Abstracts of the 9-th International Workshop on Applied Probability

18-21 June 2018,

Budapest, Hungary

Edited by László Márkus and Vilmos Prokaj

The organisation of the International Workshop on Applied Probability is initiated and supervised by the **IWAP International Board**.

The IWAP International Board

Joseph Glaz (chair), University of Connecticut, USA
Jürg Hüsler, University of Bern, Switzerland
Nikolaos Limnios, Technical University of Compiegne, France
Markos Koutras, University of Piraeus, Greece
José Luis Palacios, University of New Mexico, USA; University Simón Bolívar, Venezuela

The IWAP International Board appointed the committees below.

The Scientific Programme Committee:

Co-chairs:
Joseph Glaz, University of Connecticut, USA
László Márkus, Eötvös Loránd University, Hungary

Jogesh Babu, Pennsylvania State University, USA
Narayanaswamy Balakrishnan, McMaster University, Canada
Viktor Beneš, Charles University, Prague, Czech Republic
Alexander Bulinski, Lomonosov Moscow State University, Russia
Krzysztof Dębicki, University Wroclaw, Poland
Serkan Eryilmaz, Atilim University, Ankara, Turkey
Enkelejd Hashorva, HEC Lausanne, Switzerland
Remco van der Hofstad, Technische Universiteit Eindhoven, The Netherlands
Blanka Horváth, Imperial College, London, UK
Ron Kenett, Samuel Neaman Institute, Technion, Israel and KPA Ltd., Raanana, Israel
Claude Lefevre, University Libre de Bruxelles, Belgium
Haijun Li, Washington State University, USA
Wendy Lou, University of Toronto, Canada
Nitis Mukhopadhyay, University of Connecticut, USA
Mark Podolskij, Aarhus University, Denmark
Cristian Preda, Polytech Lille, France
Marco Scarsini, LUISS Rome, Italy
Christos H. Skiadas, Technical University of Crete, Greece
Evgeny Spodarev, Ulm University, Ulm, Germany
Peter Spreij, University Amsterdam, The Netherlands
Jianfeng Yao, University of Hong Kong, Hong Kong
Elena Yarovaya, Lomonosov Moscow State University, Russia

The Local Organising Committee:

Chair:
László Márkus, Eötvös Loránd University, Hungary

Miklós Arató, Eötvös Loránd University, Hungary
Ágnes Backhausz, Eötvös Loránd University, Hungary
Sándor Baran, University of Debrecen, Hungary
István Fazekas, University of Debrecen, Hungary
György Michaletzky, Eötvös Loránd University, Hungary
Tamás Móri, Eötvös Loránd University, Hungary
Vilmos Prokaj, Eötvös Loránd University, Hungary
Katalin Varga, Hungarian National Bank, Hungary
András Zempléni, Eötvös Loránd University, Hungary

Invited Session Organizers

Sándor Baran University of Debrecen, Hungary
Felix Belzunce, University of Murcia, Spain
Viktor Beneš, Charles University Prague, Czech Republic
Mark Brown, Columbia University, USA
Alexander Bulinski, Lomonosov Moscow State University, Russia
Ji Hwan Cha, Ewha Womans University, Seoul, South-Korea
Jie Chen, University of Massachusetts, USA
Umberto Cherubini, University of Bologna, Italy
Emilio De Santis, Sapienza University, Rome, Italy
Krzysztof Dębicki, University Wroclaw, Poland
Dipak K. Dey, University of Connecticut, USA
Antonio Di Crescenzo, University of Salerno, Italy
Luiz Duczmal, Federal University Minas Gerais, Brazil
Serkan Eryilmaz, Atilim University, Ankara, Turkey
Georgios Fellouris, University of Illinois, Urbana-Champaign, USA
Larry Goldstein, University of Southern California, USA
Priscilla Greenwood, University of British Columbia, Canada
Gholamhossein Hamedani, Marquette University, Milwaukee, USA
Enkelejd Hashorva, HEC Lausanne, Switzerland
Blanka Horváth, Imperial College, London, UK
Markos Koutras, University of Piraeus, Greece
Mei-Ling Ting Lee University of Maryland, USA
Claude Lefevre, University Libre de Bruxelles, Belgium
Regina Liu, Rutgers University, USA
Hosam Mahmoud, The George Washington University, USA
Yaakov Malinovsky, University Maryland, Baltimore County, USA
Raimondo Manca, Sapienza University, Rome, Italy
George Moustakides, Rutgers University, USA
Nitis Mukhopadhyay, University of Connecticut, USA
Jorge Navarro, University of Murcia, Spain
Enrica Pirozzi, University of Napoli Federico II., Italy
Mark Podolskij, Aarhus University, Denmark
Cristian Preda, Polytech Lille, France
Bruno Remillard, HEC Montreal, Canada
Marco Scarsini, LUISS Rome, Italy
Dmitrii Silvestrov, Stockholm University, Sweden
Fabio Spizzichino, Sapienza University, Rome, Italy
Peter Spreij, University Amsterdam, The Netherlands
Peter Straka, University New South Wales, Sydney, Australia
Thomas Taimre, University Queensland, Australia
Peter Tankov, ENSAE ParisTech, France
Venugopal V. Veeravalli, University of Illinois, Urbana-Champaign, USA
Jianfeng Yao, University of Hong Kong, Hong Kong
Elena Yarovaya, Lomonosov Moscow State University, Russia

Preface

Dear Participant, dear Reader,

It was the end of 2016 when I learned through a friend that the IWAP board is still in search for the venue and the organiser of the next meeting, due in the summer of 2018. With the organisation of one major and a few medium size conferences under my belt the first thought that stroke me was: mission doomed to failure, but then: we could do it with the team I worked together with on similar events. Decision came after thought, we applied for, and were granted the organisation of the current event. The first thing we had to make compromise was the timing: dates for other events have already been set, and the second half of June is far from ideal for conferences, the academic year is still on in many countries in Europe. Communication has also proven to be far from simple: everyone is tired of mass e-mails commercial conference organisers send out, so e-mails often land in spam folders. Personal connections and special care in writing e-mails were a major help in reaching out to potential participants.

The first major breakthrough came when the six solicited plenary speakers, all highly renowned experts of their field, agreed to come and present their lectures. Next, the Scientific Programme Committee was formed and started its work in initiating the organisation of invited sessions, and compiling the programme. When the number of invited talks - which is traditionally high in IWAP conferences - exceeded 150 my anxiety about failure definitely started to fade away. In the end, from 343 submitted abstracts 291 remain to be presented in the conference, putting the current IWAP above average in attendance.

In what traditional conferences surpass and outbeat commercial ones is the quality of talks and presented results and in that respect the abstracts of the current IWAP meeting promise a distinctive excellence. Traditional applications of probability theory appear in great numbers and variety, among them financial ones represent a particular focus of the current event. However, we also see new topics to emerge, e.g. ones related to artificial intelligence and deep learning. It is our duty to foster these new ideas by giving them proper representations.

The organisers express their deep gratitude to the IWAP International Board for granting them the right and possibility to organise the current meeting, and supporting them permanently in their work. In particular, we are greatly indebted to Joseph Glaz, head of the IWAP International Board, and co-chair of the conference for his permanent support, and active participation in the organisation, suggesting and inviting speakers. I, personally, am happy to emphasize the importance of the smooth, permanent and burden sharing working relationship with Joe during the preparations.

In the elected Scientific Programme Committee I found dedicated people who did not spare efforts in creating interesting and high quality sessions and finding the most suitable speakers to it. The Local Organising Committee provided essential help in many scientific and technical matters on the local level. In particular, let me mention Vilmos Prokaj, who supervised the submission of abstracts and edited this volume. Without his dedication this book wouldn't have been as nice and operational as it is now. Vilmos' thorough work is reflected in every page of this book, it is me to blame for all inconsistencies and errors.

In technical and financial matters Congress Ltd. provided invaluable assistance, their experience

and professionalism made a very rough path smoothly passable. My special thanks go to chief executive Judit Vermesné Stefkó, who supervised and directed all the preparations of the conference and stays on board with us managing all services until the closing words.

I gladly acknowledge the support of the Mathematical Institute and the Faculty of Science of Eötvös Loránd University. We are also grateful to the Institute of Mathematical Statistics for co-sponsoring the conference.

I am greatly indebted to all these people, and to all others not mentioned here, who this way or another helped the organisation. However, our organizational effort provides only the frames to the ultimate contribution that is yours dear Participant and Reader. Your combined performance may make the conference a success, and that serves as measure of the quality of our preparatory efforts.

I am happy to greet all of you in Budapest and hope you will experience the best traditions of Hungarian hospitality.

László Márkus
Conference co-chair

Contents

IS 6
Affine & Polynomial Proc.
Harmony lnge
Wed
17:30 - 18:00

The Volterra square-root process: Markovian structure and approximation

EDUARDO ABI JABER[*,†,§], OMAR EL EUCH[‡]

[*]Université Paris-Dauphine, Paris
[†]AXA-Investment Managers, Paris
[‡]Ecole Polytechnique, Palaiseau
email: [§]eduardo.abi.jaber@gmail.com

We characterize the Markovian and affine structure of the Volterra square-root process in terms of an infinite-dimensional adjusted forward process and specify its state space. More precisely, we show that it satisfies a stochastic partial differential equation and displays an exponentially-affine characteristic functional. As an application, we deduce an existence and uniqueness result for a Banach-space valued square-root process and provide its state space. This yields another representation of the Volterra square-root process in terms of a (possibly) infinite system of affine diffusions. We show that these representations are of particular importance in practice as they lead to new numerical approximation schemes mainly for rough volatility modeling.

The talk is based on joint work with Omar El Euch [1, 2].

References

[1] Eduardo Abi Jaber, and Omar El Euch. Markovian structure of the Volterra Heston model. HAL: hal-01716696, 2018.
[2] Eduardo Abi Jaber, and Omar El Euch. Multi-factor approximation of rough volatility models. arXiv: 1801.10359, 2018.

IS 22
Renewal & semi-Markov...II.
Rm -1.62
Tue
15:30 - 16:00

Markov Chains and Information Networks

BENARD ABOLA[*,¶], PITOS SELEKA BIGANDA[*,†], CHRISTOPHER ENGSTRÖM[*],
JOHN MANGO MAGERO[‡], GODWIN A. KAKUBA[‡], SERGEI SILVESTROV[*], DMITRII SILVESTROV[§]

[*]Mälardalen University, Västerås, Sweden
[†]University of Dar es Salaam, Dar es Salaam, Tanzania
[‡]Makerere University, Kampala, Uganda
[§]Stockholm University, Stockholm, Sweden
email: [¶]benard.abola@mdh.se

Perturbed Markov chains are popular models for description of information networks. In such models, the transition matrix $\mathbf{P}$ of an information Markov chain is usually approximated by matrix $\mathbf{P}_\varepsilon = (1-\varepsilon)\mathbf{P} + \varepsilon\mathbf{D}$, where $\mathbf{D}$ is a so-called damping stochastic matrix with identical rows and all positive elements, while $\varepsilon \in [0,1]$ is a perturbation (regularisation) parameter. We perform the detailed perturbation analysis of such models, which includes a procedure of artificial regeneration for the approximating Markov chain with the matrix of transition probabilities $\mathbf{P}_\varepsilon$ and application of ergodic theorems for perturbed regenerative processes [1], coupling methods, and methods of phase space reduction for perturbed semi-Markov processes [2], for getting effective explicit series representations for the corresponding stationary distributions $\bar{\pi}_\varepsilon$, explicit effective upper bounds in ergodic theorems for $\mathbf{P}_\varepsilon^n$, as $n \to \infty$ and $\varepsilon \to 0$, upper bounds for the deviation $|\bar{\pi}_\varepsilon - \bar{\pi}_0|$, and asymptotic expansions for $\bar{\pi}_\varepsilon$ with respect to the perturbation parameter ε. Applications to concrete information Markov chains and results of some numerical experiments are also presented.

References

[1] Gyllenberg, M. and Silvestrov, D.S (2008) Quasi-stationary phenomena in nonlinearly perturbed stochastic systems. De Gruyter Expositions in Mathematics, vol.44, ix+579 pp. Walter de Gruyter, Berlin

[2] Silvestrov, D. and Silvestrov, S. (2017) Nonlinearly Perturbed Semi-Markov Processes, Springer.

Dynamic Cournot-Nash equilibrium via causal optimal transport

IS 3
Optimal Transport ...
Harmony Inge
Wed
15:00 - 15:30

Beatrice Acciaio[*,‡], Julio Backhoff-Veraguas[†]

[*]The London School of Economics and Political Science, London, UK
[†]University of Technology of Vienna, Vienna, Austria
email: [‡]b.acciaio@lse.ac.uk

We study Cournot-Nash equilibrium problems in a dynamic setting, where each agent faces a cost that is composed by an idiosyncratic part depending on its own type and action, and a mean-field term depending on the actions distribution over all agents. We use tools from dynamic optimal transportation of non-anticipative nature in order to get a characterization of the equilibrium problems, along with existence and uniqueness results.

On the conditional quantile estimation under association

CS 19
Distributio
Rm 100/B
Tue
17:30-17:50

Adjoudj Latifa[*], Tatachak Abdelkader

Faculty of Mathematics, USTHB, Algeria
email: [*]ladjoudj@usthb.dz

In the present work, we are interested in the study of the asymptotic analysis of a kernel estimator for the conditional distribution and conditional quantile functions when the lifetime observations satisfying the association dependence in the sense of Esary et al. [1967] For this, we first establish the strong uniform consistency and then the accuracy of the estimators is checked by a simulation study.

References

Esary, J., Proschan, F. and Walkup, D. (1967) Association of random variables with applications. *Ann. Math. Statist.*, **38**, p.1466–1476.

Stability analysis of queueing systems with various rules of service and regenerative input flow

IS 32
Limits in Inventory...
Rm -1.64
Tue
17:30 - 18:00

L. Afanaseva

Department of Probability, Faculty of Mathematics and Mechanics, Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow, 119991, Russian Federation
email: l.g.afanaseva@yandex.ru

We consider a multiserver queueing system with a regenerative input flow $X(t)$. The service procedure defines the stochastic process $Y(t)$ that is the number of customers served during time interval $(0,t)$. We introduce the auxiliary process $\tilde{Y}(t)$ representing the number of served customers under condition that there are always customers in the queue. Then we formulate conditions under which there are common points of regeneration for the both processes $X(t)$ and $\tilde{Y}(t)$ and express the

traffic rate ρ with the help of rates of these processes. Basing on connection between the real service process $Y(t)$ and the auxiliary process $\tilde{Y}(t)$ we define conditions under which the system is stable if and only if $\rho < 1$. The obtained results are used for stability analysis of two models which are important from the applied point of view. The first one is a multiserver retrial queueing system with a constant retrial rate and the second one is a system with a simultaneous service of a customer by a random number of servers. The c.d.f. of service times are assumed to be phase or hyperexponential type. For the both cases the necessary and sufficient stability condition is obtained.

Acknowledgement. Work is partially supported by Russian Foundation for Basic Research grant 17-01-00468.

IS 43
Scan Statistics and Applications
Rm 100/A
Wed
17:00 - 17:30

Multidimensional discrete scan statistics with arbitrary windows

Alexandru Amărioarei
Faculty of Mathematics and Computer Science, University of Bucharest, Bucharest, Romania
National Institute of R&D for Biological Sciences, Bucharest, Romania
email: alexandru.amarioarei@fmi.unibuc.ro

The object of this talk is to extend the classical definition of the multidimensional discrete scan statistic with rectangular window shape to a more general framework in which the scanning window is constructed based on a score function. In this new framework, problems like finding the distribution of a monotone run in a sequence of i.i.d. random variables or scanning with different window shapes (rectangle, circle, ellipse or annulus) in a two-dimensional setting will be discussed. We propose several approximations for the distribution of the scan statistic and illustrate their accuracy by conducting a numerical comparison study.

Acknowledgement. This research was supported by the Romanian project Biodivers PN 18-18-0102.

CS 2
Copulas and independence
Rm 100/B
Thu
12:40 - 13:00

A New Test for Multivariate Independence

Miklós Arató
Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest
email: arato@math.elte.hu

Independence between the components of a multivariate vector is a classical problem in statistics. Blum et al. [1961] investigated the application of a Cramér-von Mises statistic based on empirical distribution functions. Deheuvels [1981], Genest and Remillard [2004] and Genest et al. [2007] have shown that powerful rank tests of multivariate independence can be developed based on empirical copula processes.

The present study suggests a new method. The proposed test is based on rank vectors, random allocation and divergence measure.

The power of our test was compared with previous methods via Monte-Carlo experiments.

References

J. R. Blum, J. KieferR, and M. Rosenblatt. Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist*, 32485–498, 1961.

P. Deheuvels. An asymptotic decomposition for multivariate distribution-free tests of independence. *J. Multivariate Anal.*, 11:102–113, 1981.

C. Genest and B. Remillard. Tests of independence and randomness based on the empirical copula process. *Test*, 13335–369, 2004.
C. Genest, J-F. Quessy and B. Remillard. Asymptotic local efficiency of Cramér–von Mises tests for multivariate independence. *Ann. Statist.*, 35(1):166–191, 2007.

From univariate to matrix variate generalized t-distributions

IS 52
Distributio
Theory
Rm -1.64
Thu
11:30 - 12:00

Olcay Arslan[*,§], Y. Murat Bulut[†], Fatma Zehra Dogru[‡]

[*]Ankara University, Ankara, Turkey
[†]Eskişehir Osmangazi University, Eskişehir, Turkey
[‡]Giresun University, Giresun, Turkey
email: [§]oarslan@ankara.edu.tr

The univariate generalized t-distribution, which was introduced by McDonald and Newey [1988], has been widely used as a robust alternative to the normal distribution for modeling data sets that may have different tail behaviors than the normal distribution. Arslan [2004] introduced the multivariate generalized t-distribution as an extension of the univariate generalized t-distribution for modeling multivariate data sets. This distribution was defined as a scale mixture of a multivariate power exponential distribution and an inverse generalized gamma distribution with a scale parameter. The generalized t-distribution, which is a member of the elliptical distributions family, is very flexible and includes several well known distributions as special or limiting cases. Recent years, the multivariate generalized t-distribution has been attracted by the researchers who are seeking an appropriate multivariate distribution that may be robust alternative to the multivariate normal distribution for modelling multivariate data sets form several different application areas. Some of these application areas are the Volatility Modeling, independent vector analysis (blind source separation) and the joint sparsity regularization in multi-task learning.

Since, the multivariate generalized t-distribution has become popular among the researchers, in this study we would like to reconsider this distribution from several different aspects that have not been considered before. We will mainly focus on the following aspects of this distribution. The parameter estimation of the multivariate generalized t-distribution will be provided. A matrix variate generalization of this distribution will be defined and some distributional properties and parameter estimation of the newly proposed matrix variate distribution will be given. Further, finite mixtures of the generalized t-distributions will be explored.

References

Arslan,O. (2004) Family of multivariate generalized t-distributions, *Journal of Multivariate Analysis, 89, 329-337.*
McDonald, J.B., and Newey, W.K. (1988) Partially adaptive estimation of regression models via the generalized t-distribution, *Econometric Theory, 4, 428-457.*

IS 53
Engineering Systems
Rm 0.99
Tue
16:30 - 17:00

On the Preventive Maintenance of Coherent Systems under Stochastic Process of Component Failures

Majid Asadi[*,†,§], Somayeh Zarezadeh[‡,¶]

[*]Department of Statistics, University of Isfahan, Isfahan 81744, Iran
[†]School of Mathematics, Institute of Research in Fundamental Sciences (IPM), P.O Box 19395-5746, Tehran, Iran
[‡]Department of Statistics, Shiraz University, Shiraz 71454, Iran
email: [§]m.asadi@sci.ui.ac.ir [¶]s.zarezadeh@shirazu.ac.ir

The survival signature is an important notion in assessing the reliability properties of a coherent system consists of multiple types of components. In this study, we consider an n-component coherent system is built up from $2 \leq L \leq n$ different batches of components in which the ith batch has m_i components of the same type, $i = 1, 2, ..., L$, such that $\sum_{i=1}^{L} m_i = n$. We assume that the components failure in the ith batch occur according to a counting process $\{N_i(t), t \geq 0\}$, $i = 1, \ldots, L$, where $N_i(t)$ denotes the number of components that fail up to time t, $t > 0$. Under this model of components failure, we obtain the survival signature based reliability of the system lifetime. Then, we address reliability and stochastic properties of the system lifetime. We also discuss the optimal preventive maintenance time of the system lifetime under different scenarios.

POSTER
Mon
18:00-19:00

A Comparative Study of Stochastic Leaky Integrate-and-Fire models based on fractional noise and time-changes

Giacomo Ascione[*], Enrica Pirozzi

Università degli Studi di Napoli Federico II, Dipartimento di Matematica e Applicazioni
email: [*]ascione.giacomo0@gmail.com

Neural modelling started in 1907 with Lapique's model (Abbott [1999]) and since then it has rapidly become a very important field (Koch and Segev [1998]), which provided new questions to mathematics and physics. One of these questions is related to the introduction of memory in stochastic differential equations. In particular, the stochastic Leaky Integrate-and-Fire (LIF) model is shown to be not so efficient in the description of particular neuronal dynamics, such cortical neurons' dynamics (Shinomoto et al [1999]). This is due to the lack of memory of the solution Markov process and it can be addressed with a lot of different tools, such as the introduction of correlated stimuli (Sakai et al [1999]) or correlated noise (Pirozzi [2017], Ascione and Pirozzi [2018]) or also the use of a time-change. Here, we focus on the memory effects generated by using the fractional white noise instead of the classical white noise. Such noise generates long-range dependent processes in absence of stimuli if the self-similarity parameter H is bigger then $1/2$ (Cheridito et al [2003]) and can preserve such behaviour in presence of a certain kind of stimuli (Ascione and Pirozzi [2018]).

Such model is taken in comparison with a LIF model in which a time-change, by means of the inverse of a stable subordinator which is independent from the starting Markov process, is made. In the latter, such process loses its Markov property, gaining dependence from the sojourn time in the previous state (Gihman and Skorohod [1975]). This particular feature is the main reason one can use such time-changes to describe adaptation in neuronal modelling, due to the power-law asymptotic behaviour of the first passage time of time-changed processes (Ascione et al. [2017]). Moreover, these processes obey some time-fractional Fokker-Planck equations. With these ideas in mind, spik-

ing times and firing rates of the two different models with memory are compared, first by means of theoretical results and then by using numerical approximations and simulation algorithms.

References

Abbott, L. F. (1999) Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain research bulletin*, **50(5-6)**, p. 303-304.

Ascione, G., Pirozzi E. and Toaldo, B. (2017), On the exit time from open sets of some semi-Markov processes. arXiv: 1709.06333

Ascione, G. and Pirozzi E. (2018), On a fractional Ornstein-Uhlenbeck process with stochastic forcing and applications. *Submitted*

Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional ornstein-uhlenbeck processes. *Electronic Journal of probability*, **8**.

Gihman, I. I. and Skorohod, A. V. (1975) *The theory of stochastic processes. II.* Die Grundlehren der Mathematischen Wissenschaften, **218**.

Koch, C. and Segev, I. (Eds.) (1998) *Methods in neuronal modeling: from ions to networks*, MIT press.

Pirozzi, E. (2017) Colored noise and a stochastic fractional model for correlated inputs and adaptation in neuronal firing. *Biological cybernetics*, p. 1-15.

Shinomoto, S., Sakai, Y. and Funahashi, S. (1999) The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. *Neural Computation*, **11(4)**, p. 935-951.

Sakai, Y., Funahashi, S. and Shinomoto, S. (1999). Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. *Neural Networks*, **12(7-8)**, p. 1181-1190.

Eigenvectors of random regular graphs

OCS 1
Large Random Graphs
Rm -1.64
Mon
15:30 - 15:50

Ágnes Backhausz[*,†,‡], Balázs Szegedy[†]

[*]ELTE Eötvös Loránd University, Budapest
[†]Alfréd Rényi Institute of Mathematics
email: [‡]agnes@cs.elte.hu

We examine the eigenvectors of the adjacency matrix of a uniformly chosen random d-regular graph. More precisely, we are interested in the empirical distribution of the eigenvectors: the distribution of a randomly chosen entry. We show that the empirical distribution of any eigenvector is close to some (maybe degenerated) Gaussian random variable with high probability as the number of vertices tends to infinity. This is analogous to the results on the eigenvector distribution of random matrices. However, in our case the degree of each vertex is fixed d, while in the previously known cases the average degree of a vertex tends to infinity. This result and its connection to the theory of graph limits will be presented in the talk.

IS 24
Reliability Analysis
Rm -1.62
Wed
16:30 - 17:00

On stochastic comparison and ageing properties of multivariate proportional hazard rate mixtures

F.G. Badía[*,‡], Hyunju Lee[†,§]

[*]Zaragoza University, Zaragoza, Spain
[†]EWHA Womans University, Seoul, Korea
email: [‡]gbadia@unizar.es [§]hyunju@ewhain.net

In this work, we study the behavior of mixtures of the multivariate proportional hazard rate whose corresponding survival function is represented as

$$\bar{F}^*(\mathbf{x}) = \mathbb{E}[\bar{F}_0(\mathbf{x})^Z], \quad \mathbf{x} \in \mathbb{R}^n,$$

where $\bar{F}_0$ is a baseline survival function and Z is a frailty random variable. This model has been widely investigated by many authors in the univariate setting, whereas, there have been only a few recent studies on this model in the multivariate setting. According to the frailty random variable Z in the model, the random environmental conditions of the lifetime of n-components with multivariate baseline survival function $\bar{F}_0$ can be described. For example, (i) if $Z = 1$, then the lifetime of n-components has the ideal environment where the corresponding baseline survival function is given by $\bar{F}_0$; (ii) if $Z > 1 (< 1)$, then that has more(less) severe random environment than the ideal environment.

