Banff International Research Station

for Mathematical Innovation and Discovery

Non-Gaussian Multivariate Statistical Models and their Applications
May 19-24, 2013

MEALS

*Breakfast (Buffet): 7:00 – 9:30 am, Sally Borden Building, Monday – Friday
*Lunch (Buffet): 11:30 am – 1:30 pm, Sally Borden Building, Monday – Friday
*Dinner (Buffet): 5:30 – 7:30 pm, Sally Borden Building, Sunday – Thursday
Coffee Breaks: As per daily schedule, in the foyer of the TransCanada Pipeline Pavilion (TCPL)
***Please remember to scan your meal card at the host/hostess station in the dining room for each meal.**

MEETING ROOMS

All lectures will be held in the lecture theater in the TransCanada Pipelines Pavilion (TCPL). An LCD projector, a laptop, a document camera, and blackboards are available for presentations.

SCHEDULE

Sunday

16:00 Check-in begins (Front Desk – Professional Development Centre - open 24 hours)
17:30-19:30 Buffet Dinner
20:00 Informal gathering in 2nd floor lounge, Corbett Hall (if desired)
Beverages and small assortment of snacks are available on a cash honor system.

Monday (chair: M. Genton)

7:00-8:45 Breakfast
8:45-9:00 Introduction and Welcome by BIRS Station Manager, TCPL
9:00-10:00 Overview 1: A. Azzalini, Skew-Symmetric Distributions in 45' (45+15)
10:00-10:30 Coffee break
10:30-11:30 Overview 2: H. Joe, Overview of copula models: dependence structures and asymmetries (45+15)
11:30-13:00 Lunch
13:00-14:00 Overview 3: K. Aas & C. Czado, Pair-copula constructions – even more flexible than copulas (45+15)
14:00-15:00 E. Giorgi, Moment-free measures for the multivariate skew-t distribution (20+10)
L. Hua, Strength of tail dependence based on conditional tail expectation (20+10)
15:00-15:30 Coffee break
15:30-17:00 Presentation of individual interests (90)
17:00-18:00 M. Jimenez-Gamero, Testing for skew-symmetric models (20+10)
A. Pewsey, Inference for skew-symmetric distributions (20+10)
18:00-19:30 Dinner

Tuesday (chair: H. Joe)

7:00-9:00 Breakfast
9:00-10:00 S. Volgushev, Empirical and sequential empirical copula processes under serial dependence and weak smoothness conditions (20+10)
M. Genton, Nonparametric identification of copula structures (20+10)
10:00-10:30 Coffee break
10:30-11:30 D. Dey, State space models for binary response data with flexible skewed link functions (20+10)
B. Liseo, Bayesian inference for the multivariate skew-normal distribution: A population Monte-Carlo approach (20+10)
11:30-13:00 Lunch
13:00-14:00 Guided Tour of The Banff Centre; meet in the 2nd floor lounge, Corbett Hall

14:00-14:15 Group Photo; meet in foyer of TCPL (photo will be taken outdoors so a jacket might be required)
14:15-15:00 C. Adcock, Discussion: pros and cons of skew-symmetric distributions and copulas (45)
15:00-15:30 Coffee break
15:30-17:00 J. Neslehova, Tests of independence for sparse contingency tables and beyond (20+10)
B. Remillard, Testing hypotheses for the copula of dynamic models (20+10)
I. Kojadinovic, An unconditional tapered multiplier bootstrap for the sequential empirical copula process under strong mixing (20+10)
17:00-18:00 Posters (60)
18:00-19:30 Dinner

Wednesday (chair: N. Balakrishnan)

7:00-9:00 Breakfast
9:00-10:00 P. Naveau, Analysis of heavy rainfall in high dimensions (20+10)
R. Loschi, Nonparametric mixtures based on skew-normal distributions: An application to density estimation (20+10)
10:00-10:30 Coffee break
10:30-11:30 R. Cooke, Tail dependence elicitation for ice sheets (20+10)
C. Adcock, Some challenges in portfolio theory and asset pricing (20+10)
11:30-13:00 Lunch
Free Afternoon (optional but highly recommended: bus tour to Johnston Canyon and Lake Louise for a fee)
17:30-19:30 Dinner
Jazz (20:00 at The Club, Theatre Complex, and 23:00 at Maclab Bistro, both at Banff Centre)

Thursday (chair: C. Field)

