ADISTA17

International Directional Statistics Workshop

Programme and Book of Abstracts

8th-9th June 2017

Roma Tre University, Rome, Italy

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Venue

Roma Tre University,
Department of Political Sciences,
Via Gabriello Chiabrera 199,
Rome

Organizing committee

Giovanna Jona-Lasinio (Sapienza Università di Roma)
Francesco Lagona (Università di Roma Tre)
Christophe Ley (Universiteit Gent)
Arthur Pewsey (Universidad de Extremadura)

Sponsors

Programme

	Day 1 (8th June)
08:30 - 09:00	Registration Coffee
09:00 - 09:15	Workshop Inauguration Giuseppe Di Battista, Francesco Guida and Francesco Lagona
09:20 - 09:40	Day 1 Inaugural talk (Chair: John Kent) Peter Jupp
09:40 - 10:20	Tests for Rotational Symmetry (Chair: Davy Paindaveine) Christophe Ley and Thomas Verdebout
10:20 - 10:50	Coffee Break
10:50 - 11:50	Applications of Directional Statistics (Chair: Domenico Marinucci) John Kent, Irene Klugkist and Lasse Holmström
11:50 - 12:00	Break
12:00 - 12:40	Poster Storming (Chair: Giovanna Jona-Lasinio) José Ameijeiras-Alonso, Davide Buttarazzi, José Chacón, Jolien Cremers, Christine Cutting, Stefania Fensore, Ulrike Genschel, Tomoaki Imoto, Pieter Jongsma, Zeynep Kalaylioglu, Yolanda Larriba, Ignacio Leguey, Sergio Luengo-Sanchez, Kees Mulder, Giovanni Porzio, Karen Sargsyan, Anna Sjöström, Ville Vuollo
12:40 - 13:50	Lunch
13:50 - 14:30	Poster Viewing
14:30 - 15:30	Statistics on the Cylinder (Chair: Antonello Maruotti) Toshihiro Abe, Kunio Shimizu and Gianluca Mastrantonio
15:30 - 15:50	Coffee Break
15:50 - 16:50	Stochastic Processes and Time Series (Chair: Shogo Kato) Eduardo García-Portugués, Aurélien Nicosia and Francesco Lagona
16:50 - 17:00	Break
17:00 - 17:40	Testing for Maxima (Chair: José Chacón) Rosa Crujeiras and Domenico Marinucci
20:00 -	Workshop Dinner

	Day 2 (9th June)
10:00 - 10:20	Day 2 Inaugural Talk (Chair: Louis-Paul Rivest) Kanti Mardia
10:20 - 11:00	Statistics on the Sphere (Chair: Thomas Verdebout) Davy Paindaveine and Shogo Kato
11:00 - 11:30	Coffee Break
11:30 - 12:30	Rotations and Orientations (Chair: Peter Jupp) Richard Arnold, Dan Nordman and Giovanna Jona-Lasinio
12:30 - 14:00	Lunch
14:00 - 15:00	Statistics on the Torus (Chair: Kunio Shimizu) Arthur Pewsey, Miguel Fernández and Stephan Huckemann
15:00 - 15:10	Break
15:10 - 15:50	Nonparametric Methods (Chair: Rosa Crujeiras) Agnese Panzera and Marco Di Marzio
15:50 - 16:20	Coffee Break
16:20 - 17:20	Regression, Longitudinal and Compositional Data Analysis (Chair: Eduardo García-Portugués) Louis-Paul Rivest, Antonello Maruotti and Janice Scealy
17:20 - 17:45	Closure of ADISTA17 and Presentation of ADISTA20 Giovanna Jona-Lasinio, Francesco Lagona, Christophe Ley, Arthur Pewsey and Rosa Crujeiras

Abstracts: Invited Speakers

A flexible cylindrical model as a combination of the Weibull and SSVM distributions for cylindrical data

Speaker: Toshihiro Abe
Department of Systems and Mathematical Science, Nanzan University, Japan
Email: abetosh@ss.nanzan-u.ac.jp

Abstract

In this talk, we consider a cylindrical model obtained by combining the sine-skewed von Mises distribution with the Weibull distribution. Our proposed model, the WeiSSVM model, has numerous good properties such as simple normalizing constant and hence very tractable density, parameter-parsimony and interpretability, maximum entropy characterization, good circular-linear dependence structure, easy random number generation thanks to known marginal and conditional distributions, and flexibility illustrated via excellent fitting abilities. As an illustrative example, our model is applied to analyses a blue periwinkle data set.

This is joint work with Christophe Ley.

Statistics of ambiguous rotations

Speaker: Richard Arnold
School of Mathematics and Statistics, Victoria University of Wellington, New Zealand
Email: richard.arnold@msor.vuw.ac.nz

Abstract

Orientations of objects in $\mathbb{R}^p$ with symmetry group K cannot be described unambiguously by elements of the rotation group $SO(p)$, but correspond to elements of the quotient space $SO(p)/K$. Specifications of probability distributions and appropriate statistical methods for such objects have been lacking – with the notable exceptions of axial objects in the plane and in $\mathbb{R}^3$.

We exploit suitable embeddings of $SO(p)/K$ into spaces of symmetric arrays to provide a systematic and intuitively appealing approach to the statistical analysis of the orientations of such objects. We firstly consider the case of *orthogonal r-frames* in $\mathbb{R}^p$, corresponding to sets of $r \leq p$ orthogonal axes in p dimensions. Probability distributions in this setting include the Watson and Bingham distributions. Using the same approach we then treat the three dimensional case of $SO(3)/K$, where K is one of the point symmetry groups.

In both cases the resulting tools include measures of location and dispersion, tests of uniformity, one-sample tests for a preferred orientation and two-sample tests for a difference in orientation. The methods are illustrated using data from earthquake focal mechanisms (fault plane and slip vector orientations) and crystallographic data with orientation measurements of distinct mineral phases in EBSD experiments.

This is joint work with Peter Jupp.

Fire seasonality identification by circular multimodality tests

Speaker: Rosa Crujeiras
Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Spain
Email: rosa.crujeiras@usc.es

Abstract

Forest fires may be caused by different factors, both climatological and anthropogenic. If fires are just considered as a natural hazard, then the number of fire seasons along a year will be just driven by climate conditions, but if human activity is not taken into consideration, this may lead to an underestimation of the number of fire seasons. In this setting, such periods can be identified by analyzing the pattern of fire ocurrences along a year, which can be viewed as a circular variable.

The statistical problem of determining the number of modes in the underlying density will provide a way of identifying significant fire seasons. This work presents a formal test for the number of modes in a circular density, based on some previous ideas from the linear space by Ameijeiras-Alonso et al. (2016). The proposal combines tools from critical bandwidth computation with an excess–mass statistic (see Silverman 1981 and Müller and Sawitzki 1991). The only alternative so far for solving this testing problem in the circular setting was the proposal by Fisher and Marron (2001), which did not provide an accurate calibration. Our method presents good calibration and power properties in simulated scenarios, and it has been extensively applied to analyze fire seasons on different regions of the Earth. For solving the issue related to applying our test systematically, in the different parts of the world, a spatially-adapted FDR correction method has been applied, based on the proposal by Benjamini and Heller (2007).

This is joint work with José Ameijeiras–Alonso and Alberto Rodríguez–Casal.

References:

1. Ameijeiras–Alonso, J., Crujeiras, R.M. and Rodríguez–Casal, A., Mode testing, critical bandwidth and excess–mass, *arXiv*, (2016).

2. Benjamini, Y., and Heller, R., False discovery rates for spatial signals, *Journal of the American Statistical Association* **102** (2007), 1272–1281.

3. Fisher, N.I. and Marron, J. S., Mode testing via the excess mass estimate, *Biometrika* **88** (2001), 419–517.

4. Müller, D. W. and Sawitzki, G., Excess mass estimates and tests for multimodality, *The Annals of Statistics* **13** (1991), 70–84.

5. Silverman, B. W., Using kernel density estimates to investigate multimodality, *Journal of the Royal Statistical Society Series B* **43** (1981), 97–99.

Nonparametric transformations for directional and shape data

Speaker: Marco Di Marzio
DMQTE, University of Chieti-Pescara, Italy
Email: marco.dimarzio@unich.it

Abstract

Regression involving data represented as points on the surface of a hypersphere has traditionally been treated using parametric families of transformations that include simple rigid rotation as an important, special case. On the other hand, nonparametric approaches have focused on modelling a scalar response trough a spherical predictor by representing the regression function as a polynomial. Therefore, they are able to address the case of estimating a spherical response component-wise. We propose very flexible (non-rigid) rotation models, in that, for each location of the manifold, a specific rotation matrix is to be estimated. To make this approach tractable, we assume continuity of the regression function that, in turn, allows for approximations of rotation matrices based on a series expansion. It is seen that higher-order approximations motivate an iterative learning procedure with connections to boosting. Extensions to general shape matching are also outlined. Both simulations and real data are used to illustrate the results.

This is joint work with Agnese Panzera and Charles Taylor.

Model selection for circular-circular models

Speaker: Miguel A. Fernández
Departamento de Estadística e Investigación Operativa, Universidad de Valladolid, Spain
Email: miguelaf@eiouva.es

Abstract

Circular-circular models, used for situations where a circular response is to be predicted using another circular variable, are frequent in the circular statistics literature as they appear in many practical situations. One of these is that of trying to relate the peak expressions of cell-cycle genes coming from two different species with possibly different cell phase lengths.

Several different models have been considered for these situations. Some of them, such as those in Downs and Mardia (2002), Rueda et al. (2009), Di Marzio et al. (2013) or Rueda et al. (2016), will be described in the talk. We will see that it is not easy to compare and select among them as some involve non-standard features. To overcome this problem, we propose a selection criteria based on the work by Ye (1998) for the Euclidean setting. We will check the performance of this criterion for the aforementioned cell-cycle data.

This is joint work with Cristina Rueda.

References:

1. M. Di Marzio, A. Panzera and C. Taylor, Non-parametric regression for circular responses, *Scandinavian Journal of Statistics* **40** (2013), 238–255.