Recently, several results related to the mixture model have been focused on stochastic comparisons between mixtures of multivariate proportional hazard rate model with the same survival function and different frailty random variables. To best of our knowledge, stochastic comparisons of the mixture models with the same frailty random variable and different baseline survival functions has not been be considered yet in the literature. In this paper, we study stochastic comparisons of the multivariate proportional hazard rate mixture models with the same frailty random variable and different baseline survival functions. Furthermore, we investigate multivariate ageing and dependence properties under sufficient conditions with respect to the baseline survival function and the frailty random variable.

Finally, the results are applied to the stochastic comparison between epoch times of a mixed non homogeneous Poisson process which follows a multivariate proportional hazard rate mixture model. In the case when the frailty random variable follows the gamma distribution, the mixed non homogeneous Poisson process is a generalized Pólya process which has been studied recently.

Acknowledgement. This work has been supported by Spanish government research project MTM2015–63978 (MINECO-FEDER).

References

Cha, J. H.(2014) Characterization of the generalized Pólya process and its applications. *Advances in applied probability* **46**, 1148-1171.

Fernández Ponce, J. M., Pellerey, F. and Rodríguez–Griñolo, M. (2016) Some stochastic properties of conditionally dependent frailty models. *Statistics* **50**, 649-666.

Gupta, R. C., and Gupta, R. D. (2009) General frailty models and stochastic orderings. *Journal of stochastic planning and inference* **139**, 3277-3287.

Misra, N., Gupta, N., and Gupta, R. D. (2009) Stochastic comparisons of multivariate frailty models. *Journal of stochastic planning and inference* **139**, 2084-2090.

Diffusion Approximation of a growing Ehrenfest Urn

CS 4
Diffusions and Random Walks
Rm -1.62
Thu
12:00 - 12:20

Srinivasan Balaji

George Washington University, Washington DC, USA
email: balaji@gwu.edu

The Ehrenfest urn is a model for the mixing of gases in two chambers. Classic research deals with this system as a Markovian model with a fixed number of balls, and derives the steady-state behavior as a binomial distribution (which can be approximated by a normal distribution). In this process we obtain the limiting diffusion process of the discrete process which counts the number of white balls at any specific time. With $n \to \infty$ we obtain the Ornstein-Uhlenbeck type diffusion process as the limit and identify the mean and variance of the process. It is to be noted that in the earlier work Balaji, Mahmoud and Zhang(2010) obtained the expression for the mean and variance and using martingale central limit theorem obtained the limiting distribution under the three phases which are growing sublinear, the linear and the superlinear phases. Here it will correspond to the behavior of the process in small t, moderate t and large t respectively. The approach can be applied for other problems including the growing Ehrenfest urn with random picking and the Coupon problem.

Phase transitions in supercritical branching random walks

OCS 3
Branching Random Walks
Rm -1.62
Mon
17:00-17:20

Daria Balashova

Lomonosov Moscow State University, Moscow, Russian Federation
email: daria.balashova@abc.math.msu.su

We consider a continuous-time symmetric supercritical branching random walk on a multidimensional lattice with a finite set of particle generation centers, i.e. branching sources (Yarovaya, E. B. [2012, 2017]). The existence of a positive eigenvalue of the evolutionary operator means the exponential growth of the first moment of the total number of particles both at an arbitrary point and on the entire lattice.

We construct a model with three branching sources of arbitrary intensities and present a branching random walk with positive or negative intensities of sources that have a simplex configuration. It is established that the amount of positive eigenvalues of the evolutionary operator, counting their multiplicity, does not exceed the amount of the branching sources with positive intensity, while the maximal eigenvalue is simple.

For branching random walk with different positive intensities of sources and arbitrary configuration for both finite and infinite variance of jumps the critical values of sources' intensities are found which allows to prove the existence of positive eigenvalues of the evolutionary operator, .

Acknowledgement. The research is supported by RFBR Grant no. 17-01-00468.

References

Yarovaya, E. B. (2012) Spectral properties of evolutionary operators in branching random walk models. Mathematical Notes, 92(1): p. 115-131, doi: 10.1134/S0001434612070139.

Yarovaya, E. B. (2017) Positive Discrete Spectrum of the Evolutionary Operator of Supercritical Branching Walks with Heavy Tails. Methodology and Computing in Applied Probability, Springer, Volume 19, Issue 4, pp 1151–116, doi: 10.1007/s11009-016-9492-9.

Gikhman I. I., Skorokhod A. V. (2004) *The Theory of Stochastic Processes II*. Reprint of the 1975 edition. Classics in Mathematics. Springer-Verlag, Berlin.

Roger A. Horn, Charles R. Johnson (2013) *Matrix analysis*. Second edition. Cambridge University Press, Cambridge.

Gradshteyn I. S., Ryzhik I. M. (2000) *Tables of Integrals, Series, and Products*. Sixth edition. Academic Press, Inc., San Diego, CA.

IS 49
Urn Models
Rm -1.64
Wed
14:30 - 16:00

"Power of Two Choices" in Negatively Reinforced Pólya Urn

Antar Bandyopadhyay[*,‡], Gursharn Kaur[†,§]

[*]Indian Statistical Institute, Delhi Centre, 7 S. J. S. Sansanwal Marg, New Delhi 110016, INDIA
[†]Indian Statistical Institute, Delhi Centre, 7 S. J. S. Sansanwal Marg, New Delhi 110016, INDIA
email: [‡]antar@isid.ac.in [§]gursharn.kaur24@gamil.com

We will consider an implementation strategy for *weighted negative reinforcement* in Pólya Urn scheme. Negatively reinforced urn schemes were first introduced by Bandyopadhyay and Kaur (2017) and Kaur (2018). Let $U_n := (U_{n,i})_{1\le i\le K}$ be the configuration of the urn at time n with a total of K colors. Starting with U_0, a non-empty urn, at every time step we will consider selecting $1 \le d \le K$ colors *with* or *without replacements* and reinforcing a selected color i with probability proportional to $w\left(\frac{U_{n,i}}{\sum_j U_{n,j}}\right)$, where the sum in the denominator inside the function w is over all the selected colors. We will assume that $w : [0,1] \longrightarrow [0,\infty)$ is a decreasing function, indicating the *negative reinforcement*, in the sense that the least proportion color is most likely to be reinforced. For $d = 1$ the model is trivial and is the random reinforcement model. For $d = K$ the model was first studied by Bandyopadhyay and Kaur (2017) for linear but deceasing w, and later by Kaur (2018) for general w. They show that for such models, the almost sure convergence to uniform vector, as well as, central limit theorem hold under mild regularity condition on w.

In this talk, We will show that for any $2 \le d \le K$ the almost sure convergence to uniform vector holds and will also discuss the asymptotic of the fluctuations around the limit under mild regularity conditions on w, similar to what was assumed in Kaur (2018). We will further show the so called "*power of two choices*" phenomenon holds here, in the sense that $d = 2$ achieves the optimal asymptotic efficiency.

IS 38
Appiled Prob. & Stat. Inference III.
Rm 100/A
Tue
15:00 - 15:30

Sequential Method for Minimizing Expected Compression Loss and Sampling Cost for Obtaining Principal Components

Swarnali Banerjee

Loyola University Chicago
email: sbanerjee@luc.edu

Principal Component Analysis (PCA) aims in reducing the dimension of data instances into its subspace using orthogonal transformations, retaining a subset of linearly uncorrelated variables or principal components. These cumulatively capture some pre-specified proportion of total variability. Thus, using PCA for dimension reduction results in a loss of information which may be quantified by the quadratic compression loss. Warmuth and Kuzmin (2008) considered the quadratic compression loss in order to develop an algorithm for online PCA with bounded regret. In order to minimize the expected quadratic compression loss, one may consider a large number of observations. However, this results in a higher sampling cost. Here, we aim to minimize both the total expected quadratic

compression loss and the sampling cost using a sequential procedure. The characteristics and performance of the procedure will be discussed and the method would be applied on a real dataset.

Acknowledgement. This is a joint with Dr. Bhargab Chattopadhyay, Indian institute of Information Technology Vadodara, India

High-Dimensional Bayesian Geostatistics

IS 46
High-dim. Bayesian Inference
Rm 0.99
Wed
14:30 - 15:00

Sudipto Banerjee

University of California Los Angeles
email: sudipto@ucla.edu

With the growing capabilities of Geographic Information Systems (GIS) and user-friendly software, statisticians today routinely encounter geographically referenced data containing observations from a large number of spatial locations and time points. Over the last decade, hierarchical spatiotemporal process models have become widely deployed statistical tools for researchers to better understand the complex nature of spatial and temporal variability. However, fitting hierarchical spatiotemporal models often involves expensive matrix computations with complexity increasing in cubic order for the number of spatial locations and temporal points. This renders such models unfeasible for large data sets. I will present a focused review of two methods for constructing well-defined highly scalable spatiotemporal stochastic processes. Both these processes can be used as "priors" for spatiotemporal random fields. The first approach constructs a low-rank process operating on a lower-dimensional subspace. The second approach constructs a Nearest-Neighbor Gaussian Process (NNGP) that ensures sparse precision matrices for its finite realizations. Both processes can be exploited as a scalable prior embedded within a rich hierarchical modeling framework to deliver full Bayesian inference. These approaches can be described as model-based solutions for big spatiotemporal datasets. The models ensure that the algorithmic complexity has n floating point operations (flops), where n is the number of spatial locations (per iteration). We compare these methods and provide some insight into their methodological underpinnings.

Acknowledgement. This work was sponsored by NSF DMS-1513654, NSF IIS-1562303 and NIH/NIEHS R01-ES027027.

Tracy-Widom limit for Spearman rho and Kendall tau

IS 54
Random Matrix Theory
Rm -1.64
Tue
14:30 - 15:00

Zhigang Bao

Hong Kong University of Science and Technology
email: mazgbao@ust.hk

We study two high-dimensional random matrix models from nonparametric statistics. They are Spearman rank correlation matrix and Kendall rank correlation matrix, which are natural multivariate extensions of Spearman rank correlation and Kendall rank correlation coefficient, respectively. When the dimension is proportional to the sample size, we establish the Tracy-Widom law for the largest eigenvalues of both matrix models.

IS 35
Environmental Models
Rm -1.63
Thu
11:00 - 11:30

Similarity-based semilocal estimation of post-processing models

Sebastian Lerch[*], Sándor Baran[†,‡]
[*]Heidelberg Institute for Theoretical Studies and Karlsruhe Institute of Technology, Germany
[†]University of Debrecen, Hungary
email: [†]baran.sandor@inf.unideb.hu

Weather forecasts are typically given in the form of forecast ensembles obtained from multiple runs of numerical weather prediction models with varying initial conditions and physics parameterizations. Such ensemble predictions tend to be biased and underdispersive and thus require statistical post-processing. In the ensemble model output statistics approach [Thorarinsdottir and Gneiting 2010], a probabilistic forecast is given by a single parametric distribution with parameters depending on the ensemble members. We propose two semilocal methods for estimating the model parameters [Lerch and Baran 2017] where the training data for a specific observation station are augmented with corresponding forecast cases from stations with similar characteristics. Similarities between stations are determined using either distance functions or clustering based on various features of the climatology, forecast errors, ensemble predictions and locations of the observation stations. In a case study on wind speed over Europe the proposed similarity-based semi-local models show significant improvement in predictive performance compared to standard regional and local estimation methods. They further allow for estimating complex models without numerical stability issues and are computationally more efficient than local parameter estimation.

Acknowledgement. Sándor Baran was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the EFOP-3.6.1-16-2016-00022 project. The project is co-financed by the European Union and the European Social Fund.

References

Thorarinsdottir, T. L. and Gneiting, T. (2010) Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. *J. Roy. Statist. Soc. Ser. A* **173**, 371–388.

Lerch, S., Baran, S. (2017) Similarity-based semi-local estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Stat.* **66**, 29–51.

OCS 7
Association and Default
Harmony Inge
Tue
17:10-17:30

Pricing Synthetic CDOs with Stochastic Correlation

Eszter Baranyi[*,†], László Márkus[*]
[*]Eötvös Loránd University, Budapest, Hungary
email: [†]WJArox@hotmail.com

The financial crisis of 2007-08 revealed that the assumptions of the earlier widely accepted standard market model of CDO pricing do not reflect reality properly. Various attempts has been made to improve pricing, one of which is the use of stochastic processes in the modeling of the correlation structure of the underlying names. This research investigates three stochastic processes as correlation structures: the Jacobi process and two transformations of the Ornstein-Uhlenbeck process by the arctangent and the hyperbolic tangent functions. In each case, the CDS spreads of the underlying names are modeled by a geometric Brownian motion, where the Wiener processes are correlated according to the stochastic correlation process. Defaults are defined as the exceedance of a given limit for the CDS spreads, for which different choices are inspected. All stochastic processes are calibrated to historical market data and the CDO tranches are priced using Monte Carlo simulation. The analysis

focuses on the iTraxx Europe Main index as the underlying basket and the traded CDO tranches for the inspected dates.

Properties and inferential issues of a bivariate version of the geometric distribution

CS 19
Distributio
Rm 100/B
Tue
17:10-17:30

Alessandro Barbiero

Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Milan, Italy
email: alessandro.barbiero@unimi.it

In recent years, the construction of bivariate (and multivariate) discrete distributions has attracted much interest, since stochastic models for correlated count data find application in many fields. Several authors have discussed the problem of constructing a bivariate version of a given univariate distribution, although there is no universally accepted criterion for producing a unique distribution which can unequivocally be called the bivariate analogue of a univariate distribution. In this paper, we revise a bivariate geometric model, introduced by Roy [1993], which is characterized by locally constant bivariate failure rates. We highlight its close relationship with Gumbel's bivariate exponential distribution [Gumbel 1960] and then we focus on four aspects of this model that have not been investigated so far: 1) Pearson's correlation and its range, 2) conditional distributions and pseudo-random simulation, 3) parameter estimation, and 4) stress-strength reliability parameter. A Monte Carlo simulation study is carried out in order to assess the performance of the different estimators proposed; an application to real data, along with a comparison with alternative bivariate discrete models, is provided as well.

References

Gumbel, E.J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.*, **55(292)**, 698-707.

Roy, D. (1993) Reliability measures in the discrete bivariate set-up and related characterization results for a bivariate geometric distribution. *J. Multivariate Anal.*, **46(2)**, 362-373.

Almost sure and L_1-growth behavior of supercritical multi-type continuous state and continuous time branching processes with immigration

CS 5
Branching Processes
Rm -1.62
Tue
16:30-16:50

Mátyás Barczy[*,§], Sandra Palau[†], Gyula Pap[‡]

[*]MTA-SZTE Analysis and Stochastics Research Group, Bolyai Institute, University of Szeged, Hungary
[†]Department of Mathematical Sciences, University of Bath, United Kingdom
[‡]Bolyai Institute, University of Szeged, Hungary
email: [§]barczy@math.u-szeged.hu

Under a first order moment condition on the immigration mechanism, we show that an appropriately scaled supercritical and irreducible multi-type continuous state and continuous time branching process with immigration (CBI process) converges almost surely. If an $x\log(x)$ moment condition on the branching mechanism does not hold, then the limit is zero. If this $x\log(x)$ moment condition holds, then we prove L_1-convergence as well. The projection of the limit on any left non-Perron eigenvector of the branching mean matrix is vanishing. If, in addition, a suitable extra power moment

condition on the branching mechanism holds, then we provide the correct scaling for the projection of a CBI process on certain left non-Perron eigenvectors of the branching mean matrix in order to have almost sure and L_1-limit. A representation of the limits is also provided under the same moment conditions. Our results generalize some recent results of Kyprianou, Palau and Ren [2017] on supercritical and irreducible multi-type CBI processes without immigration.

Acknowledgement. Mátyás Barczy is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Sandra Palau is supported by the Royal Society Newton International Fellowship and by the EU-funded Hungarian grant EFOP-3.6.1-16-2016-00008.

References

Kyprianou, A. E., Palau, S., and Ren, Y.-X. (2017) Almost sure growth of supercritical multi-type continuous state branching process. arXiv: 1707.04955

IS 39
Abraham Wald Prize Ceremony
Harmony Inge
Tue
11:00 - 11:30

Decision theoretic approach to sequential testing of multiple hypotheses

Michael Baron

American University, Washington DC, USA
email: baron@american.edu

The problem of multiple hypothesis testing often arises in sequential experiments such as sequential clinical trials with multiple endpoints, treatments, or population groups; multichannel change-point detection; acceptance sampling with different criteria of acceptance; etc. In such studies, it is necessary to make a statistical decision for each individual statement instead of combining them into one composite hypothesis.

We assume a sampling cost and a loss for each Type I and Type II error among the tested hypothesis and propose stopping rules and sequential decisions that are risk-optimal while controlling the Type I and Type II *familywise* error rates, or the familywise power, in the strong sense.

Acknowledgement. This research is supported by the U.S. National Science Foundation.

CS 6
Queuing Theory
Rm 100/A
Tue
17:30-17:50

QMCD approach for an obsolete (Q, r) inventory model with lead time and lost sales

Yonit Barron[*,‡], Opher Baron[†]

[*]Department of Industrial Engineering and Management, Ariel University, Ariel, Israel
[†]Rotman School of Management, University of Toronto, Toronto, Ontario, Canada
email: [‡]ybarron@ariel.ac.il

We consider an inventory system with state-dependent Poisson demands, general lead times, in the presence of sudden obsolescence and lost sales. The inventory control policy is the continuous-review (Q, r) policy, where $r < Q$. The system is modulated using an embedded Markov process at the replenishment points. Using the queueing and Markov chain decomposition (QMCD) approach, the distribution of the residual lead time is characterized and the stationary distribution of the inventory level is derived. We construct the expected total long-run average cost function and obtain, numerically, the optimal parameters. Our study considers an unit Poisson demand, and also generalized to batch demands. Numerical analysis, insights and conclusions are provided.

Short Time Near-the-Money Skew in Rough Fractional Volatility Models

IS 4
(Rough) Volatility Asymptotics
Rm 100/A
Tue
10:30 - 11:00

Christian Bayer[*,¶], Peter Friz[*], Archil Gulisashvili[‡], Blanka Horvath[§], Benjamin Stemper[*]

[*]Weierstrass Institute, Berlin, Germany
[†]TU Berlin, Germany
[‡]Ohio University, USA
[§]Imperial College, London, UK
email: [¶]christian.bayer@wias-berlin.de

We consider rough stochastic volatility models where the driving noise of volatility has fractional scaling, in the "rough" regime of Hurst parameter $H < 1/2$. This regime recently attracted a lot of attention both from the statistical and option pricing point of view. With focus on the latter, we sharpen the large deviation results of Forde and Zhang (2017) in a way that allows us to zoom-in around the money while maintaining full analytical tractability. More precisely, this amounts to proving higher order moderate deviation estimates, only recently introduced in the option pricing context. This in turn allows us to push the applicability range of known at-the-money skew approximation formulae from CLT type log-moneyness deviations of order $t^{1/2}$ (recent works of Alòs, León & Vives and Fukasawa) to the wider moderate deviations regime.

Acknowledgement. Support from DFG through grants BA5484/1 and FR2943/2 (C.B., P.F., B.S.), the European Research Council through grant CoG-683166 (P.F.) and the SNF through the Early Postdoc Mobility Grant 165248 (B.H.) is gratefully acknowledged.

Long-memory Gaussian processes governed by fractional Fokker-Planck equations

IS 16
Fractional Stochastic Models
Rm 100/B
Wed
15:00 - 15:30

L. Beghin

Sapienza University of Rome

It is well-known that the transition function of the Ornstein-Uhlenbeck process solves the Fokker–Planck equation. This standard setting has been recently generalized in different directions, for example, by considering the so-called α-stable driven Ornstein-Uhlenbeck, or by time-changing the original process with an inverse stable subordinator. In both cases, the corresponding partial differential equations involve fractional derivatives (of Riesz and Riemann–Liouville types, respectively) and the solution is not Gaussian. We consider here a new model, which cannot be expressed by a random time-change of the original process: we start by a Fokker-Planck equation with a time-derivative replaced by a different fractional differential operator. The resulting process is Gaussian and, in the stationary case, exhibits a long-range dependence.

IS 28
Stochastic Precedence
Rm -1.62
Wed
15:00 - 15:30

A new preference type order for the stochastic dominance of dependent random variables

Félix Belzunce*, Carolina Martínez-Riquelme†
Universidad de Murcia, Facultad de Matemáticas (Campus de Espinardo), Murcia (Spain)
email: *belzunce@um.es †carolina.martinez7@um.es

The usual stochastic order is a tool to compare the magnitude of two random variables. However, this criteria only takes into account the marginal distributions of the two random variables, and does not take into account their possible dependence. An alternative, in the dependent case, is the precedence order, but this criteria is not very informative, given that it reduces all the information of the bivariate random vector in just two numbers. In this talk, we present from an applied point of view, a new criteria of stochastic dominance that takes into account the dependence structure of the two random variables involved in the comparison. Relationships with some existing criteria, closure properties and applications are also given.

Acknowledgement. The authors want to acknowledge the support received by the Ministerio de Economía, Industria y Competitividad under grant MTM2016-79942-P (AEI/FEDER, UE).

IS 20
Stochastic Geometry
Rm -1.63
Wed
15:30 - 16:00

Modeling of 3D random tessellations with interactions of cell characteristics

Viktor Benes
Charles University, Faculty of Mathematics and Physics, Prague
email: benesv@karlin.mff.cuni.cz

The paper Dereudre and Lavancier [2011] presents simulation methods and estimation for Gibbs-Voronoi and Delaunay tessellations in the plane. Our aim is to generalize this paper in some ways. First instead of the plane we are in the three-dimensional Euclidean space, with a hope to apply later the tessellations to the fitting of polycrystalline materials microstructures. Secondly besides the Voronoi tessellation also another model is used, namely the power tessellation. The interactions of cell characteristics are stated in the energy function, typically the pair potential is defined based on the volumes of neighbouring cells. We use also triplet potentials based on characteristics of cells sharing a common edge. The hard existence and uniqueness problems are not solved. Our experience is based on computationally demanding simulations of parametric models with varying parameters. They enable also to study estimation properties for small samples. This is a joint work with Filip Seitl.

References

D. Dereudre and F. Lavancier. Practical simulation and estimation for Gibbs Delaunay-Voronoi tessellations with geometric hardcore interaction. *Comput. Statist. Data Anal.*, 55(1):498–519, 2011. doi: 10.1016/j.csda.2010.05.018.

A New Procedure for Public Health Monitoring

IS 51
Run and scan models
Rm 0.87
Wed
15:30 - 16:00

Sotirios Bersimis[*,‡], Athanasios Sachlas[*], Polychronis Economou[†]

[*]University of Piraeus, Piraeus, Greece
[†]University of Patras, Rio Achaia, Greece
email: [‡]sbersim@unipi.gr

In order to timely and efficiently detect disease outbreaks we should take into account both spatial and temporal dimensions. Of interest are global changes in the number of new disease events on time and/or hotspots of disease events which may evolve into outbreaks. A key assumption in biosurveilance is that under normal conditions events are uniformly distributed in the plane and there are not too many changes over time. In this work, we propose a new two-step monitoring procedure with which we monitor the number of disease events through statistical process monitoring techniques and the spatial distribution of disease events via scan statistics.

References

Anderson, N.H. and Titterington, D.M. (1997) Some Methods for Investigating Spatial Clustering, with Epidemiological Applications. *Journal of Royal Statistical Society, Series A*, **160**, p.87–105.

Bersimis, S. and Economou, P. (2017) The use of length-biased distributions in statistical monitoring. *Australian & New Zealand Journal of Statistics*, **59(2)**, p.155-167.

Bersimis, S., Sachlas, A. and Sparks, R. (2017) Performance Monitoring and Competence Assessment in Health Services. *Methodology and Computing in Applied Probability*, **19(4)**, p.1169-1190.

Bersimis, S., Koutras, M.V. and Papadopoulos, G. (2014) Waiting Time for an Almost Perfect Run and Applications in Statistical Process Control. *Methodology and Computing in Applied Probability*, **16(1)**, p.207-222.

Fricker, R.D., Jr. and Chang, J.T. (2008) A Spatio-temporal Methodology for Real-time Biosurveillance. *Quality Engineering*, **20**, p.465–477.

Kulldorff, M. (1997) A Spatial Scan Statistic. *Communications in Statistics - Theory and Methods*, **26(6)**, p.1481–1496.

Woodall, W.H., Brooke Marshall, J., Joner Jr, M.D. and Fraker, S.E. and Abdel-Salam, G. (2008) On the use and evaluation of prospective scan methods for health-related surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **171(1)**, p.223-237.

A probabilistic approach to spectral analysis of growth-fragmentation equations

PLENARY
Harmony Inge
Thu
14:30 - 15:30

Jean Bertoin

Institute of Mathematics, Zurich
email: jean.bertoin@math.uzh.ch

The growth-fragmentation equation describes a system of growing and dividing particles, and arises in models of cell division, protein polymerisation and even telecommunications protocols. Several important questions about the equation concern the asymptotic behaviour of solutions at large times: at what rate do they converge to zero or infinity, and what does the asymptotic profile of the solutions look like? Does the rescaled solution converge to its asymptotic profile at an exponential speed? These questions have traditionally been studied using analytic techniques such as entropy methods or splitting of operators. In this talk, we present a probabilistic approach to the study of this asymptotic behaviour. We use a Feynman–Kac formula to relate the solution of the growth-fragmentation equation to the semigroup of a Markov process, and characterise the rate of

decay or growth in terms of this process. We then identify the Malthus exponent and the asymptotic profile in terms of a related Markov process, and give a spectral interpretation in terms of the growth-fragmentation operator and its dual.