7:00-9:00 Breakfast
9:00-10:00 C. Jones, Alternative skew-symmetric distributions (20+10)
E. Valdez, Multivariate negative binomial models for insurance claim counts (20+10)
10:00-10:30 Coffee break
10:30-12:00 A. Nikoloulopoulos, Factor copula models for item response data (20+10)
T. Yoshiba, Skew t-copula and its estimation: For application to risk aggregation (20+10)
D. Kurowika, Joint density of correlations in correlation matrix with chordal sparsity patterns (20+10)
12:00-13:30 Lunch
13:30-14:30 A. McNeil, Discussion: Multivariate scenario sets in the non-Gaussian world (60)
14:30-15:10 A. Azzalini: The new "sn" package for R (30+10)
15:10-15:40 Coffee break
15:40-16:20 I. Kojadinovic: A brief presentation of the R copula package (30+10)
16:20-17:00 E. Brechmann: Statistical inference of vine copulas using the R-package VineCopula (30+10)
17:00-17:30 Discussion
17:30-19:30 Dinner
Jazz (20:00 at The Club, Theatre Complex, and 23:00 at Maclab Bistro, both at Banff Centre)

Friday

7:00-9:00 Breakfast
9:00-10:00 Discussion (60)
10:00-10:30 Coffee break
10:30-11:30 Discussion (60)
11:30-13:30 Lunch

Checkout by 12 noon.

** 5-day workshop participants are welcome to use BIRS facilities (BIRS Coffee Lounge, TCPL and Reading Room) until 3 pm on Friday, although participants are still required to checkout of the guest rooms by 12 noon. **

Non-Gaussian Multivariate Statistical Models and their Applications
May 19-24, 2013

TALKS ABSTRACTS

Speaker: Kjersti Aas (Norwegian Computing Center, Norway) & Claudia Czado (Technische Universität München, Germany)

Title: Pair-Copula Constructions – Even more Flexible than Copulas

Abstract: A copula is a multivariate distribution with standard uniform marginal distributions. While the literature on copulas is substantial, most of the research is still limited to the bivariate case. However, some years ago hierarchical copula-based structures were proposed as an alternative to the standard copula methodology. One of the most promising of these structures is the pair-copula construction. Pair-copula constructions are also called regular vines. The modeling scheme is based on a decomposition of a multivariate density into a cascade of pair copulae, applied on original variables and on their conditional and unconditional distribution functions. Each pair copula can be chosen arbitrarily and the full model exhibit complex dependence patterns such as asymmetry and tail dependence. In this talk we will give an introduction to pair-copula constructions and apply the methodology to a 19-dimensional financial data set. Significant recent advances in this field will also be reviewed.

Speaker: Chris J. Adcock (The University of Sheffield, UK)

Title: Pros and Cons of Skew-Symmetric Distributions and Copulas – Discussion

Abstract: The use of non-normal multivariate distributions generally presents challenges, which are both theoretical and empirical. Different classes of such distributions offer different opportunities and problems. The aim of this discussion is twofold. First it is to agree a list of headings with which non-normal multivariate distributions may be characterised. Secondly it is to discuss the pros and cons of skew-symmetric distributions and copulas under these headings. An obvious example is the computation of tail probabilities, for which the differences in the two types of distributions are clear. An output from this discussion could be the re-invigoration of existing streams of research. It could include a list of potential research projects.

Speaker: Chris J. Adcock (The University of Sheffield, UK)

Title: Some Challenges in Portfolio Theory and Asset Pricing

Abstract: The foundations of modern finance are Markowitz' theory of portfolio selection, the Capital Asset Pricing Model (CAPM) of Sharpe, Lintner and Mossin, the option pricing formula due to Black, Scholes and Merton and, perhaps to a lesser extent, Ross' Arbitrage Pricing Theory (APT). In their traditional form, these foundations depend to a considerable extent on the assumption that the returns on risky financial assets follow a multivariate normal distribution or, in some cases, a multivariate elliptically symmetric distribution. Some of the foundations continue to hold, with suitable extensions,

if instead returns follow some skew-symmetric symmetric distributions. This presentation covers some of the challenges that face those who would seek to extend portfolio theory and asset pricing under other non-normal multivariate distributions. The talk, which presents the author's experiences based both on his research and professional experience, covers statistical issues and those which are more specific to finance.

Speaker: Adelchi Azzalini (University of Padua, Italy)

Title: Skew-Symmetric Distributions in 45'

Abstract: The phrase "skew-symmetric distributions" refers to a mechanism for the generation of families of distributions starting from a given symmetric "base" distribution. The talk aims at presenting a few key formal facts in a non-formal way and discussing the qualitative aspects. Specific illustrations will focus on some noteworthy special cases.

Speaker: Adelchi Azzalini (University of Padua, Italy)

Title: The New "sn" Package for R

Abstract: The "sn" package has existed for about 15 years on the R repository, to deal with skew-normal and skew-t distributions. The package is now being completely re-organized, with large portions re-written, both to incorporate recent advances in research and to (finally!) accomplish the R programming standards.