2. T. Downs and K. V. Mardia, Circular regression, *Biometrika* **89** (2002), 683–697.

3. C. Rueda, M. A. Fernández and S. D. Peddada, Estimation of parameters subject to order restrictions on a circle with application to estimation of phase angles of cell-cycle genes, *Journal of the American Statistical Association* **104** (2009), 338–347.

4. C. Rueda, M. A. Fernández, S. Barragán, K. V. Mardia and S. D. Peddada, Circular Piecewise Regression with Applications to Cell-Cycle Data, *Biometrics* **72** (2016), 1266–1274.

5. J. Ye, On measuring and correcting the effects of data mining and model selection, *Journal of the American Statistical Association* **93** (1998), 120–131.

Toroidal diffusions and protein structure evolution

Speaker: Eduardo García-Portugués
Department of Statistics, Carlos III University of Madrid, Spain
Email: edgarcia@est-econ.uc3m.es

Abstract

We introduce stochastic models for continuous-time evolution of angles and develop their related estimation. We focus on modeling protein ergodic diffusions with stationary distributions equal to well-known distributions from directional statistics, since these can be regarded as toroidal analogues of the Ornstein-Uhlenbeck process. Their likelihood function is a product of transition densities without analytical expression, but that can be calculated by solving the Fokker-Planck equation numerically through adequate schemes. We propose three approximate likelihoods that are computationally less demanding, one of them being a specific approximation to the transition density of the wrapped normal process (Garcá-Portugués et al. 2017). A simulation study compares, in dimensions one and two, the approximate transition densities to the exact ones and investigates the empirical performance of the approximate likelihoods. Due to its proved tractability and accuracy, the wrapped normal process is employed as a building block for the development of a protein structure evolution model, Evolutionary Torus Dynamic Bayesian Network (ETDBN, Golden et al. 2017), the evolutionary extension of TorusDBN (Boomsma et al. 2008). This model couples sequence and structure evolution, allowing for both "smooth" conformational changes and "catastrophic" conformational jumps. The model is comparatively more realistic than previous stochastic models, providing new insights into the relationship between sequence and structure evolution.

This is joint work with Michael Sørensen, Kanti V. Mardia, Thomas Hamelryck, Michael Golden and Jotun Hein.

References:

1. W. Boomsma, K. V. Mardia, C. C. Taylor, J. Ferkinghoff-Borg, A. Krogh and T. Hamelryck, A generative, probabilistic model of local protein structure, *Proceedings of the National Academy of Sciences* **105** (2008), 8932–8937.

2. E. García-Portugués, M. Sørensen, K. V. Mardia, and T. Hamelryck, Langevin diffusions on the torus: estimation and applications, *arXiv*, (2017).

3. M. Golden, E. García-Portugués, M. Sørensen, K. V. Mardia, T. Hamelryck and J. Hein, A generative angular model of protein structure evolution, *arXiv*, (2017).

Analyzing infant head flatness and asymmetry using kernel density estimation of directional surface data from a craniofacial 3D model

Speaker: Lasse Holmström
Research Unit of Mathematical Sciences, University of Oulu, Finland
Email: lasse.holmstrom@oulu.fi

Abstract

Human skull deformation is analyzed using the distribution of head normal vector directions computed from a 3D image (Vuollo et al. 2016). Round head shape corresponds to a uniform distribution of normal vector directions and for a symmetrical head, the distribution exhibits symmetry across the plane that divides the head into left and right halves. Severity of head flatness and asymmetry can then be quantified by suitable functionals of the kernel estimate of the normal vector direction density. The density estimates can be visualized using a 2D contour plot of an area-preserving projection of the estimate to a planar disk. The numerical integration needed in evaluating the density functionals can be performed efficiently by employing an evenly spaced Fibonacci grid on the unit sphere. Using image data from 99 infants and clinical deformation ratings made by experts, our approach is compared with some recently suggested methods. The results show that the proposed method performs competitively.

This is joint work with Ville Vuollo, Henri Aarnivala, Virpi Harila, Tuomo Heikkinen, Pertti Pirttiniemi and Arja Marita Vakama.

References:

1. V. Vuollo, L. Holmström, H. Aarnivala, V. Harila, T. Heikkinen, P. Pirttiniemi, and A. M. Valkama, Analyzing infant head flatness and asymmetry using kernel density estimation of directional surface data from a craniofacial 3D model, *Statistics in Medicine* **35** (2016), 4891–4904. Available on-line at `http://dx.doi.org/10.1002/sim.7032`.

Principal component analysis on the torus

Speaker: Stephan Huckemann
Institut für Mathematische Stochastik, University of Göttingen, Germany
Email: huckeman@math.uni-goettingen.de

Abstract

Principal Component Analysis (PCA) on the Torus is challenging. In the canonical flat torus geometry, there are dense geodesics and hence every data set can be arbitrarily well approximated by geodesics, i.e. by the first PC. In order to obtain a useful set of geodesics, one may either restrict geodesics from winding around, thereby reducing the dimension of the descriptor space, or do tangent space PCA, which, in effect, imposes a new topology on the torus, that destroys cyclic data structure. We propose an intermediate approach, namely imposing only a new geometry on the torus, but not changing the topology. Specifically the new geometry is that of a stratified sphere and in reward the rich toolbox of principal nested spheres (PNS) analysis can be adapted. We measure effectively the statistical performance of our method by dimension reduction and faithfulness in terms of preserving previously known RNA structure. With this modified PNS, also the recently developed asymptotics for backward nested descriptor families become available allowing for inference.

This is joint work with Benjamin Eltzner and Kanti Mardia.

Invariance properties and statistical inference for circular data

Speaker: Giovanna Jona-Lasinio
Department of Statistical Sciences, "Sapienza" University of Rome, Italy
Email: giovanna.jonalasinio@uniroma1.it

Abstract

Observations on the circle should be treated and analyzed with caution as statistical inference may strongly depend on the chosen reference system. We define two invariance properties as well as necessary and sufficient conditions to avoid inferential problems and misinterpretation of parameter estimates for any circular distribution. We further investigate the effects of all possible changes in the reference system on existing and widely used continuous and discrete circular distribution. The construction of invariant distributions that should be used in practice to avoid misleading results is discussed. Numerical examples on artificial and real data are presented to corroborate and illustrate the theoretical results.

This is joint work with Gianluca Mastrantonio and Antonello Maruotti.

Probability integral transforms in directional statistics

Speaker: Peter Jupp
School of Mathematics and Statistics, University of St Andrews, UK
Email: pej@st-andrews.ac.uk

Abstract

The standard method of transforming a continuous distribution on the line to the uniform distribution on $[0, 1]$ is the probability integral transform. Such a transform can be defined also for distributions on the circle, although it is not unique. A form of probability integral transform is introduced for distributions with continuous positive density on compact Riemannian manifolds (such as spheres, projective spaces or rotation groups), $\mathcal{X}$. It is a continuous mapping of $\mathcal{X}$ to itself that transforms the distribution into the uniform distribution. It is based on the usual probability integral transform along each geodesic through a given point. Although the mapping is not unique, there are 'almost canonical' choices. Applications include (i) decomposition of distributions on product spaces into marginal distributions and copulae, (ii) derivation of tests of goodness of fit from tests of uniformity. The construction can be extended to some other sample spaces, including simplices and simply-connected spaces of non-positive curvature.

This is joint work with Alfred Kume.

Möbius transformation and a multivariate Cauchy family on the sphere

Speaker: Shogo Kato
Institute of Statistical Mathematics, Japan
Email: skato@ism.ac.jp

Abstract

We discuss some properties of a spherical extension of the wrapped Cauchy family on the circle (Kent and Tyler 1988, McCullagh 1996). The family on the sphere is briefly considered in probability theory (Dunau and Sénateur 1988), but its statistical properties have not been considered before. It is seen that the family is closed under the Möbius transformation on the sphere and that there is a similar induced transformation on the parameter space. Some properties of a marginal distribution of the spherical family such as certain moments and a closure property associated with the real Möbius group are obtained. It is shown that the spherical family is connected to a multivariate Cauchy family on the Euclidean space via stereographic projection. Maximum likelihood estimation for the spherical family is considered; closed-form expressions for the maximum likelihood estimators are available when the sample size is not greater than three, and the unimodality holds for the maximized likelihood functions.

This is joint work with Peter McCullagh.

References:

1. Dunau, J.-L. and Sénateur, H. Une caracterisation du type de la loi de Cauchy-conforme sur $\mathbb{R}^n$, *Probability Theory and Related Fields* **77** (1988), 129–135.

2. Kent, J. T. and Tyler, D. E. Maximum likelihood estimation for the wrapped Cauchy distribution, *Journal of Applied Statistics* **15** (1988), 247–254.

3. McCullagh, P. Möbius transformation and Cauchy parameter estimation, *The Annals of Statistics* **24** (1996), 787–808.

Tracking space debris: the need for directional statistics

Speaker: John Kent
Department of Statistics, University of Leeds, UK
Email: j.t.kent@leeds.ac.uk

Abstract

There are many thousands of pieces of debris in orbit around the earth, and these objects pose a real threat to satellite operations. Tracking the debris is a challenging problem, even under the simplifying assumption of Keplerian dynamics (i.e. each object follows an exact elliptical orbit about the earth). Techniques from directional statistics have important insights to contribute to the tracking problem, both for accuracy and speed. In particular, a new version of the Fisher Bingham distribution on the sphere has been developed (Kent et al. 2016) to describe data closely concentrated near an arc of a great circle. This distribution can be used to model the propagated angular position of a space object, given some initial uncertainty in its location and velocity.

References:

1. J. T. Kent, I. Hussein, and M. K. Jah. Directional Distributions in Tracking of Space Debris, *in Proceedings of the 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany*, pp. 2081–2086, IEEE (2016).

2. J. T. Kent, S. Bhattacharjee, I. Hussein, and M. K. Jah, Orbital error propagation analysis using directional statistics for space objects, *in Proceedings of the 27th AAS/AIAA Space Flight Mechanics Meeting, San Antonio, Texas, February 5-9, 2017*, Paper AAS 17-390 (2017).

3. J. T. Kent, S. Bhattacharjee, I. Hussein, and M. K. Jah, Angles-only data association using directional discriminant analysis, *in Proceedings of the 27th AAS/AIAA Space Flight Mechanics Meeting, San Antonio, Texas, February 5-9, 2017*, Paper AAS 17-393 (2017).

