performance of the two currently used methods. Results show that the new method increases the model fit in case of ordinal data.

Bayesian inference for modeling of protein folding data

Poster session

Samuel Wiqvist¹, Umberto Picchini¹ and Julie Lyng Forman²

¹Lund University (Sweden), ²University of Copenhagen (Denmark)

Modelling of protein folding is an important topic in biophysics. The inference problem for the model parameters is however often challenging, because of the size of the data, the nonlinear dynamics and the presence of local perturbations requiring computer intensive Monte Carlo methods. It is therefore of interest to construct a simple model for protein folding, useful for both frequentist and Bayesian inference. We introduced a double-well potential stochastic differential equation model with additive red-noise that has a decreased computational complexity, allowing us to generate data from the model more rapidly compared to an earlier model. But the inference problem for this model is still challenging since the likelihood function is intractable. However, by utilizing recent developments for Monte Carlo based algorithms we could resort to particle

Monte Carlo based methods to sample from the posterior distribution of the parameters or from a surrogate model of the likelihood function based on Gaussian processes. Simulation studies show that the proposed double-well potential stochastic differential equation model fits the protein folding data well. However, the inference problem still constitutes a challenging problem due to the computational complexity of the particle Monte Carlo based algorithms and the properties of the proposed model.

Keywords: Protein folding, double-well potential stochastic differential equation, intractable likelihood functions, particle Markov chain Monte Carlo

Inference for the driving Lévy process of continuous-time moving average processes

Jeannette H.C. Woerner¹

¹Technical University of Dortmund (Germany)

Thursday 27
10:30-12:30
Lecture room 1
(Sali 1)

Continuous-time moving average processes, defined as integrals of a deterministic kernel function integrated with respect to a two sided Lévy process, provide a unifying framework to different types of processes, including the popular examples of fractional Brownian motion and fractional Lévy processes on the one side and Ornstein-Uhlenbeck processes on the other side. The whole class of processes especially allows for a combination of a given correlation structure with an infinitely divisible marginal distribution as it is desirable for applications in finance, physics and hydrology.

So far inference for these processes is mainly concerned with estimating parameters entering the kernel function which is responsible for the correlation structure. We now consider the estimating problem for the driving Lévy process. We will provide two methods working on different sets of conditions, one is based on a suitable integral transform, the other on the Mellin transform.

(Partly based on joint work with D. Belomestny and V. Panov.)

graphics.SDA: the R base graphics companion for exploratory symbolic data analysis

Han-Ming Wu

National Taipei University (Taiwan)

Tuesday 25
14:00-16:00
Lecture room 10
(Sali 10)

Exploratory data analysis (EDA) serves as a preliminary yet essential tool for summarizing the main characteristics of a data set before appropriate statistical modeling can be applied. Quite often, EDA employs the traditional graphical techniques such as the boxplot, histogram and scatterplot and are equipped with various dimension reduction methods and computer-aided interactive functionalities. EDA has been used to explore different data types. Examples were the cases of the survival data, the time series data, the functional data and the longitudinal data. Conventionally, these data set were tabulated by a table with p columns corresponding to p variables. Each subject is measured by a single numerical value for each variable. Nowadays the collected data keeps getting bigger and complex. The description of data was no longer stored by a form of a single value but the intervals, histograms and/or distributions. These are examples of the so-called symbolic data. This study intends to develop an R package as the companion of R base graphics package for symbolic data. Two dimension reduction methods, the principal component analysis (PCA) and the sliced inverse regression (SIR), are also extended and used to reveal the insight structure of symbolic objects embedded in the high-dimensional space. SDA supplies various data descriptions and has great capacity for big data. As a consequence, exploratory symbolic data analysis as a tool that supports the efficient, effective and practical exploration of symbolic data sets is needed.

Thursday 27
10:30-12:30
Lecture room 10
(Sali 10)

A permutation test for the two-sample right-censored model

Grzegorz Wyłupek
University of Wrocław (Poland)

The paper presents a novel approach to solving a classical two-sample problem with right-censored data. As a result, an efficient procedure for verifying equality of the two survival curves is developed. It generalizes a well known standard, that is, the log-rank test in a natural manner. First, the system of Laguerre polynomials in combination with the weighted log-rank statistics are used to define the components, which span the space of the general alternatives. Next, an auxiliary statistic, which is the sum of squares of the finite numbers of those components is defined. Finally, a flexible selection rule chooses the number of summands from the data at hand. Under the null hypothesis, the new test statistic has an asymptotic chi-square distribution with one degree of freedom, while the corresponding test is consistent against essentially any alternative. An extensive simulation study shows that the new testing procedure improves upon classical solutions and popular recent developments of the considered testing problem.