Acknowledgement. Based on a joint work with Alex Watson, Manchester University

IS 38
Appiled Prob. & Stat. Inference III.
Rm 100/A
Tue
15:30 - 16:00

Sequentially Estimating the Required Optimal Observed Number of Tagged Items with Bounded Risk in the Recapture Phase Under Inverse Binomial Sampling

Nitis Mukhopadhyay[*,‡], Debanjan Bhattacharjee[†]

[*]University of Connecticut, Storrs, USA
[†]Utah Valley University, Orem, USA
email: [‡]nitis.mukhopadhyay@uconn.edu

Estimation of a closed population size (N) under inverse binomial sampling consists of four basic steps: First, one captures t items, then tag these t items, followed by releasing the t tagged items back to the population. Then, one draws an item from the population one-by-one until s tagged items are recaptured where s is fixed in advance. In the recapturing stage (fourth step), items are normally drawn with replacement. But, without replacement sampling will not impact much if N is large. Under squared error loss (SEL) as well as weighted SEL, we propose sequential methodologies to come up with bounded risk point estimators of an optimal choice of s; leading to an appropriate sequential estimator of N. The sequential estimation methodologies are supplemented with appropriate first-order asymptotic properties which are followed by extensive data analyses.

References

Anscombe, F. J. (1952) Large-Sample Theory of Sequential Estimation. *Proceedings of Cambridge Philosophical Society*, **48**, p. 600-607.

Chow, Y. S. and Robbins, H. (1965) On the Asymptotic Theory of Fixed Width Sequential Confidence Intervals for the Mean. *Annals of Mathematical Statistics* **36**, p. 457-462.

Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997) *Sequential Estimation*, New York: Wiley.

Mukhopadhyay, N. and Solanky, T. K. S. (1994) Multistage Selection and Ranking Procedures, New York: Dekker.

Sen, P. K. and Ghosh, M. (1981) Sequential Point Estimation of Estimable Parameters Based on U-Statistics. *Sankhya, Series A* **43**, p. 331-344.

IS 3
Optimal Transport ...
Harmony Inge
Wed
14:30 - 15:00

Reduced-form framework under model uncertainty

Francesca Biagini[*], Yinglin Zhang

Department of Mathematics, LMU Munich, Theresienstraße, 39, 80333 Munich, Germany
Department of Mathematics, University of Oslo, Box 1053, Blindern, 0316, Oslo, Norway
email: [*]biagini@math.lmu.de

In this talk we introduce a sublinear conditional expectation with respect to a family of possibly nondominated probability measures on a *progressively enlarged* filtration. In this way, we extend the classic reduced-form setting for credit and insurance markets to the case under model uncertainty, when we consider a family of priors possibly mutually singular to each other. These results close the gap between robust framework for financial market, which is recently studied in an intensive way,

and the one for credit and insurance markets, which is limited in the present literature only to some very specific cases.

Comparison of PageRank of Perturbed Markov chains of First and Second Order

OCS 4
Markov & Semi-Markov Models
Rm -1.62
Tue
12:20 - 12:40

Pitos Seleka Biganda[*,†,§], Benard Abola[†], Christopher Engström[†], John Mango Magero[‡], Godwin Kakuba[‡], Sergei Silvestrov[†]

[*]University of Dar es Salaam, Dar es Salaam, Tanzania
[†]Mälardalen University, Västerås, Sweden
[‡]Makerere University, Kampala, Uganda
email: [§]pitos.biganda@mdh.se

PageRank is a widely-used hyperlink-based algorithm to estimate the relative importance of nodes in networks [3]. Since many real world networks are large sparse networks, this makes efficient calculation of PageRank complicated. Moreover, one needs to escape from dangling effects in some cases as well as slow convergence of the transition matrix P. Primitivity adjustment with a damping parameter $\epsilon \in (0, \epsilon_0]$ (for fixed $\epsilon_0 \simeq 0.15$) is one of the essential procedure that is known to ensure convergence of P [2]. If ϵ is large, P looses information because there will be shift of information to teleportation matrix [1]. In this talk, we aim to formulate PageRank problem as a first and second order Markov chains perturbation problem. The convergence of the two set-ups will be compared in terms of number of iterations k and ϵ on different graph structures using numerical experiments.

References

[1] Silvestrov, D. and Silvestrov, S. (2017) *Nonlinearly Perturbed Semi-Markov Processes*. Springer.
[2] Langville, A. N. and Meyer, C. D. (2011) *Google's PageRank and beyond: The science of search engine rankings*. Princeton University Press.
[3] Brin, S. and Page, L. (1998) The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7), 107-117.

A general central limit theorem and subsampling variance estimator for α-mixing point processes.

IS 19
Spatial Point Processes
Rm -1.63
Wed
10:30 - 11:00

Christophe A. N. Biscio[*,†], Rasmus Waagepetersen[*]

Aalborg University, Denmark[*]
email: [†]christophe@math.aau.dk

Central limit theorems for multivariate summary statistics of α-mixing spatial point processes have usually been established using either the so-called Bernstein's blocking technique or an approach based on Bolthausen [1982]. It is characteristic that essentially the same theorems have been (re)-invented again and again for different specific settings and statistic considered: Moreover, although there exists estimates in some particular cases, the asymptotic variance is usually unknown or difficult to compute.

In this talk, we present a unified framework based on Bolthausen [1982] to state, once and for all, a general central limit theorem for α-mixing point process that applies to general non-stationary settings and is also applicable to non-parametric kernel estimators depending on a band width converging to

zero. In particular, we argue why this approach is more suitable that the one using Bernstein's blocking technique. We believe this can save a lot of work and tedious repetitions in future applications of α-mixing point processes.

Finally, we present a subsampling estimator of the asymptotic variance in central limit theorems. Our estimator is very flexible and model free. We illustrate its use in connection to confidence interval of estimators obtained by composite likelihood method for several non stationary point processes that may be regular or clustered.

Acknowledgement. Christophe A.N. Biscio and Rasmus Waagepetersen are supported by The Danish Council for Independent Research | Natural Sciences, grant DFF – 7014-00074 "Statistics for point processes in space and beyond", and by the "Centre for Stochastic Geometry and Advanced Bioimaging", funded by grant 8721 from the Villum Foundation.

References

Erwin Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10(4):1047–1050, 1982.

IS 12
Efficient Monte Carlo Simulation
Rm 100/B
Thu
10:30 - 11:00

On the Total Variation Convergence of Monte Carlo Splitting

Zdravko Botev
UNSW Sydney, Australia

It is well-known that Monte Carlo splitting can be used to simulate approximately from difficult high-dimensional distributions, similar to MCMC sampling. In this talk we provide some theoretical analysis of the convergence of the output states of the splitting algorithm. In particular, we derive the rate of convergence of the total variation discrepancy and find a nontrivial upper bound that depends on quantities estimable from the simulation output. We argue that, in contrast to traditional MCMC convergence analysis, which only provides a qualitative convergence assessment, splitting provides us with a quantitative convergence assessment. For example, splitting makes it possible to estimate the length of the "burn-in" period necessary to achieve a prespecified total variation discrepancy.

CS 10
Change Point Detection
Rm 100/A
Thu
12:00 - 12:20

A Change-Point Model for Detecting Heterogeneity in Ordered Survival Responses

Olivier Bouaziz[*,‡], Grégory Nuel[†,§]
[*]MAP5 (CNRS 8145), Université Paris Descartes, Sorbonne Paris Cité
[†]LPSM (CNRS 8001), Sorbonne Université, Paris.
email: [‡]olivier.bouaziz@parisdescartes.fr [§]gregory.nuel@math.cnrs.fr

In survival analysis it is quite common that heterogeneity between patients results in various survival response distributions. This heterogeneity can be controlled through known covariates (such as date of birth, age at diagnosis, gender, treatment, co-exposure, BMI, etc.) using regression-type models such as the Cox proportional hazard model and by performing stratified analyses or by incorporating a random effect in a frailty model. Other types of heterogeneous dataset arise when the incidence rate changes over the calendar time in a cohort study and specific models like age-period-cohort have been extensively studied to take into account this kind of heterogeneity. While theses

models have proved to be most useful, it is however likely that unaccounted latent heterogeneity remains in the survival signal. This might be due for example to an unknown interaction between a treatment and some exposure, or to some unaccounted heterogeneity of the disease itself (for example an unknown cancer sub-type). For instance, age at diagnosis might be associated with a higher chance to receive a new treatment or BMI might be associated with a specific exposure.

In this talk, we suggest a new approach considering survival heterogeneity as a breakpoint model in an ordered sequence of survival responses. The survival responses might be ordered according to any numerical covariate (ties are possible) like age at diagnosis, BMI, etc. The basic idea being that heterogeneity will be detected as soon as it is associated with the chosen covariate. In such a model, we aim at two objectives: first we want to estimate the hazard rates and the proportional factors in each homogenous region through a Cox model. Secondly, we want to accurately provide the number and location of the breakpoints. A constrained Hidden Markov Model (HMM) method was suggested in the context of breakpoint analysis, see Luong et al. [2013]. This method allows to perform a full change-point analysis in a segment-based model (one parameter by segment) providing linear EM estimates of the parameter and a full specification of the posterior distribution of change points. In this talk we adapt this method to the context of survival analysis, where the estimation is performed through the EM algorithm to provide update of the hazard rate estimates and the posterior distribution at each iteration step. The method was published in Bouaziz and Nuel [2017]

The method will be illustrated on the dataset on diabetic patients from the Steno Memorial hospital in Copenhagen which can be found in Andersen et al. [1993], where the event times are ordered with respect to the calendar time of disease onset. On this dataset, the years of disease onset of the patients range from 1933 to 1972. A two breakpoint model is found from our method and survival functions and hazard ratios are estimated on each three segment. Our results clearly indicate a general medical improvement over time for Danish diabetic patients.

References

P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding. *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993.

Olivier Bouaziz and Grégory Nuel. A change-point model for detecting heterogeneity in ordered survival responses. *Statistical methods in medical research*, 2017.

T. M. Luong, Y. Rozenholc, and G. Nuel. Fast estimation of posterior probabilities in change-point analysis through a constrained hidden markov model. *Computational Statistics and Data Analysis*, 68:129–140, 2013.

Copula-link functions for regression models

Taoufik Bouezmarni

Université de Sherbrooke, Sherbrooke, Canada
email: Taoufik.Bouezmarni@USherbrooke.ca

CS 2
Copulas and independence
Rm 100/B
Thu
12:00 - 12:20

The paper proposes a new link functions, based on copulas, for regression models that joint a binary or multinomial response with a set of covariates. The idea consists in writing the conditional probabilities in terms of conditional copula and marginal distributions. The new links generalize the logit, probit and robit models and allow more flexible models. For estimating the proposed model, we consider a parametric family for the copula and we use either a parametric or a nonparametric estimators for the marginal distributions. The asymptotic properties of these estimators are established and a finite sample simulations showing the performance of the proposed procedure, compared to logit and probit link functions, are presented. Finally, we apply the new method for analyzing the burn injury data

Acknowledgement. Natural Sciences and Engineering Research Council of Canada

IS 53
Engineering Systems
Rm 0.99
Tue
17:30 - 18:00

Weighted k-out-of-n system with three-state components

Ali Riza Bozbulut*, Serkan Eryilmaz

Atilim University, Ankara
email: *ali.bozbulut@atilim.edu.tr

Consider a system consisting of n components. Each component can be in one of three possible states, e.g. perfect functioning, partial working and complete failure at any time. The weight/contribution of a component when it is in perfect functioning state is greater than the weight/contribution when it is in partial working state. The system is assumed to be in a working state when the total weight of components is at least k. This study is concerned with reliability analysis of such a system when the time spent by components in a perfect functioning state and the time spent in a partially working state are statistically dependent.

CS 5
Branching Processes
Rm -1.62
Tue
16:50-17:10

An algorithmic approach to the extinction of branching processes with countably many types

Peter Braunsteins*,§, Geoffrey Decrouez†,¶, Sophie Hautphenne*,‡,‖

*The University of Melbourne, Melbourne
†National Research University, Higher School of Economics, Moscow
‡EPFL, Lausanne
email: §p.braunsteins@student.unimelb.edu.au ¶ggdecrouez@hse.ru
‖sophiemh@unimelb.edu.au

We consider the extinction events of Galton-Watson processes with countably infinitely many types. In particular, we construct truncated and augmented Galton-Watson processes with finite but increasing sets of types. A pathwise approach is then used to show that, under some sufficient conditions, the corresponding sequence of extinction probability vectors converges to the global extinction probability vector of the Galton-Watson process with countably infinitely many types. Besides giving rise to a family of new iterative methods for computing the global extinction probability vector, our approach paves the way to new global extinction criteria for branching processes with countably infinitely many types.

IS 37
Applied Prob. & Stat. Inference II.
Rm 100/A
Mon
16:30 - 17:00

Long memory, fractional integration, and regime switches in dependent time series

Boris Brodsky

CEMI RAS, Moscow
email: bbrodsky@yandex.ru

In this report the problem of long memory, fractional integration, and regime switchings for dependent time series is considered. We consider both classic retrospective case pionered by Hurst in 1950s and a new sequential problem statement. The following hypotheses are tested: H_0 - stationatity of data, H_1 - regime switchings, and H_2 - fractional integration of data. It is well known that

the presence of structursl changes in data seriously hinders testing of H_0 and H_2 and discrimination between them. For detection of structural changes and fractional integration, a nonparametric test is proposed based upon the Kolmogorov statistic. We demonstrate that the probabilities of wrong deisions about the type of non-stationarity tend to zero both for retropective and sequentisl case under ψ -weak dependence conditions. For sequential case, the normalized delay times under H_1 and H_2 are considered, as well as discrimination between a structural change and fractional integration hypotheses. Results of a simulation study of the proposed method for different types of nonstationarities are presented.

Taylor's Law via Ratios, for Some Distributions with Infinite Mean

IS 55
Applied Probability
Rm 0.99
Wed
12:00 - 12:30

Mark Brown

Department of Statistics, Columbia University
email: mb2484@columbia.edu

Taylor's law (TL) originated as an empirical pattern in ecology. In many sets of samples of population density, the variance of each sample was approximately proportional to a power of the mean of that sample. In a family of nonnegative random variables, TL asserts that the population variance is proportional to a power of the population mean. TL, sometimes called fluctuation scaling, holds widely in physics, ecology, finance, demography, epidemiology, and other sciences, and characterizes many classical probability distributions and stochastic processes such as branching processes and birth-and-death processes. We demonstrate analytically for the first time that a version of TL holds for a class of distributions with infinite mean. These distributions and the associated TL differ qualitatively from those of light-tailed distributions. Our results employ and contribute to methodology of Albrecher and Teugels (2006) and Albrecher, Ladoucette and Teugels (2010). This work opens a new domain of investigation for generalizations of TL.

This work is joint with Professors Joel Cohen and Victor de la Pena.

Limit behavior of some applied probability models

IS 32
Limits in Inventory...
Rm -1.64
Tue
17:00 - 17:30

Ekaterina Bulinskaya

Lomonosov Moscow State University, Moscow
email: ebulinsk@yandex.ru

In order to investigate some real process one needs an appropriate mathematical model. The same model can arise in different applications, e.g., in insurance, inventory and many others. Moreover, for the process under consideration there exist a lot of models describing it more or less precisely. The main aim of presentation is the study of limit behavior of the processes describing the capital of insurance company or the inventory level of a supply system. The results are used for optimization of systems' performance in the framework of reliability and cost approaches. To establish the stability of optimal control to small fluctuations of system parameters we use the local and global methods of sensitivity analysis such as Sobol's decomposition, FAST or Monte Carlo filtering. For evaluation of underlying processes perturbations we employ various probability metrics, e.g., Kolmogorov uniform metric and Kantorovich-Wasserstein one. Several continuous-time and discrete-time models are studied. Thus, we establish the asymptotically optimal inventory control for a periodic-review system for the case of unreliable supplier under incomplete information. A new dividend policy is proposed for the continuous-time dual systems with reinsurance and investment.

Acknowledgement. The research was partially supported by the Russian Foundation for Basic Research grant 17-01-00468.

IS 47
Entropy Estimates & Appl.
Rm -1.62
Mon
11:00 - 11:30

Statistical estimation of the entropy and applications

Alexander Bulinski
Lomonosov Moscow State University, Moscow, Russia
email: bulinski@yandex.ru

The concept of entropy belongs to the principle ones in Physics and Mathematics. We discuss various definitions of entropy and consider different approaches to its statistical estimation. The asymptotic behavior of the Kozachenko - Leonenko estimates for the (differential) Shannon entropy is studied when the number of i.i.d. vector-valued observations tends to infinity. The asymptotic unbiasedness and L^2-consistency of the estimates are established in Bulinski, A. and Dimitrov, D. [2018]. The conditions employed involve the analogues of the Hardy - Littlewood maximal function. It is shown that the results are valid in particular for the entropy estimation of any nondegenerate Gaussian vector. The new estimates (see Bulinski, A. and Kozhevin, A. [2018]) of the conditional Shannon entropy are introduced in the framework of the model describing a discrete response variable depending on a vector of d factors having a density w.r.t. the Lebesgue measure in $\mathbb{R}^d$. Such models include for instance the famous logistic regression. The limit behavior of the proposed estimates is considered. The obtained results are applied to the feature selection problem which is important, e.g., for medical and biological investigations. Namely, we demonstrate that our estimates can be used for identification of the relevant factors (features) having an impact on the response variable under consideration. The applications to detection of the material inhomogeneities are provided as well.

Acknowledgement. The work is supported by the Russian Science Foundation under grant 14-21-00162 and performed at the Steklov Mathematical Institute of Russian Academy of Sciences.

References

Bulinski, A. and Dimitrov, D. (2018) Statistical estimation of the Shannon entropy. arXiv: 1801.02050
Bulinski, A. and Kozhevin, A. (2018) Statistical estimation of the condition entropy. (to appear).

IS 53
Engineering Systems
Rm 0.99
Tue
18:00 - 18:30

On distorted representations for inactivity time of systems under double monitoring

Jorge Navarro[*], Camilla Calì[†,‡]
[*]Universidad de Murcia, Murcia, Spain
[†]Universita' di Napoli Federico II, Napoli, Italy
email: [‡]camilla.cali@unina.it

The analysis of coherent systems is one of the most relevant topics in reliability theory. In real life situation the monitoring of a system can be scheduled at different times. Under these periodical inspections, the information about the system can be different and can be affected by the condition of the components of the system at the inspection points. We tipically may know that the system was working at a time t_1 but that is broken at another time $t_2 > t_1$. Here we are interested in the inactivity time of the system, $(t_2 - T|t_1 < T < t_2)$. We obtain representations for the reliability function of such inactivity time through distortion functions. Similar representations are obtained under other

assumptions with partial information about component failures at times t_1 and t_2. We consider general coherent systems studying the inactivity time of a coherent system formed by components with possibly dependent components. The representations obtained are used to compare stochastically the inactivity times under different assumptions.

Acknowledgement. C. Calì is a member of the INdAM Research group GNAMPA.

References

Navarro, J. and Durante, F. (2017) Copula-based representations for the reliability of the residual lifetimes of coherent systems with dependent components. *Journal of Multivariate Analysis*, **158**, p.87–102.

Navarro, J., Pellerey, P. and Longobardi, M. (2017) Comparison results for inactivity times of k-out-of-n and general coherent systems with dependent components. *Test*, **26**, p.822–846.

Poursaeed, M.H. and Nematollahi, A.R. (2010) On mean past and mean residual life under double monitoring. *Communications in Statistics - Theory and Methods*, **37**, p.1119–1133.

Inference for Partial Differential Equations defined over complex domains

IS 45
Inference for Complex Data
Rm -1.64
Wed
12:30 - 13:00

Michelle Carey[*,†,‡], James O. Ramsay[*]

[*]University College Dublin, Dublin, Ireland
[†]McGill University, Montreal, Canada
email: [‡]michelle.carey@ucd.ie

Geo spatial data are observations of a process that are collected in conjunction with reference to their geographical location. This type of data is abundant in many scientific fields, some examples include: population census, social and demographic (health, justice, education) and environmental (atmospheric and oceanographic) data. They are often distributed over irregularly shaped spatial domains with complex boundaries that may contain interior holes. Modelling approaches must account for the spatial dependence over these irregular domains as well as describing the temporal evolution.

Dynamic systems modelling has a huge potential in statistics, as evidenced by the amount of activity in functional data analysis. Many seemingly complex forms of functional variation can be more simply represented as a set of differential equations.

In this talk, I will present a class of semi parametric regression models with differential regularization in the form of PDEs. This methodology is called Data2PDE "Data to Partial Differential Equations". Data2PDE characterizes spatial processes that evolve over complex geometries in the presence of uncertain and incomplete observations and prior knowledge regarding the physical principles of the process characterized by a PDE.

IS 43
Scan Statistics and Applications
Rm 100/A
Wed
16:30 - 17:00

Off-line multiple change-point detection with reproducing kernels

Alain Celisse
University of Lille, France
email: celisse@math.univ-lille1.fr

In this talk we discuss the change-point detection problem when dealing with complex data in the off-line (batch) context. Our goal is to present a new procedure involving reproducing kernels and allowing to detect abrupt changes arising in the full distribution of the observations along the time (and not only in their means). The two-stage procedure we introduce is based first on dynamic programming, and second on a new l_0-type penalty derived from a non-asympsotic model selection result applying to vectors in a reproducing kernel Hilbert space. We will illustrate the practical behavior of our kernel change-point procedure on a wide range of simulated data.

IS 24
Reliability Analysis
Rm -1.62
Wed
17:30 - 18:00

Multivariate Lifetime Modeling Based on Dynamic Shock Models

Ji Hwan Cha[*,‡], German Badía[†]
[*]Ewha Womans University, Seoul, Rep. of Korea
[†]Universidad de Zaragoza, Zaragoza, Spain
email: [‡]jhcha@ewha.ac.kr

In this paper, we stochastically model positively dependent multivariate lifetime distributions based on stochastically dependent dynamic shock models. In the first part, we consider a shock model with delayed failures. This shock model will be used to construct a class of absolutely continuous multivariate distributions. Explicit parametric forms for the multivariate reliability functions are suggested. Multivariate ageing properties and dependence structures of the class are discussed as well. In the second part, we obtain two types of absolutely continuous multivariate exponential distributions based on further generalized shock models.

OCS 2
Non-Markov Models in Finance
Harmony Inge
Tue
12:20 - 12:40

On fixed gain recursive estimators with discontinuity in the parameters

Huy N. Chau[*], Chaman Kumar[†], Miklós Rásonyi[*], Sotirios Sabanis[‡]
[*]MTA Alfréd Rényi Institute of Mathematics, Budapest
[†]Indian Institute of Technology, Roorkee
[‡]University of Edinburgh

In this talk we estimate the tracking error of a fixed gain stochastic approximation scheme. We are interested in a recursion of the form

$$\theta_{t+1} = \theta_t + \gamma H(\theta_t, X_{t+1}), \quad t \in \mathbb{N},$$

starting from some guess θ_0. The underlying process X_t is not assumed Markovian, a mixing condition is required instead. Furthermore, the updating function $H(\cdot, \cdot)$ may be discontinuous in the parameter. Some simulations show that the theoretical estimate is in accordance with numerical results. Financial applications are also introduced.

Acknowledgement. Chaman Kumar were supported by the School of Mathematics, University of Edinburgh, United Kingdom. Huy N. Chau and Miklós Rásonyi were supported by the "Lendület" Grant LP2015-6 of the Hungarian Academy of Sciences. Sotirios Sabanis gratefully acknowledges the support of the Royal Society through the IE150128 grant. We have made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF), see http://www.ecdf.ed.ac.uk. This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1, in the framework of a "small research group", during the summer of 2016.

Conditional Scan Statistics for Detecting a Local Change in Population Mean and Variance for Normal Data

IS 42
Scan Stats – Methods...
Rm 100/A
Thu
11:30 - 12:00

Jie Chen[*,‡], Joseph Glaz[†]

[*]University of Massachusetts Boston, MA, USA
[†]University of Connecticut, Storrs, CT 06269-4120, U.S.A
email: [‡]jie.chen@umb.edu

Let $X_1, \ldots, X_M$ be a sequence of independent and identically distributed (iid) normal observation with mean μ and variance σ^2, where μ and σ^2 are both unknown. In this talk, we investigate the performance of fixed, multiple and variable window conditional scan statistics in detecting a local change in population mean or population variance. Under the null hypothesis of randomness, the distribution of the iid normal observations given the total sum of squares, follows a join uniform distribution on the $(M-1)$ dimensional sphere with radius equal to the square root of the sum squares of the observations. Since the exact distribution of these conditional scan statistics is unknown, we use a Monte Carlo simulation to evaluate their probabilities. We investigate the performance of multiple window scan statistics using minimum p-value statistics and variable window scan statistics using generalized likelihood ratio tests. Numerical results are presented to compare the performance of these conditional scan statistics based on the accuracy of achieving the prescribed significance level and the power. When the location and length of the window where a change in the population mean or population variance has occurred is unknown, both multiple and variable window scan statistics perform well.

A credit derivative on simultaneous defaults

IS 10
Copula Functions
Rm 100/B
Wed
10:30 - 11:00

Umberto Cherubini

University of Bologna
email: umberto.cherubini@unibo.it

The paper addresses the problem of estimating singularities in joint default distributions. The reference model is the Gumbel-Marshall-Olkin model, that is an extension of the Marshall-Olkin distribution in which the hidden shocks are linked by a Gumbel copula. We show how to estimate the singular part of the copulas, that is the probability of simultaneous events, and how to use the estimation to design a credit derivative product on simultaneous events. The estimation strategy is based on the non parametric analysis of the dependence structure, through the Kendall tau, and the validation is performed by estimating the intensity of the singular part.