Bayesian analysis of circular data in social and behavioural sciences

Speaker: Irene Klugkist
Department of Social and Behavioural Sciences, Section Methodology and Statistics,
Utrecht University, The Netherlands
Email: i.klugkist@uu.nl

Abstract

Although circular data may arise more often in fields like biology and environmental sciences, also in the social and behavioural sciences researchers are occasionally confronted with circular measurements. In this presentation we will discuss an example from a large sociological survey that contains data from a so-called circumplex scale. Data from a circumplex, like Leary's scale or the interpersonal cirumplex, are often seen in fields like psychology and educational sciences. However, the analyses of resulting data are not always taking the circular structure of the data into account. With our work we aim to make it easier and more accessible for social scientists to perform circular analyses on such data.

For a multiple regression model with linear predictors and a circular outcome variable we will present results of both the intrinsic and the embedding approach within the Bayesian framework. Recent developments in the Bayesian intrinsic and embedding approaches will be summarized and illustrated in the context of a subset of data from the European Social Survey. We will compare the two approaches with respect to ease of computation, flexibility, and interpretation of results. Advantages of using the Bayesian approach will be illustrated by giving examples of estimating parameters, functions of parameters, and measures for model fit.

This is joint work with Jolien Cremers and Kees Mulder.

Dynamic mixtures of copulas for cylindrical time series

Speaker: Francesco Lagona
Department of Political Sciences, University of Roma Tre, Italy
Email: francesco.lagona@uniroma3.it

Abstract

Cylindrical time series are bivariate series of angles and intensities that often arise in environmental research (Bulla et al. 2012). The statistical analysis of these series has been overlooked due to the special topology of the support on which the measurements are taken (the cylinder), and to the difficulties in modeling the cross-correlations between angular and linear measurements over time. Further complications arise from the skewness and the multimodality of the marginal distributions of the data.

By extending previous works on the analysis of cylindrical time series data (Abe and Ley 2017, Lagona et al. 2015), a novel mixture model is introduced to accommodate these features. The model is based on a mixture of copula-based cylindrical distributions, whose parameters evolve according to a latent Markov chain. The cylindrical distributions are specified by integrating a von Mises distribution and a Weibull distribution by means of a circular copula. A computationally efficient Expectation-Maximization algorithm is described to estimate the parameters. A parametric bootstrap routine is introduced to compute confidence intervals. These methods are illustrated to segment cylindrical time series of wave heights and directions in the Adriatic sea.

References:

1. T. Abe and C. Ley, A tractable, parsimonious and flexible model for cylindrical data, with applications, *Econometrics and Statistics*, to appear.

2. J. Bulla, F. Lagona, A. Maruotti and M. Picone, A multivariate hidden Markov model for the identification of sea regimes from incomplete skewed and circular time series, *Journal of Agricultural, Biological and Environmental Statistics* **17** (2012), 544–567.

3. F. Lagona, M. Picone and A. Maruotti, A hidden Markov model for the analysis of cylindrical time series, *Environmetrics* **26** (2015), 534–544.

Skew-rotationally-symmetric distributions and tests for rotational symmetry

Speaker: Christophe Ley
Department of Applied Mathematics, Computer Science and Statistics, Universiteit Gent, Belgium
Email: christophe.ley@ugent.be

Abstract

Most commonly used distributions on the unit hypersphere $\mathcal{S}^{p-1}$, $p \geq 2$, assume that the data are rotationally symmetric about some direction $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$. However, there is empirical evidence that this assumption often fails to describe reality. We present in this talk a new class of skew-rotationally-symmetric distributions on $\mathcal{S}^{p-1}$ that enjoy numerous good properties. The new distributions are called "skew-rotationally-symmetric" (Ley and Verdebout 2017). We discuss the Fisher information structure of the model and derive efficient inferential procedures. In particular, we obtain the first semi-parametric test for rotational symmetry about a known direction. We also propose a second test for rotational symmetry, obtained through the definition of a new measure of skewness on the hypersphere. We investigate the finite-sample behavior of the new tests through a Monte Carlo simulation study.

This is joint work with Thomas Verdebout.

References:

1. Ley, C. and Verdebout, T., Skew-rotationally-symmetric distributions on unit spheres and related efficient inferential procedures, *Journal of Multivariate Analysis*, to appear.

Magic of score matching estimators for directional distributions

Speaker: Kanti V. Mardia
Department of Statistics, School of Mathematics, University of Leeds, UK; Also Visiting Professor, University of Oxford
Email: k.v.mardia@leeds.ac.uk

Abstract

One of the major problems for maximum likelihood estimation in the well-established directional models is that the normalising constants can be difficult to evaluate. A general method of "score matching estimation" will be presented on a compact oriented Riemannian manifold. The method for the Euclidean case has been studied by Forbes and Lauritzen (2014), and Hyvarinen (2005, 2007a,b). Important applications of our work include von Mises-Fisher, Bingham and joint models on the sphere, the multivariate von Mises model on torus and related spaces. The estimator will be shown to be consistent and asymptotically normally distributed under mild regularity conditions. Further, it is easy to compute as a solution of a linear set of equations and requires no knowledge of the normalizing constant. Several cases will be treated, both analytic and numerical, to demonstrate its good performance. Practical examples will be given from various scientific fields including molecular biology. The most common and general approximate method for directional distributions is using the saddle-point approach, where the estimates are comparatively easier to compute than the maximum likelihood estimators but still involves non-linear optimisation.

This is joint work with Arnab Laha and John Kent (Mardia et al. 2016).

References:

1. Forbes, P. G. M. and Lauritzen, S., Linear estimating equations for exponential families with applications to Gaussian linear concentrations, *Linear Algebra and its Applications* (2014), to appear.

2. Hyvarinen, A., Estimation of non-normalized statistical models by score matching, *Journal of Machine Learning Research* **6** (2005), 695–709.

3. Hyvarinen, A., Some extensions of score matching, *Computational Statistics and Data Analysis* **51** (2007a), 2499–2512.

4. Hyvarinen, A., Connections between score matching, contrastive divergence, and pseudo-likelihood for continuous valued variables, *IEEE Transactions on Neural Networks* **18** (2007b), 1529–1531.

5. Mardia, K.V., Kent, J.T., and Laha, A.K., Score matching estimators for directional distributions, *arXiv*, (2016).

Multiple testing of local maxima for detection of peaks on the (celestial) sphere

Speaker: Domenico Marinucci
Department of Mathematics, University of Rome Tor Vergata, Italy
Email: marinucc@mat.uniroma2.it

Abstract

We present a topological multiple testing scheme for detecting peaks on the sphere under isotropic Gaussian noise, where tests are performed at local maxima of the observed field filtered by the spherical needlet transform. Our setting is different from the standard Euclidean/large same asymptotic framework, yet highly relevant to realistic experimental circumstances for some important areas of application in astronomy. More precisely, we shall focus on cases where a single realization of a smooth isotropic Gaussian random field on the sphere is observed, and a number of well-localized signals are superimposed on such background field. The proposed algorithms, combined with the Benjamini-Hochberg procedure for thresholding p-values, provide asymptotic strong control of the False Discovery Rate (FDR) and power consistency as the signal strength and the frequency of the needlet transform get large. This novel multiple testing method can be applied, for instance, to point-source detection in Cosmic Microwave Background radiation (CMB) data.

This is joint work with Dan Cheng, Valentina Cammarota, Yabebal Fantaye and Armin Schwartzman.

References:

1. Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D., Asymptotics for Spherical Needlets, *Annals of Statistics* **37** (2009), 1150–1171.

2. Cammarota, V., Marinucci, D., Wigman, I., On the distribution of the critical values of random spherical harmonics, *Journal of Geometric Analysis* **4** (2016), 3252–3324.

3. Cheng,D., Cammarota, V., Fantaye, Y., Marinucci, D., Schwartzman, A. Multiple testing of local maxima for detection of peaks on the (celestial) sphere, *arxiv*, (2016).

4. Cheng, D., Schwartzman, A., Distribution of the height of local maxima of Gaussian random fields, *Extremes* **18** (2015), 213–240.

5. Marinucci, D., and Peccati, G. (2011), *Random Fields on the Sphere. Representation, Limit Theorem and Cosmological Applications*, Cambridge University Press, Cambridge (2008).

6. Planck Collaboration Planck 2013 results. I. Overview of products and scientific results, *Astronomy and Astrophysics* **571** (2014), idA1.

7. Schwartzman, A., Gavrilov, Y., and Adler, R. J. , Multiple testing of local maxima for detection of peaks in 1D, *Annals of Statistics* **39** (2011), 3290–3319.

Robust model-based clustering for longitudinal circular data

Speaker: Antonello Maruotti
Dipartimento di Scienze Economiche, Politiche e delle Lingue Moderne, Libera Università Maria Ss. Assunta, Italy
Email: a.maruotti@lumsa.it

Abstract

In this work, we propose two general hidden Markov model-based approaches, namely the noise and the trimming approaches, for clustering longitudinal circular data contaminated by atypical observations that may affect parameter estimates and the recovering of the clustering structure. Hidden Markov models are the state of the art in the analysis of (clustered) longitudinal data (Zucchini et al. 2016), and have been increasingly used in circular data analysis (Maruotti et al. 2016). With respect to the noise approach, we consider hidden Markov fittings with the addition of a further component to capture observations that are not consistent with the projected normal hidden Markov model. This is done by adding a uniform conditional density on the circle or representing the noise component by a fixed constant. The trimming approach introduces a robust estimation procedure based on the choice of a representative trimmed subsample. While the noise approach tries to fit the atypical observations in the model, the trimming approach attempts to discard them completely.

References:

1. A. Maruotti, A. Punzo, G. Mastrantonio and F. Lagona, A time-dependent extension of the projected normal regression model for longitudinal circular data based on a hidden Markov heterogeneity structure, *Stochastic Environmental Research and Risk Assessment* **30** (2016), 1725–1740.

2. W. Zucchini, I. L. MacDonald and R. Langrock, *Hidden Markov Models for Time Series: An Introduction Using R*, 2nd ed., Chapman & Hall (2016).