Speaker: Eike Brechmann (Technische Universität München, Germany)

Title: Statistical Inference of Vine Copulas using the R-Package VineCopula

Abstract: In this talk, the R-package VineCopula will be presented and illustrated by means of an example data set. The package contains functions for statistical inference of vine copulas and tools for exploratory data analysis and selection of bivariate copulas as building blocks of vine copulas. Vine trees can be selected using a tree-by-tree approach. Models can be estimated by joint maximum likelihood estimation and standard errors of model parameters are provided. Algorithms for sampling and illustrating vine copulas are also included.

Speaker: Roger M. Cooke (Resources for the Future, Netherlands)

Title: Tail Dependence Elicitation for Ice Sheets

Abstract: The behavior of the East Antarctic, West Antarctic and Greenland Ice Sheets is one of the major uncertainties regarding the impacts of climate change. The geological record shows several instances of rapid instabilization and break up which current models cannot emulate. The ice sheet contribution to sea level rise [mm/yr] is calculated yearly as *Discharge* + *Runoff* - *Accumulation*. Uncertainty in the contribution to sea level rise has been repeatedly assessed using structured expert judgment (Bamber, J.L., and Aspinall, W.P., (2012) An expert judgement assessment of future sea 1evel rise from the ice sheets, Nature Climate Change, PUBLISHED ONLINE: January 6, 2012 | DOI: 10.1038/NCLIMATE1778.http://www.nature.com/nclimate/journal/vaop/ncurrent/full/nclimate1778.html). Factors such as loss of ice shelves which buffer ice sheets, loss of sea ice, changing ocean currents and moisture content of the atmosphere may jointly influence these factors both within and between ice

sheets. As part of an elicitation of European ice sheet experts, an attempt was made to capture these dependences in a simplified vine copula representation of the joint distribution, and to capture departures from Gaussian dependence. Ice sheet experts quickly became familiar with the protocol for dependence elicitation and were able to quantify significant departures from independence and from Gaussian dependence. The results are provisional as a US elicitation is planned for June 2013. The talk will focus on methods and preliminary dependence elicitation results.

Speaker: Dipak K. Dey (University of Connecticut, USA)

Title: State Space Models for Binary Response Data with Flexible Skewed Link Functions

Abstract: State space models (SSM) for binary time series data using a flexible skewed link functions are introduced in this presentation. Commonly used probit and logit links are prone to link misspecification because of their fixed skewness. Here we introduce three flexible links as alternatives, they are generalized extreme value (GEV) link, symmetric power logit (SPLOGIT) link and scale mixture of normal (SMN) link. Markov chain Monte Carlo (MCMC) methods for Bayesian analysis of SSM with these links are implemented using the JAGS package, freely available software. Model comparison relies on the deviance information criterion (DIC). The flexibility of the proposed model is illustrated to measure effects of deep brain stimulation (DBS) on attention of a macaque monkey performing a reaction-time task (Smith et al., 2009). Empirical results showed that the flexible links provide better fitting over the usual probit and logit links. Joint work with C. A. Abanto-Valle and Xun Jiang.

Speaker: Marc G. Genton (King Abdullah University of Science and Technology, Saudi Arabia)

Title: Nonparametric Identification of Copula Structures

Abstract: We propose a unified framework for testing a variety of assumptions commonly made about the structure of copulas, including symmetry, radial symmetry, joint symmetry, associativity and Archimedeanity, and max-stability. Our test is nonparametric and based on the asymptotic distribution of the empirical copula process. We perform simulation experiments to evaluate our test and conclude that our method is reliable and powerful for assessing common assumptions on the structure of copulas, particularly when the sample size is moderately large. We illustrate our testing approach on two datasets. Joint work with Bo Li.

Speaker: Emanuele Giorgi (Lancaster University, UK)

Title: Moment-Free Measures for the Multivariate Skew-t Distribution

Abstract: The main features that are used to describe a multivariate probability distribution are its location, dispersion, structure of dependence, asymmetry and kurtosis. These are generally associated to the moments of the distribution if these are defined. In the case of the multivariate skew-t distribution, the moments of order greater than or equal to the respective degrees of freedom do not exist. To overcome such a problem, we propose a set of measures that are defined regardless of the value of the degrees of freedom. We use measures based on the halfspace depth function in order to summarize location and kurtosis. The halfspace depth function can be viewed as a multivariate generalization of the rank method in the univariate case and it leads to the definition of a multivariate median with desirable properties, such as affine invariance and robustness. We investigate the properties of the halfspace depth function for the skew-t distribution and its relation to radial symmetry. For the particular case of one degree of freedom, we show that the skew-t distribution is

radially symmetric with center of radial symmetry given by the halfspace median. Moreover, in this case, the elliptical shape of the halfspace depth contours eases the computation of a multivariate measure of kurtosis. As for dispersion and structure of dependence, we use measures based on notions of concordance (e.g. Kendall's tau and Spearman's rho) and quantile. Finally, asymmetry is quantified by computing the Bowley index for each of the marginal distributions.