Monday 24
14:00-16:00
Lecture room 12
(Sali 12)

A two-stage approach for estimating the parameters of an age-group epidemic model from incidence data

Rami Yaari
University of Haifa and The Gertner Institute for Epidemiology and Health Policy Research (Israel)

Age-dependent dynamics is an important characteristic of many infectious diseases. Age-group epidemic models describe the infection dynamics in different age-groups by allowing to set distinct parameter values for each. However, this type of models are highly nonlinear and may have a large number of unknown parameters. Thus, parameter estimation of age-group models, while becoming a fundamental issue for both the scientific research and policy making in infectious diseases, is not a trivial task in practice. In this talk, we examine the estimation of the so called next-generation matrix using incidence data of a single entire outbreak, and extend the approach to deal with recurring outbreaks. Unlike previous studies, we do not assume any constraints regarding the structure of the matrix. A novel two-stage approach is developed, which allows for efficient parameter estimation from both statistical and computational perspectives. Simulation studies corroborate the ability to estimate accurately the parameters of the model for several realistic scenarios. The model and estimation method are applied to real data of influenza-like-illness in Israel. The parameter estimates of the key relevant epidemiological parameters and the recovered structure of the estimated next-generation matrix are in line with results obtained in previous studies.

Tuesday 25
14:00-16:00
Lecture room 10
(Sali 10)

Bayesian change-point detection for ordinal data with North Atlantic tropical cyclones

Guosheng Yin
The University of Hong Kong (China)

From a Bayesian point of view, we analyze the North Atlantic tropical cyclone data to provide quantitative evidence for this deduction. By reframing the original change-point detection problem in a Bayesian variable selection context, we propose a latent probit model approach to estimating both the number and locations of change-points with multinomial data. We then conduct model selection via such an approach using a reversible-jump Markov chain Monte Carlo sampler due to changes in dimensionality. Under this framework, the posterior distribution of the number of change-points, the marginal posterior probability of each location to be a change-point, and the credible intervals of the regression coefficients, can all be obtained conveniently using posterior samples. We conduct simulation studies to assess our method, and as an illustration apply it to the North Atlantic tropical cyclone record.

Detection of imprinting effects for quantitative traits on X chromosome

Kexin Yu

The University of Hong Kong (China)

Tuesday 25
10:30-12:30
Lecture room 7
(Sali 7)

Genomic imprinting is an epigenetic phenomenon in which the expression of an allele copy depends on its parental origin. This mechanism has been found to play an important role in many complex diseases. Statistical tests for imprinting effects have been developed for more than fifteen years, but they are only suitable for autosomes. It is not until recently that the parental-asymmetry test on the X chromosome (XPAT) has been proposed to test for imprinting effects. However, this test can only be used for qualitative traits. Therefore, in this article, we propose a class of PAT-type tests to test for imprinting for quantitative traits on the X chromosome in the presence of association, namely, Q-XPAT(c), Q-1-XPAT(c) and Q-C-XPAT(c), where c is a constant. These methods can accommodate complete and incomplete nuclear families with an arbitrary number of daughters. Extensive simulation studies demonstrate that the proposed tests control the size well under the null hypothesis of no imprinting effects and are powerful under various family structures. Moreover, by setting the inbreeding coefficient in females to be nonzero and using the assortative mating pattern in simulations, the proposed tests are shown to be valid under Hardy-Weinberg disequilibrium.

Variational diagonal discriminant analysis with model selection

Weichang Yu and John Ormerod

University of Sydney (Australia)

Thursday 27
10:30-12:30
Lecture room 12
(Sali 12)

Discriminant analysis (DA) is widely used as a predictive tool in diverse topics. However, DA using full covariance matrix estimates is only possible when P (number of predictors) is smaller than n (sample size). We propose a Bayesian diagonal discriminant analysis with binary model selection parameters. The diagonal covariance structure reduces the model selection problem to performing multiple independent hypothesis tests and outperforms the general covariance structure in prediction error as shown by Bickel and Lavina (2004) if all true pairwise correlations are 0. By choosing a novel prior structure (known as Cake priors) and collapsed variational Bayes, we show (under stated assumptions) that the model selection type I and II errors approach 0 as n approaches infinity. Finally, the model is extended to semisupervised setting. Through simulated and real datasets, we demonstrate that our proposed model yields lower prediction error and computational cost in comparison to existing solutions.

Fréchet means and Procrustes analysis in Wasserstein space

Yoav Zemel and Victor M. Panaretos

École polytechnique fédérale de Lausanne (Switzerland)

Tuesday 25
14:00-16:00
Lecture room 5
(Sali 5)

We consider two statistical problems at the intersection of functional and non-Euclidean data analysis: the determination of a Fréchet mean in the Wasserstein space of multivariate distributions; and the optimal registration of deformed random measures. We elucidate how the two problems are linked, each being in a sense dual to the other. We first study the finite sample version of the problem. Exploiting the tangent bundle structure of Wasserstein space, we deduce the Fréchet mean via gradient descent. We show that this is equivalent to a Procrustes analysis for the registration maps, thus only requiring successive solutions to pairwise optimal coupling problems. We then study the population version of the problem, focussing on inference and stability: in practice, the data are i.i.d. realisations from a law on Wasserstein space, and indeed their observation is discrete, where one observes a proxy finite sample or point process. We construct regularised nonparametric estimators, and prove their consistency for the population mean, and uniform consistency for the population Procrustes registration maps. Our methods are intrinsic to the Wasserstein geometry, since Hessian-type methods from manifold statistics do not apply to the Wasserstein space.