The joint projected and skew normal; a distribution for hyper-cylindrical data

Speaker: Gianluca Mastrantonio
Department of Mathematics, Polytechnic of Turin, Italy
Email: gianluca.mastrantonio@polito.it

Abstract

We introduce a multivariate circular-linear distribution obtained by combining the projected and the skew normal. We show the flexibility of our proposal, its property of closure under marginalization and how to quantify multivariate dependence. Due to a non-identifiability issue that our proposal inherits from the projected normal, a computational problem arises. We overcome it in a Bayesian framework, adding suitable latent variables and showing that posterior samples can be obtained with a post-processing of the estimation algorithm output. Under specific prior choices, this approach enables us to implement a Markov chain Monte Carlo algorithm relying only on Gibbs steps, where the updating of the parameters is done as if we were working with a multivariate normal likelihood. The proposed approach can be also used with the projected normal. On simulated examples we show the ability of our algorithm in recovering the parameter values and to solve the identification problem. Then the proposal is used in a real data example, where the turning-angles (circular variables) and the log step-lengths (linear variables) are jointly modelled.

A general random walk model on the plane using von Mises distribution

Speaker: Aurélien Nicosia
Université Laval, Canada
Email: aurelien.nicosia.1@ulaval.ca

Abstract

In this talk, we propose a general directional random walk model to describe the movement of an animal that takes into account features of the environment. A consensus angular regression model with von Mises errors models the direction. This regression model lets the dependent circular angle depend on multiple explanatory angles and their interactions with real explanatory variables. A hidden process structure enables modeling situations where the animal exhibits various movement behaviors. The model is fitted using the EM algorithm. We illustrate its use by modeling the movement of an animal in Canada's boreal forest.

This is joint work with Louis-Paul Rivest and Thierry Duchesne.

Bayes inference with equivalence classes of 3-D orientations

Speaker: Dan Nordman
Department of Statistics & Department of Industrial and Manufacturing Systems Engineering, Iowa State University, USA
Email: dnordman@iastate.edu

Abstract

Experiments in materials science often collect data which are equivalence classes of crystallographically symmetric orientations. These intend to represent how lattice structures of particles are orientated relative to a reference coordinate system. Motivated by a materials science application, this talk describes inference about parametric probability models for equivalence classes of 3-D rotations (i.e., a type of unlabeled orientation data). Probability distributions on such equivalence classes are induced by a flexible, popular family for symmetric random rotations. We develop one-sample Bayes inference for the parameters in these models, and compare this methodology to some likelihood-based approaches. We also contrast the parametric analysis of unlabeled orientation data with other analyses that proceed as if the data have been pre-processed into actual orientation data.

This is joint work with Steve Vardeman.

On the asymptotic non-null behavior of high-dimensional spherical location tests

Speaker: Davy Paindaveine
Département de Mathématique and ECARES, Université libre de Bruxelles, Belgium
Email: dpaindav@ulb.ac.be

Abstract

In the Fisher–von Mises–Langevin model, we consider the high-dimensional version of the spherical location testing problem, that is, we want to test the null hypothesis that the modal location θ coincides with a given value θ_0 on the p-dimensional unit sphere, with p large. It is well-known that, in standard situations where the dimension p and the underlying concentration κ are fixed, the Watson test is asymptotically optimal. In Paindaveine and Verdebout (2017), we investigated whether or not this extends to the case where the concentration $\kappa = \kappa_n$ goes to zero as n goes to infinity. In the present work, we consider, as in Cutting et al. (2017), the high-dimensional setup where the dimension p_n goes to infinity with n. We allow the concentration κ_n to behave in a completely free way with n, which offers a complete spectrum of problems ranging from arbitrarily challenging to arbitrarily easy problems. We identify five different regimes, depending on the convergence/divergence properties of κ_n, that yield different limiting asymptotic experiments. Asymptotically optimal tests are obtained in each regime.

This is joint work with Yvik Swan and Thomas Verdebout.

References:

1. C. Cutting, D. Paindaveine, and T. Verdebout, Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives, *Annals of Statistics* (2017), to appear.

2. D. Paindaveine and T. Verdebout, Inference on the mode of weak directional signals: a Le Cam perspective on hypothesis testing near singularities, *Annals of Statistics* (2017), to appear.

Nonparametric estimating equations for circular densities and their derivatives

Speaker: Agnese Panzera
DiSIA, University of Florence, Italy
Email: a.panzera@disia.unifi.it

Abstract

We propose estimating equations whose unknown parameters are the values taken by a circular density and its derivatives at a point. Specifically, we solve equations which relate local versions of population trigonometric moments with their sample counterparts. Major advantages of our approach are: higher order bias without asymptotic variance inflation, and closed form for the estimators. We provide theoretical results along with simulation experiments.

This is joint work with Marco Di Marzio, Stefania Fensore and Charles Taylor.

Circulas: the circular analogues of copulas

Speaker: Arthur Pewsey
Department of Mathematics, Escuela Politécnica, University of Extremadura, Spain
Email: apewsey@unex.es

Abstract

My talk will begin with a brief introduction to the circular analogues of copulas, coined *circulas* by Jones et al. (2015), before concentrating specifically on a particular bivariate class usually ascribed to Wehrly and Johnson (1980). The latter is appealing for various reasons but generally does not lead to bivariate circular models with attractive shapes when the marginals are not circular uniform. A notable exception is the bivariate wrapped Cauchy distribution proposed and developed in Kato and Pewsey (2015). Simulation methods and the parametric bootstrap goodness-of-fit testing approach of Pewsey and Kato (2016) will be outlined and applied in the modelling of wind directions recorded at a Texan weather station.

This is joint work with Shogo Kato and Chris Jones.

References:

1. M. C. Jones, A. Pewsey and S. Kato, On a class of circulas: copulas for circular distributions, *Annals of the Institute of Statistical Mathematics* **67** (2015), 843–862.

2. S. Kato and A. Pewsey, A Möbius transformation-induced distribution on the torus, *Biometrika* **102** (2015), 359–370.

3. A. Pewsey and S. Kato, Parametric bootstrap goodness-of-fit testing for Wehrly–Johnson bivariate circular distributions, *Statistics and Computing* **26** (2016), 1307–1317.

4. T. E. Wehrly and R. A. Johnson, Bivariate models for dependence of angular observations and a related Markov process, *Biometrika* **67** (1980), 255–256.

Regression models for angular responses

Speaker: Louis-Paul Rivest
Départment de Mathématiques et de Statistique, Université Laval, Canada
Email: Louis-Paul.Rivest@mat.ulaval.ca

Abstract

This presentation investigates hierarchical data structures where units, indexed by j, are nested within groups, indexed by i. The goal is to derive a model for dependent angle y_{ij} in terms of vector valued explanatory variables (x_{ij}, z_{ij}) where x is a vector of explanatory angles and z is a vector of continuous variables possibly impacting the relationship between y and x. We also want to model the dependency between the units in the same group. First, following Rivest et al. (2016), a parametric expression for the conditional mean direction $\mu(x, z)$ of y given (x, z) is proposed. The explanatory angles x are seen as the directions of targets that might attract y and $\mu(x, z)$ is the direction of a weighted sum of the unit vectors corresponding to x where the weights possibly involve z. Then the dependence is modeled using group specific random effects having a von Mises distribution, leading to an angular generalization of Battese et al. (1988). A closed form expression for the density of the angles in a group is derived and several measures of the correlation within a group are presented. Several estimators of the parameters are investigated and an example dealing with repeated measures of a sand hopper orientation, discussed in Lagona (2016) and Maruotti (2016), is presented.

This is joint work with Shogo Kato.

References:

1. G. E. Battese, R. M. Harter, and W. A. Fuller, An error-components models for prediction of county crop areas using survey and satellite data, *Journal of the American Statistical Association* **83** (1988), 28–36.

2. F. Lagona, Regression analysis of correlated circular data based on the multivariate von Mises distribution, *Environmental and Ecological Statistics* **23** (2016), 89–116.

3. A. Maruotti, Analyzing longitudinal circular data by projected normal models: a semi-parametric approach based on finite mixture models, *Environmental and Ecological Statistics* **23** (2016), 257–277.

4. L.-P. Rivest, T. Duchesne, A. Nicosia,and D. Fortin, A general angular regression model for the analysis of data on animal movement in ecology, *Journal of the Royal Statistical Society Series C* **66** (2016), 445–463.

Estimation and inference in directional mixed models for compositional data

Speaker: Janice Scealy
Research School of Finance, Actuarial Studies and Statistics, Australian National University, Australia
Email: janice.scealy@anu.edu.au

Abstract

Compositional data are vectors of proportions defined on the unit simplex and this type of constrained data occur frequently in applications. It is also possible for the compositional data to be correlated due to the clustering or grouping of the observations. We propose a new class of mixed model for compositional data based on the Kent distribution for directional data, where the random effects also have Kent distributions. The advantage of this approach is that it handles zero components directly and the new model has a fully flexible underlying covariance structure. One useful property of the new directional mixed model is that the marginal mean direction has a closed form and is interpretable. The random effects enter the model in a multiplicative way via the product of a set of rotation matrices and the conditional mean direction is a random rotation of the marginal mean direction. For estimation we apply a quasi-likelihood method which results in solving a new set of generalised estimating equations and these are shown to have low bias in typical situations. For inference we use a nonparametric bootstrap method for clustered data which does not rely on estimates of the shape parameters (shape parameters are difficult to estimate in Kent models). The new approach is shown to be more tractable than the traditional approach based on the log-ratio transformation.

This is joint work with Alan Welsh.