Speaker: Lei Hua (Northern Illinois University, USA)

Title: Strength of Tail Dependence based on Conditional Tail Expectation

Abstract: Tail order is a quantity to describe the degree of dependence in joint tails. In addition to the usual tail dependence, the tail order also covers intermediate tail dependence, tail quadrant independence and even tail negative dependence. Strength of dependence in the tails can also be captured through the conditional distribution of one random variable given the other one being large. We study how these two ways of measuring tail dependence influence each other, especially the behavior of the boundary conditional cumulative distribution function for various bivariate copula families. Also we study the tail behavior of two forms of conditional tail expectation (CTE). In this talk, three levels of strength of tail dependence will be discussed, and they can be described by the tail behavior of CTEs: asymptotically linear, sub-linear and constant. For each of these three levels, we investigate the tail behavior of CTEs for marginal distributions belonging to maximum domain of attraction of Frechet and Gumbel, respectively, and for copula families with different tail behavior.

Speaker: María Dolores Jiménez-Gamero (Universidad de Sevilla, Spain)

Title: Testing for Skew-Symmetric Models

Abstract: Skew-symmetric (ss) models are semiparametric models for continuous random vectors. Starting with a symmetric probability density function (pdf), a ss family is generated by multiplying the symmetric pdf by a skewing function. A characterization of this skewing function let us propose a test for testing goodness of fit to a ss model. The test statistic is based on comparing two estimators of the population pdf: a consistent kernel-based estimator and another estimator which is built by assuming that the ss in the null hypothesis is true. The proposed test shares some properties with other tests based on comparing a kernel-based estimator of the population pdf to an estimator derived under the null hypothesis, like the ones in Hall (J. Multivar. Anal. 14, 1984, 1--16) and Fan (Econometric Theory, 10, 1994, 316--356), and other are new.

Speaker: Harry Joe (University of British Columbia, Canada)

Title: Overview of Copula Models: Dependence Structures and Asymmetries

Abstract: The overview will discuss asymmetries and tail behavior, and the diagnostics for them that make inadequate the use classical multivariate statistics based on multivariate normal. Although copulas are multivariate uniform distributions, plots of multivariate data and copula densities are more informative after transforms to standard normal margins. Some classes of multivariate copulas will be summarized together with some of their dependence and tail properties. Copula models that can be used for high-dimensional data (dimensions of 50-100 and beyond) are based on vines or pair-copula constructions. It will be shown how vines extend multivariate normal distributions after the d-dimensional correlation matrix is reparametrized in terms of $d-1$ correlations and $(d-1)(d-2)/2$ partial correlations that are algebraically independent (i.e., parameter space is $(-1,1)^{d(d-1)/2}$ with no positive definite matrix constraint). With the appropriate reparametrization to partial correlations,

the copula version of many methods based on multivariate normal can be obtained. These include common and structured (confirmatory) factor analysis, latent variable models, structured covariance models, and factor models with residual dependence.

Speaker: Chris Jones (The Open University, UK)

Title: Alternative Skew-Symmetric Distributions

Abstract: I will remind the gathering about some families of univariate skew-symmetric distributions that are competitors/complements to the one that I assume will dominate proceedings prior to this talk. Each family has its pros and its cons. Two of those other families and the "Azzalini" skew-symmetric family form a trio with interesting interconnections. I will probably also make some brief remarks on their bivariate extensions and interactions with copulas.

Speaker: Ivan Kojadinovic (Université de Pau et des Pays de l'Adour, France)

Title: An Unconditional Tapered Multiplier Bootstrap for the Sequential Empirical Copula Process under Strong Mixing

Abstract: Two key ingredients to carry out inference on the unknown copula of multivariate observations are the empirical copula process and an appropriate resampling scheme to obtain replicates of the latter. Among the existing resampling techniques used for i.i.d. observations, the multiplier bootstrap introduced in the seminal paper of Rémillard and Scaillet (2009) frequently appears to lead to inference procedures with the best finite-sample properties. Bücher and Ruppert (2013) recently proposed an extension of this technique to strictly stationary strongly mixing observations by adapting the tapered muliplier bootstrap of Bühlmann (1993) to the empirical copula process. The main result of the work to be presented is an unconditional and sequential version of the tapered multiplier bootstrap of Bühlmann whose validity requires substantially weaker conditions on the rate of decay of the strong mixing coefficients and slightly less constrained multipliers. The obtained result can also be regarded as a partial extension of the unconditional multiplier central limit theorem to the case of strongly mixing observations. The resulting generalization of the tapered multiplier bootstrap of Bühlmann is then adapted to the sequential empirical copula process thereby allowing to transpose to the strongly mixing setting all of the existing multiplier tests on the unknown copula, including nonparametric tests for change-point detection. Joint work with Axel Bücher.
.