Wednesday 26
10:30-12:30
Lecture room 12
(Sali 12)

Joint models for time-to-event and bivariate longitudinal data

Mariangela Zenga and Marcella Mazzoleni
University of Milano-Bicocca (Italy)

The joint models analyse the effect of a longitudinal covariate onto the risk of an event. They are composed of two sub-models, the longitudinal and the survival sub-model. In this paper the focus is on the case in which the longitudinal sub-model is bivariate, considering more than one longitudinal covariate. For the longitudinal sub-model a multivariate mixed model can be proposed. Whereas for the survival sub-model, a Cox proportional hazards model is proposed, considering jointly the influence of both the longitudinal covariates onto the risk of the event. The purpose of the paper is to implement an estimation method that is able to deal with the computational problem given by the introduction of other covariates and the increase of the number of parameters that must be estimated in a model that is already highly computationally demanding.

Monday 24
14:00-16:00
Lecture room 8
(Sali 8)

A new robust regression model: Type II multivariate t distribution with applications

Chi Zhang
The University of Hong Kong (China)

Motivated by a real data analysis, we propose a new kind of multivariate t (MVT) distribution by allowing variable degrees of freedom for each univariate component. Compared with the classical multivariate t distribution, it is more flexible in model specification which can be used to deal with the variant amounts of tail weights on marginals in multivariate data modeling. In particular, it could include components following the multivariate normal distribution when the corresponding degrees of freedom approach to infinity, and it contains the product of independent t-distributions as a special case. Subsequently, it is extended to the regression model as the joint distribution of the error term. Important distributional properties are explored and useful statistical methodologies are developed. The specified structure in the new MVT distribution has the advantage of its flexibility to better capturing the characteristic of data, as exemplified by both simulation studies and real data applications.

Monday 24
14:00-16:00
Lecture room 10
(Sali 10)

Envelopes for censored quantile regression

Yue Zhao¹, Zhihua Su² and Ingrid Van Keilegom¹
¹KU Leuven (Belgium), ²University of Florida (USA)

Quantile regression has emerged as a powerful tool for survival analysis with censored data. In this article, we propose an efficient estimator for the coefficient in quantile regression with censored data using the envelope model. First introduced in Cook et al. (2010), the envelope model uses dimension reduction techniques to identify material and immaterial components in the data, and forms the estimator of the regression coefficient based only on the material component, thus reducing the variability of the estimation. We will derive asymptotic properties of the proposed estimator and demonstrate its efficiency gains as compared to the traditional estimator for quantile regression with censored data. Recent advances in algorithms for the envelope model allow for efficient implementation of the proposed method. The strength of our proposed method is demonstrated via simulation studies.

Higher-order Berry-Esseen inequalities and accuracy of the bootstrap

Mayya Zhilova

Georgia Institute of Technology (USA)

Wednesday 26
10:30-12:30
Lecture room 6
(Sali 6)

In this talk, we study higher-order accuracy of a bootstrap procedure for approximation in distribution of a smooth function of a sample average in high-dimensional non-asymptotic framework. Our approach is based on Berry-Esseen type inequalities which extend the classical normal approximation bounds. These results justify in non-asymptotic setting that bootstrapping can outperform Gaussian (or chi-squared) approximation in accuracy with respect to both dimension and sample size. In addition, the presented results lead to improvements of accuracy of a weighted bootstrap procedure for general log-likelihood ratio statistics (under certain regularity conditions). The theoretical results are illustrated with numerical experiments on simulated data. The talk is based on the paper arXiv:1611.02686.

Quantile double autoregression

Qianqian Zhu

The University of Hong Kong (China)

Monday 24
10:30-12:30
Lecture room 8
(Sali 8)

This paper proposes a novel conditional heteroscedastic time series model which is driven by functionally dependent autoregressive coefficients. This new model includes the commonly used double autoregressive model as a special case, where both the location and scale structures are incorporated in double autoregressive forms. However, due to a simple but nontrivial transformation, it enjoys much more flexible parameter structures without positive restrictions on the scale coefficients. The double autoregressive coefficients are monotone functions of a single random variable which is related to quantile levels in view of conditional quantiles of the process, and hence the associated quantile double autoregressive model is able to capture different structures for conditional quantiles of a time series. The strict stationarity of the proposed quantile double autoregressive process is discussed under mild conditions, and a self-weighted conditional quantile estimator is considered for the quantile double autoregressive model. The strong consistency and asymptotic normality of the proposed estimator are obtained with only fractional moments of the process, which makes the model possible to handle heavy-tailed data. Simulation experiments are carried out to assess the finite sample performance of the proposed methodology, and an empirical example is presented to illustrate the usefulness of the new model.