Probability distributions for cylindrical data

Speaker: Kunio Shimizu
School of Statistical Thinking, The Institute of Statistical Mathematics, Japan
Email: k-shmz@ism.ac.jp

Abstract

A probability distribution for cylindrical data was proposed in Johnson and Wehrly (1978), whose circular marginal distribution of Θ is a wrapped Cauchy distribution, but the linear marginal distribution of X does not follow a familiar distribution. The conditional distribution of Θ given $X = x$ is a von Mises distribution, and the conditional distribution of X given $\Theta = \theta$ is an exponential distribution. Recently the WeiSSVM, Weibull and sine-skewed von Mises, model as an extension was given in Abe and Ley (2017) and it was used in Lagona et al. (2015) to propose a hidden Markov model with an illustration on cylindrical time series of wave heights and directions. In this study we provide a Pareto-type probability distribution on the cylinder by using a gamma mixture for the WeiSSVM distribution with zero skewing parameter. The marginal distribution of Θ is a wrapped Cauchy distribution and the marginal distribution of X can be interpreted as a mixture distribution. The conditional distribution of Θ given $X = x$ belongs to the t-distribution on the circle (cf. Shimizu and Iida 2002) and the family of Jones and Pewsey (2005). The conditional distribution of X given $\Theta = \theta$ is a transformed generalized Pareto distribution. Illustrative examples are given for periwinkle data and for magnitude and epicenter data for foreshocks during 72 hours before the 2011 Great East Japan Earthquake.

This is joint work with Tomoaki Imoto and Toshihiro Abe.

References:

1. T. Abe and C. Ley, A tractable, parsimonious and flexible model for cylindrical data, with applications, *Econometrics and Statistics* (2017), to appear.

2. R. A. Johnson and T. E. Wehrly, Some angular-linear distributions and related regression models, *Journal of the American Statistical Association* **73** (1978), 602–606.

3. M. C. Jones and A. Pewsey, A family of symmetric distributions on the circle, *Journal of the American Statistical Association* **100** (2005), 1422–1428.

4. F. Lagona, M. Picone, and A. Maruotti, A hidden Markov model for the analysis of cylindrical time series, *Environmetrics* **26** (2015), 534–544.

5. K. Shimizu and K. Iida, Pearson type VII distributions on spheres, *Communications in Statistics-Theory and Methods* **31** (2002), 513–526.

On some tests for rotational symmetry

Speaker: Thomas Verdebout
ECARES and Département de Mathématique, Université Libre de Bruxelles (ULB), Belgium
Email: tverdebo@ulb.ac.be

Abstract

Most commonly-used distributions on unit spheres $\mathcal{S}^{p-1} := \{\mathbf{v} \in \mathbb{R}^p : \mathbf{v}'\mathbf{v} = 1\}$, $p \geq 2$, assume that the data are rotationally symmetric about some direction $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$. The problem of testing for rotational symmetry has been relatively well studied in the circular ($p = 2$) case; see e.g. Pewsey (2002) and Ley and Verdebout (2014) and the references therein. Surprisingly, it has been less investigated in the $p \geq 3$ case; Ley and Verdebout (2017) very recently provided some test in the $\boldsymbol{\theta}$-specified case. In this paper, we propose new tests of rotational symmetry both in the $\boldsymbol{\theta}$-specified and $\boldsymbol{\theta}$-unspecified cases. We show that our new tests enjoy many desirable properties. In particular they are shown to be locally and asymptotically optimal against new classes of distributions on $\mathcal{S}^{p-1}$. Our results are illustrated via Monte Carlo simulations and on a real data set related to the study of protein structures.

This is joint work with Eduardo García-Portugués and Davy Paindaveine.

References:

1. Ley, C. and Verdebout, T. Simple optimal tests for circular reflective symmetry about a specified median direction, *Statistica Sinica* **14** (2014), 1319–1340.

2. Ley, C. and Verdebout, T., Skew-rotationally symmetric distributions on unit spheres and related efficient inferential procedures, *Journal of Multivariate Analysis*, to appear.

3. Pewsey, A. Testing circular symmetry, *Canadian Journal of Statistics* **30** (2002), 591–600.

Abstracts: Poster Presenters

Testing circular reflective symmetry on the cracks in cemented femoral components

Speaker: José Ameijeiras-Alonso
Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Spain
Email: jose.ameijeiras@usc.es

Abstract

Knowing whether the symmetry, or more precisely reflective symmetry, is present in a circular sample is a matter of interest when one wants to know if the data is equally distributed at the left and the right of some value (the centre of symmetry). Testing circular symmetry is not only important per se: it is one of the most frequently encountered simplifying hypotheses, whose rejection leads to the exploration of models more complex than the classic symmetric ones (like the von Mises, wrapped normal, wrapped Cauchy or cardioid). On the circle, one of the best known asymmetric models is the k–sine–skewed family of distributions (see, e.g., Abe and Pewsey 2011) thanks to its flexibility and simplicity.

The objective of this work is to present a new proposal for testing reflective symmetry when the central direction is unknown. This new proposal is based on the Le Cam theory of asymptotic experiments. The distribution of the test statistic will be obtained under the null hypothesis of reflective symmetry and under local k–sine–skewed alternatives. Also, it will be seen that this test is optimal (in the maximin sense) against these alternatives. This work extends previous results from Ley and Verdebout (2014), where the central direction was supposed to be known. The obtained test will be applied to investigate if the cracks produced in a femur after a hip replacement (see Mann et al. 2003) are symmetrical around some unknown centre in two regions of the bone (close and far from the hip).

This is joint work with Christophe Ley, Arthur Pewsey and Thomas Verdebout.

References:

1. T. Abe, and A. Pewsey. Sine–skewed circular distributions, *Statistical Papers* **52** (2011), 683–707.

2. C. Ley and T. Verdebout. Simple optimal tests for circular reflective symmetry about a specified median direction, *Statistica Sinica* **24** (2014), 1319–1339.

3. K. A. Mann, S. Gupta, A. Race, M. A. Miller and R. J. Cleary. Application of circular statistics in the study of crack distribution around cemented femoral components, *Journal of Biomechanics* **36** (2003), 1231–1234.

Graphical tools for exploratory directional data analysis

Speaker: Davide Buttarazzi
Department of Economics and Law, University of Cassino and Southern Lazio, Italy
Email: d.buttarazzi@unicas.it

Abstract

The box-and-whisker plot (Tukey 1977) and the density strip plot (Jackson 2008) are effective tools to explore and compare distributions on the line. Their counterpart within the directional setting is introduced.

On the one hand, a rotationally and zero-direction invariant circular boxplot turns out to be particularly appealing for visualizing unimodal circular distributions. By means of a simulation study, a procedure for non-linear rescaling of the whiskers is obtained. To facilitate comparison of distributions, a proposal of visual display of 'parallel' circular boxplots is also provided.

On the other hand, the circular density strip plot well serves the exploration and comparison of both unimodal and multimodal distributions. Static and dynamic visualizations are proposed. In the latter case, groups comparison is enhanced by embedding circular density strip plots on the surface of an interactive and dynamic cylinder. In addition, the use of circular density strip plots for assessing goodness-of-fit is shown.

All graphical proposals are illustrated through some real and simulated data examples. The routines for drawing circular boxplots and circular density strip plots are developed in the `R` software.

This is joint work with Giovanni Porzio.

References:

1. C. H. Jackson, Displaying uncertainty with shading, *The American Statistician* **62** (2008), 340–347.

2. J. W. Tukey, *Exploratory Data Analysis*. MA: Addison-Wesley, Reading (UK) (1977).

Estimation of the integrated squared density of a circular random variable

Speaker: José E. Chacón
Departamento de Matemáticas, Universidad de Extremadura, Spain
Email: jechacon@unex.es

Abstract

Given the density f of a circular random variable Θ, the goal of this communication is to study the estimation of $\psi_0 = \int_0^{2\pi} f(\theta)^2 d\theta$. This problem has attracted much interest for linear data, and applications and motivation for its study can be found in the introductory sections of Aldershof (1991), Sheather et al. (1994) or, more recently, Giné and Nickl (2008). For circular variables, however, the problem has received far less attention. Nevertheless, a detailed description of the asymptotic performance of a kernel-type estimator can be extracted as a particular case of a more general result, which is valid for arbitrary data dimensionality and higher order derivatives, contained in Di Marzio et al. (2011). But this result shows that the optimal bandwidth for estimating $\psi_r = \int_0^{2\pi} f^{(r)}(\theta) f(\theta) d\theta$ depends on ψ_{r+2}, so that the problem becomes somehow cyclic. In the linear case, the usual solution consists of using a multi-stage estimation procedure, starting with an initial approximation based on some reference distribution. Unfortunately, for circular data, the common choice of the von Mises distribution as the reference distribution is known to entail some difficulties, for example in the presence of antipodal modes, see Oliveira et al. (2012). Here we present a new estimator for θ_0 which, despite being based on kernel smoothing, does not require a bandwidth choice or reference distributions, and even so it is asymptotically efficient.

References:

1. B. K. Aldershof, *Estimation Of Integrated Squared Density Derivatives*. Ph.D. thesis, University of North Carolina at Chapel Hill (1991).

2. M. Di Marzio, A. Panzera and C. C. Taylor, Kernel density estimation on the torus, *Journal of Statistical Planning and Inference* **141** (2011), 2156–2173.

3. E. Giné and R. Nickl, A simple adaptive estimator of the integrated square of a density, *Bernoulli* **14** (2008), 47–61.

4. M. Oliveira, R. M. Crujeiras and A. Rodríguez-Casal, A plug-in rule for bandwidth selection in circular density estimation, *Computational Statistics and Data Analysis* **56** (2012), 3898–3908.

5. S. J. Sheather, T. P. Hettmansperger and M. R. Donald, Data-based bandwidth selection for kernel estimators of the integral of $f^2(x)$, *Scandinavian Journal of Statistics* **21** (1994), 265–275.

Circular modelling of circumplex measurements for teacher behavior

Speaker: Jolien Cremers
Department of Methodology and Statistics, Utrecht University, The Netherlands
Email: j.cremers@uu.nl

Abstract

One of the disciplines in which circular data may be encountered is the educational sciences, where they measure teachers positions on the interpersonal circumplex. However, in the literature this type of data is not yet analyzed using circular statistics even though theoretically it is assumed that the data are circular. This results in a discrepancy between the questions that researchers are interested in and the questions that can be answered with the (linear) methods that are currently employed by the researchers who have longitudinal circular data.