Speaker: Ivan Kojadinovic (Université de Pau et des Pays de l'Adour, France)

Title: A Brief Presentation of the R Copula Package

Abstract: The copula package (http://CRAN.R-project.org/package=copula) is a package for the R statistical system. It contains routines for dealing with frequently encountered copula families, mostly of the elliptical, Archimedean and extreme value type. Methods for density and c.d.f. evaluation, random number generation, perspective and contour plots are available. In addition to fitting routines, the package can be used to carry out various types of tests such as independence, exchangeability and goodness-of-fit tests. The package is distributed under the free copyleft GPL3 license in the hope that it will be useful, but *without any warranty*. It is co-developed with Marius Hofert, Martin Maechler, and Jun Yan.

Speaker: Dorota Kurowicka (Delft University of Technology, Netherlands, and Nanyang Technological University, Singapore)

Title: Joint Density of Correlations in Correlation Matrix with Chordal Sparsity Patterns

Abstract: We extend the methodology of generating random correlation matrix of Joe, (2006) and Lewandowski et al. (2009) by introducing a partial correlation expansion of the determinant of a correlation matrix which is more general than the partial correlations on a regular vine used in Lewandowski et al. (2009). This generalization allows to formulate the partial correlation expansion of determinant for a correlation matrix with chordal sparsity pattern. For such a partially specified correlation matrix we find a uniform density of unspecified correlations. This leads to a closed form formula for the volume of the space of correlation matrices with specified correlations corresponding to a chordal graph.

Speaker: Brunero Liseo (Sapienza Universita di Roma, Italy)

Title: Bayesian Inference for the Multivariate Skew-Normal Distribution: A Population Monte-Carlo Approach

Abstract: Frequentist and likelihood methods of inference based on the multivariate skew-normal model encounter several technical difficulties with this model. In spite of the popularity of this class of densities, there are no broadly satisfactory solutions for estimation and testing problems. A general population Monte Carlo algorithm is proposed which: 1) exploits the latent structure stochastic representation of skew-normal random variables to provide a full Bayesian analysis of the model and 2) accounts for the presence of constraints in the parameter space. The proposed approach can be defined as weakly informative, since the prior distribution approximates the actual reference prior for the shape parameter vector. Results are compared with the existing classical solutions and the practical implementation of the algorithm is illustrated via a simulation study and a real data example. A generalization to the matrix variate regression model with skew-normal error is also presented. Joint work with Antonio Parisi**.**

Speaker: Rosangela Loschi (Universidade Federal de Minas Gerais, Brazil)

Title: Nonparametric Mixtures based on Skew-Normal Distributions: An Application to Density Estimation

Abstract: This work addresses the density estimation problem using non-parametric Bayesian approach. It is considered hierarchical mixture models where the uncertainty about the mixing measure is modeled using the Dirichlet process. The main goal is to build more flexible models for density estimation. Extensions of the normal mixture model via Dirichlet process previously introduced in the literature are twofold. Firstly, Dirichlet mixtures of skew-normal distributions are considered, say, in the first stage of the hierarchical model, the normal distribution is replaced by the skew-normal one. We also assume a skew-normal distribution as the center measure in the Dirichlet mixture of normal distributions. Some important results related to Bayesian inference in the location-scale skew-normal family are introduced. In particular, we obtain the stochastic representations for the full conditional distributions of the location and skewness parameters. The algorithm introduced by MacEachern and Mueller in 1998 is used to sample from the posterior distributions. The models are compared considering simulated data sets. Finally, the well-known Old Faithful Geyser data set is analyzed using the proposed models and the Dirichlet mixture of normal distributions. The model based on Dirichlet mixture of skew-normal distributions captured the data bimodality and skewness shown in the empirical distribution. Joint work with Caroline Vieira and Denise Duarte.

Speaker: Alexander J. McNeil (Heriot-Watt University, UK)

Title: Multivariate Scenario Sets in the Non-Gaussian World (Discussion)

Abstract: There are numerous ways of creating multivariate scenario sets based on multivariate probability distributions. The idea is that these sets should contain the most plausible values and exclude the most extreme values generated by the distribution. The terminology "scenario set" comes from applications in financial risk modelling where the elements of the set have interpretations as future values of financial risk factors. Scenario sets can be constructed from density contours but, while this is natural for elliptical distributions, it is not natural for other kinds of distribution (e.g. multivariate discrete distributions). An alternative method uses the concepts of half-space trimming and depth to define convex scenario sets, which can be thought of as higher-dimensional generalisations of quantiles. I will review this method with some examples and then explain how further scenario sets can be defined by replacing quantiles with other measures of "tail risk" such as the so-called expected shortfall or tail conditional expectation. I will give an idea of why these ideas are important in financial risk management but also highlight the computational difficulties of using the ideas in practice. My aim is to generate discussion of the different ways of creating scenario sets and their pros and cons. In particular I would like to explore how the ideas could be implemented in practice for non-Gaussian distributions in higher dimensions.