Author index

- Özdemir, Şenay, 62, 107, 141
Çankaya, Emel, 77
Çetin, Meral, 140
- Achcar, Jorge, 27, 47
Alsmeyer, Gerold, 62
Amro, Lubna, 27
Anastasiou, Andreas, 28
Anatolyev, Stanislav, 28
Anderson, Eric C., 127
Anevski, Dragi, 109
Angelov, Angel, 29
Antolin, Roberto, 66
Arab, Idir, 29, 105
Archimbaud, Aurore, 103
Arcos Cebrián, Antonio, 29
Arslan, Olcay, 62, 77, 107, 140, 141
Artemiou, Andreas, 30
Asin, Nicolas, 30
Aston, John, 30
Athey, Susan, 147
Aubin, Jean-Baptiste, 31, 88
Audrino, Francesco, 138
Aue, Alexander, 144
Augugliaro, Luigi, 31
- Bacallado, Sergio, 31
Bachoc, François, 32, 145
Baek, Changryong, 113
Balakrishnan, Sivaraman, 149
Barigozzi, Matteo, 32, 46
Barker, Richard, 38
Baroni, Enrico, 81
Barreyre, Clémentine, 32
Barthel, Nicole, 33
Basiri, Shahab, 33
Bastien, Béangère, 59
Batsidis, Apostolos, 34
Battagin, Mara, 66
Bauer, Benedikt, 34
Beering, Carina, 34
Behr, Merle, 35
Belitser, Eduard, 103
Benaych-Georges, Florent, 35
Berenguer Rico, Vanessa, 35
- Bernard, Carole, 49
Berrendero, José R., 140
Beskos, Alexandros, 109
Betancourt, Michael, 151
Bethere, Liga, 36
Betken, Annika, 36
Bharath, Karthik, 36
Bibinger, Markus, 37
Bissantz, Nicolai, 46
Boaz, Nadler, 119
Bodnar, Taras, 98, 108
Bogdan, Małgorzata, 37
Bongiorno, Enea, 31, 60
Bonnet, Anna, 37
Boot, Tom, 38
Bossert, Frank, 130
Bran, Paula, 38
Brandes, Dirk-Philip, 38
Braumann, Alexander, 39
Bravington, Mark V., 127
Breunig, Christoph, 39
Broderick, Tamara, 39
Browne, William, 130
Brzozowska-Rup, Katarzyna, 40
Bucchia, Beatrice, 151
Bundesen, Claus, 51
But, Anna, 63
Butucea, Cristina, 143
Bühlmann, Peter, 124
- Cai, Diana, 39
Cai, Xia, 139
Campbell, Trevor, 39
Candes, E. J., 37
Cattelan, Manuela, 40
Chacón, José E., 41
Chakraborty, Anirvan, 41
Chan, Hsin-Yu, 91
Chan, Ngai Hang, 42
Chandrasekaran, Venkat, 42
Chang, ChihHao, 42
Chang, Yuan-Chin, 43
Charkhi, Ali, 43
Charlton, Christopher, 130
Chavez-Demoulin, Valerie, 43

Chen, Li-Ching, 44
 Chen, May-Ru, 44
 Chen, Vivian Yi-Ju, 44, 92
 Chen, Yuh-Ing, 45
 Chen, Zheng, 45
 Cheplyukova, Irina, 45
 Chiquet, Julien, 117
 Cho, Haeran, 46
 Chown, Justin, 46
 Christoffersen, Benjamin, 46
 Claeskens, Gerda, 47
 Cleveland, Matthew, 66
 Coelho-Barros, Emilio, 47
 Comte, Fabienne, 94
 Corander, Jukka, 111
 Cortese, Giuliana, 47
 Couillet, Romain, 35
 Cousido-Rocha, Marta, 48
 Cribben, Ivor, 48
 Crisan, Dan, 109
 Cuevas, Antonio, 140
 Curato, Imma, 38, 48
 Czado, Claudia, 33, 49, 84, 100

Düker, Marie-Christine, 52
 Dai, Hongsheng, 114
 Dattner, Itai, 49
 Davis, Richard, 65
 Dawabsha, Mohammed, 120
 De Backer, Mickaël, 49
 de Gunst, Mathisca, 128
 de la Pena, Victor, 50
 de Uña-Álvarez, Jacobo, 48
 Deane, C., 116
 Dehling, Herold, 50
 Deijfen, Maria, 50
 Dette, Holger, 46, 108
 Diaz, Ivan, 104
 Dillenbourg, Pierre, 43
 Ditlevsen, Susanne, 51
 Dombry, Clement, 51
 Dominicy, Yves, 64
 Dong, Yuexiao, 30
 Dubhashi, Devdatt, 51
 Dudek, Anna, 52
 Dunker, Fabian, 85
 Duval, Céline, 97
 Dyckerhoff, Rainer, 100