We fit a Bayesian longitudinal model for circular data to a dataset collected by Mainhard et al. (2011) in which secondary school teachers position on the interpersonal circumplex was measured during the first 16 weeks of the schoolyear. This model was originally developed by Nuñez-Antonio and Gutiérrez-Peña (2014) and is a mixed-effects model based on a projected normal distribution. We present new measures for interpreting fixed effects of predictor variables on a circular scale and investigate how to interpret circular random effects in this model. Additionally we can employ the Bayesian framework for developing measures for model fit and hypothesis tests.

This is joint work with Irene Klugkist.

References:

1. M.T. Mainhard, M. Brekelmans, P. den Brok and T. Wubbels, The development of the classroom social climate during the first months of the school year, *Contemporary Educational Psychology* **36** (2011), 190–200.

2. G. Nuñez-Antonio and E. Gutiérrez-Peña, A Bayesian model for longitudinal circular data based on the projected normal distribution, *Computational Statistics and Data Analysis* **71** (2014), 506–519.

Testing uniformity on high-dimensional spheres against symmetric and asymmetric spiked alternatives

Speaker: Christine Cutting
Département de Mathématique, Université libre de Bruxelles, Belgium
Email: ccutting@ulb.ac.be

Abstract

The problem of testing uniformity on high-dimensional unit spheres is considered. We are primarily interested in non-null issues and focus on spiked alternatives. We show that such alternatives lead to two Local Asymptotic Normality (LAN) structures. The first one is for a fixed spiked direction θ and allows to derive locally asymptotically optimal tests under specified θ. The second one relates to the unspecified-θ problem and allows us to identify locally asymptotically optimal invariant tests. Interestingly, symmetric and asymmetric spiked alternatives lead to very different optimal tests, based on sample averages and sample covariance matrices, respectively. Most of our results allow the dimension p to go to infinity in an arbitrary way as a function of the sample size n.

This is joint work with Davy Paindaveine and Thomas Verdebout.

References:

1. C. Cutting, D. Paindaveine and T. Verdebout, Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives, *The Annals of Statistics*, to appear.

2. A. Onatski, M. J. Moreira and M. Hallin, Asymptotic power of sphericity tests for high-dimensional data, *The Annals of Statistics* **41** (2013), 1204–1231.

3. L. Rayleigh, On the problem of random vibrations and random flights in one, two and three dimensions, *Philosophical Magazine* **37** (1919), 321–346.

Local discriminant analysis for directional data

Speaker: Stefania Fensore
DMQTE, University of Chieti-Pescara, Italy
Email: stefania.fensore@unich.it

Abstract

We propose nonparametric methods to address the problem of classification for directional data. In particular proposed methods rely on local regression and kernel density estimation for data lying on the d-dimensional unit sphere. Firstly we introduce a local supervised classification estimating a rotation that superimposes the set of the observations to their labels. Then we propose a local logistic regression introducing several periodic link functions. Finally we adapt the KDE discrimination to the directional setting. We provide asymptotic theory for the proposed methods along with simulation results.

This is joint work with Marco Di Marzio, Agnese Panzera and Charles Taylor.

A discordancy test for isotropic random rotations

Speaker: Ulrike Genschel
Department of Statistics, Iowa State University, USA
Email: ulrike@iastate.edu

Abstract

We propose two test statistics that can be used to test for discordant observations in the rotation group $SO(3)$. The first is a novel statistic that accounts for the topology of the $SO(3)$ space, while the second is an extension of a statistic proposed previously for data on the unit hypersphere. Neither test statistic has distribution available in closed form. Therefore, we propose two methods to approximate the p-value for the test. The first is an upper bound based on the Bonferroni inequality and is applicable in a wide range of scenarios, while the second is a parametric bootstrap that requires a distributional assumption for the data under the null hypothesis. We compare the power of the tests based on each statistic and p-value approximation method in an empirical study and in a dataset from the material sciences.

Construction of circular distribution by using real-valued analytic function

Speaker: Tomoaki Imoto
School of Management and Information, University of Shizuoka, Japan
Email: imoto0923@gmail.com

Abstract

Many methods for constructing circular distributions are known in the literature, i.e., conditioning, maximizing entropy, Möbius transformation, stereographic projection, wrapping and so on. In my presentation, we consider the method for constructing a circular distribution by using a real-valued analytic function which controls the absolute value of trigonometric moment about the derived distribution. If the based analytic function includes only one parameter, we can construct four-parameter circular distribution and, through the re-parametrization, interpret the parameters by the first and second trigonometric moments. Especially, when the constructed distribution is unimodal, it can be directly parametrized in terms of parameters individually measuring location, concentration, skewness and kurtosis. The derived probability density function and its characteristic function are expressed by tractable forms, and therefore we can straightforwardly estimate the parameters by method of moments, estimation based on the characteristic function and maximum likelihood estimation.

Fitting mixtures of von Mises components with a reversible jump MCMC sampler

Speaker: Pieter Jongsma
Methoden & Statistiek, Faculteit Sociale Wetenschappen, Universiteit Utrecht, The Netherlands
Email: p.s.jongsma@uu.nl

Abstract

To model the distribution of circular data, a mixture model with von Mises components can be used. The Expectation Maximization (EM) algorithm has been developed for fitting mixtures of von Mises components (Banerjee et al. 2005). In a Bayesian framework, Markov chain Monte Carlo (MCMC) sampling provides a method for estimating the model parameters. In situations where the number of components in the mixture is not known, these methods can not be applied. The dimensionality of the parameter space depends on the number of components and so a method is required capable of estimating parameters across parameter spaces. Reversible jump algorithms (Richardson and Green 1997) have been developed that allow an MCMC sampler to jump across the differences in dimensions.

A reversible jump MCMC sampler for fitting a mixture of an unknown number of von Mises components is presented. The performance of the sampler has been investigated in a simulation study. Results show that for the presented scenarios, the number of simulated components is estimated adequately. For small samples ($n \leq 100$), there is high uncertainty about the number of components. At larger sample sizes ($500 \leq n \leq 10000$) the estimation of the number of components is generally accurate. Application of the sampler to music listening data shows interpretable results that correspond with intuition.

This is joint work with Kees Mulder and Irene Klugkist.

References:

1. A. Banerjee and I. S. Dhillon and J. Ghosh, and S. Sra, Clustering on the Unit Hypersphere using von Mises-Fisher Distributions, *Journal of Machine Learning Research* **6** (2005), 1345–1382.

2. S. Richardson, and P. Green, On Bayesian Analysis of Mixtures with Unknown Number of Components, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **59** (1997), 731–792.

Bayesian longitudinal circular data modeling

Speaker: Zeynep Kalaylioglu
Department of Statistics, Orta Doğu Teknik Üniversitesi, Türkiye
Email: kzeynep@metu.edu.tr

Abstract

Our research is motivated by a recent medical study that aims to estimate the general fetal head rotation trajectory during the first stage of normal labour adjusted for maternal characteristics and environmental factors. A rather primitive manual method for determining the rotation has recently been replaced by an ultrasound technology that can precisely measure the fetal's head angle. The particular challenge with such data are the model selection procedures that could objectively assess the models when outcome data are longitudinal and directional. We consider a Bayesian random effects model on the circle and review and comment on the current model selection technology used in Bayesian analysis of circular data. We then focus our attention on criteria based on minimizing a predictive loss and propose new extensions to a current method by Ravindiran and Ghosh (2011). We use extensive Monte Carlo simulation studies controlled for the sample size and intraclass correlation to study the performances of these criteria under various realistic longitudinal settings. We used several quantities to evaluate the model selection criteria such as proportion of selecting the true model and a ratio that measures the strength of the particular selection. Simulations reveal a noticeable gain in performance achieved by the proposed method. A conventional longitudinal data set (sandhopper data by D'elia et al. 2001) is used to further compare the Bayesian model selection methods for circular data (ours, Nunez-Antonio et al. (2011), Ravindiran and Ghosh (2011), DIC, DIC_4). This research hopes to address and contribute to the model selection in directional data, a rather fertile area for methodological and theoretical development, while the demand increases with the advancing technology as seen in our motivating data set.

This is joint work with Onur Camli.

References:

1. A. D'elia, C. Borgioli and F. Scapini, Orientation of sandhoppers under natural conditions in repeated trials: an analysis using longitudinal directional data, *Estuarine, Coastal and Shelf Science* **53** (2001), 839–847.

2. G. Nunez-Antonio, E. Gutierrez-Pena and G. Escarela, A Bayesian regression model for circular data based on the projected normal distribution, *Statistical Modelling* **11** (2011), 185–201.

3. P. Ravindiran and S.K. Ghosh, Bayesian analysis of circular data using wrapped distributions, *Journal of Statistical Theory and Practice* **5** (2011), 547–561.

Modelling biological rhythms using order restricted inference

Speaker: Yolanda Larriba
Departamento de Estadística e Investigación Operativa, Universidad de Valladolid, Spain
Email: yolanda.larrriba@uva.es

Abstract

Many biological processes, such as cell cycle, circadian clock or menstrual cycles, are governed by oscillatory systems consisting of numerous components that exhibit periodic patterns over time. Modelling these rhythms is a challenge in literature since usually the sampling density is low, the number of periods is generally two and the underlying signals adopt a wide range of temporal patterns, see Larriba et al. (2016). Several authors proposed parametric functions of time, such as the sinusoidal function, to model these signals. However these parametric functions might be too rigid for data derived from cell-cycle or circadian clock. Among these, a common shape of interest to a biologist is the circular up-down-up signal with a unique peak (U) and a unique trough (L) within each period. The shape of these signals is entirely described by mathematical inequalities among their components which allow to establish a relationship between the euclidean and the circular space using circular isotonic regression.

In this work we state this connection between the euclidean and the circular space based on circular isotonic regression, formulate the ML isotonic regression estimator under circular up-down-up constraints and assess its computational advantages to calculate the circular isotonic regression estimator (CIRE), see Rueda et al. (2009). Results are shown both on simulations and on real data.

This is joint work with Cristina Rueda, Miguel A. Fernández and Shyamal D. Peddada.

References:

1. Y. Larriba, C. Rueda, M. A. Fernández and S. D. Peddada, Order restricted inference for oscillatory systems for detecting rhythmic signals, *Nucleic Acids Research* **44** (2016), DOI:10.1093/nar/gkw771.