IS 31
Actuarial Risk Models II
Rm 0.99
Tue
14:30 - 15:00

Partially Schur-constant models with actuarial applications

Anna Castañer[*,¶], M. Mercè Claramunt[†,‖], Claude Lefèvre[‡,**], Stéphane Loisel[§,††]

[*]Universitat de Barcelona, Barcelona, Spain
[†]Universitat de Barcelona, Barcelona, Spain
[‡]Université Libre de Bruxelles, Bruxelles, Belgium and ISFA, Université Lyon 1, Lyon, France
[§]ISFA, Université Lyon 1, Lyon, France
email: [¶]acastaner@ub.edu [‖]mmclaramunt@ub.edu [**]clefevre@ulb.ac.be
[††]stephane.loisel@univ-lyon1.fr

This paper introduces a new dependence model for a random vector that generalizes the standard Schur-constant model. The vector is formed of m subvectors which satisfy a property of partial exchangeability, hence its appellation of partially Schur-constant. Two different representations are obtained which point out the simple dependence structure underlying the model. Several other distributional properties are derived, including the correlations inside and between the subvectors. As an illustration, the model is applied to a risk management for insurance-reinsurance networks.

Acknowledgement. The authors received support from the *Cátedra ICEA-UB de Seguros y Fondos de Pensiones*.

References

Castañer, A. and Claramunt, M.M. (2017) Equilibrium distributions and discrete Schur-constant models. HAL: hal-01593552.

Castañer, A., Claramunt, M.M., Lefèvre, C. and Loisel, S. (2015) Discrete Schur-constant models. *Journal of Multivariate Analysis*, **140**, p.343-362.

Lefèvre, C. and Loisel, S. (2013) Markov property in discrete Schur-constant models. *Journal of Applied Probability*, **50**, p.827-847.

Lefèvre, C., Loisel, S. and Utev, S. (2017) Markov property in discrete Schur-constant models. *Methodology and Computing in Applied Probability*, doi: 10.1007/s11009-017-9564-5.

CS 8
Financial Risk and Valuation
Rm 100/B
Tue
12:20 - 12:40

Kyle equilibrium under random price pressure

José Manuel Corcuera

Department of Mathematics and Computer Science, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585, E-08007 Barcelona
email: jmcorcuera@ub.edu

We study the equilibrium in the model proposed by Kyle in 1985 and extended to the continuous time setting by Back in 1992. The novelty of this paper is that we consider a framework where the price pressure can be random. We also allow for a random release time of the fundamental value of the asset. This framework includes all the particular Kyle models proposed in the literature. The results enlighten the equilibrium properties shared by all these models and guide the way of finding equilibriums in this context.

Acknowledgement. This work is supported by the Spanish grant MTM2016-76420-P.

References

Back, K. (1992) Insider trading in continuous time. *The Review of Financial Studies*, **5** (3), 387–409.
Kyle, A. S. (1985) Continuous auctions and insider trading. *Econometrica*, **53** (6), 1315–1335.

Ergodicity of a Level Dependent Quasi-Birth and Death (LDQBD) Process via the Generalized Inverse

CS 5
Branching Processes
Rm -1.62
Tue
17:30-17:50

James D. Cordeiro[*,¶], Jeffrey P. Kharoufeh[†], Mark E. Oxley[‡], Manil T. Mohan[§]

[*]University of Dayton, Dayton, Ohio, USA
[†]University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[‡]Air Force Institute of Technology, Dayton, Ohio, USA
[§]Indian Statistical Institute, Bombay, India
email: [¶]jcordeiro1@udayton.edu

Level dependent quasi-birth-and-death (LDQBD) processes extend the classical level independent quasi-birth-and-death (QBD) processes by allowing the rows of the infinitesimal generator matrix to vary with levels of the state space. Due to their prevalence in broad classes of queueing models, an analytic condition for determining the ergodicity of the embedded LDQBD process is of significant interest. In this presentation, we employ Foster-Lyapunov drift conditions to determine a necessary and sufficient criterion for the positive recurrence of an irreducible, discrete-time LDQBD process whose transition matrix converges over block rows. The Markov group inverse, which is described by Meyer [1975], is used to find a positive Lyapunov potential function that satisfies the sufficiency criteria of Fayolle *et al.* [1995] for both the recurrence and positive recurrence of Markov chains with countable state spaces. Necessity is provided by a similar drift approach, supplemented by theorems of alternatives for linear systems of equations.

References

Fayolle, G., Malyshev, V.A., & Menshikov, M.V. 1995. *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge, UK: Cambridge University Press.

Meyer, Carl D. 1975. The role of the group generalized inverse in the theory of finite Markov chains. *SIAM Review*, **17**(3), 443–464.

Non-parametric statistics for marked inhomogeneous point processes

IS 19
Spatial Point Processes
Rm -1.63
Wed
11:30 - 12:00

O. Cronie[*,‡], M. N. M. van Lieshout[†,§]

[*]Umeå University, Umeå, Sweden
[†]CWI, Amsterdam, The Netherlands
email: [‡]ottmar.cronie@umu.se [§]marie-colette.van.lieshout@cwi.nl

In this talk we discuss summary statistics for inhomogeneous, intensity-reweighted moment stationary marked point processes. The new statistics are based on the n-point correlation functions of the underlying point process and reduce to cross J- and nearest neighbour distance distribution functions when stationarity holds. We explore the relationships between the various functions and discuss their explicit forms under specific model assumptions. We further derive ratio-unbiased minus sampling estimators for our statistics and illustrate their use on a data set of wildfires. If time permits, we will also briefly discuss intensity estimation.

References

O. Cronie and M. N. M. van Lieshout. Summary statistics for inhomogeneous marked point processes. *Ann. Inst. Statist. Math.*, 68(4):905–928, 2016. doi: 10.1007/s10463-015-0515-z.

IS 6
Affine & Polynomial Proc.
Harmony lnge
Wed
16:30 - 17:00

Infinite dimensional polynomial processes

Christa Cuchiero
University of Vienna
email: christa.cuchiero@univie.ac.at

Motivated from high and infinite dimensional problems in mathematical finance, we consider infinite dimensional polynomial processes taking values in certain space of measures or functions. We have two concrete applications in mind: first, modeling high or even potentially infinite dimensional financial markets in a tractable and robust way, and second analyzing stochastic Volterra processes, which recently gained popularity through rough volatility models and ambit processes. The first question leads to probability measure valued polynomial diffusions and the second one to Markovian lifts of polynomial Volterra processes of general jump diffusion type. For both cases we provide existence and uniqueness results and a moment formula.

IS 22
Renewal & semi-Markov...II.
Rm -1.62
Tue
14:30 - 15:00

Insurance contracts for hedging wind power uncertainty

Guglielmo D'Amico[*,§], Filippo Petroni[†,¶], Fulvio Gismondi[‡,‖]
[*]1Department of Pharmacy, University "G. d'Annunzio" of Chieti-Pescara, Italy
[†]Department of Business and Economics, University of Cagliari, Italy
[‡]Department of Economic and Business Science, "Guglielmo Marconi" University, Italy
email: [§]g.damico@unich.it [¶]fpetroni@unica.it [‖]f.gismondi@unimarconi.it

Wind energy is assuming even more importance in the production of electricity. The share of production due to wind is continuously increasing in time although there are still relevant problems that affect this industry. The most important limitation for a further development of the wind energy industry concerns the variability of the wind speed phenomenon. The problem of the wind speed volatility has been approached mainly by energy storage systems; that is, by storing a surplus of energy to be used for compensating an eventual future deficit of production. More recently an insurance contract between the wind energy producer and a dispatchable energy producer has been proposed as a mean to manage the uncertainty of the wind speed. In this paper we extend previous results involving the use of insurance contracts by considering the dependence existing between electricity prices and wind energy production. The dependence structure is modeled using an appropriate copula function and we show the the impact of this dependence on the fair premium that the wind power supplier has to pay in order to hedge the risk of inadequate output of electricity at any time. Recursive type equations are obtained for the prospective mathematical reserves of the insurance contract and for their higher order moments. The model and the validity of the results are illustrated through a numerical example.

Empirical L^2-distance test statistic for discretely observed SDEs

CS 3
SDEs and Semi-martingales
Rm 100/B
Mon
17:00-17:20

Alessandro De Gregorio[*,‡], Stefano Maria Iacus[†]

[*]"Sapienza" University of Rome, Rome, Italy
[†]University of Milan, Milan, Italy
email: [‡]alessandro.degregorio@uniroma1.it

In this talk we introduce a new type of test statistic for simple null hypothesis on one-dimensional ergodic diffusion processes sampled at discrete times. We deal with a quasi-likelihood approach for stochastic differential equations (i.e. local gaussian approximation of the transition functions) and define a test statistic by means of the empirical L^2-distance between quasi-likelihoods. We prove that the introduced test statistic is asymptotically distribution free; namely it weakly converges to a χ^2 random variable. Furthermore, we study the power under local alternatives of the parametric test. We show by the Monte Carlo analysis that, in the small sample case, the test based on the empirical L^2-distance seems to perform better than other tests proposed in literature.

References

De Gregorio, A. and Iacus S.M. (2018) Empirical L^2-distance test statistics for ergodic diffusions. *Statistical Inference for Stochastic Processes*, online

Kessler, M. (1997) Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, **24**, 211–229.

Kitagawa, H., Uchida, M. (2014) Adaptive test statistics for ergodic diffusion processes sampled at discrete times, *Journal of Statistical Planning and Inference*, **150**, 84-110.

Performance of the smallest-variance-first rule in appointment sequencing

CS 12
Epidemiological and Medical Applications
Rm 100/A
Mon
17:40-18:00

Madelon de Kemp

University of Amsterdam
email: m.a.dekemp@uva.nl

A classical problem in appointment scheduling, with applications in health care, concerns the determination of the patients' arrival times that minimize an objective function that is a weighted sum of mean waiting times and mean idle times. Part of this problem is the sequencing problem, which focuses on ordering the patients. We assess the performance of the smallest-variance-first (SVF) rule, which sequences patients in order of increasing variance of their service durations. While it was known that SVF is not always optimal, many papers have found that it performs well in practice and simulation. We give theoretical justification for these observations by proving quantitative upper bounds on the ratio between the objective value corresponding to the SVF rule and to the optimal sequence. When the interarrival times are equal to the mean service times, the ratio is bounded by 2. In the limiting regime that the number of patients grows large, the ratio approaches 1, so the smallest-variance-first rule is asymptotically optimal. We also find an upper bound on the ratio when the interarrival times can be chosen optimally. Our results appear to be the first of this type in the appointment scheduling literature.

IS 55
Applied Probability
Rm 0.99
Wed
12:30 - 13:00

On an approach to boundary crossing

Victor de la Pena
Columbia University
email: vp@stat.columbia.edu

In this talk I will introduce an approach to boundary crossing by processes that was inspired by decoupling inequalities. The basic principle is that one can decouple the stopping time from the process and obtain sharp lower bounds. Joint work with Mark Brown, Michael Klass and Tony Sit.

IS 28
Stochastic Precedence
Rm -1.62
Wed
15:30 - 16:00

Qualitative analysis of "Penney-type" games through stochastic precedence

Emilio De Santis
University of Rome Sapienza, Department of Mathematics, Rome, Italy
email: desantis@mat.uniroma1.it

We consider a new type of two-players stochastic game based on a finite set of random variables $X_1, \ldots, X_n$. This construction results in a zero-sum game which can be seen as a modified version of a Penney game.

We aim to establishing whether the game is favorable to the first or to the second player, avoiding calculations that could potentially be extremely complicated. On this purpose, we develop a qualitative analysis of the game that puts us in a position to avoid computation of Nash's equilibria. A basic tool in our derivations is the appropriate construction of a complete oriented graph determined by stochastic precedence relations within any pair of the random variables $X_1, \ldots, X_n$.

IS 15
Extremes of Gaussian Processes
Rm 100/B
Wed
17:00 - 17:30

Approximation of sojourn times of Gaussian processes

Krzysztof Dębicki[¶], Enkelejd Hashorva[†], Peng Liu[†], Zbigniew Michna[‡], Xiaofan Peng[§]
[*]Mathematical Institute, University of Wrocław, Poland
[†]Department of Actuarial Science, University of Lausanne, Switzerland
[‡]Department of Mathematics, Wrocław University of Economics, Poland
[§]School of Mathematical Sciences, University of Electronic Science and Technology of China, China
email: [¶]debicki@math.uni.wroc.pl

We derive exact tail asymptotics of sojourn time above the level $u \in \mathbb{R}$

$$\mathbb{P}\Big\{v(u) \int_0^T \mathbb{I}(X(t) > u)dt > x\Big\}, \quad x \geq 0$$

as $u \to \infty$, where X is a (non centered) Gaussian process with continuous sample paths, $v(u)$ is a positive function and $T \in (0, \infty]$.

Additionally, we analyze asymptotic distributional properties of

$$\tau_u(x) := \inf\Big\{t \geq 0 \,:\, v(u) \int_0^t \mathbb{I}(X(s) > u)ds > x\Big\},$$

as $u \to \infty$, $x \geq 0$, where $\inf \emptyset = \infty$.

We illustrate the obtained findings by detailed analysis of the case when X is a Gaussian process with stationary increments and linear drift.

The obtained results complement and extend classical findings of Berman [1992].

References

Dębicki, K., Hashorva, E., Peng, X., and Michna, Z. (2017) Approximation of sojourn times of Gaussian processes. arXiv: 1712.04770

Berman, S. M. *Sojourns and Extremes of Stochastic Processes*. The Wadsworth & Brooks/Cole Statistics/Probability Series, Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1992.

Estimating the size of hidden populations from register data in heterogeneous and open populations

CS 1
Biological Statistics
Rm 100/A
Wed
12:00 - 12:20

Birgit Debrabant

Epidemiology and Biostatistics, Department of Public Health, University of Southern Denmark
email: bdebrabant@health.sdu.dk

Prevalence estimates of infectious diseases are important in many contexts. However, it is not always possible to simply count the infected individuals when incubation times are long or diagnosis is complicated or lengthy. Consequently there is a need for methods that can estimate the size of the corresponding hidden populations.

One possible approach uses registration frequencies and estimates the population size based on an underlying probability model. Such an approach was recently presented by [LW14], who on the basis of binomial removal sampling propose a maximum likelihood estimate for the size of the hidden population. The method assumes all individuals in the target distribution to have a constant probability p for being diagnosed during a given time interval and the population to be closed, i. e. individuals do not enter or exit. However, in many realistic settings such as for HCV-infections these assumptions are invalid, since the probability to be diagnosed is not constant for heterogeneous populations and depends on the duration of the infection. We discuss extensions of the approach of [LW14] to encompass settings with heterogeneous p and populations where individuals can exit.

[LW14] A. Ledberg and P. Wennberg. "Estimating the size of hidden populations from register data". In: BMC Medical Research Methodology 14.1 (2014).

Analysis of multilevel Monte Carlo using the Milstein discretisation

CS 3
SDEs and Semi-martingales
Rm 100/B
Mon
17:20-17:40

Kristian Debrabant[*,§], Michael B. Giles[†,¶], Andreas Roessler[‡,‖]

[*]Department of Mathematics and Computer Science, University of Southern Denmark
[†]Mathematical Institute, University of Oxford
[‡]Institute of Mathematics, University of Lübeck
email: [§]debrabant@imada.sdu.dk [¶]mike.giles@maths.ox.ac.uk
[‖]roessler@math.uni-luebeck.de

Using a simple Monte Carlo method with a numerical discretisation with first order weak convergence, to achieve a root-mean-square error of $\mathcal{O}(\epsilon)$ would require $\mathcal{O}(\epsilon^{-2})$ independent paths, each

with $\mathcal{O}(\epsilon^{-1})$ timesteps, giving a computational complexity which is $\mathcal{O}(\epsilon^{-3})$. However, Giles' multilevel Monte Carlo (MLMC) approach (Giles [2008]), which combines the results of simulations with different numbers of timesteps, reduces the cost to $\mathcal{O}(\epsilon^{-2})$ under certain circumstances.

In this presentation we analyse the efficiency of the MLMC approach for different options and scalar SDEs using the Milstein discretisation, determining or bounding the order of convergence of the variance of the multilevel estimator, and hence the computational complexity of the method.

References

Giles, Michael B. 2008. Multilevel Monte Carlo path simulation. *Oper. Res.*, **56**(3), 607–617.

Giles, Michael B., Debrabant, Kristian, & Rößler, Andreas. 2018. *Analysis of multilevel Monte Carlo using the Milstein discretisation*. Preprint.

CS 9
Stability of financial systems
Harmony Inge
Tue
17:50-18:10

Capital reserve management for a multi-dimensional risk model

Gertruda Antoinette Delsing

University of Amsterdam

email: g.a.delsing@uva.nl

This research focusses on the calculation of ruin probabilities and capital reserves for a multi-dimensional risk model.

We consider the evolution of the capital reserves of a company with multiple lines of business facing dependent risks in finite time. The individual reserve of these lines of business is modeled by means of a Cramér–Lundberg model with constant incoming premiums and outgoing claims that arrive according to a Poisson process. To allow for common environmental factors that impact all lines of business, the claim inter-occurrence times as well as the claim sizes have been made dependent on a single environmental process. Considering a fixed environmental process over time, we present a novel Bayesian approach to calibrate the latent environmental state distribution based on observations concerning the claim processes. For a time changing latent environmental process, approximations can be found for the joint ruin probability of different sublines dependent on the rate of change of the latent environmental process. For a fast changing environmental process, for example, one can approximate the distribution by its stationary distribution and use a functional central limit theorem result to approximate the probability of ruin.

We then point out how to determine the optimal initial capital of the different business lines under specific constraints on the ruin probability of subsets of business lines. Combining the Bayesian updating approach with the optimization method to determine capital, we have developed an easy-to-implement approach to capital risk management in a multi-dimensional risk model.

Acknowledgement. The research has been performed together with Erik Winands, Michel Mandjes and Peter Spreij.

Stochastic clearing in discrete-time queue and some results

POSTER
Mon
18:00-19:00

Mustafa Demircioglu

TELIN SMACS, Ghent, Belgium
email: mustafa.demircioglu@ugent.be

In this poster, the phenomenon disaster and its effect on discrete-time queueing system, especially GI-GI-1 model is considered. Computers, tablets, smart phones, watches or any other telecommunication equipments are often confronted with this stochastic clearing. Moreover it occurs frequently in social media such as photo sharing, microblog or instant messaging. Disaster or queue flushing occurs by infection of a virus on a computer system or resetting order in communication systems, and deleting the all information in the systems. The communication systems with disaster has a stochastic characteristic, and it has an impact on the performance of our discrete-time queue system.

Bayesian Variable Selection with Application to High Dimensional EEG Data by Local Spatial Modeling

IS 46
High-dim. Bayesian Inference
Rm 0.99
Wed
15:00 - 15:30

Dipak Kumar Dey

University of Connecticut
email: dipak.dey@uconn.edu

Due to the immense technological advances, very often we encounter data in high-dimensions. Any set of measurements taken at multiple time points for multiple subjects leads to data of more than two dimensions (matrix of covariates for each subject). In this talk, we present a Bayesian method for binary classification of subject-level responses by building binary regression models using latent variables along with the structured spike and slab priors. We also study the scaled normal priors on the parameters, as they cover a large family of distributions. Due to the computational complexity, we build many local (at different time points) models and make predictions using the temporal structure between the local models. We perform variable selection for each of these local models. If the variables are locations, then the variable selection can be interpreted as spatial clustering. We show the results of a simulation study and also present the performance of these models on multi-subject neuroimaging (EEG) data.

On a growth model similar to Gompertz and Korf laws, and a related birth-death process

IS 26
Growth & Contagion Models
Rm 0.99
Mon
11:00 - 11:30

Antonio Di Crescenzo

Dipartimento di Matematica, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, I-84084 Fisciano (SA), Italy
email: adicrescenzo@unisa.it

A general model of population growth can often be described by an ordinary differential equation of the form

$$\frac{dN(t)}{dt} = \xi(t)N(t), \quad t > 0,$$

where $\xi(t)$ is a time-dependent growth rate. The suitable choices

$$\xi(t) = \xi_G(t) = \alpha e^{-\beta t} \quad \text{and} \quad \xi(t) = \xi_K(t) = \alpha t^{-(\beta+1)}$$

yield the differential equations of the Gompertz growth and of the Korf growth, respectively (see Gompertz [1825] and Korf [1939]). The aim of this contribution is to describe a suitable deterministic growth model which arises from an alternative choice of the growth rate, given by

$$\xi(t) = \alpha(1+t)^{-(\beta+1)}, \quad t > 0.$$

We elucidate the main features of this model, including the correction factor, the relative growth rate, the inflection point, the maximum specific growth rate, the lag time and the threshold crossing problem.

Furthermore, in analogy with Tan [1986] we investigate a stochastic counterpart of the considered model, consisting in a linear time-inhomogeneous birth-death process whose mean behaves as the deterministic model. We focus on the transition probabilities, the moments and the population ultimate extinction probability of this process. Attention is also given to the special case of a simple birth process, which better mimics the deterministic growth model. We also address some issues concerning a possible extension of the considered model.

Most of this contribution is based on a joint work with Serena Spina (cf. Di Crescenzo and Spina [2016]).

Acknowledgement. The author is member of Gruppo Nazionale per il Calcolo Scientifico (GNCS-INdAM).

References

Di Crescenzo, A. and Spina, S. (2016) Analysis of a growth model inspired by Gompertz and Korf laws, and an analogous birth-death process. *Math. Biosci.*, **282**, p.121-134. doi: 10.1016/j.mbs.2016.10.005.

Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos. Trans. R. Soc. Lond.*, **155**, p.513-583. JSTOR: 107756

Korf, V. (1939) Prìspevek k matematickrè formulaci vzrustovèho zàkona lesnìch porostu [contribution to mathematical definition of the law of stand volume growth]. *Lesnickà pràce*, **18**, p.339-379.

Tan, W.Y. (1986) A stochastic gompertz birth-death process. *Stat. Prob. Lett.*, **4**, p.25-28. doi: 10.1016/0167-7152(86)90034-9

POSTER
Mon
18:00-19:01

Statistical Postprocessing Models of Ensemble Forecasts for Calibrating the Temperature in Santiago de Chile

Mailiu Díaz[*,§], Orietta Nicolis[*,¶], Sándor Baran[†], Julio Cesar Marin[‡]

[*]Department of Statistics, University of Valparaíso, Valparaíso, Chile
[†]Faculty of Informatics, University of Debrecen, Debrecen, Hungary
[‡]Department of Meteorology, University of Valparaíso, Valparaíso, Chile
email: [§]mailiu.diaz@postgrado.uv.cl [¶]orietta.nicolis@uv.cl

Numerical weather prediction (NWP) models provide information on high spatial and temporal resolutions, however, the outputs of these models are subject to systematic and random model errors which adversely impact forecast quality and the operational usefulness. The same applies for ensemble forecasts obtained by multiple runs of these models with different initial conditions or parametrizations, so statistical post-processing approaches have been developed to improve the calibration and prediction accuracy. The main aim of our research is to explore some spatial and spatio-temporal post-processing models to calibrate the ensemble forecasts for surface temperature.

A number of 9 Weather Research Forecast (WRF) simulations are run with 2 nested domains centered over Santiago city to create a forecast ensemble, where the members differ either in the physical parametrization of the underlying land-surface model, or planetary boundary layer, or microphysics used. The corresponding hourly observations are downloaded from 19 stations of the Servicio Meteorologico de la Armada de Chile y Sistema de Informacion Nacional de Calidad del Aire for the period between October 1st, 2017 and January 31st, 2018 every hour. Using appropriate scoring rules we show that compared to the raw ensemble forecasts statistical post-processing significantly improves the calibration of probabilistic and accuracy of point forecasts.

Acknowledgement. This work was supported by National Commission for Scientific and Technological Research (CONICYT).

Entropy Analytics of Mortality Data Series

CS 7
Actuarial Applications
Rm 0.99
Mon
17:40-18:00

Yiannis Dimotikalis[*,‡], Christos H. Skiadas[†]

[*]Dept of Business Administration, T.E.I. of Crete, Greece
[†]ManLab, Technical University of Crete, Greece
email: [‡]y.dimotikalis@gmail.com

In this paper the max entropy principle applied to human mortality life table data. The entropy framework is presented and adjusted for five years age groups life tables. By entropy constrained optimization the exponential, discrete Geometric distribution, is max entropy distribution subject to expected value constraint. The Binomial distribution appears as max entropy distribution subject to expected value and variance constraint asymptotically. Those max entropy distributions curves on entropy chart calculated by constrained nonlinear optimization numerically. For the evolution in time of mortality data series, Theil index of inequality and Kullback-Leibler divergence used to measure the divergence from unconstrained max entropy Uniform distribution and convergence to Binomial distribution. The analysis of some European countries, USA and Japan mortality data series shows that the asymptotic convergence to Binomial distribution is evident and occurs in different decade of 20th century for considered data series. The implications of those findings to human mortality discussed, and some future tasks suggested.

On a non-Markov process drifted by a compound Poisson process with applications in neuronal modeling

IS 18
Stochastic Processes & Appl. II.
Rm -1.63
Tue
14:30 - 15:00

Giuseppe D'Onofrio[*,‡], Lubomir Kostal[*], Enrica Pirozzi[†]

[*]Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 14220 Prague 4, Czech Republic
[†]Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II, Via Cintia, Complesso Monte S. Angelo, 80126, Napoli, Italy
email: [‡]giuseppe.donofrio@fgu.cas.cz

Stochastic differential equations (SDEs) play a key role in the description of fluctuating phenomena belonging to different areas of applied mathematics (Oksendal [2014]). In this talk we focus on the following type of SDE: $dX(t) = [a(t)X(t) + b(t) + Z(t)]dt + \sigma dW$, where the drift is characterized by a stochastic process $Z(t)$ independent of $X(t)$. The solution $X(t)$ of such an equation is a non-markovian process characterized by a mean function dependent on $Z(t)$. The problem of the first-passage-time (FPT) of the solution through a boundary is addressed using a suitable Gauss-Markov (GM) process

approximating the process under study. This strategy allows us to use some theoretical results on FPT of GM processes (see for instance Buonocore et al. [2011], Buonocore et al. [2015], D'Onofrio and Pirozzi [2016]) in order to give some approximations of the FPT probability density function of the original process $X(t)$. These findings consitute a general theoretical result themselves, but they can also be specialized in the modeling context.