Early, Jeffrey J., 62
 Eckle, Konstantin, 52
 Edelman, Dominic, 53
 Ekström, Magnus, 29
 El Gouch, Anouar, 49, 96
 Elsener, Andreas, 53
 Engelke, Sebastian, 51

Erdugan, Funda, 53
 Erhardt, Tobias, 84
 Espericueta, Dora Erika, 66
 Etheridge, Alison, 54

Favaro, Stefano, 31, 54
 Fearnhead, Paul, 114
 Fechner, Zywilla, 130
 Feng, Yuanhua, 89
 Fiedler, Jochen, 54
 Fiocco, Marta, 120, 152
 Fitzner, Robert, 50
 Fokianos, Konstantinos, 55
 Forman, Julie Lyng, 152
 Fortini, Sandra, 55
 Fosdick, Bailey, 101
 Francq, Christian, 56
 Franks, Jordan, 56
 Frej, Michał, 114
 Friedrich, Sarah, 56
 Frostig, Ron, 106
 Fryzlewicz, Piotr, 28, 46
 Fujimori, Kou, 57

Güney, Yeşim, 62, 107, 141
 Gégout-Petit, Anne, 59
 Gadrach, Tamar, 57
 Gamboa, Fabrice, 145
 Ganatsiou, Chrysoula, 58
 Gao, Xu, 106
 Gasbarra, Dario, 93
 Gassiat, Elisabeth, 37
 Gaunt, R., 116
 Geffray, Ségolen, 58
 Gelman, Andrew, 151
 Gerardin, Émilie, 143
 Gerds, Thomas Alexander, 121
 Giovannelli, Alessandro, 115
 Girard, Stéphane, 132
 Girolami, Mark, 59
 Goia, Aldo, 1, 60
 Golalizadeh, Mousa, 60
 Gonen, Serap, 66
 González Hernandez, Javier, 61
 González, Miguel, 62
 Gorjanc, Gregor, 66
 Grabarczyk, Peter, 148
 Groll, Andreas, 142
 Guadalupe, Tzintzun, 118
 Gueudin, Aurélie, 59
 Gugushvili, Shota, 61, 129
 Guillaumin, Arthur, 62
 Guillou, Armelle, 132
 Guntuboyina, Aditya, 149
 Gutiérrez Pérez, Cristina, 62

Härkänen, Tommi, 63

Hambuckers, Julien, 63
 Hansen, Niels Richard, 99
 Hart, Jeffrey D., 48
 Haukka, Jari, 63
 Heikkilä, Matias, 64
 Heikkilä, Mikko, 64
 Hein, Anne, 64
 Heinrich, Lothar, 65
 Heiny, Johannes, 65
 Helander, Sami, 65
 Helske, Jouni, 56
 Hernández-Molinar, Raúl, 66
 Hernández-García, Vicente, 66
 Herring, William, 66
 Hickey, John, 66
 Hirose, Yuichi, 66
 Holevo, Alexander S., 67
 Holmes, Chris, 35
 Honda, Toshio, 67
 Honkela, Antti, 64
 Hossain, Shaikh Tanvir, 72
 Hsu, Hsiang-Ling, 67
 Hušková, Marie, 69
 Huang, Shih-Feng, 68
 Huang, Xifen, 68
 Huckemann, Stephan, 69
 Hudecová, Šárka, 69
 Huser, Raphael, 69
 Hwang, Ruey-Ching, 70
 Hwang, Yi-Ting, 70
 Hyttinen, Antti, 70

 Ibragimov, Rustam, 71
 Ilmonen, Pauliina, 64, 65, 92, 147
 Imbens, Guido W., 147
 Iyer, Srikanth K., 139

 Jach, Agnieszka, 71
 Jansen, Maarten, 71
 Jasra, Ajay, 109
 Javed, Farrukh, 72
 Jeantheau, Thierry, 143
 Jenko, Janez, 66
 Jentsch, Carsten, 34, 72
 Jermyn, Ian, 73
 Jiang, Ci-Ren, 73
 Jiang, Zhenyu, 95
 Jiménez-Gamero, M. Dolores, 34
 Johannes, Jan, 94
 Johansen, Adam, 114
 Johansen, Søren, 102
 Jorswieck, Eduard, 74
 Josse, Julie, 117