2. C. Rueda, M. A. Fernández and S. D. Peddada, Estimation of parameters subject to order restrictions on a circle with application to estimation of phase angles of cell cycle genes, *Journal of the American Statistical Association* **104** (2009), 338–347.

Hybrid mutual information

Speaker: Ignacio Leguey
Department of Artificial Intelligence, Universidad Politécnica de Madrid, Spain
Email: ig.leguey@upm.es

Abstract

Measuring the mutual dependence between two linear variables has been studied at length in Rényi (1959a,b) and Lloyd (1962), among many others. Mutual information (Shannon 1949, Cover and Thomas 2012) between two linear variables is a general measure that determines the similarity between the joint distribution and the product of their marginal distributions. For directional statistics, the circular mutual information was recently proposed in Leguey et al. (2016). This is suitable when the underlying paired distributions follow bivariate wrapped Cauchy distributions (Kato and Pewsey 2015), whose marginals and conditionals belong to the univariate wrapped Cauchy family.

Here we go one step further by presenting the hybrid mutual information, which allows to express in a closed form the mutual information measure between a circular-linear or a linear-circular pair of variables regardless of the marginal distribution of each variable.

This is joint work with Shogo Kato, Concha Bielza and Pedro Larrañaga.

References:

1. T. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons (2012).

2. S. Kato and A. Pewsey, A Möbius transformation-induced distribution on the torus, *Biometrika* **102** (2015), 359–370.

3. I. Leguey, C. Bielza and P. Larrañaga, Tree-structured Bayesian networks for wrapped Cauchy directional distributions, *Lecture Notes in Computer Science*, Springer, **9868** (2016), 207–216.

4. S.P. Lloyd, On a measure of stochastic dependence, *Theory of Probability & Its Applications* **7** (1962), 301–312.

5. A. Rényi, On measures of dependence, *Acta Mathematica Academiae Scientiarum Hungarica* **10** (1959a), 441–451.

6. A. Rényi, On the dimension and entropy of probability distributions, *Acta Mathematica Academiae Scientiarum Hungarica* **10** (1959b), 193–215.

7. C.E. Shannon and W. Weaver *The Mathematical Theory of Communication*, Urbana: University of Illinois Press **29** (1949).

Directional-linear data clustering using structural expectation-maximization algorithm

Speaker: Sergio Luengo-Sanchez
Department of Artificial Intelligence, Universidad Politécnica de Madrid, Spain
Email: sluengo@fi.upm.es

Abstract

The study of plethora of phenomena requires the measurement of their magnitude and direction as in meteorology (Carta et al. 2009), rhythmometry, medicine or demography (Batschelet 1981, Batschelet et al. 1973). Probabilistic clustering of this data is typically tackled by means of mixtures of Gaussians (Fraley and Raftery 2002, McLachlan and Basford 1988, Melnykov and Maitra 2010), although they tend to underperform due to their inability to handle periodicity of directional data. To address this problem several distributions have been proposed to cluster bivariate cylindrical data (Carta et al. 2009, Gatto and Jammalamadaka 2007, Mardia and Sutton 1978, Qin et al. 2010) and multivariate data having one circular variable (Roy et al. 2014).

Recently, an approach (Luengo-Sanchez et al. 2016) based on exploiting the conditional independence assumptions encoded by a Bayesian network enables efficient clustering of multivariate directional-linear data, distributed as Gaussian and von Mises respectively, even when there is more than one directional variable by means of the structural expectation-maximization algorithm (Friedman 1997). However, strong constraints on the structure of the Bayesian network must be imposed.

Here we propose measures of divergence and distance among clusters, as Kullback-Leibler divergence and Bhattacharyya distance, for the previous model to evaluate the quality of the clustering outcomes and we extend the model by relaxing the structural constraints to include relations of dependence between directional variables and Gaussians. We present an application for neuroscience to cluster dendritic spines according to a set of morphological features that combine directional and linear variables.

This is joint work with Concha Bielza and Pedro Larrañaga.

References:

1. E. Batschelet, *Circular Statistics in Biology*, Academic Press, London (1981).

2. E. Batschelet, D. Hillman, M. Smolensky and F. Halberg, Angular-linear correlation coefficient for rhythmometry and circannually changing human birth rates at different geographic latitudes, *International Journal of Chronobiology* **1** (1973), 183–202.

3. J. A. Carta, P. Ramirez and C. Bueno, A joint probability density function of wind speed and direction for wind energy analysis, *Energy Conversion and Management* **49** (2009), 1309–1320.

4. C. Fraley and A. E. Raftery, Model-based clustering, discriminant analysis and density estimation, *Journal of the American Statistical Association* **97** (2002), 611–631.

5. R. Gatto and S. R. Jammalamadaka, The generalized von Mises distribution, *Statistical Methodology* **4** (2007), 341–353.

6. N. Friedman, Learning belief networks in the presence of missing values and hidden variables, *in Proceedings of the Fourteenth International Conference of Machine Learning (ICML 1997)*, D. H. Fisher (eds.), pp. 125–133, Morgan Kaufmann, Nashville (1997).

7. S. Luengo-Sanchez, C. Bielza and P. Larrañaga, Hybrid Gaussian and von Mises model-based clustering, *in ECAI 2016: 22nd European Conference on Artificial Intelligence*, G. A. Kaminka, M. Fox and P. Bouquet (eds.), pp. 855–862, IOS Press, The Hague (2016).

8. K. V. Mardia and T. W. Sutton, A model for cylindrical variables with applications, *Journal of the Royal Statistical Society Series B* **40** (1978), 229–233.

9. G. McLachlan and K. E. Basford, *Mixture Models: Inference and Applications to Clustering*, Marcel Dekker, New York (1988).

10. V. Melnykov and R. Maitra, Finite mixture models and model-based clustering, *Statistics Surveys* **4** (2010), 80–116.

11. X. Qin, S. J. Zhang and D. X. Yan, A new circular distribution and its application to wind data, *Journal of Mathematics Research* **2** (2010), 12–17.

12. A. Roy, S. K. Parui and U. Roy, SWGMM: A semi-wrapped Gaussian mixture model for clustering of circular-linear data, *Pattern Analysis and Applications* **19** (2014), 631–645.

Bayesian tests for circular isotropy

Speaker: Kees Mulder
Department of Methodology and Statistics, Utrecht University, The Netherlands
Email: K.T.Mulder@uu.nl

Abstract

Circular data are data measured in angles or directions, which occur in a wide variety of scientific fields. An initial hypothesis of interest is that of isotropy, that is, circular uniformity. Frequentist methods for assessing the isotropy null hypothesis exist, but do not allow the user faced with an insignificant result to distinguish lack of power and support for the null hypothesis. Bayesian hypothesis tests are developed here which solve this issue. Priors based on a conjugate prior and the Jeffreys prior are developed. Two alternative hypotheses are considered, the von Mises distribution and a kernel density alternative, which acts as an omnibus test. Assessing the performance of the tests using different priors, it is shown that they are powerful and perform well, and allow more elaborate conclusion than the classical tests of isotropy.

This is joint work with Irene Klugkist.

A class of depth functions for directional data

Speaker: Giovanni C. Porzio
Department of Economics and Law, University of Cassino and Southern Lazio, Italy
Email: porzio@unicas.it

Abstract

Directional data are constrained to lie on circles, spheres or hyperspheres. They naturally arise when measurements are directions, axes, rotations, clock or calendar measures, and are of some interest in multivariate analysis as well. A specific feature of directional data is the lack of a natural ordering. Consequently, data depth functions can be particularly useful in this field, given that they aim to provide an inner-outward ranking of the data. However, few depth measures have been proposed and investigated within this setting. Furthermore, those most commonly used and studied are computationally expensive. For this reason, this work introduces a class of distance-based depth functions for directional data, for which several properties are established. Within the class, two cases are deeply investigated: the arc distance and the cosine distance depths. Their computational costs, efficiency and robustness as location estimators are evaluated and compared with the angular simplicial and Tukey's depth functions through a simulation study. In addition, in order to deal with multimodal directional densities, this work also introduces the corresponding class of distance-based local depth functions. The latter is eventually exploited to define a plot able to reveal if a directional distribution is unimodal.

This is joint work with Giuseppe Pandolfo.

An extension of Kuiper's test for bivariate circular data

Speaker: Rachael Quill
School of Physical, Environmental and Mathematical Sciences, University of New South Wales (UNSW) Canberra, Australia
Bushfire and Natural Hazards Cooperative Research Centre, Australia
Email: rachael.quill@gmail.com

Abstract

Kolmogorov-Smirnov (KS) style tests are well-known and commonly applied across numerous fields. The KS test allows for efficient and distribution-free comparison of two univariate linear distributions, while Kuiper's test (Kuiper 1960) provides an adaptation which allows for comparison of circular distributions. To compare two bivariate datasets, the KS test has been extended to allow for sufficiently distribution-free comparison for all 'practical' cases (Peacock 1983). While the asymptotic behaviour and power of the univariate KS style tests have been well studied (e.g. Stephens 1970), the asymptotic behaviour of the extended KS test has been analysed by a limited number of authors.

Joint distributions between prevailing wind directions and those observed on the surface are used to represent the response of wind flows to the underlying terrain. In understanding the variation of wind fields across complex landscapes, it is of interest to compare these bivariate circular datasets under differing environmental conditions. However, a bivariate extension of Kuiper's test for the comparison of such toroidal distributions does not appear to exist in the literature. Quill et al. (2017) have therefore proposed an equivalent extension of Kuiper's test to that of the extended KS test. The power of this proposed test is considered under changing distribution structures, while the power curves are also evaluated in relation to those of both the univariate Kuiper's test and the bivariate KS test.

References:

1. Kuiper, N. H., Tests concerning random points on a circle, *Indagationes Mathematicae (Proceedings)* **63** (1960), 38–47.

2. Peacock, J. A., Two-dimensional goodness-of-fit testing in astronomy, *Monthly Notices of the Royal Astronomical Society* **202** (1983), 615–627.

3. Quill, R., Sharples, J. J. and Sidhu, L. A., Sensitivity analysis of Kolmogorov–Smirnov style statistics for univariate and bivariate data, *Journal of Applied Statistics* (Under Review).