We propose a neuronal model for the description of the firing activity of a neuron embedded in a network and subject to the inputs coming from the surrounding neurons. This corresponds to the case in which $Z(t) = \sum_{i=1}^{N(t)} Z_i$ with Z_i independent and identically distributed random variables. If $N(t)$ is assumed to be a Poisson process independent on $X(t)$ then $Z(t)$ is a compound Poisson process. The stochastic nature of the drift in the model equation is due to the stochastic flux of the inputs originated from the other neurons.

We also analyze the amount of information that can be reliably exchanged through the network, studying mutual information and channel capacity (see for instance Kostal et al. [2013], Kostal and D'Onofrio [2017]). Particular attention will be given to the changes of these quantities with respect to the proportion of inhibitory and excitatory neurons constituting the network.

Acknowledgement. This work was supported by the Institute of Physiology RVO:67985823, by the Czech Science Foundation project 17-06943S and by Gruppo Nazionale per il Calcolo Scientifico (GNCS-INdAM).

References

A. Buonocore, L. Caputo, E. Pirozzi, and L. M. Ricciardi. The first passage time problem for Gauss-diffusion processes: Algorithmic approaches and applications to LIF neuronal model. *Methodology and Computing in Applied Probability*, 13(1):29–57, Mar 2011. ISSN 1573-7713. doi: 10.1007/s11009-009-9132-8.

A. Buonocore, L. Caputo, G. D'Onofrio, and E. Pirozzi. Closed-form solutions for the first-passage-time problem and neuronal modeling. *Ricerche di Matematica*, 64(2):421–439, Nov 2015. ISSN 1827-3491. doi: 10.1007/s11587-015-0248-6.

G. D'Onofrio and E. Pirozzi. Successive spike times predicted by a stochastic neuronal model with a variable input signal. *Mathematical Biosciences and Engineering*, 13(3):495–507, 2016. doi: 10.3934/mbe.2016003.

L. Kostal and G. D'Onofrio. Coordinate invariance as a fundamental constraint on the form of stimulus-specific information measures. *Biological Cybernetics*, Aug 2017. ISSN 1432-0770. doi: 10.1007/s00422-017-0729-7.

L. Kostal, P. Lansky, and M. D. McDonnell. Metabolic cost of neuronal information in an empirical stimulus-response model. *Biological Cybernetics*, 107(3):355–365, Jun 2013. ISSN 1432-0770. doi: 10.1007/s00422-013-0554-6.

B. Oksendal. *Stochastic Differential Equations*. Springer, 2014. ISBN 9783662028483.

CS 6
Queuing Theory
Rm 100/A
Tue
17:50-18:10

Towards Optimality in Parallel Job Scheduling

Benjamin Berg[*], Jan-Pieter Dorsman[†,‡], Mor Harchol-Balter[*]

[*]Carnegie Mellon University, Pittsburgh, Pennsylvania, USA [†]University of Amsterdam, Amsterdam, The Netherlands

email: [‡]j.l.dorsman@uva.nl

Running jobs in parallel is an excellent way to reduce their response time in many computer systems. In applications such as jobs running on multiple cores, one must address the question of how many cores to assign to each job, also called the level of parallelization. In making this choice, one is confronted with the following trade-off. Parallelizing an individual job across multiple cores reduces the response time of that individual job. In practice, however, the speedup a job experiences is sublinear and concave in its level of parallelization, leading to an inefficient use of resources and additional system load. Hence, while a higher level of parallelization may decrease an individual job's response time, it may have a deleterious effect on overall response time.

In this presentation, we will study this tradeoff using queueing theory methods. In particular, when all jobs share the same size distribution and the same concave speedup curve, we will argue that using a fixed level of parallelization is near-optimal. We will also consider the case where jobs may follow different speedup curves, in which case finding a good scheduling policy is even more challenging.

Multiple source scan statistics via multi-criteria analysis

IS 44
Inference for Large Datasets
Rm 0.99
Wed
17:30 - 18:00

Luiz H. Duczmal[*,¶], Alexandre C.L. Almeida[†], Andre L.F. Cancado[‡], Fabio R. da Silva[§]

[*]Federal University of Minas Gerais
[†]Federal University of Sao Joao del Rei
[‡]University of Brasilia
[§]Centro Federal de Educacao Tecnologica de Minas Gerais, Brazil
email: [¶]duczmal@gmail.com

Multiple data sources provide more reliable information regarding the emergence of potential health threats, compared to a single source. Spatial Scan Statistics have been adapted to study multivariate data sources, but only ad hoc procedures have been devised to address the problem of selecting the most likely cluster and computing its significance. Here, information from multiple data sources of disease surveillance is incorporated to achieve more coherent spatial cluster detection using tools from multi-criteria analysis. The best cluster solutions are found by maximizing two objective functions simultaneously, based on the concept of dominance. To evaluate the statistical significance of solutions, a statistical approach based on the concept of attainment function is used. The multi-criteria approach has several advantages: the representation of the evaluation function for each data source is clear, and does not suffer from an artificial, and possibly confusing mixture with the other data source evaluations; it is possible to attribute, in a rigorous way, the statistical significance of each candidate cluster; and it is possible to analyze and pick-up the best cluster solutions, as given naturally by the non-dominated set. The methodology is illustrated with real datasets.

Acknowledgement. L. Duczmal's research was supported by CNPq and FAPEMIG.

Statistical post-processing of hydrological forecasts using Bayesian model averaging

POSTER
Mon
18:00-19:00

Sándor Baran[*], Stephan Hemri[†], Mehrez El Ayari[*,‡]

[*]University of Debrecen, Hungary [†]Federal Office of Meteorology and Climatology MeteoSwiss, Zürich, Switzerland.
email: [‡]mehrez.enit@gmail.com

Nowadays, weather services typically produce ensemble forecasts, which consist of multiple runs of numerical weather prediction models that differ in initial conditions and/or the numerical representation of the atmosphere. Plugging these predictions into state-of-the-art hydrological models result in hydrological ensemble forecasts, however, the hydrological ensemble often does not account for hydrological uncertainties, like, e.g. the estimate of the soil moisture content, and often inherits the lack of calibration and/or bias of the atmospheric ensemble prediction system, so it requires statistical post-processing.

Bayesian model averaging (BMA) is a statistical method for post-processing forecast ensembles [Raftery *et al.* 2005], in order to create calibrated predictive probability density functions (PDFs). The

BMA predictive PDF of the future weather quantity is the mixture of the individual PDFs corresponding to the ensemble members and the weights and model parameters are estimated using forecast ensembles and validating observations from a given training period. We introduce a BMA model for calibrating hydrological ensemble forecasts using doubly truncated normal distributions as conditional PDFs, which generalizes the wind speed model of Baran [2014]. The method is applied to Box-Cox transformed forecasts of runoff at gauge Kaub of river Rhine [Hemri and Klein 2017], and the post-processed forecasts outperform the raw ensemble both in calibration and in accuracy of predictions.

Acknowledgement. Sándor Baran acknowledges the support of the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

Baran, S. (2014) Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components. *Comput. Stat. Data. Anal.* **75**, 227–238.

Hemri, S. and Klein, B. (2017) Analog based post-processing of navigation-related hydrological ensemble forecasts. *Water Resour. Res.* **53**, 9059–9077.

Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.* **133**, 1155–1174.

CS 13
Filtering
Rm -1.63
Tue
17:30-17:50

Robust estimation of ship speed through water using Kalman filtering.

Christopher Engström

Division of Applied Mathematics, Mälardalen University, Västerås, Sweden
email: christopher.engstrom@mdh.se

To optimally control a ship in regards its to fuel optimization, roll reduction and trip planning it is important to have a very reliable estimation of the ship's speed since the optimization uses system identification methods which work to detect small speed variations. While a GPS estimation of the ship speed is fairly accurate, the main interest is in the ships speed through the water since it truly shows the ship's performance with regards to water currents, winds and other environmental factors. However the sensors measuring speed through water are much less reliable. Optimization applications require an estimation which is expected to be robust to disturbances and sensor errors.

In this paper we seek a solution to this problem by using two sensors measuring different velocities to give a better estimation of both, in particular in the presence of outliers and sensor failure resulting in missing data. Two sensors will be considered, one measuring the ship speed through the use of GPS (SOG) and one measuring the speed of the ship through the water by measuring the speed through the propeller(STW). These two signals are modeled using a three state Kalman filter, representing SOG, STW and the difference between the two (bias term).

Results are validated with data from real ship logs obtained at Qtagg AB illustrating how this filter can account for when data is missing or there are large outliers in the data. Finally we show some further improvements and generalizations: how it can be applied to three or more signals as well as how to estimate relevant model variances reliably.

Scaled Least Squares Estimator for GLMs in Large-Scale Problems

IS 48
Stein's Method
Rm -1.64
Mon
11:30 - 12:00

Murat A Erdogdu

Microsoft Research, New England
email: erdogdu@cs.toronto.edu

We study the problem of efficiently estimating the coefficients of generalized linear models (GLMs) in the large-scale setting where the number of observations n is much larger than the number of predictors p, i.e. $n \gg p \gg 1$. We show that in GLMs with random (not necessarily Gaussian) design, the GLM coefficients are approximately proportional to the corresponding ordinary least squares (OLS) coefficients. Using this relation, we design an algorithm that achieves the same accuracy as the maximum likelihood estimator (MLE) through iterations that attain up to a cubic convergence rate, and that are cheaper than any batch optimization algorithm by at least a factor of $O(p)$. We provide theoretical guarantees for our algorithm, and analyze the convergence behavior in terms of data dimensions. Finally, we demonstrate the performance of our algorithm through extensive numerical studies on large-scale real and synthetic datasets, and show that it achieves the highest performance compared to several other widely used optimization algorithms.

Distributions of runs in a sequence of dependent trials with a change point

IS 51
Run and scan models
Rm 0.87
Wed
14:30 - 15:00

Serkan Eryilmaz

Atilim University, Ankara, Turkey
email: serkan.eryilmaz@atilim.edu.tr

Runs and run related statistics are of great interest in probability due to their wide range of applications. Distributions of runs have been extensively studied under various distributional assumptions on a given sequence $\{X_i\}_{i=1}^n$ of discrete valued trials. This study is concerned with the derivation of exact distributions of various run statistics when $\{X_i\}_{i=1}^n$ is a sequence of n dependent trials such that the first n_1 in $\{X_i\}_{i=1}^n$ are of type 1 and follow an exchangeable joint distribution denoted by $\mathcal{L}_1$, and the last n_2 elements in $\{X_i\}_{i=1}^n$ are of type 2 and follow an exchangeable joint distribution denoted by $\mathcal{L}_2$, where $n_1 + n_2 = n$. That is, the elements of $\{X_i\}_{i=1}^n$ are fully dependent but the dependence structure changes after X_{n_1}.

References

Eryilmaz, S. (2018) On success runs in a sequence of dependent trials with a change point. *Statist. Probab. Lett.*, **132**, p.91–98.

IS 32
Limits in Inventory...
Rm -1.64
Tue
16:30 - 17:00

Convergence Theorems for Varying Probabilities and their Applications to MDPs and Inventory Control

Eugene A. Feinberg*, Pavlo O. Kasyanov, Yan Liang
Stony Brook University, Stony Brook, NY, USA
email: *eugene.feinberg@stonybrook.edu

The talk discusses Fatou's lemma, the monotone convergence theorem, and Lebesgue's dominated convergence theorem for varying probabilities. In particular, we study the case of weakly converging probabilities. In general, convergence theorems may not hold for such problems, but they hold under stronger assumptions on convergence of functions. We formulate such assumptions, which are stronger than pointwise convergence of functions, and show their applicability to Markov Decision Processes and inventory control. In particular, the dominated convergence theorem for weakly converging probabilities is useful for approximating average-cost optimal policies by discount-optimal policies with the discount factor being closed to 1. The dominated convergence theorem provides useful tools for proving two groups of results for average-cost problems: (i) the validity of average-cost optimality equations, and (ii) establishing structural properties of optimal policies.
Acknowledgement. This research was partially supported by NSF grant CMMI-1636193.

IS 40
Sequential Methods
Rm 100/A
Wed
14:30 - 16:00

Efficient Importance Sampling in Sequential Multiple Testing

Georgios Fellouris*, Yanglei Song
University of Illinois at Urbana-Champaign
email: *fellouri@illinois.edu

We consider the problem of estimating via Monte Carlo simulation the misclassification probabilities of various sequential multiple testing procedures. A prototypical example of such a procedure is an extension of Wald's SPRT that stops when all log-likelihood ratio statistics are simultaneously above a positive threshold or below a negative threshold. When the log-likelihood ratio statistics are random walks, we propose an importance sampling algorithm and establish its logarithmic efficiency as the error probabilities of interest go to 0. This optimality property is extended beyond the random-walk case when the local log-likelihood ratio statistics satisfy a stability and a symmetry condition.
Acknowledgement. This work was supported by the U.S. National Science Foundation under Grants DMS 1737962 and CIF 1514245.

IS 28
Stochastic Precedence
Rm -1.62
Wed
14:30 - 15:00

On some applications of the stochastic precedence order in reliability

Maxim Finkelstein
University of the Free State, Bloemfontein
email: FinkelM@ufs.ac.za

The stochastic precedence order (STPRO) did not attract much attention in the literature so far. One of the reasons for that is that distinct from other well-known orders (e.g., usual stochastic order, hazard rate order, likelihood ratio order, etc.) it does not possess the transitivity property. However, it should be noted that due to its definition, it can be the most natural order in numerous reliability applications. Therefore, for illustration, in this talk, we consider three applications, where using this

order is efficient and leads to results that either cannot be obtained or were not yet obtained using other orders. The first application deals with the so-called multiple shocks stress/strength problem. The shocks affecting a system arrive in accordance with the NHPP. Each shock is described by its random magnitude (i.i.d.) called the stress The strength of a system is also random. A failure upon as shock occurs when the stress exceeds the strength. We derive the probability that this will not happen for multiple stresses occurring in accordance with the NHPP and analyze the obtained probability. Specifically, we show that the unobserved strength is increasing in the sense of the STPRO with the number of survived shocks (similar to Bayesian updating) The second application considers an open problem of optimal operational sequence for the 1 out of n system with warm standby. Using the virtual age concept and the accumulated exposure model, we show that the components should be activated in accordance with the increasing sequence of their lifetimes. Lifetimes of the components and the system are compared with respect to the stochastic precedence order. Only specific cases of this optimal problem were considered in the literature previously. We believe that general results were not obtained due to the non-appropriate choice of the corresponding stochastic ordering. While considering the third application, we show that, for a k-out-of-n system, redundancy at the component level is superior to that at the system level with respect to the stochastic precedence order. Cases of active and cold redundancy are considered. Similar results for other stochastic orders were intensively discussed in the literature, however not for the case of cold redundancy and the stochastic precedence order.

Unbiased estimators of weighted Voronoi cell characteristics

IS 20
Stochastic Geometry
Rm -1.63
Wed
14:30 - 15:00

Daniela Flimmel

Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University

Let $\mathcal{P}_m$ be a unit stationary independently marked Poisson point process on $\mathbb{R}^d$ with marks in $\mathbb{R}_+$. We are dealing with three types of tessellation generated by $\mathcal{P}_m$: the Voronoi tessellation, the Laguerre tessellation and the additively weighted model. Let $W_\lambda := [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$ be the d-dimensional window of volume $\lambda > 0$. We consider only the cells that are fully contained in the window W_λ. We study the limit behaviour of the Horvitz-Thompson type statistics

$$T_\lambda(\mathcal{P}_m) = \sum_{(x,r)\in\mathcal{P}_m} \frac{\lambda\xi((x,r),\mathcal{P}_m)}{\mathrm{Vol}(W_\lambda \ominus C((x,r),\mathcal{P}_m))}\mathbf{1}\{C((x,r),\mathcal{P}_m) \subset W_\lambda\},$$

where $C((x,r),\mathcal{P}_m)$ is the cell generated by $(x,r) \in \mathcal{P}_m$ and $W_\lambda \ominus C((x,r),\mathcal{P}_m)$ stands for the erosion of W_λ by the cell $C((x,r),\mathcal{P}_m)$. The score ξ is assumed to have the form $\xi((x,r),\mathcal{P}_m) := h(C((x,r),\mathcal{P}_m))$, where h is a translation invariant function. Let Z be a typical cell of the tessellation. Then $\lambda^{-1}T_\lambda(\mathcal{P}_m)$ is an unbiased estimator of $\mathbb{E}h(Z)$. It can be shown that the scores ξ are exponentially stabilizing under each considered type of tessellation and hence with some additional moment assumptions, the stabilization methods (see Schreiber, T. [2010] or Lachièze-Rey, R. et al. [2017]) can be used to prove the asymptotic normality of $\lambda^{-1}T_\lambda(\mathcal{P}_m)$.

References

Lachièze-Rey, R. and Schulte, M. and Yukich, J. E. (2017) Normal approximation for stabilizing functionals. Preprint. arXiv: 1702.00726

Schreiber, T. (2010) Limit theorems in stochastic geometry. In *New Perspectives in Stochastic Geometry*, W. S. Kendall and I. Molchanov (eds.). Oxford University Press, Oxford, 111-144.

CS 15
State Space and Markov Models
Rm 0.99
Thu
11:50 - 12:10

Yaglom limits for R-transient chains with non-trivial Martin boundary

R. D. Foley[*,‡], D. R. McDonald[†]
[*]Georgia Institute of Technology,Atlanta, Georgia U.S.A.
[†]The University of Ottawa, Ottawa, Ontario Canada
email: [‡]david.r.mcdonald@gmail.com

To quote John Maynard Keynes,"The long run is a misleading guide to current affairs. In the long run we are all dead". The study of quasi-stationary behaviour of an absorbing Markov chain was initiated by Yaglom who gave the limiting conditional behavior given non-absorption.

Here we construct a simple example of an R-transient Markov chain on a countable state space $S \cup \{\delta\}$ where δ is absorbing. The transition matrix K on S is irreducible and strictly substochastic. We determine the Yaglom limit, that is, the limiting conditional behavior given non-absorption. Each starting state $x \in S$ results in a different Yaglom limit. Each Yaglom limit is an R^{-1}-invariant quasi-stationary distribution where R is the convergence parameter of K. Yaglom limits that depend on the starting state are related to a nontrivial R^{-1}-Martin boundary. We give an intuitive idea of why Yaglom limits can depend on the initial distribution.

Acknowledgement. Professor Foley's research was supported in part by Foley's NSF Grant CMMI-0856489. Professor McDonald's research was supported in part by NSERC Grant A4551.

CS 16
Spatial structures
Rm -1.63
Wed
12:40 - 13:00

Jumps clusters of discrete-time Hawkes processes and forecasting shocks

Rachele Foschi[*,‡], Francesca Lilla[†], Cecilia Mancini[†]
[*]Department of Management and Economics, University of Pisa, via C. Ridolfi 10, 56124 Pisa
[†]Department of Management and Economics, University of Florence, via delle Pandette 9, 50127 Firenze
email: [‡]rachele.foschi@unipi.it

In this work we study variables and events that can be exploited to forecast catastrophic episodes.

Our motivation is forecasting large depreciations of financial assets, given the observation of clusters of previous jumps in the series of the prices. Assuming a jump diffusion model for the log returns of a given asset and based on the Lévy modulus of continuity (Mancini, C. [2009], Figueroa-López, J.E. and Mancini, C. [2017]), we fix a threshold function such that the standardized returns falling below the threshold are likely to contain no jumps. When on the contrary such a threshold is exceeded the model is likely to have generated one jump within the interval relative to the computed return.

We believe that during periods preceding catastrophic episodes the probability of a jump immediately increases after the occurrence of one or more previous jumps. That led us to filter out and study the point process counting the number of the occurred jumps. We found in fact that such a point process is well described by a Hawkes model, which is characterized by the ability of capturing self-excitation mechanisms.

A Hawkes process $\{N_t\}_{t\geq 0}$, with $\mathcal{F}_t = \sigma(N_s, 0 \leq s < t)$, is a simple counting process such that

$$P(N_{t+dt} - N_t = 1|\mathcal{F}_t) = \lambda_t dt + o(dt).$$

If we consider an exponential decay of self-excitation, the intensity

$$\lambda_t \equiv \lim_{dt\to 0} \frac{1}{dt} P(N_{t+dt} - N_t = 1|\mathcal{F}_t)$$

can be written as

$$\lambda_t = \lambda_0 + \alpha \sum_{T_i<t} e^{-\beta(t-T_i)},$$

with λ_0, α, β positive, deterministic and constant parameters, which makes explicit the fact that the occurrence of a jump at t (or some close jumps) induces an increase in the probability of having some other jumps immediately after.

Since the intensity depends on the realization of the jump times until t, it is a stochastic function. In view of the recursive formula provided in Rasmussen, J.G. [2011], we can state some properties of the process $\{\bar{\lambda}_{T_n}\}_{n\in\mathbb{N}}$ of the unconditional intensities computed at the jump times and some properties of the interarrival times of the Hawkes process.

Even if, theoretically, Hawkes models are designed in continuous time, our application requires a time discretization, induced by the shortest length of the observation interval δ, in our case $\delta = 5$ minutes. However, the discrete time allows us to provide an expression for some conditional probabilities that cannot be explicitly computed in continuous time. More precisely, we computed the probability of jump in the shortest interval, conditional on some relevant configurations of jump times in the immediate past, and the probability of these configurations.

We are also interested in the sizes of jumps, in fact we have in mind that the parameters may depend in some way on them and that a significant connection also exists cross-sectionally between the small jumps and the big ones. We can illustrate different mechanisms for the interaction between sizes and jump intensity. For our aims, the most appropriate generalization is the one known in the literature as multivariate Hawkes process.

A multivariate Hawkes process $\mathbf{N}$ results from the superposition of different interacting simple counting processes $\mathbf{N}^i$, each one with its own jump intensity. The intensity of a multivariate Hawkes process is therefore a vector λ_t, whose components λ_t^i consist in a baseline constant term, a self-excitation one, and possibly some cross-excitation terms. These last terms represent the influence of any univariate process $\mathbf{N}^j$, $j \neq i$, on the process $\mathbf{N}^i$ and therefore the interactions among the different components of the multivariate process.

In the framework described above, we separately model the jumps according to their sizes using the two sub-processes $\mathbf{N}^b$ of the big jumps and $\mathbf{N}^s$ of the small ones, with intensities given by

$$\begin{cases} \lambda_t^b = \lambda_0^b + \alpha_{bb} \sum_{T_i^b<t} e^{-\beta_{bb}(t-T_i^b)} + \alpha_{bs} \sum_{T_i^s<t} e^{-\beta_{bs}(t-T_i^s)} \\ \lambda_t^s = \lambda_0^s + \alpha_{ss} \sum_{T_i^s<t} e^{-\beta_{ss}(t-T_i^s)} + \alpha_{sb} \sum_{T_i^b<t} e^{-\beta_{sb}(t-T_i^b)} \end{cases}$$

A catastrophic event is represented by the occurrence of a big, and negative, jump. We empirically found that the occurrence of a cluster of small jumps is significantly related to the occurrence of a following big jump in our dataset. By applying the same discretization defined for the univariate processes, we quantified the increase in probability of the occurrence of a big jump given a cluster of small jumps.

Analysis of both the theoretical and the practical implications of the found expressions are still on going.

References

Bacry, E., Mastromatteo, I. and Muzy, J.F. (2015) Hawkes processes in finance, *Market Microstructure and Liquidity*, **1(1)**, 1550005

Chavez -Demoulin, V., Davison, A.C. and McNeil A.J. (2005) Estimating value-at-risk: a point process approach. *Quantitative Finance* **5(2)**, 227-234.

Daley, D.J. and Vere-Jones, D. (2003) An Introduction to the Theory of Point Processes. Springer.

Mancini, C. (2009) Non-parametric Threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, **36**, 270-296

Figueroa-López, J.E. and Mancini, C. (2017) Optimum thresholding using mean and conditional mean square error. arXiv: 1708.04339

Figueroa-López, J.E. and Nisen, J. (2013) Optimally thresholded realized power variations for Lévy jump diffusion models. *Stochastic Processes and their Applications* **123(7)**, 2648-2677

Hawkes, A.G. (1971) Spectra of some self-exciting and mutually exciting point processes, *Biometrika*, **58(1)**, 83-90

Rasmussen, J.G. (2011) Temporal point processes: the conditional intensity function. Draft, http://people.math.aau.dk/ jgr/teaching/punktproc11/tpp.pdf

Zhuang, J., Ogata, Y. and Vere-Jones, D. (2002) Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistica Association* **97(458)**, 369-380

IS 13
Lévy Processes & Appl.
Harmony Inge
Mon
17:00 - 17:30

Around GGC-random variables and HCM-functions

Sonia Fourati

INSA Rouen, France

The history of GGC variables began with the question asked by Olaf Thorin " is the Log Normal ditribution Indefinitely divisible?". He proved that not only is the LogNormal distribution ID, but actually it belongs to a class, called later "Generalized Gamma Convolutions distributions (GGC)", which are defined as convolutions of Gamma distributions and their limits. Later on, Bondesson introduced "HCM functions" which are Laplace transforms of GGC-r.v. if they are completly monotonous, and, if they are probability densities, are the probability density of GGC random variables. More recently, L.Bondesson showed that sums and products of independent GGC random variables are still GGC random variables.