 Kalisch, Markus, 111
 Kalliovirta, Leena, 74
 Kang, Jing, 74

 Kangasrääsio, Antti, 75
 Karagöz, Derya, 75
 Karagrigoriou, Alex, 76
 Kardaras, Kostas, 141
 Karjalainen, Joonas, 76
 Karvonen, Toni, 77
 Kaski, Samuel, 64
 Kattuman, Paul, 77
 Kawka, Rafael, 148
 Kaya, Mutlu, 77
 Kechagias, Stefanos, 113
 Kharin, Yuriy, 146
 Killiches, Matthias, 78
 Kirichenko, Alisa, 78
 Kiriliouk, Anna, 79
 Klasen, Stephan, 85
 Kleinsmith, Deon, 79
 Knapik, Bartek, 128
 Kneib, Thomas, 63, 79, 142
 Koenig, Claudia, 80
 Koivunen, Visa, 33
 Kolar, Ana, 80
 Komjathy, Julia, 81
 Konietschke, Frank, 142
 Kosorok, Michael R., 81
 Koyuncu, Nursel, 75
 Kozyrev, Vladislav, 51
 Kpogbezan, Gino Bertrand, 82
 Krampe, Jonas, 82
 Kranis, Andreas, 66
 Kraus, Andrea, 82
 Kraus, Daniel, 49, 83
 Kraus, David, 83
 Kreiss, Jens-Peter, 39, 82
 Kreutz, Clemens, 83, 84
 Kreuzer, Alexander, 84
 Kristensen, Dennis, 84
 Krivobokova, Tatyana, 85, 126
 Kroll, Martin, 85
 Kuhn, Estelle, 86
 Kulik, Rafal, 85
 Kumar, Divakar, 114
 Kuo, Kun-Lin, 86
 Kurosawa, Takeshi, 86
 Kurtek, Sebastian, 36
 Kyllingsbæk, Søren, 51
 Kypraios, Theodore, 131, 132

 Lévy-Leduc, Céline, 37
 Laga, Hamid, 73
 Landon, Joshua, 108
 Langrock, R., 63
 Latukka, Arto, 126
 Lauritzen, Steffen, 87
 Le Corff, Sylvain, 87
 Le Guével, Ronan, 87

Le, Huiling, 88
 Leśkow, Jacek, 89
 LeCun, Yann, 88
 Ledoux, Michel, 103
 Lee, Daniel, 151
 Lee, Young Jun, 84
 Leeb, Hannes, 123
 Leoni, Samuela, 88
 Leri, Marina, 89
 Leskelä, Lasse, 76
 Letmathe, Sebastian, 89
 Leucht, Anne, 34
 Li, Bing, 146
 Li, Fan, 90, 98
 Li, Guodong, 90
 Li, H., 149
 Li, Housen, 152
 Li, Kang, 51
 Li, Pai-Ling, 91
 Li, Wai Keung, 124
 Li, Weiming, 91
 Licari, Federica, 91
 Lietzén, Niko, 92
 Lillo, Rosa, 92
 Lilly, Jonathan M., 62
 Lin, Kuo-Chin, 92
 Lin, Liang-Ching, 93
 Lin, Yi-Chen, 93
 Lin, Zhenhua, 101
 Ling, Nengxiang, 95
 Liu, Jia, 93
 Liu, Qingfen, 94
 Liu, Zhang, 94
 Loizeau, Xavier, 94
 Loubes, Jean-Michel, 145
 Lu, Zudi, 95
 Luguzis, Artis, 95

 Mäkinen, Taneli, 98
 Müller, Hans-Georg, 1, 101
 Müller, Ursula, 101
 Maathuis, Marloes, 111
 Mabon, Gwennaëlle, 95
 Maistre, Samuel, 96
 Malefaki, Sonia, 96
 Malov, Sergey, 96
 Mandjes, Michel, 128
 Mariadassou, Mahendra, 117
 Mario, Tarumoto, 118
 Mariucci, Ester, 97
 Marra, Giampiero, 97
 Martínez, Rodrigo, 62
 Matilainen, Markus, 98
 Mattei, Alessandra, 91
 Mazucheli, Josmar, 47
 Mazur, Stepan, 72, 98

 Mazzoleni, Marcella, 156
 McCormack, Tyler, 101
 McElroy, Tucker, 71
 Mealli, Fabrizia, 91
 Meinshausen, Nicolai, 139
 Meintanis, Simos, 69
 Mercatanti, Andrea, 98
 Meulman, Jacqueline, 152
 Meyer, Marco, 34
 Mielniczuk, Jan, 114
 Mikkelsen, Frederik Riis, 99
 Mikosch, Thomas, 65
 Mileham, Alan, 66
 Milovic, Ivana, 99
 Moritz, Jirak, 100
 Moulines, Éric, 105, 117
 Mozharovskyi, Pavlo, 100
 Mueller, Dominik, 100
 Munk, A., 149
 Munk, Axel, 35, 80, 126, 135
 Murphy, Brendan, 101

 Nagler, Thomas, 84
 Nagy, Stanislav, 102
 Nason, Guy, 102
 Navarro, Fabien, 122
 Naveau, Philippe, 79
 Ng, James, 101
 Nielsen, Bent, 35, 102
 Nishiyama, Yoichi, 57
 Nkurunziza, Séverien, 102
 Nordhausen, Klaus, 92, 98, 103, 104
 Nourdin, Ivan, 103
 Nurushev, Nurzhan, 103
 Nyberg, Henri, 103