Speaker: Philippe Naveau (CNRS, France)

Title: Analysis of Heavy Rainfall in High Dimensions

Abstract: One of the main objectives of statistical climatology is to extract relevant information hidden in complex spatial-temporal climatological datasets. In impact studies, heavy rainfall are of primary importance for risk assessment linked to floods and other hydro- logical events. At an hourly time scale, precipitation distributions often strongly differ from Gaussianity. To identify spatial patterns, most well-known statistical techniques are based on the concept of intra and inter clusters variances (like the k-means algorithm or PCA's) and such approaches based on deviations from the mean may not be the most appropriate strategy in our context of studying rainfall extremes. One additional difficulty resides in the dimension of climatological databases of hourly recordings that may gather measurements from hundreds or even thousands of weather stations during many decades. A possible avenue to fill up this methodological gap resides in taking advantage of multivariate extreme value theory, a well-developed research field in probability, and to adapt it to the context of spatial clustering. In this talk, we propose and study two step algorithm based on this plan. Firstly, we adapt a Partitioning Around Medoids (PAM) clustering algorithm proposed by Kaufman to weekly maxima of hourly precipitation. This provides a set of homogenous spatial clusters of extremes of reasonable dimension. Secondly, we fine-tune our analysis by fitting a Bayesian Dirichlet mixture model for multivariate extremes within each cluster. We compare and discuss our approach throughout the analysis of hourly precipitation recorded in France (Fall season, 92 stations, 1993-2011). Joint work with A. Sabourin, E. Bernard, O. Mestre, and M. Vrac.

Speaker: Johanna G. Neslehova, (McGill University, Canada)

Title: Tests of Independence for Sparse Contingency Tables and Beyond

Abstract: New statistics are proposed for testing the hypothesis that arbitrary random variables are mutually independent. These tests are consistent and well-behaved for any type of data, even for sparse contingency tables and tables whose dimension depends on the sample size. The statistics are Cram_ér-

von Mises and Kolmogorov-Smirnov type functionals of the empirical checkerboard copula. The asymptotic behavior of the corresponding empirical process will be characterized and illustrated; it will also be shown how replicates from the limiting process can be generated using a multiplier bootstrap procedure. As will be seen through simulations, the new tests are considerably more powerful than those based on the Pearson chi squared, likelihood ratio, and Zelterman statistics often used in this context.

Speaker: Aristidis K. Nikoloulopoulos (University of East Anglia, UK)

Title: Factor Copula Models for Item Response Data

Abstract: Factor or conditional independence models based on copulas are proposed for multivariate discrete data such as item response. The factor copula models have interpretations of latent maxima/minima (in comparison with latent means) and can lead to more probability in the joint upper and/or lower tail compared with factor models based on the discretized multivariate normal distribution (or multidimensional normal ogive model). Details on maximum likelihood estimation of parameters for the factor copula model are given, as well as analysis of the behavior of the log-likelihood. Our general methodology is illustrated with several item response data sets and it is shown that there is a substantial improvement on existing models both conceptually and in fits to data. Joint work with Harry Joe.

Speaker: Arthur Pewsey (University of Extremadura, Spain)

Title: Inference for Skew-Symmetric Distributions

Abstract: In my talk I will present a review of some of the issues that can arise when performing inference for skew-symmetric distributions. Those issues condition the parametrizations we might use, the strategies we might adopt when carrying out certain forms of inference, and the interpretation we give to the results obtained. Much of the discussion will centre on Azzalini's univariate skew-normal distribution employed as a case study familiar to us all.

Speaker: Bruno Remillard (HEC Montreal, Canada)

Title: Testing Hypotheses for the Copula of Dynamic Models

Abstract: The asymptotic behavior of the empirical copula constructed from residuals of stochastic volatility models is studied. It is shown that if the stochastic volatility matrix is diagonal, then the empirical copula process behaves like if the parameters were known, a remarkable property. However, this is not true in general. Applications for goodness-of-fit and detection of structural change in the copula of the innovations are discussed.