I will answer some questions rised by this author as : suppose X is GGC, is X^2 also GGC? What is the additionnal property needed for having $1/X$ GGC? Moreover, if S is an α- stable random variable ($\alpha \in]0,1]$) then the density of $S^{\frac{1-\alpha}{\alpha}}$ (if $\alpha \leq 1/2$), $S^{\frac{-\alpha}{1-\alpha}}$ (if $\alpha > 1/2$), is an HCM function (up to the multiplication with the function $e^{\delta x}$).

Acknowledgement. This work is part of the Class project, co-financed by the European Union with the European regional development fund (ERDF) and by the Normandie Regional Council.

PLENARY
Harmony Inge
Tue
9:00 - 10:00

Rough path analysis of rough volatility

Peter Friz

TU Berlin, Weierstraß-Institut Berlin

email: p.k.friz@gmail.com

We will explain that the analysis of rough volatility models becomes a natural application area of rough paths and regularity structures. As a specific application, we will consider the option pricing problem in the moderate regime.

Acknowledgement. Joint work with P. Gassiat and P. Pigato

Analyzing the performance of nonparametric tests to compare the fist-order distribution of spatial point processes

CS 16
Spatial structures
Rm -1.63
Wed
12:00 - 12:20

Isabel Fuentes-Santos[*,§], Wenceslao Gonzalez-Manteiga[‡,¶], Jorge Mateu[‖]

[*]Marine Research Institute, Spanish National Research Council, Vigo, Spain
[†]Department of Statistics, Mathematical Analysis and Optimization. University of Santiago de Compostela, Spain
[‡]Department of Mathematics, University Jaume I, Castellón, Spain
email: [§]isafusa@gmail.com [¶]wenceslao.gonzalez@usc.es [‖]mateu@mat.uji.es

A common question in the analysis of multitype spatial point processes is whether two types of events have the same spatial structure. This question can arise in different scientific contexts such as the analysis of environmental risks, testing whether wildfires with different causes have the same spatial distribution, or in epidemiology, where the spatial pattern of disease cases is compared with the population at risk. A natural way to test whether two spatial point patterns have the same spatial structure should be through comparison of first-order intensity functions.

This issue has been addressed from different perspectives. Zhang and Zhuang [2017] test the proportionality between the first-order intensities of two spatial point processes through a Kolmogorov–Smirnov test with asymptotic calibration. Fuentes-Santos *et al* [2017b] proposed a nonparametric test based on the square discrepancy between the densities of event locations, the test statistic is estimated by kernel smoothing and a bootstrap procedure is needed to calibrate te test. Finally, following the proposal of Fuentes-Santos *et al* [2017a] to test spatio-temporal separability, we propose a regression test that checks whether the ratio between the first-order intensities under comparison is spatially invariant, which is calibrated through a permutation test.

This work compares the performance of the three tests outlined above in the analysis of simulated Poisson and non-Poisson point processes, as well as in the comparison of wildfire patterns in Galicia (NW Spain).

Acknowledgement. This work has been supported by Projects MTM2016-78917-R and MTM2016-76969-P from the Spanish Ministry of Economy and Competitivity and FEDER funding, grant P1-1B2015-60 from Bancaja Foundation, and IAP network StUDyS from the Belgian Science Policy.

References

Fuentes-Santos, I., González-Manteiga, W., Mateu, J. (2017) A first-order ratio-based nonparametric separability test for spatiotemporal point processes. *Envirometrics* (avaliable online 6-11-2017)

Fuentes-Santos, I., González-Manteiga, W., Mateu, J. (2017) A nonparametric test for the comparison of first-order structures of spatial point processes. *Spatial Statistics*, **22(2)**: 240-260.

Zhang, T., Zhuang, R. (2017) Testing proportionality between the first-order intensity functions of spatial point processes. *Journal of Multivariate Analysis*, **155**, 72-82.

IS 4
(Rough) Volatility Asymptotics
Rm 100/A
Tue
11:00 - 11:30

Short-term at-the-money asymptotics under stochastic volatility models

Masaaki Fukasawa
Osaka University

A small-time Edgeworth expansion of the density of an asset price is given under a general stochastic volatility model, from which asymptotic expansions of put option prices and at-the-money implied volatilities follow. A limit theorem for at-the-money implied volatility skew and curvature is also given as a corollary. The rough Bergomi model is treated as an example.

CS 6
Queuing Theory
Rm 100/A
Tue
16:30-16:50

An Exact and Efficient Solution to the Queuing System GI/Geo^{a,b}/c

Mohan Chaudhry, Jing Gai*
Royal Military College of Canada, Kingston, ON, Canada
email: *Jing.Gai@rmc.ca

In this paper, we extend the techniques we have used to analyze the continuous queueing system GI/M^{a,b}/c to the discrete system GI/Geo^{a,b}/c, a complex multiple-server steady-state bulk-service queueing system with general independent interarrival times and geometrically distributed service times adhering to a quorum. The introduction of this quorum increases the complexity and leads to a model which has not been studied previously. We obtain the probability distribution at the pre-arrival epoch and outside observer's observation epoch for the early-arrival system.

IS 45
Inference for Complex Data
Rm -1.64
Wed
11:30 - 12:00

Robust Model-Based Clustering with Determinants and Shapes Constraints

Luis Angel García-Escudero*,†, Agustín Mayo-Iscar*
*University of Valladolid, Valladolid, Spain
email: †lagarcia@eio.uva.es

Maximizing (classification and mixture) likelihoods is the approach often applied in model-based clustering. Under standard normal assumptions, these likelihood maximizations are mathematically ill-posed problems without any constraint on the components' scatter matrices. Additionally, non-interesting or "spurious" solutions are often wrongly detected by traditional CEM and EM algorithms designed for them. Establishing an upper bound on the ratio between the largest and smallest determinants for the components' scatter matrices is apparently a sensible way to overcome these problems. Unfortunately, this type of constraints, although affine equivariant, does not always solve these degeneracy issues. A methodology is proposed that allows the simultaneous control of the determinants and shape matrices. The proposed methodology is based on (almost) affine equivariant constraints and it serves to address these troubles. A very flexible robust methodology can be obtained by allowing trimming a fixed fraction of most outlying observations. A computationally feasible algorithm is proposed for both non-robust and robust approaches.

Acknowledgement. Research by partially supported by the Spanish Ministerio de Economía y Competitividad y fondos FEDER, grant MTM2017-86061-C2-1-P, and by the Consejería de Educación de la Junta de Castilla y León, grant VA005P17.

On Search Games

IS 56
Stochastic Game Theory
Rm -1.64
Tue
10:30 - 11:00

Tristan Garrec

Toulouse School of Economics
email: tristan.garrec@ut-capitole.fr

In a search game, two players, a searcher and a hider, act on a search space. The searcher typically intends to minimize the time needed to locate the mobile or immobile hider. The searcher choses a trajectory in the search space and detects the hider when he is sufficiently close to him. We will present variants of search games for which we compute the value and optimal strategies of the players depending on the geometry of the search space.

Multivariate spatial scan statistics for spatial correlated data

IS 43
Scan Statistics and Applications
Rm 100/A
Wed
17:30 - 18:00

Michael Genin[*,§], Lionel Cucala[†], Caroline Lanier[‡], Florent Occelli[‡]

[*]Santé Publique: épidémiologie et qualité des soins EA 2694, Université Lille, France.
[†]Institut Montpelliérain Alexander Grothendieck, Université de Montpellier, France.
[‡]IMPact de l'Environnement Chimique sur la Santé humaine EA 4483, CHU Lille, Institut Pasteur de Lille, Université Lille 2, France.
email: [§]michael.genin@univ-lille.fr

A new spatial scan statistic is proposed for multivariate data indexed in space. Like many other scan methods, it relies on a generalized likelihood ratio but it also takes into account the correlations between variables. This spatial scan test seems to be more powerful than the independent version, whatever the level of correlation between variables. We apply this method to a data set recording the metal pollutants in the area of Lille, France.

References

Kulldorff, M., Huang, L., Pickle, L. and Duczmal, L. (2006). An elliptic spatial scan statistic. *Statistics in Medicine* **25,** 3929–3943.

Kulldorff, M., Huang, L. and Konty, K. (2009). A scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics*, **8:** 58.

Kulldorff, M. (2007) Multivariate spatial scan statistics for disease surveillance. *Statistics in Medicine*, **26**, p.1824-1833.

IS 29
Group Testing
Rm -1.63
Wed
17:00 - 17:30

Smart elements in combinatorial group testing problems

Dániel Gerbner*, Máté Vizer
Alfréd Rényi Institute of Mathematics, Budapest, Hungary
email: *gerbner.daniel@renyi.mta.hu

In group testing problems the questioner needs to find some special elements in a finite underlying set X by testing subsets of X. Here we consider a new variant, where the elements are smart - each of them knows the answers to the tests that contain it, and is able to deduce information from them. We study several models depending on what the elements should or should not figure out. We only consider the combinatorial version of them, thus we are interested in the worst case query complexity when we know there are exactly d defective elements, and focus on the non-adaptive version, where all the queries have to be asked in advance (i.e. the questioner has to ask a family of subsets of X).

The most natural model is where each element should be able to identify the only defective one. Using classical results of extremal set theory we prove that if $\mathcal{F}_n$ solves this on an n-element underlying set and has minimal cardinality, then

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n|}{\log_2 n} = \log_{(3/2)} 2.$$

We also consider related models inspired by secret sharing models, where the elements should share information among them to find the defectives, and the adaptive versions of the different models are investigated.

IS 6
Affine & Polynomial Proc.
Harmony Inge
Wed
17:00 - 17:30

Moment explosions in the rough Heston model

Stefan Gerhold*, Christoph Gerstenecker, Arpad Pinter
TU Wien
email: *sgerhold@fam.tuwien.ac.at

We show that the moment explosion time in the rough Heston model [Euch and Rosenbaum 2016] is finite if and only if it is finite for the classical Heston model. Upper and lower bounds for the explosion time are established, as well as an algorithm to compute the explosion time (under some restrictions). We show that the critical moments are finite for all maturities. For negative correlation, we apply our algorithm for the moment explosion time to compute the lower critical moment. The analysis is based on the fractional Riccati equation and an equivalent Volterra integral equation, which are satisfied by the characteristic exponent.

References

Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. 2016, arXiv: 1609.02108.

Bivariate Kumaraswamy Models via Modified FGM Copulas: Properties and Applications

IS 52
Distributio
Theory
Rm -1.64
Thu
11:00 - 11:30

Indranil Ghosh

University of North Carolina, Wilmington
email: ghoshi@uncw.edu

A copula is a useful tool for constructing bivariate and/or multivariate distributions. In this article, we consider a new modified class of FGM (Farlie–Gumbel–Morgenstern) bivariate copula for constructing several different bivariate Kumaraswamy type copulas and discuss their structural properties, including dependence structures. It is established that construction of bivariate distributions by this method allows for greater flexibility in the values of Spearman's correlation coefficient, ρ and Kendall's τ.

A unifying framework for entropy-based goodness-of-fit tests, with R-package and application to DNA replication.

IS 34
Advances in Bio-statistics
Rm -1.63
Mon
16:00 - 16:30

Valérie Girardin[*,¶], Justine Lequesne[†,§], Philippe Regnault[‡,‖]

[*]Laboratoire de Mathématiques Nicolas Oresme, UMR 6139 CNRS, Université de Caen Normandie, Campus II, BP5186, 14032 Caen, France
[†]Service de Recherche Clinique, Centre Henri Becquerel, Rouen, rue d'Amiens, CS11516, 76038 Rouen Cedex1, France
[‡]Laboratoire de Mathématiques de Reims, EA 4535, Université de Reims Champagne-Ardenne, BP 1039, 51687 Reims Cedex 2, France
email: [§]justine.lequesne@chb.unicancer.fr [¶]valerie.girardin@unicaen.fr
[‖]philippe.regnault@univ-reims.fr

The goodness-of-fit tests based on Shannon entropy introduced in [4] and the goodness-of-fit tests based on relative entropy introduced in [3] are here unified as a unique procedure. Both tests are known to have good power properties and to lead to straightforward computations. While Vasicek's procedure is widely used in the literature, Song's has remained more confidential.

Mathematical justification of both tests is detailed here [1], leading to show their equivalence for testing parametric composite null hypothesis of maximum entropy distributions. For testing any other distribution, the tests based on relative entropy are still reliable goodness-of-fit tests whereas the tests based on entropy become mere tests of entropy level. Moreover, for simple null hypothesis, only the former ones are of use.

In addition, we have developed a package called *vsgoftest* for the statistical software R [2]. It provides an easy implementation of these goodness-of-fit tests for numerous families of maximum entropy distributions, including, e.g., Pareto, Fisher, Weibull distributions.

The methodology is applied to a real dataset of a DNA replication process, issued from a collaboration with biologists. The objective is to validate an experimental protocol to detect chicken cell lines for which the spatio-temporal program of DNA replication is not correctly executed. We propose a two-step approach through entropy-based tests. First, a Fisher distribution with non-integer parameters is retained as reference, and then the experimental protocol is validated.

References

[1] Valérie Girardin and Justine Lequesne. Entropy-based goodness-of-fit tests—a unifying framework: Application to DNA replication. *Communications in Statistics - Theory and Methods*, 2017. doi: 10.1080/03610926.2017.1401084.

[2] Justine Lequesne and Philippe Regnault. Goodness-of-fit tests based on entropy : package vsgoftest. Work in progress.

[3] Kai-Sheng Song. Goodness-of-fit tests based on Kullback-Leibler discrimination information. *IEEE Transactions on Information Theory*, 48(5):1103–1117, 2002.

[4] Oldrich Vasicek. A test for normality based on sample entropy. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38:54–59, 1976.

CS 15
State Space and Markov Models
Rm 0.99
Thu
10:30 - 10:50

Closed Form Expressions for Rescaled Entropy Rates. Application to Markov Chains

Valérie Girardin[*,§], Loick Lhote[†,¶], Philippe Regnault[‡,‖]

[*]Laboratoire de Mathématiques Nicolas Oresme, UMR 6139 CNRS, Université de Caen Normandie, Campus II, BP5186, 14032 Caen, France
[†]ENSICAEN, GREYC, UMR 6072, Université de Caen Normandie, BP5186, 14032 Caen, France
[‡]Laboratoire de Mathématiques de Reims, FRE 2011, Université de Reims Champagne-Ardenne, BP 1039, 51687 Reims Cedex 2, France
email: [§]valerie.girardin@unicaen.fr [¶]loick.lhote@unicaen.fr
[‖]philippe.regnault@univ-reims.fr

We study entropy rates of random sequences for general entropy functionals including the classical Shannon and Rényi and the more recent Tsallis and Sharma-Mittal ones.

We obtain an explicit formula for the classical time averaged entropy rate for a large class of entropy functionals, as soon as the process satisfies a regularity property known in dynamical systems theory as the quasi-power property. Independent and identically distributed sequences of random variables naturally satisfy this property. Markov chains are proven to satisfy it too, under simple explicit conditions on their transition probabilities. All the entropy rates under study are thus shown to be either infinite or zero except at a threshold where they are equal to Shannon or Rényi entropy rates up to a multiplicative constant; see [1].

Based on rescaling by some suitable sequence instead of the number of time units, the usual notion of entropy rate is then extended to define and determine meaningful generalized entropy rates. A suitable rescaling is naturally induced by the asymptotic behavior of the sequence of marginal entropies. Closed form formulas are obtained as soon as it behaves like powers of some analytical functions. A wide class of countable Markov chains is proven to satisfy this property; see [2].

The rates are first expressed in terms of Perron-Frobenius eigenvalues of perturbations of the transition matrices. This leads to a classification of generalized entropy functionals into five exclusive types. Then, a weighted expression is obtained in which the associated Perron-Frobenius eigenvectors play the same role as the stationary distribution in the well-known weighted expression of Shannon entropy rate. Finally, all terms are shown to bear a meaning in terms of dynamics of an auxiliary absorbing Markov chain; see [4] and [3].

References

[1] Gabriela Ciuperca, Valerie Girardin, and Loïck Lhote. Computation and estimation of generalized entropy rates for denumerable Markov chains. *IEEE Transactions on information theory*, 57(7):4026–4034, 2011.

[2] Valérie Girardin and Loïck Lhote. Rescaling entropy and divergence rates. *IEEE Transactions on Information Theory*, 61(11):5868–5882, 2015.
[3] Valérie Girardin, Loïck Lhote, and Philippe Regnault. Different closed-form expressions for generalized entropy rates of Markov chains. Submitted.
[4] Philippe Regnault, Valérie Girardin, and Loïck Lhote. Weighted closed form expressions based on escort distributions for Rényi entropy rates of Markov chains. In *International Conference on Geometric Science of Information*, pages 648–656. Springer, 2017.

Robust Scan Statistics for Detecting a Local Change in Population Mean for Normal Data

IS 42
Scan Stats – Methods...
Rm 100/A
Thu
10:30 - 11:00

Qianzhu Wu*, Joseph Glaz

University of Connecticut, Storrs
email: *joseph.glaz@uconn.edu

In this talk the performance of robust scan statistics based on moving medians, for one and two dimensional normal data, will be discussed. When a local change in the population mean has not occurred and outliers are not present in the data, approximations are derived for the tail probabilities of fixed window scan statistics based on moving medians. The performance of proposed robust scan statistics are evaluated and compared to the performance of scan statistics based on moving sums. Numerical results based on a simulation study indicate that in presence of outliers the scan statistics based on moving medians outperform the scan statistics based on moving sums, in terms of achieving more accurately the specified probability of type I error. The performance of multiple window scan statistics based on moving medians, for detecting a local change in the population mean for one or two dimensional normal data, in presence of outliers, when the size of the window where a change has occurred is unknown, is investigated as well.

The collision spectrum of Λ-coalescents

CS 18
Theoretical Probability
Rm -1.63
Mon
17:00-17:20

Alexander Gnedin*,§, Alexander Iksanov†, Alexander Marynych†, Martin Möhle‡

*Queen Mary, University of London, UK
†Taras Shevchenko National University of Kyiv, Ukraine
‡University of Tübingen, Germany
email: §a.gnedin@qmul.ac.uk

Λ-coalescents introduced in 1995 by J. Pitman and S. Sagitov model the evolution of a coalescing system in which any number of blocks randomly sampled from the whole may merge into a larger block. For the coalescent restricted to initially n singletons we study the collision spectrum $(X_{n,k} : 2 \le k \le n)$, where $X_{n,k}$ counts, throughout the history of the process, the number of collisions involving exactly k blocks. Our focus is on the large n asymptotics of the joint distribution of bluethe $X_{n,k}$'s, as well as on functional limits for the bulk of the spectrum for simple coalescents. Similarly to the previous studies of the total number of collisions X_n, the asymptotics of the collision spectrum largely depends on the behaviour of the measure Λ in the vicinity of 0. We obtain results on the collision spectrum in all situations, where the limit distribution of X_n has been previously known Gnedin, A., Iksanov, A. and Marynych, A. [2014]. In particular, for beta(a, b)-coalescents different types of limiting distributions occur depending on whether $0 < a \le 1$, $1 < a < 2$, $a = 2$ or $a > 2$.

References

Gnedin, A., Iksanov, A. and Marynych, A. (2014) Λ-coalescents: A survey. *J. Appl. Probab.*, **51A**, 23–40.
Gnedin, A., Iksanov, A., Marynych, A. and Möhle, M. (2017) The collision spectrum of Λ-coalescents. arXiv: 1708.03938

IS 10
Copula Functions
Rm 100/B
Wed
11:00 - 11:30

Convolution autoregressive processes

Fabio Gobbi
University of Bologna
email: fabio.gobbi@unibo.it

We introduce a modified version of a standard first-order autoregressive process where we allow for a dependence structure between the state variable Y at the time $t-1$ and the next innovation. We call this model convolution-based AR(1) process (C-AR(1)). We analyze the moment and temporal dependence properties of the new model. After proving that the OLS estimator does not consistently estimate the autoregressive parameter, we provide a proof that our C-AR(1) process is beta-mixing.

IS 51
Probability models
Rm 0.87
Wed
11:00 - 11:30

Covering, Packing, and Generalizations of Classical Probability Problems

Anant Godbole[*,¶], Thomas Grubb[†], Kyutae Han[‡], Bill Kay[§]
[*]East Tennessee State University, Johnson City, USA
[†]University of California, San Diego, USA
[‡]University of California, Los Angeles, USA
[§]Emory University, Atlanta, USA.
email: [¶]godbolea@mail.etsu.edu

Using standard methods (due to Janson, Stein-Chen, and Talagrand) from probabilistic combinatorics, we explore the following general theme: As one progresses from each member of a family of objects $\mathcal{A}$ being "covered" by at most one object in a random collection $\mathcal{C}$, to being covered at most λ times, to being covered at least once, to being covered at least λ times, a hierarchy of thresholds emerge. The best example of this is the classical "balls in boxes" situation. We will then see how the results change according to the context, and level of dependence introduced. Examples will be from extremal set theory; combinatorics; and additive number theory. In fact we will argue that many of the results run in parallel with the classical birthday, coupon collector, and other problems from classical applied probability. Of special note will be what we call the log log phenomenon.

Regularized estimation of hazard interaction effects for age-period-cohort analysis

CS 12
Epidemiological and Medical Applications
Rm 100/A
Mon
17:00-17:20

Vivien Goepp[*,‡], Grégory Nuel[†], Olivier Bouaziz[*]

[*]MAP5 (CNRS 8145), Université Paris Descartes, Paris, France
[†]LPSM (CNRS 8001), Sorbonne Université, Paris, France
email: [‡]vivien.goepp@parisdescartes.fr

In epidemiological studies with variable age at onset, a typical quantity of interest is the incidence of a disease, for example the cancer incidence. In such studies, the data are usually reported in the form of registers which contain the number of observed cases and the number of individuals at risk to contract the disease. These data are usually highly heterogeneous in terms of dates of birth (the cohort) and with respect to the calendar time (the period) and appropriate estimation methods are needed.

The model of choice in this context is the age-period-cohort model [Clayton and Schifflers 1987b]. But it suffers from well-known unidentifiability issues due to the relation period = age + cohort [Holford 1983, Carstensen 2007]. The three two-factor models age-cohort, age-period, and period-cohort offer to infer the effects of two of the three variables and are comparatively simpler [Clayton and Schifflers 1987a]. However, both these models have a strong limitation. Since they assume the same effect of one variable for all values of the other variables, they do not allow for interactions between effects to be estimated.

In this talk, we propose a new model that extends the two-factor models and includes an interaction term in order to take into account possible interactions. For instance, the age-cohort model is extended such that a term of interaction between age and cohort is estimated jointly with the marginal effects:

$$\log(\lambda(a,c)) = \alpha(a) + \beta(c) + \delta(a,c),$$

where $\lambda(a,c)$ is the hazard rate at age a and cohort c. The age and cohort effects are estimated by α and β respectively and δ is the interaction term.

This model is fitted on discretized values of the age and cohort variables. To overcome classical overfitting issues, $\delta(a,c)$ is subject to regularization over its first order differences. Estimation is then performed by maximizing the penalized likelihood:

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) - \frac{\text{pen}}{2} \sum_{a,c} v_{a,c} \left(\delta(a,c) - \delta(a-1,c)\right)^2 + w_{a,c} \left(\delta(a,c) - \delta(a,c-1)\right)^2$$

We present two types of regularization of the interaction term. A first possibility is to set $v_{a,c} = w_{a,c} = 1$, which corresponds to penalizing over the L_2 norm of the first order differences of $\delta(a,c)$. It provides a smooth estimate of the interaction term. Another possibility is to iteratively adapt the values of the weights using the previously estimated values of $\delta(a,c)$. This procedure, called adaptive ridge [Rippe and Meulman and Eilers 2012, Frommlet and Nuel 2016], allows for an approximation of the L_0 norm using continous functions, which makes inference computationally tractable. Regularizing on the L_0 norm of the first order differences of $\delta(a,c)$ provides a piecewise constant estimation of the interaction term, which allows for precise detection of shift points in the hazard rate. The choice of the penalty constant is made using a bayesian criterion.

Our new method is validated on simulations and applied to cancer incidence data.

References

B. Carstensen. Age–period–cohort models for the lexis diagram. *Statistics in medicine*, 2007.

D. Clayton and E. Schifflers. Models for temporal variation in cancer rates. i: age–period and age–cohort models. *Statistics in medicine*, 1987.

D. Clayton and E. Schifflers. Models for temporal variation in cancer rates. ii: age–period–cohort models. *Statistics in medicine*, 1987.

F. Frommlet and G. Nuel. An adaptive ridge procedure for L0 regularization. *PloS one*, 2016.

T. Holford. The estimation of age, period and cohort effects for vital rates. *Biometrics*, 1983.

R. Rippe, J. Meulman, and P. Eilers Visualization of genomic changes by segmented smoothing using an L0 penalty. *PloS one*, 2012.

IS 30
Actuarial Risk Models I
Rm 0.99
Mon
15:30 - 16:00

Fraud risk assessment within blockchain transactions

Pierre-Olivier Goffard

University of California, Santa Barbara
email: pierre.olivier.goffard@gmail.com

The probability of successfully spending twice the same bitcoins is considered. A double-spending attack consists in issuing two transactions transferring the same bitcoins. The first transaction, from the fraudster to a merchant, is included in a block of the public chain. The second transaction, from the fraudster to himself, is recorded in a block that integrates a private chain, exact copy of the public chain up to substituting the fraudster-to-merchant transaction by the fraudster-to-fraudster transaction. The double-spending hack is completed once the private chain reaches the length of the public chain, in which case it replaces it. The growth of both chains are modeled by two independent counting processes. The probability distribution of the time at which the malicious chain catches up with the honest chain, or equivalently the time at which the two counting processes meet each other, is studied. The merchant is supposed to await the discovery of a given number of blocks after the one containing the transaction before delivering the goods. This grants a head start to the honest chain in the race against the dishonest chain.