 O'Neill, Philip, 131, 132
 Oesting, Marco, 51
 Ogburn, Elizabeth L., 104
 Oja, Hannu, 98, 104
 Okhrin, Ostap, 138
 Okhrin, Yarema, 33
 Okimoto, Yusuke, 64
 Olhede, Sofia, 105
 Olhede, Sofia C., 62
 Oliveira, Paulo Eduardo, 29, 105
 Ollila, Esa, 33, 134
 Olsson, Jimmy, 105
 Ombao, Hernando, 106
 Omelka, Marek, 106
 Ommane, Younes, 107
 Opitz, Thomas, 69
 Ormerod, John, 155
 Ospina, L., 116
 Otto, Peter, 107
 Ozekici, Suleyman, 108

Palmes, Christian, 108
 Panaretos, Victor M., 41, 155
 Parolya, Nestor, 98, 108
 Parrilo, Pablo, 42
 Pastukhov, Vladimir, 109
 Patilea, Valentin, 109
 Paulin, Daniel, 109
 Pauly, Markus, 121, 142
 Pavlenko, Tatjana, 110
 Pavlov, Yuri, 45
 Pchelintsev, Evgenii, 110
 Peccati, Giovanni, 103
 Peneva, Iliana, 110
 Perkovic, Emilija, 111
 Pessia, Alberto, 111
 Petrone, Sonia, 55
 Petropoulos, Constantinos, 111
 Philippe, Anne, 87
 Picchini, Umberto, 112, 152
 Picek, Jan, 112
 Pipiras, Vladas, 113
 Pircalabelu, Eugen, 113
 Platis, Agapios, 113
 Pokarowski, Piotr, 114
 Pollock, Murray, 114
 Praskova, Zuzana, 114
 Preston, Simon, 115
 Prochenka, Agnieszka, 114
 Proietti, Tommaso, 115
 Proksch, Katharina, 115
 Putter, Hein, 120

 Racine-Poon, Amy, 151
 Radice, Rosalba, 97
 Radulovic, Dragan, 116
 Railavo, Juha, 93
 Rantala, Sakke, 65
 Rastogi, Manoj Kumar, 111
 Reager, John, 42
 Reichmann, Lena, 72
 Reinert, Gesine, 116
 Reiss, Markus, 95
 Richter, Wolf-Dieter, 116
 Robert, Christian, 117
 Roberts, Gareth, 114
 Robin, Genevieve, 117
 Robin, Stéphane, 117
 Robinson, Peter M., 118
 Rockova, Veronika, 118
 Rodrigues, Eliane, 118
 Roldán-Nofuentes, Jose Antonio, 119
 Román-Román, Patricia, 121
 Romo, Juan, 132
 Rooks, Brian Thomas, 119
 Rootzén, Holger, 79
 Ros-Freixedes, Roger, 66

 Rosenblatt, Jonathan, 119
 Rueda, Maria del Mar, 29
 Rueten-Budde, Anja, 120
 Ruiz-Castro, Juan Eloy, 120
 Ruiz-Gazen, Anne, 103
 Rytgaard, Helene Charlotte, 121

 Sabatti, C., 37
 Samlis, Christos, 96
 Sant'Ana, Istoni, 121
 Sardy, Sylvain, 117
 Sattler, Paavo, 121
 Saumard, Adrien, 122
 Savage, Richard S., 127
 Schauer, Moritz, 61
 Schefzik, Roman, 122
 Schindler, Martin, 112
 Schmitz, Hendrik, 89
 Schneider, Ulrike, 123
 Schnurr, Alexander, 50
 Schofield, Matthew, 38
 Schwartzman, Armin, 123
 Segers, Johan, 79
 Senitschnig, Nina, 123
 Sepponen, Anne-Mari, 126
 Shah, Nihar, 149
 Shah, Rajen, 124, 139
 Sharipov, Olimjon Sh., 151
 Sharma, Kshitij, 43
 Sharpnack, James, 124
 Shen, Keren, 124
 Shi, Yaojie, 59
 Shimizu, Kana, 64
 Shin, Seung-Jun, 30
 Sieling, H., 149
 Sigrist, Fabio, 125
 Sillanpää, Mikko J., 125
 Silvennoinen, Annastiina, 137
 Silvestrini, Andrea, 98
 Simkova, Tereza, 125
 Singer, Marco, 126
 Sinisalo, Alina, 126
 Siouris, George-Jason, 126
 Skaug, Hans J., 127
 Skinner, Jim, 127
 Skolkova, Alena, 128
 Sofrygin, Oleg, 104
 Sohn, A., 63
 Sollie, Birgit, 128
 Somavilla, Adriana, 66
 Sommerfeld, Max, 135
 Soyer, Refik, 108
 Sporysheva, Polina, 55
 Spreij, Peter, 61, 129
 Srivastava, Anuj, 73, 129
 Stattin, Oskar, 110