Speaker: Emiliano Valdez (University of Connecticut, USA)

Title: Multivariate Negative Binomial Models for Insurance Claim Counts

Abstract: It is no longer uncommon these days to find the need in actuarial practice to model claim counts from multiple types of coverage, such as the ratemaking process for bundled insurance contracts. Since different types of claims are conceivably correlated with each other, the multivariate count regression models that emphasize the dependency among claim types are more helpful for

inference and prediction purposes. Motivated by the characteristics of an insurance dataset, we investigate alternative approaches to constructing multivariate count models based on the negative binomial distribution. A classical approach to induce correlation is to employ common shock variables. However, this formulation relies on the NB-I distribution which is restrictive for dispersion modeling. To address these issues, we consider two different methods of modeling multivariate claim counts using copulas. The first one works with the discrete count data directly using a mixture of max-id copulas that allows for flexible pair-wise association as well as tail and global dependence. The second one employs elliptical copulas to join continuitized data while preserving the dependence structure of the original counts. The empirical analysis examines a portfolio of auto insurance policies from a Singapore insurer where claim frequency of three types of claims (third party property damage, own damage, and third party bodily injury) are considered. The results demonstrate the superiority of the copula-based approaches over the common shock model. Finally, we implemented the various models in loss predictive applications. Joint work with Peng Shi.

Speaker: Stanislav Volgushev (Ruhr-Universitaet Bochum, Germany)

Title: Empirical and Sequential Empirical Copula Processes under Serial Dependence and Weak Smoothness Conditions

Abstract: The empirical copula process plays a central role for statistical inference on copulas. The main purpose of this talk is to give an overview of recent results on the asymptotic properties of this process. In the first part of the talk, we focus on process convergence with respect to the supremum norm. We also briefly discuss properties of sequential copula processes. It is known that in some interesting examples, weak convergence of the copula process with respect to the supremum norm does not hold. In the second part of the talk, we introduce a new metric that allows to handle most of those cases. Joint work with Axel Buecher and Johan Segers.

Speaker: Toshinao Yoshiba (Bank of Japan, Japan)

Title: Skew t-Copula and its Estimation: For Application to Risk Aggregation

Abstract: Correlation structure of risk factors matters to financial portfolio risk management. When the risk factors are specified by some assets' return, risk managers have to care about lower tail dependence of those factors rather than upper. On the other hand, it is practical to specify overall nontail dependence by sample linear or rank correlation matrix. Based on this background, we focus on the application of skew t-copula. Skew t-copula is defined by a multivariate skew t-distribution and its marginal. As indicated in Kotz and Nadarajah (2004), various types of multivariate skew t-distribution have been proposed. That implies various types of skew t-copula can exist. Three types of skew t-copula are known so far. The first type was mentioned in Demarta and McNeil (2005), which is based on multivariate version of generalized hyperbolic (GH) skew t-distribution proposed in Barndorff-Nielsen (1977) (See also Aas and Haff (2006)). The second type was constructed in Smith et al. (2012) based on Sahu et al. (2003). Kollo and Pettere (2010) tried to construct the third type based on Azzalini and Capitanio (2003). This paper is constructed as follows. First, we improve the skew t-copula approach proposed in Kollo and Pettere (2010) and derive log-likelihood function to estimate the parameters by maximum likelihood estimation (MLE). Second, we indicate that MLE for the skew t-copula requires fast and accurate quantile calculation for univariate skew t-distribution. Third, we show the difference between upper and lower tail dependence of the skew t-copula citing Fung and Seneta (2010) and Bortot (2010). Finally, we introduce a method of moment approach for the parameter estimation of the skew t-copula.

POSTERS ABSTRACTS

Speaker: Marcia D´Elia Branco (University of São Paulo, Brazil)

Title: Bayesian Sensitivity Analysis under a Skew-Normal Class of Prior Distributions using Concentration Function

Abstract: An important question in statistical inference is the possible mis-specification of the probabilistic model. Bayesian robustness is concerned with the impact of the model specification on the posterior inference. Considering a normal likelihood and a general class of prior distributions with different shapes for the location parameter, we study the behavior of the posterior class of distribution. As a divergence measure for the posterior distributions we consider the concentration function proposed in Fortini and Ruggeri (1995). The usual conjugate Bayesian property for the normal model is extended here for a general class of skew-normal distributions. We obtain closed expressions for the concentration function between the normal and skew-normal distributions, that is helpful to study the robustness of the posterior distributions. Moreover, a careful study of the behavior of these measures are developed through a simulation study. Joint work with Luciana Graziela de Godoi.

Speaker: Eike Brechmann (Technische Universität München, Germany)

Title: Joint Modeling of Multivariate Return Periods

Abstract: The notion of a return period is common, e.g., in hydrology and engineering. In a multivariate framework, return periods of dependent variables can be studied using copulas. As recently discussed in the literature, a multivariate return period can then be quantified using the Kendall distribution function of a copula. We propose a methodology to specify the interdependence of different multivariate return periods in a hierarchical setup. The proposed methodology then allows the study of joint extreme events. In an application, we analyze extreme sea levels in England. Multivariate return periods are characterized using asymmetric Tawn copulas.