IS 7
Optimisation & Machine Learning
Harmony Inge
Wed
11:30 - 12:00

Lévy forward price approach for multiple yield curves in presence of persistently low and negative interest rates

Ernst Eberlein, Christoph Gerhart, Zorana Grbac*

Department of Mathematical Stochastics, University of Freiburg
Department of Quantitative Finance, University of Freiburg
Laboratoire de Probabilités, Statistique et Modélisation, University Paris Diderot
email: *grbac@math.univ-paris-diderot.fr

In this talk we present a framework for discretely compounding interest rates which is based on the forward price process approach. This approach has a number of advantages, in particular in the current market environment. Compared to the classical Libor market models, it allows in a natural way for negative interest rates and has superb calibration properties even in the presence of persistently low rates. Moreover, the measure changes along the tenor structure are simplified significantly. This property makes it an excellent base for a post-crisis multiple curve setup. Two variants for multiple curve constructions will be discussed. As driving processes we use time-inhomogeneous Lévy

processes, which lead to explicit valuation formulas for various interest rate products using well-known Fourier transform techniques. Based on these formulas we present calibration results for the two model variants using market data for caps with Bachelier implied volatilities.

Mexican hat coupling of quasi-cycle oscillators produces quasi-patterns

IS 33
Neuron Models, Neural Fields
Rm -1.63
Mon
12:00 - 12:30

Priscilla (Cindy) Greenwood

University of British Columbia, Vancouver
email: pgreenw@math.ubc.ca

A family of stochastic processes has quasi-cycle oscillations if its otherwise-damped oscillations are sustained by noise. Such a family forms the reaction part of a stochastic reaction-diffusion system when we insert a local Mexican Hat-type, difference of Gaussians, coupling on a one-dimensional or two-dimensional lattice. We find spatial patterns of oscillating quasi-cycles that resemble Turing patterns, called quasi-patterns. Specific properties of these patterns, such as local phase synchronization, can be predicted from the parameters of the reaction and of the Mexican Hat coupling. When the damping parameters of the reaction and diffusion parts are small and balanced, phase synchronization vanishes but amplitude patterns persist. These results extend our knowledge of the behaviour of coupled neural field equations and its dependence on stochastic fluctuations. This is joint work with Lawrence M Ward.

Comparison of queueing systems with various rules of service and regenerative input flow

OCS 3
Branching Random Walks
Rm -1.62
Mon
17:20-17:40

S. Grishunina

Department of Probability, Faculty of Mathematics and Mechanics, Lomonosov Moscow State University, Moscow, Russian Federation
Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, Moscow, Russian Federation
email: svetagri@live.ru

We study stability conditions of the multiserver queueing system in which each customer requires a random number of servers simultaneously. The input flow is supposed to be a regenerative one and service time is the same on all occupied servers and equals a constant τ. Discipline is FIFO. We define an auxiliary service process $Z(t)$ that is the number of completed services by all m servers during the time interval $(0, t)$ under the assumption that there are always customers in the system. Then we construct the sequence of common regeneration points for the regenerative input flow and the auxiliary service process. It allows us to deduce the stability condition of the model under consideration. We compare this stability condition with the stability conditions of the systems where service time has an exponential, phase-type or hyper-exponential distribution.

Acknowledgement. Work is partially supported by Russian Foundation for Basic Research grant 17-01-00468.

IS 1
Stochastic Methods in Finance
Harmony lnge
Tue
14:30 - 15:00

Integration by parts on the law of the modulus of the Brownian bridge

Martin Grothaus*, Robert Vosshall

TU Kaiserslautern, Kaiserslautern, Germany
email: *grothaus@mathematik.uni-kl.de

We prove an infinite dimensional integration by parts formula on the law of the modulus of the Brownian bridge from 0 to 0. The main motivation for all this is the construction of an SPDE whose invariant measure would be the law of the reflecting Brownian bridge, a problem which is still open despite the recent fantastic advances in very difficult SPDEs, thanks to regularity structures and, or paraproducts. It seems that the SPDE which motivates this integration by parts formula is even more difficult than KPZ, since it contains a local time which is not covered by the new theories yet.

IS 33
Neuron Models, Neural Fields
Rm -1.63
Mon
11:30 - 12:00

Deterministic Principles and Stochastic Solutions to Estimate Synaptic Conductances

Antoni Guillamon

Departament de Matemàtiques, Universitat Politècnica de Catalunya.
email: antoni.guillamon@upc.edu

We will discuss on the mathematical problem of estimating the time courses of excitatory and inhibitory conductances of currents arriving to a specific neuron, knowing only the output (membrane voltage) presented by the neuron. Due to the characteristics of most accepted neuron models, this problem entails challenges that can be partially addressed by tools from bifurcation theory of nonlinear dynamical systems and stochastic and probabilistic methods (mainly, MLE-based and Bayesian inference). We will show recent progresses obtained with different collaborators (see references) achieved by combining both approaches; they allow to improve estimations when the neuron is not spiking and to address a framework to deal with the more difficult problem of estimating the synaptic conductances when the neuron is spiking.

Acknowledgement. This work has been partially funded by the Spanish grant MINECO/FEDER MTM2015-71509-C2-2-R and the Catalan grant AGAUR 2017SGR-1049.

References

Catalina Vich, Rune W. Berg, Antoni Guillamon, Susanne Ditlevsen (2017). Estimation of synaptic conductances in presence of nonlinear effects caused by subthreshold ionic currents. *Frontiers in Computational Neuroscience* **11**, 69.

Antoni Guillamon, Rafel Prohens, Antonio E. Teruel, Catalina Vich (2017). Estimation of synaptic conductances in the spiking regime for the McKean neuron model. *SIAM J Applied Dynamical Systems* **16**, 1397–1424.

Pau Closas, Antoni Guillamon (2017). Sequential estimation of intrinsic activity and synaptic input in single neurons by particle filtering with optimal importance density. *EURASIP Journal on Advances in Signal Processing*, **65**, 1–22.

Catalina Vich, Antoni Guillamon (2015). Dissecting estimation of conductances in subthreshold regimes. *Journal of Computational Neuroscience* **39**, 271–287.

Controlled Bisexual Branching Processes with Random Control Function

CS 5
Branching Processes
Rm -1.62
Tue
17:50-18:10

Cristina Gutiérrez[*,†,‡], Miguel González[*,†,§], Rodrigo Martínez[*,†,¶]

Department of Mathematics, University of Extremadura, Badajoz (Spain)[*]
Instituto de Computación Científica Avanzada, University of Extremadura, Badajoz (Spain) [†]
email: [‡]cgutierrez@unex.es [§]mvelasco@unex.es [¶]rmartinez@unex.es

The standard bisexual branching process introduced by Daley [1968] is a discrete-time process which models the evolution of a two-sex population in which there are females and males that form couples in order to produce offspring. This process becomes extinct or grows to infinity without any restriction. This fact makes the model unsuitable for applications to real biological situations.

In this work, we assume that the total population size is limited by the environment and cannot grow indefinitely. With the aim of studying the evolution of such population, we introduce a controlled bisexual branching process with random control function at mating time. This function will depend on the number of couples initially formed in the generation and on an associated survival probability which will be related to the carrying capacity of the environment. In this way, we can control the total number of couple which can produce offspring.

For this model, we show the behaviour of the process in long term by means of simulations and we present some results about the extinction of the population.

Acknowledgement. The research was supported by the grant IB16103 (Junta de Extremadura and Fondo Europeo de Desarrollo Regional, UE).

References

D.J. Daley(1968) Extinction conditions for certain bisexual Galton-Watson branching processes. Z. Wahrscheinlichkeitsth. **9**, 315–322

Characterizations of Probability Distribution Via the Concept of Sub-Independence

IS 52
Distributio
Theory
Rm -1.64
Thu
10:30 - 11:00

G.G. Hamedani

Department of Mathematics, Statistics and Computer Science Marquette University Milwaukee, WI 53201-1881
email: gholamhoss.hamedani@marquette.edu

Limit theorems as well as other well-known results in probability and statistics are often based on the distribution of the sums of independent (and often identically distributed) random variables rather than the joint distribution of the summands. Therefore, the full force of independence of the summands will not be required. In other words, it is the convolution of the marginal distributions which is needed rather than the joint distribution of the summands which in the case of independence, is the product of the marginal distributions. In this talk, we establish characterizations of the probability distribution based on the concepts of sub-independence, conditionally sub-independence and max-sub-independence.

OCS 7
Association and Default
Harmony lnge
Tue
16:30-16:50

How much should you trust your Least Squares Method as a risk manager?

Norbert Hári

Morgan Stanley, Budapest, Hungary
email: Norbert.Hari@morganstanley.com

In a Monte Carlo framework the least squares method (LSM) is a widely used and very efficient approximation technique when one would like to compute the value of an early exercise option for pricing, or future exposures in a counterparty risk setting. In the latter we are not only interested in calculating the expectation of future exposures (i.e. Expected Positive Exposure (EPE)), for risk management purposes we also heavily rely on quantiles (i.e. Potential Future Exposure (PFE)) of the distribution. In this paper we compare the future exposure distribution of equity options calculated by the LSM to alternatives. We consider the novel stochastic grid bundling method (SGBM) and analytical solutions if available, and investigate the exposure metric (EPE, PFE) differences between the methods for selected option maturities and moneyness levels.

IS 14
Excursions of Lévy Processes
Rm 100/B
Mon
11:00 - 11:30

High Excursion Probabilities of Correlated Brownian motion

Krzysztof Dębicki, Enkelejd Hashorva*, Dmitry Korshunov, Zbigniew Michna

University of Wrocław
University of Lausanne
Lancanster University
Wrocław University of Economics
email: *enkelejd.hashorva@unil.ch

Let $X(t) = (X_1(t) \dots X_d(t)), t \in [0,T]$ be a separable vector-valued random processes and define

$$p_a := \mathbb{P}\{\exists t \in [0,T] : X_1(t) > a_1 u, \dots X_d(t) > a_d u\}$$

for given constants $a_1, \dots, a_d$ and $u > 0$. The probability p_u appears in various investigations related for instance with the calculation of simultaneous ruin probability or the conjunction probability. When X is Gaussian, logarithmic asympototics of p_u is known under general assumptions on the covariance function of X. In the special case that the components of X are independent, exact asymptotics of p_u as $u \to \infty$ can be derived for both stationary and non-stationary X. The crucial tool here is the Gordon inequality, which is however not available in case that X has dependent components. Due to the lack of appropriate tools in the vector-valued settings, currently in the literature there are no theoretical developments that cover the extremes of vector-valued Gaussian processes. In this talk, we shall discuss first what are the main difficulties in the general d-dimensional setup and then focus on the tractable case of the correlated Brownian motion. We shall give explicit approximations of p_u for both the finite and infinite time horizon. Further we shall discuss the approximation of the Shepp-statistics and that of the ruin times.

Stochastic Modeling of Water Quality Seasonality to Estimate Riverine Phytoplankton Biomass

CS 1
Biological Statistics
Rm 100/A
Wed
12:20 - 12:40

István Gábor Hatvani[*,¶], József Kovács[†,‖], Péter Tanos[‡,**], Gábor Várbíró[§,††]

[*]Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, MTA, Budapest, Hungary
[†]Eötvös Loránd University, Department of Physical and Applied Geology, Budapest, Hungary
[‡]Szent István Universty, Department of Informatics, Gödöllő, Hungary
[§]Danube Research Institute, Department of Tisza River Research, Centre for Ecological Research, MTA, Budapest, Hungary
email: [¶]hatvaniig@gmail.com [‖]kevesolt@geology.elte.hu [**]tanospeter@gmail.com
[††]varbiro.gabor@okologia.mta.hu

The development of models to estimate environmental variables is especially important in aquatic systems, e.g. where measurements of primary production are lacking. Morlet wavelet spectrum- and multiple linear regression analyses were conducted on 15 water quality variables sampled at 14 monitoring sites along the Hungarian section of the River Tisza and 4 sites from artificial tributary channels for 1993-2005. Interestingly, the presence of annual periodicity was found to vary over space and time. In general, an increase was observed in the company of higher trophic states of the river heading downstream. Water quality variables were combined into 7 different groups (nutrients, ions, etc.) and their periodicity indices were computed. By exploiting their spatial distribution along the studied river section, multiple regression models were derived. In the models the periodicity indices of the groups of water quality variables served as independent-, and the proxy of primary production (chlorophyll-a) served as the dependent variable. In fine, an improved model was constructed which was capable of explaining about half (adjusted $R^2 = 0.5$) of the variance of primary production in the study area.

Acknowledgement. The authors are grateful for the support of the Bolyai János Scholarship and the Hungarian Ministry of Human Capacities (NTP-NFTÖ-17B-B-0608).

The role of migration in speciation: linking micro- and macro-evolution

CS 5
Branching Processes
Rm -1.62
Tue
17:10-17:30

Pablo Duchen[*,§], Sophie Hautphenne[†,‡,¶], Laurent Lehmann[*,‖], Nicolas Salamin[*,**]

[*]University of Lausanne, Lausanne
[†]EPFL, Lausanne
[‡]The University of Melbourne, Melbourne
email: [§]pduchen@gmail.com [¶]sophiemh@unimelb.edu.au [‖]laurent.lehmann@unil.ch
[**]nicolas.salamin@unil.ch

The process of speciation is of key importance in evolutionary biology because it shapes overall macroevolutionary patterns. Still, any speciation process starts at the microevolutionary level, for instance, when two diverging populations drift towards different phenotypic optima. The speed at which these optima are reached is controlled by 1) the selection strength, which pushes a mean trait towards an optimum, and 2) ongoing migration that pulls the mean phenotype away from that optimum. Traditionally, phenotypic evolution with selection has been modelled by Ornstein-Uhlenbeck (OU) processes at the macroevolutionary level, but these models have ignored the role of migration within species.

Here, our goal is to reconcile the processes of micro and macroevolution by modelling migration during speciation. More precisely, we introduce an OU model where migration happens between two subpopulations and this migration decreases over time as it happens during speciation. We then use this model to study the evolution of trait means along a phylogeny, as well as the way phenotypic disparity between species changes with successive epochs.

IS 23
Extremes and Bursts
Rm -1.62
Thu
10:30 - 11:00

Extreme events of observations separated by random waiting times in the presence of dependencies

Katharina Hees[*,§], Peter Scheffler[†,¶], Peter Straka[‡,‖]

[*]Faculty of Statistics, TU Dortmund University
[†]Department of Mathematics, University of Siegen
[‡]School of Mathematics and Statistics, UNSW Sydney
email: [§]Hees@statistik.uni-dortmund.de [¶]Scheffler@mathematik.uni-siegen.de
[‖]p.straka@unsw.edu.au

Classical Extreme Value theory deals with the extreme values of observations of iid random variables $X_1, X_2, \ldots$ and admits that these events occur at intervals of fixed time length or with exponentially distributed inter-arrival times. Recently a new model called Continuous Time Random Maxima, also called max-renewal process or shock process, has been introduced. This model generalizes the temporal dynamics of extreme value theory by assuming random waiting times between observations. If the waiting times are heavy-tailed, this theory provides a model for 'bursty' events. In this talk, we will consider such a model in the case where the observations and the waiting times are dependent. We will have a look at the scaling limit of the process that tracks the maximum as well the process of the exceedance times. Furthermore, we want to discuss, how these limits change in the presence of dependencies in the sequence of waiting times as well as in the sequence of magnitudes.

References

Hees, K., Nayak, S. and Straka, P. (2018) Inference for Continuous Time Random Maxima with Heavy-Tailed Waiting Times. arXiv: 1802.05218

Hees, K. and Scheffler, H. P. (2017) Coupled Continuous Time Random Maxima. *Extremes*, p.1-25.

Hees, K., and Scheffler. H.P. (2016) On joint sum/max stability and sum/max domains of attraction. arXiv: 1606.03109v2

CS 12
Epidemiological and Medical Applications
Rm 100/A
Mon
17:20-17:40

A simulator for chronic disease progression

Ayman Hijazy[*], András Zempléni

Eötvös Loránd University, Budapest, Hungary
email: [*]hijaayman@gmail.com

An individual who is susceptible to a chronic disease naturally progresses from being disease free to being asymptomatic (preclinical) . This progression is modelled by assuming that the time spent in the disease free and the asymptomatic states are random variables following specified distributions. Early detection may occur if screening takes place before the development of symptoms. The parameters to be estimated are those regarding sensitivity of screening, the preclinical intensity (the

probability of the disease to onset in a given short time interval) and the time spent in the preclinical state.

To get data is hard and costly in such medical scenarios, so a simulator is definitely useful for checking the properties of the proposed estimation methods. That is what we have carried out: based on assumed distributions we gave confidence intervals for estimators and have analyzed the effects of misspecified models.

References

Zelen, M., and Feinleib, M. (1969) On the Theory of Screening for Chronic Diseases. *Biometrika*, **56(3)**, pp. 601–614, doi: 10.2307/2334668

Wu, D., Rosner, G. L. and Broemeling, L. (2005), MLE and Bayesian Inference of Age-Dependent Sensitivity and Transition Probability in Periodic Screening. *Biometrics*, **61**, pp. 1056–1063. doi: 10.1111/j.1541-0420.2005.00361.x

Rare-event probabilities in space-time models for wireless networks

IS 12
Efficient Monte Carlo Simulation
Rm 100/B
Thu
11:00 - 11:30

Christian Hirsch[*,‡], Benedikt Jahnel[†], Paul Keeler[†], Robert I. A. Patterson[†]

[*]Aalborg University, Denmark
[†]WIAS Berlin, Germany
email: [‡]hirsch@math.lmu.de

We present a space-time wireless network model based on a Poisson point process. Participants either communicate directly with each other or via relays. We investigate the probability that an atypically large proportion of users experiences bad quality of service. This situation could be caused either by a lack of connectivity due to strong interference or by a lack of capacity at the relays. We show that as the system size grows, the undesirable outcomes become rare and satisfy a large deviation principle. Ad-hoc importance-sampling schemes show promising signs towards improving the estimation error of rare-event probabilities, so that further research in the direction of more structured approaches has the potential of yielding fast and accurate Monte Carlo estimators.

Estimation of geodesic tortuosity and constrictivity in stationary random closed sets

IS 20
Stochastic Geometry
Rm -1.63
Wed
15:00 - 15:30

Matthias Neumann[*,§], Christian Hirsch[†], Jakub Staněk[‡], Viktor Beneš[‡], Volker Schmidt[*]

[*]Ulm University, Germany
[†]Aalborg University, Denmark
[‡]Charles University in Prague, Czech Republic
email: [§]matthias.neumann@uni-ulm.de

We investigate the problem of estimating geodesic tortuosity and constrictivity as two structural characteristics of stationary random closed sets. They are of central importance for the analysis of transportation paths in materials science. Loosely speaking, geodesic tortuosity measures the windedness of paths whereas the notion of constrictivity captures the appearance of bottlenecks resulting from narrow passages. We first provide mathematically precise definitions of these quantities and introduce appropriate estimators. Second, we show strong consistency of these estimators for growing sampling windows. In order to apply our estimators to real datasets, the extent of edge effects

needs to be controlled. This is illustrated in the example of a multi-phase material model that has been investigated in the materials science literature.

Acknowledgement. The work of MN, JS, VB, VS was funded by the German Science Foundation (DFG) and the Czech Science Foundation (GACR, project number 17-00393J). The work of CH was funded by LMU Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative.

IS 2
Optimal Transport ...
Rm 100/B
Mon
16:30 - 17:00

Learning rough volatility

Blanka Horvath

Imperial College London
email: b.horvath@imperial.ac.uk

Calibration time being the bottleneck for models with rough volatility, we present ways for substantial speed-ups, along every step of the calibration process: In a first step we describe a powerful numerical scheme (based on functional central limit theorems) for pricing a large family of rough volatility models. In a second step we discuss various machine learning methods that significantly reduce calibration time for these models. By simultaneously calibrating several (classical and rough) models to market data, we re-confirm as a byproduct of our calibration results, that volatility is rough, calibration performance being best for very small Hurst parameters in a multitude of market scenarios.

IS 19
Spatial Point Processes
Rm -1.63
Wed
11:00 - 11:30

Disagreement percolation for Gibbs point processes

Christoph Hofer-Temmel[*], Pierre Houdebert[†]

[*]NLDA & CWI, Den Helder & Amsterdam, The Netherlands
[†]Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

In this talk I will present a result proving uniqueness of Gibbs measure at low activities, for general interactions satisfying natural conditions.

This result is in particular applicable for all finite-range interactions and for geometric interactions such as the Continuum Random Cluster model or Quermass-interaction model with unbounded radii.

To prove this result we are using a disagreement percolation technique which is controling the influence of the boundary, "placing it" into a percolation cluster of a dominating Poisson Boolean model.

A Purely Sequential Minimum Risk Point Estimation Methodology Based on Gini's Mean Difference for a Normal Mean

IS 37
Applied Prob. & Stat. Inference II.
Rm 100/A
Mon
15:30 - 16:00

Jun Hu

University of Connecticut - Storrs, USA
email: jun.hu@uconn.edu

We have developed a purely sequential methodology for problems associated with minimum risk point estimation for the normal mean μ when the variance σ^2 is assumed unknown. A new stopping rule is constructed by replacing the sample standard deviation with an appropriate multiple of *Gini's mean difference* (GMD) in defining the conditions for boundary crossing. Asymptotic first-order and second-order properties associated with this new estimation strategy have been investigated. These are followed by summaries obtained from extensive sets of simulations by drawing samples from (i) normal universes or (ii) mixture-normal universes where samples may be reasonably treated as observations from a normal universe in a large majority of simulations. We also include illustrations using the "horticulture data". Overall, we empirically feel confident that our newly developed GMD-based methodology is more robust for practical purposes when we compare them with the one based on sample standard deviation, especially when suspect outliers may be expected.

Acknowledgement. This is joint work with Prof. Nitis Mukhopadhyay.

References

Mukhopadhyay, N. and Hu, J. (2017) Confidence intervals and point estimators for a normal mean under purely sequential strategies involving Gini's mean diference and mean absolute deviation. *Sequential Analysis*, **36(2)**, 210–239.

Using Big Data in Survival Analysis

IS 34
Advances in Biostatistics
Rm -1.63
Mon
15:30 - 16:00

Catherine Huber-Carol

MAP5 University Paris Descartes, Paris, France
email: catherine.huber@parisdescartes.fr

Big Data are now available in many fields of knowledge, comprising Medicine and Biology. Though it is very appealing to provide the huge data bases that are now available in order to try to solve prevention or medication problems, many difficulties arise. In classical statistics, appropriate data are collected in order to solve a specific problem. For example, to test whether a new treatment induces a better survival than the usual one, a number of patients are involved, whose risk factors are observed. The attribution of treatments is done through randomization or matching in order to ensure representativeness of the target population. When using Big Data, the number of factors and the number of people are both huge. There may be an inhomogeneity among them, a high degree of associations between the factors, irrelevance of some of them, and lack of important ones. Two related ways are available to extract the relevant information from Big Data: data reduction (through Singular Value Decomposition, Principal Component Analysis, …) and shrinkage of usual statistical procedures (Ridge Regression, Lasso, …). This introduces a close link between statistical learning and machine learning, which is based on algorithms more than on probabilistic assumptions.

IS 11
Monte Carlo, Importance Sampling...
Rm 100/B
Tue
15:00 - 15:30

Min-max representations of viscosity solutions of Hamilton-Jacobi equations and applications in rare-event simulation

Henrik Hult
Department of Mathematics, KTH

In this paper a duality relation between the Mané potential and the action functional is derived in the context of convex and state-dependent Hamiltonians. The duality relation is used to obtain min-max representations of viscosity solutions of first order evolutionary Hamilton-Jacobi equations. These min-max representations naturally suggest classes of subsolutions of Hamilton-Jacobi equations that arise in the theory of large deviations. The subsolutions, in turn, are good candidates for designing efficient rare-event simulation algorithms. This is joint work with Boualem Djehiche and Pierre Nyquist.

IS 4
(Rough) Volatility Asymptotics
Rm 100/A
Tue
11:30 - 12:00

Pathwise moderate deviations for option pricing

Antoine Jacquier
Imperial College London
email: a.jacquier@imperial.ac.uk

We provide a unifying treatment of pathwise moderate deviations for models commonly used in financial applications, and for related integrated functionals. Suitable scaling allows us to transfer these results into small-time, large-time and tail asymptotics for diffusions, as well as for option prices and realised variances. In passing, we highlight some intuitive relationships between moderate deviations rate functions and their large deviations counterparts; these turn out to be useful for numerical purposes, as large deviations rate functions are often difficult to compute. Joint work with Konstantinos Spiliopoulos (Boston University).

IS 15
Extremes of Gaussian Processes
Rm 100/B
Wed
17:30 - 18:00

Ruin probability for correlated Brownian motions

Krzysztof Dębicki[*], Enkelejd Hashorva[†], Lanpeng Ji[‡,§], Tomasz Rolski[*]
[*]Mathematical Institute, University of Wroclaw, Wroclaw, Poland
[†]Department of Actuarial Science, University of Lausanne, Lausanne, Switzerland
[‡]School of Mathematics, University of Leeds, Leeds, United Kingdom
email: [§]l.ji@leeds.ac.uk

Consider $X(t) - \mu t$, $t \geq 0$, a correlated d-dimensional Brownian motion with drift, where $X(t) = AB(t)$, $A \in \mathbb{R}^{d\times d}$ is a non-singular matrix, $B(t) = (B_1(t), \ldots, B_d(t))^\top$, $t \geq 0$ is a standard d-dimensional Brownian motion with independent coordinates and $\mu = (\mu_1, \ldots, \mu_d)^\top \in \mathbb{R}^d$.