Steele, Fiona, 130
 Steffens, Britta, 130
 Stelzer, Robert, 48, 130
 Stenlund, David, 131
 Stepanova, Natalia, 110, 131
 Stockdale, Jessica, 131, 132
 Strzalkowska-Kominiak, Ewa, 132
 Stupfler, Gilles, 132
 Stypka, Oliver, 133, 148
 Su, Nan-Cheng, 134
 Su, W., 37
 Su, Yu-Ru, 150
 Su, Zhihua, 156
 Sultana, Farha, 111
 Suzuki, Nobuhiro, 86
 Sykulski, Adam M., 62
 Szabó, Botond, 144

 Tabakan, Gulin, 53, 134
 Tabassum, Muhammad Naveed, 134
 Tameling, Carla, 135
 Tang, Yaohua, 135
 Tardivel, Patrick, 135, 136
 Teisseyre, Paweł, 137
 Telkmann, Klaus, 137
 Teräsvirta, Timo, 137
 Teterova, Anastasija, 138
 Tewes, Johannes, 138, 151
 Thacker, Debleena, 139
 Thanei, Gian-Andrea, 139
 Thibaud, Emeric, 69
 Tian, Yubin, 139
 Timmer, Jens, 83, 84
 Tjøstheim, Dag, 95
 Toka, Onur, 140
 Torrecilla, José Luis, 1, 140
 Torres-Ruiz, Francisco, 121
 Trabs, Mathias, 37
 Trapani, Lorenzo, 32
 Treue, Stefan, 51
 Tripathi, Yogesh Mani, 111
 Trippa, Lorenzo, 31
 Trujillo, Manuel, 29
 Truquet, Lionel, 140
 Tsagris, Michail, 115
 Tsimbalyuk, Alexandra, 141
 Tsou, Tsung-Shan, 141
 Tuuç, Yetkin, 62, 107, 140, 141
 Turmon, Michael, 42
 Tyler, David E., 104

 Uhler, Caroline, 142
 Umlauf, Nikolaus, 142
 Umlauft, Maria, 142
 Uriel, Leclaire, 143

 Valeinis, Jānis, 36, 95, 143

 Valkovska, Jelena, 143
 Van Bever, Germain, 65
 van de Geer, Sara, 53
 van de Wiel, Mark, 82, 143
 van Delft, Anne, 144
 van den Berg, E., 37
 van der Hofstad, Remco, 81
 van der Laan, Mark J., 104
 van der Meulen, Frank, 61, 129
 van der Pas, Stéphanie, 144
 van der Vaart, Aad, 82, 144
 Van Keilegom, Ingrid, 49, 96, 109, 156
 van Wieringen, Wessel, 82
 van Zanten, Harry, 78
 Vanhems, Anne, 144
 Varin, Cristiano, 40
 Vehtari, Aki, 151
 Velasco, Carlos, 118
 Veliyev, Bezirgen, 145
 Venet, Nil, 145
 Vieu, Philippe, 60
 Vihola, Matti, 56
 Viitasaari, Lauri, 145, 147
 Vimond, Myriam, 146
 Virta, Joni, 104, 146
 Voloshko, Valeriy, 146
 Voutilainen, Marko, 147

 Wadsworth, Jennifer, 79
 Wager, Stefan, 147
 Wagner, Martin, 133, 148
 Wahl, Martin, 95
 Wainwright, Martin, 149
 Waldmann, Patrik, 149
 Walther, Guenther, 149
 Wan, Phyllis, 150
 Wang, Jane-Ling, 150
 Wang, Yibo, 131
 Wang, Yueleng, 102
 Wang, Yuxiao, 106
 Wann, Ellen, 106
 Washbrook, Elizabeth, 130
 Weber, Sebastian, 151
 Wegner, A., 116
 Wendler, Martin, 151
 Werner, Frank, 80, 152
 Westling, Ted, 101
 Willems, Sanne, 152
 Willsky, Alan, 42
 Wiqvist, Samuel, 152
 Woerner, Jeannette H.C., 153
 Wood, Andy, 115
 Woolliams, John, 66
 Wu, Han-Ming, 153
 Wyłupek, Grzegorz, 154

 Xie, Qian, 73

Xu, Cong, 150
Xu, Xiaochuan, 116

Yaari, Rami, 154
Yan, Yin-Jhen, 44
Yang, Tse-Chuan, 44
Yao, Fang, 101
Yao, Jianfeng, 91, 124
Yao, Qingsong, 94
Yin, Guosheng, 154
Yu, Kexin, 155
Yu, Philip L.H., 135

Yu, Weichang, 155

Zakoian, Jean-Michel, 56
Zemel, Yoav, 155
Zenga, Mariangela, 156
Zhang, Chi, 156
Zhang, Qiang, 95
Zhang, Rongmao, 42
Zhao, Yue, 156
Zhilova, Mayya, 157
Zhu, Qianqian, 157