Speaker: Hyoung-Moon Kim (Konkuk University, Korea)

Title: Mixtures of Skewed Kalman Filters

Abstract: Normal state-space models are prevalent, but to increase the applicability of the Kalman filter, we propose mixtures of skewed, and extended skewed, Kalman filters. To do so, the closed skew-normal distribution is extended to a scale mixture class of closed skew-normal distributions. Some basic properties are derived and a class of closed skew-t distributions is obtained. Our suggested family of distributions is skewed and has heavy tails too, so it is appropriate for robust analysis. Our proposed special sequential Monte Carlo methods use a random mixture of the closed skew-normal distributions to approximate a target distribution. Hence it is possible to handle skewed and heavy tailed data simultaneously. These methods are illustrated with numerical experiments. Joint work with Duchwan Ryu, Bani K. Mallick, and Marc G. Genton.

Speaker: Pavel Krupskii (University of British Columbia, Canada)

Title: Factor Copula Models for Multivariate Data

Abstract: General conditional independence models for d observed variables, in terms of p latent variables, will be presented in terms of bivariate copulas that link observed data to latent variables. The representation is called a factor copula model and the classical multivariate normal model with a correlation matrix having a factor structure is a special case. The factor copula model can handle multivariate data with tail dependence and tail asymmetry, properties that the multivariate normal copula does not possess. It is a good choice for modeling high-dimensional data as a parametric form can be specified to have O(d) dependence parameters instead of O(d^2) parameters.

Speaker: Farouk S. Nathoo (University of Victoria, Canada)

Title: A Skew-t Space-Varying Regression Model for Spectral Analysis of Resting State Brain Activity

Abstract: It is known that in many neurological disorders such as Down syndrome, main brain rhythms shift their frequencies slightly, and characterizing the spatial distribution of these shifts is of interest. This paper reports on the development of a Skew-t mixed model for the spatial analysis of resting state brain activity in healthy controls and individuals with Down syndrome (DS). Time series of oscillatory brain activity are recorded using magnetoencephalography (MEG), and spectral summaries are examined at multiple spatial locations across the scalp. We focus on the mean-frequency of the power spectral density, and use space-varying regression to jointly examine associations with age, gender and Down syndrome across several scalp regions. Spatial smoothing priors are incorporated based on a multivariate Markov random field, and the markedly non-Gaussian nature of the spectral response variable is accommodated by the use of a Skew-t distribution. A range of models representing different assumptions on the association structure and response distribution are examined. Our analysis suggests region-specific differences between healthy controls and individuals with Down syndrome, particularly in the left and right temporal regions of the brain, and produces smoothed maps indicating the scalp topography of the estimated differences. Joint work with Salimah Ismail, Wenqi Sun, Arif Babul, Alexader Moiseev, Mirza Faisal Beg, Naznin Virji-Babul.

Speaker: Bruno Scarpa (University of Padua, Italy)

Title: Fitting Age-Specific Fertility Rates by a Flexible Generalized Skew-Normal Probability Density Function

Abstract: Mixture probability density functions have recently been proposed to describe some fertility patterns characterized by a bimodal shape. These functions are adequate when the fertility pattern is actually bimodal, but appear to be less useful when the shape of age-specific fertility rates is unimodal. A further model is proposed, based on flexible skew-symmetric probability density functions. This model is able to fit symmetric and skew patterns as well as to reflect humps observed in some fertility patterns. It is both more parsimonious than mixture distributions and more flexible, showing a good fit with several shapes (bimodal or unimodal) of fertility patterns. Empirical evaluation of the proposed model and comparisons with other functions to the Italian data, from 1952 to 2003 and US data from 1933 to 2006 are also discussed. Joint work with Stefano Mazzuco.

Speaker: Ying Sun (University of Chicago, USA)

Title: Functional Median Polish with Climate Applications

Abstract: We propose functional median polish, an extension of univariate median polish, for one-way and two-way functional analysis of variance (ANOVA). The functional median polish estimates the

functional grand effect and functional main factor effects based on functional medians in an additive functional ANOVA model assuming no interaction among factors. A functional rank test is used to assess whether the functional main factor effects are significant. The robustness of the functional median polish is demonstrated by comparing its performance with the traditional functional ANOVA fitted by means under different outlier models in simulation studies. The functional median polish is illustrated on various applications in climate science, including one-way and two-way ANOVA when functional data are either curves or images. Specifically, Canadian temperature data, U.S. precipitation observations and outputs of global and regional climate models are considered, which can facilitate the research on the close link between local climate and the occurrence or severity of some diseases and other threats to human health. Joint work with Marc G. Genton.