4. Stephens, M. A., Use of the Kolmogorov–Smirnov, Cramer-von Mises and related statistics without extensive tables, *Journal of the Royal Statistical Society Series B (Methodological)* **32** (1970), 115–122.

From high to low dimension: connections between different PCA-like approaches for circular data

Speaker: Karen Sargsyan
Institute of Biomedical Sciences, Academia Sinica, Taiwan
Email: karsar@ibms.sinica.edu.tw

Abstract

Principal Component Analysis (PCA) enjoyed massive success in applications where original multidimensional linear data looks intractable, but most of the variation could be expressed just by using a few dimensions. As an example, to analyse Molecular Dynamics trajectories one can use "essential" coordinates obtained by applying PCA to the highly dimensional data composed of atomic coordinates. This approach still poses some challenges in analysis, as number of variables remains high. One may consider to use angles instead of Cartesian coordinates, which is itself an effective reduction of dimensionality. This led to several proposals of PCA style analysis for circular data such as dPCA (Altis et al. 2007), GeoPCA (Sargsyan et al. 2012) and Torus PCA (Eltzner et al. 2017). While original motivation came from the field of Structural Biology, PCA methods could be applied in various contexts. Therefore, it is important to clarify properties of dPCA, GeoPCA and Torus PCA, as well as their relationship. I will demonstrate several new interesting connenctions among them and illustrate their behaviour on new data sets of biological importance.

References:

1. A. Altis, P. H. Nguyen, R. Hegger, G. Stock, Dihedral angle principal component analysis of molecular dynamics simulations, *The Journal of Chemical Physics* **126** (2007), 244111.

2. K. Sargsyan, J. Wright and C. Lim, GeoPCA: a new tool for multivariate analysis of dihedral angles based on principal component geodesic, *Nucleic Acids Research* **40** (2012), e25.

3. B. Eltzner, S. Huckemann, K. V. Mardia, Torus Principal Component Analysis with an application to RNA structures, *arXiv*, (2017).

The lattice generalized von Mises distribution

Speaker: Anna Sjöström
Department of Statistics, Lund University, Sweden
Email: anna.sjostrom@stat.lu.se

Abstract

The Lattice Generalized von Mises distribution is presented, which allows for modelling multimodal discrete circular data, where no further restrictions are put on data. In the case of circular data, a wide range of continuous distributions are described in the literature, whereas the discrete distributions presented are fewer. Nevertheless, the need for discrete circular distributions is not insignificant, since circular data are often observed on discrete outcome spaces, for example in the form of temporal or spatial data. For continuous circular data, the Generalized von Mises distribution (GvM) allows for modelling higher orders of modality. Due to this fact, the GvM is a natural starting point in the process of modelling unrestricted discrete circular multimodal data, leading to what we call the Lattice Generalized von Mises distribution. This is a flexible probability model which is able to handle multimodality and asymmetry between modes. The pdf of the continuous GvM typically includes types of modality in a sequential manner. In the proposed model, we allow for including some but not necessarily all types of modality of a sequence. We demonstrate the model for a data set concerning bird migration behavior, where counts on circle sectors are measured over a time interval.

A scale space approach for spherical density estimation

Speaker: Ville Vuollo
Department of Orthodontics, Oral Health Sciences, University of Oulu, Finland
Medical Research Center, Oulu University Hospital, Finland
Research Unit of Mathematical Sciences, University of Oulu, Finland
Email: ville.vuollo@oulu.fi

Abstract

Statistical scale space methods are useful tools for exploring significant underlying structures in data (Holmström and Pasanen 2016). SiZer (SIgnificant ZERo crossing of the derivatives) has become a seminal tool in this field (Chaudhuri and Marron 1999). In SiZer the features of the function of interest are examined using simultaneously many different levels of smoothing and the results are summarized as easily interpretable color maps. Inspired by CircSiZer (see Oliveira et al. 2014), the SiZer tool for circular data, we extended the SiZer approach to spherical data. The novel method, SphSiZer, can be used to explore the structure of a probability density function defined on the two-dimensional sphere. The scale-dependent features of the density are visualized by projecting the statistically significant gradients on a planar contour plot of the density function in a manner similar to the scale space analysis of a euclidean bi-variate density function proposed in Godtliebsen et al. (2002). A movie can be used to present the results for a large number of scales. We apply our method to the analysis of samples of surface unit normal vectors of an infant head. Flat regions of the head lead to local maxima in the density function of normal directions thus presenting a potentially useful context in which to explore density features with SphSiZer. The performance of the method is also checked with simulated data.

This is joint work with Lasse Holmström.

References:

1. P. Chaudhuri and J. S. Marron, SiZer for exploration of structures in curves, *Journal of the American Statistical Association* **94** (1999), 807–823.

2. F. Godtliebsen, J. S. Marron and P. Chaudhuri, Significance in scale space for bivariate density estimation, *Journal of Computational and Graphical Statistics* **11** (2002), 1–21.

3. L. Holmström and L. Pasanen, Statistical scale space methods, *International Statistical Review* DOI: 10.1111/insr.12155 (2016).

4. M. Oliveira, R. Crujeiras and A. Rodríguez-Casal, CircSiZer: an exploratory tool for circular data, *Environmental and Ecological Statistics* **21** (2014), 143–159.

List of participants

Name	Institution	Email	Page of Abstract
Abe Toshihiro	Nanzan University, Nagoya, Japan,	abetosh@ss.nanzan-u.ac.jp	7
Ameijeiras-Alonso José	Universidad de Santiago de Compostela, Spain,	jose.ameijeiras@usc.es	36
Arnold Richard	Victoria University of Wellington, New Zealand	richard.arnold@vuw.ac.nz	8
Buttarazzi Davide	Università di Cassino, Italy	d.buttarazzi@unicas.it	37
Chacón José Enrique	University of Extremadura, Spain	jechacon@unex.es	38
Craens Domien	Universiteit Gent, Belgium	domien.craens@ugent.be	
Cremers Jolien	Utrecht University, Holland	j.cremers@uu.nl	39
Crujeiras Rosa	Universidad de Santiago de Compostela, Spain	rosa.crujeiras@usc.es	9
Cutting Christine	Université libre de Bruxelles, Belgium	ccutting@ulb.ac.be	40
Di Marzio Marco	Universitá di Chieti-Pescara, Italy	dimarzio@dmqte.unich.it	10
Fensore Stefania	Università di Pescara, Italy	stefania.fensore@unich.it	41
Fernández Miguel	Universidad de Valladolid, Spain	miguelaf@eio.uva.es	41
García-Portugués Eduardo	Universidad Carlos III de Madrid, Spain	edgarcia@est-econ.uc3m.es	12
Genschel Ulrike	Iowa State University, USA	ulrike@iastate.edu	42
Holmström Lasse	University of Oulu, Finland	lasse.holmstrom@oulu.fi	13
Huckemann Stephan	University of Göttingen, Germany	huckeman@math.uni-goettingen.de	14
Imoto Tomoaki	University of Shizuoka, Japan	imoto@u-shizuoka-ken.ac.jp	43
Jona-Lasinio Giovanna	Sapienza Università di Roma, Italy	giovanna.jonalasinio@uniroma1.it	15
Jongsma Pieter	Utrecht University, Holland	p.s.jongsma@uu.nl	44
Jupp Peter	University of St. Andrews, UK	pej@st-andrews.ac.uk	16
Kalaylioglu Zeynep	Middle East Technical University, Turkey	kzeynep@metu.edu.tr	45
Kato Shogo	Institute for Statistical Mathematics, Tokyo, Japan	skato@ism.ac.jp	17
Kent John	University of Leeds, UK	j.t.kent@leeds.ac.uk	18
Klugkist Irene	Utrecht University, Netherlands	i.klugkist@uu.nl	19
Lagona Francesco	Università di Roma Tre, Italy	lagona@uniroma3.it	20
Larriba Yolanda	Universidad de Valladolid, Spain	yolanda.larriba@uva.es	46
Leguey-Vitoriano Ignacio	Universidad Politécnica de Madrid, Spain	ig.leguey@upm.es	47
Ley Christophe	Universiteit Gent, Belgium	christophe.ley@ugent.be	21
Luengo Sergio	Universidad Politécnica de Madrid, Spain	sluengo@fi.upm.es	48
Mardia Kanti	University of Leeds, UK	k.v.mardia@leeds.ac.uk	22
Marinucci Domenico	University of Rome 2, Italy	marinucci@mat.uniroma2.it	23
Maruotti Antonello	LUMSA, Italy	a.maruotti@lumsa.it	24
Mastrantonio Gianluca	Politécnico di Torino, Italy	mastrantonio.gianluca@gmail.com	25
Mulder Kees	Utrecht University, Holland	k.t.mulder@uu.nl	50
Nicosia Aurélien	University of Laval, Canada	nicosia.aurelien@gmail.com	26
Nordman Dan	Iowa State University, USA	dnordman@iastate.edu	27
Paindaveine Davy	Université libre de Bruxelles, Belgium	dpaindav@ulb.ac.be	28
Panzera Agnese	Università di Firenze, Italy	a.panzera@disia.unifi.it	29
Pewsey Arthur	Universidad de Extremadura, Spain	apewsey@unex.es	30
Picone Marco	ISPRA, Italy	marco.picone@isprambiente.it	
Porzio Giovanni	Università di Cassino, Italy	porzio@unicas.it	51
Ranalli Monia	Universitá di Roma Tre, Italy	monia.ranalli@uniroma3.it	
Rivest Louis-Paul	University of Laval, Canada	louis-paul.rivest@mat.ulaval.ca	31
Sargsyan Karen	Academia Sinica, Taipei	karsar@ibms.sinica.edu.tw	53
Scealy Janice	Australian National University, Australia	janice.scealy@anu.edu.au	32
Shimizu Kunio	Institute for Statistical Mathematics, Japan	k-shmz@ism.ac.jp	33
Simone Rosaria	Universitá di Napoli Federico II, Italy	rosaria.simone@unina.it	
Sjöström Anna	Lund University, Sweden	anna.sjostrom@stat.lu.se	54
Verdebout Thomas	Université libre de Bruxelles, Belgium	tverdebout@gmail.com	34
Vuollo Ville	University of Oulu, Finland	ville.vuollo@oulu.fi	55