In this talk we are concerned about the asymptotics of probability that in infinite-time horizon, the process $X(t) - \mu t$, $t \geq 0$, starting at point $x_u = (-\alpha_1 u, \ldots, -\alpha_d u)^\top$ with $\alpha_i \in \mathbb{R}$, $1 \leq i \leq d$, $u > 0$,

enters the cone $\prod_{i=1}^{d}[0,\infty)$, that is

$$\begin{aligned} P(u) &= \mathbb{P}_{x_u}\Big\{\exists t \geq 0,\, X(t) - \mu t \in \textstyle\prod_{i=1}^{d}[0,\infty)\Big\} \\ &= \mathbb{P}\big\{\exists t \geq 0,\, \cap_{i=1}^{d}\{X_i(t) - \mu_i t > \alpha_i u\}\big\}, \quad u \to \infty. \end{aligned}$$

Our main result shows that

$$P(u) \sim C_{I,K}\mathcal{H}_I u^{\frac{1-m}{2}} e^{-\frac{g_m}{2}u}, \quad u \to \infty,$$

where $C_{I,K} > 0$, $g_m, m \in \mathbb{N}$ are known constants and $\mathcal{H}_I$ is a multidimensional counterpart of the celebrated Pickands constant that appears in the extreme value theory of Gaussian random fields.

One of the findings of this paper is that the set of indices $\{1,\ldots,d\}$ of the vector-process X can be partitioned into three subsets I, J, K. The index set I determines m, g_m and $\mathcal{H}_I$ in the asymptotics, whereas both I and K determine the constant $C_{I,K}$. Moreover, the set J, whenever non-empty, contains indices that do not play any role in our asymptotic consideration.

On the double boundary non-crossing probability for a class of compound risk processes with applications

CS 9
Stability of financial systems
Harmony Inge
Tue
18:10-18:30

Vladimir K. Kaishev

Cass Business School, City, University of London
email: v.kaishev@city.ac.uk

We present explicit formulas and a numerically efficient method for computing the probability that a non-decreasing, pure jump stochastic risk process will not exit the strip between two time-dependent, possibly discontinuous boundaries, within a finite-time interval. Our stochastic process is a compound risk process with risk events arriving according to an arbitrary point process with conditional stationary independent increments (PPCSII), and risk severities assumed independent identically distributed random variables. The class of PPCSII is rather rich covering point processes with independent increments (among which non-homogeneous Poisson processes and negative binomial processes), doubly stochastic Poisson processes (i.e. Cox processes) and mixed Poisson processes (among which processes with the order statistics property). This makes our framework and results generally applicable for a broad range of models arising in insurance, finance, queuing, economics, physics, astronomy and many other fields. We present extensive (numerical) examples of such applications in statistics (e.g. computing Kolmogorov-Smirnov distribution), ruin theory, inventory management and exotic option pricing.

This talk is based on joint work with Dimitrina Dimitrova, Zvetan Ignatov and Senren Tan.

IS 39
Abraham Wald Prize Ceremony
Harmony Inge
Tue
11:30 - 12:00

Locally most powerful sequential rank tests

Jan Kalina

Institute of Information Theory and Automation of the Czech Academy of Sciences, Prague
Institute of Computer Science of the Czech Academy of Sciences, Prague
email: kalina@cs.cas.cz

Sequential ranks are defined as ranks of such observations, which have been observed so far in a sequential design. This paper studies hypotheses tests based on sequential ranks for different situations. The locally most powerful sequential rank test is derived for the hypothesis of randomness against a general alternative, including the two-sample difference in location or regression in location as special cases for the alternative hypothesis. Further, the locally most powerful sequential rank tests are derived for the one-sample problem and for independence of two samples. All these tests are derived for a fixed sample size and the results bring arguments in favor of existing tests.

In addition, we propose a sequential testing procedure based on these statistics of the locally most powerful tests. Principles of such sequential testing are explained on the two-sample Wilcoxon test based on sequential ranks.

Acknowledgement. The work is supported by the Czech Science Foundation project 17-07384S.

PLENARY
Harmony Inge
Mon
9:30 - 10:30

Competing diffusive particle systems and models of large equity markets

Ioannis Karatzas

Columbia University, New York
email: ikaratzas@icloud.com

We introduce and study stable multidimensional diffusions interacting through their ranks. These interactions give rise to invariant measures which are in broad agreement with stability properties observed in large equity markets over long time-periods. The models we develop assign growth rates and variances that depend on both the name (identity) and the rank (according to capitalization) of each individual asset. Such models are able realistically to capture certain critical features of the observed stability of the capital distribution over the past century, all the while being simple enough to allow for rather detailed analytical study.

The methodologies used in this study touch upon the question of triple points for systems of competing diffusive particles; in particular, some choices of parameters may permit triple (or higher-order) collisions to occur. We show, however, that such multiple collisions have no effect on any of the stability properties of the resulting system. This is accomplished through a detailed analysis of collision local times.

The models have connections with the analysis of Queueing Networks in heavy traffic, and with competing particle systems in Statistical Mechanics (e.g., Sherrington-Kirkpatrick model for spin-glasses). Their hydrodynamic-limit behavior is governed by generalized porous medium equations with convection, and the fluctuations around these limits by appropriate linear stochastic partial differential equations of parabolic type with additive noise; whereas limits of a different kind display phase transitions and are governed by Poisson-Dirichlet distributions. We survey briefly recent progress on some of these fronts, and suggest open problems for further study.

Numerical stochastic model of air temperature and relative humidity periodically correlated joint time-series

OCS 5
Environmental Statistics
Rm -1.63
Tue
12:00 - 12:20

Nina Kargapolova

Novosibirsk State University and Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk
email: nkargapolova@gmail.com

Weather and climate significantly affect many environmental processes. For example, crop yield depends on the level of solar radiation and air temperature, regimes of precipitation influence on a river runoff. It is not always possible to use only real meteorological observations for solution of environmental problems. Rare and extreme weather events often have a significant impact on the processes that occur, but the statistical properties of these events (for example, the average frequency of occurrence, average duration, etc.) are estimated unreliably on basis of real data. In this regard, often, instead of data from real meteorological observations, simulated trajectories of meteorological processes are used. For this purpose, the models should reproduce the behavior of real processes as accurately as possible. In recent decades a lot of scientific groups all over the world work at development of so-called "stochastic weather generator". At its core, "generators" are software packages that allow numerically simulate long sequences of random numbers having statistical properties, repeating the basic properties of real meteorological series (for example, one-dimensional distribution and correlation function). Most often series of surface air temperature, daily minimum and maximum temperatures, precipitation and solar radiation are simulated.

In this talk a stochastic model of air temperature and relative humidity joint time-series is presented. The model is based on long-term observations at meteorological stations, where weather elements were measured every 3 hours. Month-long time-series are considered as a periodically correlated random process, the period of which is equal to 1 day. To describe the correlation structure of time-series, sample correlation matrices are used. Instead of sample one-dimensional distributions mixtures of two Gaussian and beta-distributions are used. Parameters of the mixtures are time-dependent periodic functions whose values are determined from real data. Proposed model let to simulate time-series of temperature and humidity, taking into account the daily periodicity of real processes. On the basis of simulated trajectories some properties of unfavorable combinations of temperature and humidity are studied.

Acknowledgement. This work was supported by the Russian Foundation for Basis Research (grant No 18-01-00149-a), Russian Foundation for Basis Research and Government of Novosibirsk region (grants No 18-41-540003-r-a, 18-41-543006-r-mol-a), the President of the Russian Federation (grant No MK-659.2017.1).

Omega-killed Markov additive processes

CS 14
Stopping and First Exit
Rm 0.87
Wed
12:40 - 13:00

Adam Kaszubowski

University of Wrocław, Poland
email: adam.kaszubowski@math.uni.wroc.pl

We will consider the exit problems for Markov additive process (X, J) exponentially killed with respect to the intensity $\omega_i(x)$, which is positive bounded function. We assume that intensity $\omega_i(x)$ depends on the state i of the Markov chain J and position of the Lévy process X. In classical approach we set $\omega_i(x) = q$, for every $i \in J$, which can be understood as a killing by independent exponentially distributed random variable with the parameter q. Furthermore, all identities will be given in terms

of (ω)-scale matrices, which are extending general one-dimensional scale functions. Finally, we will apply derived results to the dividends problem as well as present numerical results for particular examples of $\omega_i(x)$. My presentation is based on joint work with I. Czarna (Wroclaw University of Science and Technology), Shu Li (University of Illinois) and Z. Palmowski (Wroclaw University of Science and Technology)

IS 29
Group Testing
Rm -1.63
Wed
17:30 - 18:00

When the lie depends on the target

Gyula O.H. Katona
Rényi Institute
email: ohkatona@renyi.hu

The following model is considered. There is exactly one unknown element in the n-element set. A question is a partition of S into three classes: (A, L, B). If $x \in A$ then the answer is "yes" (or 1), if $x \in B$ then the answer is "no" (or 0), finally if $x \in L$ then the answer can be either "yes" or "no". In other words, if the answer "yes" is obtained then we know that $x \in A \cup L$ while in the case of "no" answer the conclusion is $x \in B \cup L$. The mathematical problem is to minimize the minimum number of questions under certain assumptions on the sizes of A, B and L. This problem has been solved under the condition $|L| \geq k$ by the author and Krisztián Tichler in previous papers for both the adaptive and non-adaptive cases. In this paper we suggest to solve the problem under the conditions $|A| \leq a, |B| \leq b$. We exhibit some partial results for both the adaptive and non-adaptive cases. We also show that the problem is closely related to some known combinatorial problems. Let us mention that the case $b = n - a$ has been more or less solved in earlier papers.

IS 11
Monte Carlo, Importance Sampling...
Rm 100/B
Tue
15:30 - 16:00

Adaptive Importance Sampling for Multilevel Monte Carlo Euler method

Mohamed Ben Alaya, Kaouther Hajji, Ahmed Kebaier*
University Paris 13.
email: *kebaier@math.univ-paris13.fr

This paper focuses on the study of an original combination of the Multilevel Monte Carlo method introduced by Giles [2008] and the popular importance sampling technique. To compute the optimal choice of the parameter involved in the importance sampling method, we rely on Robbins-Monro type stochastic algorithms. On the one hand, we extend our previous work Ben Alaya et al. [2015] to the Multilevel Monte Carlo setting. On the other hand, we improve Ben Alaya et al. [2015] by providing a new adaptive algorithm avoiding the discretization of any additional process. Furthermore, from a technical point of view, the use of the same stochastic algorithms as in Ben Alaya et al. [2015] appears to be problematic. To overcome this issue, we employ an alternative version of stochastic algorithms with projection (see e.g. Laruelle et al. [2013]). In this setting, we show innovative limit theorems for a doubly indexed stochastic algorithm which appear to be crucial to study the asymptotic behavior of the new adaptive Multilevel Monte Carlo estimator. Finally, we illustrate the efficiency of our method through applications from quantitative finance.

References

M. Ben Alaya, K. Hajji, and A. Kebaier. Importance sampling and statistical Romberg method. *Bernoulli*, 21(4): 1947–1983, 2015. ISSN 1350-7265. doi: 10.3150/14-BEJ622.

M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008. ISSN 0030-364X. doi: 10.1287/opre.1070.0496.

S. Laruelle, C. Lehalle, and G. Pagès. Optimal posting price of limit orders: learning by trading. *Math. Financ. Econ.*, 7(3):359–403, 2013.

Semi-static and sparse variance-optimal hedging

IS 9
Hedging, Model Fitting, Estimation
Rm 100/B
Tue
11:00 - 11:30

Martin Keller-Ressel

TU Dresden
email: martin.keller-ressel@tu-dresden.de

We consider hedging a financial derivative by a 'semi-static' strategy composed of a dynamic position in one asset and a static (buy-and-hold) position in other assets. We give general representations of the optimal strategy and the hedging error under the criterion of variance-optimality and provide tractable formulas using Fourier-integration in case of the Heston model. We also consider the problem of optimally selecting a sparse semi-static hedging strategy, i.e. a strategy which only uses a small subset of available hedging assets. The developed methods are illustrated in an extended numerical example where we compute a sparse semi-static hedge for a variance swap using European options as static hedging assets.

Acknowledgement. Joint work with Paolo Di Tella and Martin Haubold.

A New and Pragmatic Approach to the GIx/Geo/c/N Queues Using Roots

CS 6
Queuing Theory
Rm 100/A
Tue
16:50-17:10

James Kim*, Mohan Chaudhry

Royal Military College of Canada
email: *s25412@rmc.ca

A simple and complete solution to determine the distributions of queue lengths at different observation epochs for the model GIX/Geo/c/N is presented. In the past, various discrete-time queueing models, particularly the multi-server bulk-arrival queues with finite-buffer have been solved using complicated methods that lead to incomplete results in a non-explicit form. The purpose of this paper is to present a simple derivation for the model GIX/Geo/c/N that leads to a complete solution in an explicit form. The same method can also be used to solve the GIX/Geo/c/N queues with heavy-tailed distributions. The roots of the underlying characteristic equation form the basis for all distributions of queue lengths at different time epochs. All queue-length distributions are in the form of geometric terms.

IS 51
Run and scan models
Rm 0.87
Wed
15:00 - 15:30

Distributions of some run-related patterns in random sequences

Yong Kong
Yale University, New Haven, CT, USA
email: yong.kong@yale.edu

Runs and pattern statistics have found successful applications in various fields. Many earlier results in the field were obtained by traditional methods, which are usually *ad hoc*, not easy to generalize, and when dealing with complicated patterns in multi-state or multiset systems, become intractable. Here we use a systematic approach that is inspired by methods in statistical physics [Kong 2006]. In this approach the study of run and pattern distributions is decoupled into two easy independent steps. In the first step, elements of each object (usually represented by its generating function) are considered in isolation without regards of elements of the other objects. In the second step, formulas in matrix or explicit forms combine the results from the first step into a whole multi-object system with potential nearest neighbor interactions. By considering only one kind of object each time in the first step the complexity arising from the simultaneous interactions of elements from multiple objects is avoided. In essence the method builds up a higher level generating function for the whole system by using the lower level of generating functions from individual objects. Two examples, whose properties of systems without repeated elements (symmetric permutation group) are well-studied, will be discussed in the context of arbitrary multisets (where each element can appear multiple times). The first example is the joint distributions of rises, falls, and runs [Kong 2018a]. The second example is the distributions of successions (also called 2-sequences) [Kong 2018b]. For both examples, explicit formulas for the generating function of the whole system, expectation, and (co)variance will be given.

References

Yong Kong. Distribution of runs and longest runs: A new generating function approach. *Journal of the American Statistical Association*, 101:1253–1263, 2006.

Yong Kong. Joint distribution of rises, falls, and number of runs in random sequences. *Communications in Statistics - Theory and Methods*, 2018a. doi: 10.1080/03610926.2017.1414261.

Yong Kong. Distributions of successions of arbitrary multisets. *manuscript*, 2018b.

IS 15
Extremes of Gaussian Processes
Rm 100/B
Wed
16:30 - 17:00

On Erdös–Révész type laws of the iterated logarithm

Kamil Marcin Kosiński
University of Wrocław, Wrocław, Poland
email: kamil.kosinski@gmail.com

Let X be any stochastic process with almost surely (a.s.) continuous sample paths. For any non-decreasing function f, define the last passage time of X over f as $\xi_f(t) = \sup\{s : 0 \leq s \leq t, X(s) > f(s)\}$. Assume that X crosses over f infinitely many times so that $\lim_{t\to\infty} \xi_f(t) = \infty$ and $\limsup_{t\to\infty}(\xi_f(t) - t) = 0$ a.s.. We say that X exhibits an Erdös–Révész type law of the iterated logarithm if

$$\liminf_{t\to\infty} \frac{\xi_f(t) - t}{h(t)\log\log t} = -1 \quad \text{a.s.}, \tag{1}$$

for some function h.

During the talk we will present new results obtained in the recent literature based on the seminal work by Shao, Q.M. [1992]. In particular, we shall cover the case when X is a stationary Gaussian process or its order statistic process, reflected fractional Gaussian process or a more general Gaussian driven storage process. We shall present a criterion for which (1) holds, for a given f, and how h can be recovered from the tail asymptotics of X above level f.

Acknowledgement. Research of K.M. Kosiński was conducted under scientific Grant No. 2014/12/S/ST1/00491 funded by National Science Centre.

References

Dębicki, K. and Kosiński, K.M. (2017) An Erdös–Révész type law of the iterated logarithm for reflected fractional Brownian motion. *Extremes*, **20**, p.729-749

Dębicki, K. and Kosiński, K.M. (2018) An Erdös–Révész type law of the iterated logarithm for order statistics of a stationary Gaussian process. *J. Theor. Probab.*, **31**, p.579-597

Kosiński, K.M. and Liu, P. (2018) Sample path properties of reflected Gaussian processes arXiv: 1711.01165

Shao, Q.M. (1992) An Erdös–Révész type law of the iterated logarithm for stationary Gaussian processes. *Probab. Theory Relat. Fields*, **94**, p.119-133

Exact distribution of random order statistics and applications

IS 51
Probability models
Rm 0.87
Wed
11:30 - 12:00

Markos V. Koutras*, Vasileios M. Koutras†

Department of Statistics and Insurance Science, School of Finance and Statistics, University of Piraeus, Greece
email: *mkoutras@unipi.gr †vkoutras@icloud.com

In the present work we study the exact distribution of order statistics coming from a sample of random variables, with random sample size. Some new results are provided for the exact distribution of the r-th largest observation of the sample, and several interesting properties are developed when the sample size belongs to wide classes of discrete distributions such as the family of power series distributions, the class of exchangeable Bernoulli mixtures etc. Finally, we illustrate how the stochastic model under study can be exploited for modeling problems arising in financial risk management.

Acknowledgement. The work presented in this paper has been partially funded by the National Matching Funds 2016-2017 of the Greek Government,and more specifically by the General Secretariat for Research and Technology (GSRT), related to EU project "ISMPH: Inference for a Semi-Markov Process using Hazards Specification" (GA No 329128).

OCS 5
Environmental Statistics
Rm -1.63
Tue
12:20 - 12:40

Optimal groupings and inhomogeneity detection in environmental sciences and beyond: examples of the combined cluster and discriminant analysis (CCDA) method

József Kovács[*], Solt Kovács[†,‖], István Gábor Hatvani[‡], Péter Tanos[§], Norbert Magyar[¶], Anita Erőss[*]

[*]Department of Physical and Applied Geology, Eötvös Loránd University, Hungary
[†]Seminar für Statistik, Department of Mathematics, ETH Zürich, Switzerland
[‡]Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, MTA, Hungary
[§]Department of Mathematics and Informatics, Szent István University, Hungary
[¶]Department of Methodology, Budapest Business School, Hungary
email: [‖]solt.kovacs@stat.math.ethz.ch

Some research questions in environmental sciences sparked the development of new statistical methods of general applicability. An example is the combined cluster and discriminant analysis (CCDA) method [1]. CCDA can be used for finding the optimal grouping of sampling sites in a monitoring network [2], as well as for detecting very small differences between sub-groups of sampling sites [1, 3, 4]. Obtaining such information about the system can be particularly useful for the optimization and planning of current monitoring networks. While in environmental sciences sampling sites (or even periods of time) are the sources of data, CCDA could be used far beyond, in any setup, where multiple multivariate samples are available from different sources and the goal is to group these sources optimally or to detect inhomogeneities. The implementation of the CCDA method is available as a package in R [5].

References

[1] J. Kovács, S. Kovács, N. Magyar, P. Tanos, I.G. Hatvani and A. Anda. Classification into homogeneous groups using combined cluster and discriminant analysis, *Environmental Modelling and Software*, 57, 52–59, 2014.

[2] J. Kovács and A. Erőss. Statistically optimal grouping using combined cluster and discriminant analysis (CCDA) on a geochemical database of thermal karst waters in Budapest, *Applied Geochemistry*, 84, 76–86, 2017.

[3] J. Kovács, S. Kovács, I.G. Hatvani, N. Magyar, P. Tanos, J. Korponai and A.P. Blaschke. Spatial Optimization of Monitoring Networks on the Examples of a River, a Lake-Wetland System and a Sub-Surface Water System, *Water Resources Management*, 29(14), 5275–5294, 2015.

[4] P. Tanos, J. Kovács, S. Kovács, A. Anda and I.G. Hatvani. Optimization of the monitoring network on the River Tisza (Central Europe, Hungary) using combined cluster and discriminant analysis, taking seasonality into account, *Environmental Monitoring and Assessment*, 187(9), Paper 575, 14 p., 2015.

[5] S. Kovács, J. Kovács and P. Tanos. Package `ccda` in R: Combined Cluster and Discriminant Analysis, pp. 1-6, 2014.

Change point detection for high-dimensional linear regression and its applications for covariance matrices

CS 10
Change Point Detection
Rm 100/A
Thu
12:20 - 12:40

Solt Kovács[*,†], Peter Bühlmann[*]

[*]Seminar für Statistik, Department of Mathematics, ETH Zürich, Switzerland
email: [†]solt.kovacs@stat.math.ethz.ch

We pursue the goal of high-dimensional ($p > n$) covariance matrix estimation for data with abrupt structural changes. We try to detect these changes and estimate the covariance matrices in the resulting segments. Our approaches closely follow the proposal of Leonardi and Bühlmann [2016] for change point detection in the case of high-dimensional linear regression. We consider the therein proposed estimator in more general setups and propose estimation procedures for covariance matrices based on this regression estimator, as well as another procedure, which is the analogy of the regression estimator, but directly for the case of covariance matrices. We present theoretical results, simulations for the comparison of these proposals, advantages and disadvantages, as well as an illustration of the developed methodology on a real-life example of stock returns.

References

Leonardi, F. and Bühlmann, P. (2016) Computationally efficient change point detection for high-dimensional regression. arXiv: 1601.03704

Feature selection and the mutual information estimation

IS 47
Entropy Estimates & Appl.
Rm -1.62
Mon
11:30 - 12:00

Alexey Kozhevin

Lomonosov Moscow State University, Moscow, Russia
email: kozhevin.alexey@gmail.com

We employ the new statistical estimates of the conditional entropy proposed in Bulinski, A. and Kozhevin, A. [2018] for models comprising the widely used logistic regression. Namely, we concentrate on the estimation of mutual information (for two random vectors). Then the important applications for feature selection are discussed. We compare the proposed approach with previous ones (see, e.g., Coelho, F. et al. [2016], Gao, W. et al. [2018]). In particular, the XOR-model introduced in Bulinski, A. and Kozhevin A. [2017] is also considered. The computer simulations in the framework of the logistic regression with Gaussian predictors show the advantages of the developed feature selection method.

Acknowledgement. The work is supported by the Russian Science Foundation under grant 14-21-00162 and performed at the Steklov Mathematical Institute of Russian Academy of Sciences.

References

Bulinski, A. and Kozhevin, A. (2017) New version of the MDR method for stratified samples. *Statistics, Optimization and Information Computing*, **5:1**, p. 1-18.

Bulinski, A. and Kozhevin, A. (2018) Statistical estimation of the condition entropy (to appear).

Coelho, F. et al. (2016) A mutual information estimator for continuous and discrete variables applied to Feature Selection and Classification problems. *International Journal of Computational Intelligence Systems.*, **9:4**, p.726-733

Gao, W. et al. (2018) Estimating mutual information for discrete-continuous mixtures. arXiv: 1709.06212

CS 10
Change Point Detection
Rm 100/A
Thu
12:40 - 13:00

Quickest drift change detection in Lévy-type force of mortality model

Michał Krawiec

University of Wrocław, Poland
email: michal.krawiec@math.uni.wroc.pl

In the talk I will give solution to the quickest change-point detection problem for the Lévy process consisting of both continuous and jump part. We will consider Bayesian framework with exponential a priori distribution of the change point and an optimality criterion based on probability of false alarm and expected delay of the detection. Our approach is based on optimal stopping theory and it is followed by numerical analysis. We will use this theoretical results to analyse Polish life tables and to model force of mortality in population with drift changing in time.

The talk is based on joint work with Zbigniew Palmowski and Łukasz Płociniczak from the Wrocław University od Science and Technology.

CS 13
Filtering
Rm -1.63
Tue
16:30-16:50

Continuous time linear filtering of general Gaussian processes in Hilbert spaces.

Vít Kubelka

Charles University, Prague, Czech Republic
email: kubelka@karlin.mff.cuni.cz

The Kalman-Bucy Filter is a continuous time counterpart to the discrete time linear Kalman Filter. Therefore, it deals with dynamical system described by stochastic differential equations. First, the continuous time linear filtering problem will be introduced and some results on Kalman - Bucy filter for a general Gaussian signal in finite - dimension will be recalled. Afterwards, an extension for signal with values in a Hilbert space and finite - dimensional observation process will be shown and some interesting examples will be discussed, e.g. the signal processes described by linear stochastic partial differential equations driven by Fractional Brownian motion.

CS 19
Distributio
Rm 100/B
Tue
16:50-17:10

Asymptotic distributions of interaction information

Mariusz Kubkowski[*,†,‡], Jan Mielniczuk[*,†]

[*]Institute of Computer Science, Polish Academy of Sciences, Poland
[†]Warsaw University of Technology, Faculty of Mathematics and Information Science, Poland
email: [‡]m.kubkowski@ipipan.waw.pl

Interaction Information is one of the most promising measures of interaction strength having many desirable properties. However, its use for interaction detection was hindered by the fact that apart from the simple case of overall independence, asymptotic distribution of its estimate was not known. In the contribution we discuss asymptotic distributions of its empirical versions which are needed for formal testing of interactions. We show that for trivariate qualitative vector standardized empirical interaction information converges to the normal law unless the distribution coincides with its Kirkwood approximation.
More specifically, we consider three-dimensional qualitative variable (X_1, X_2, Y) with distribution

$P = P_{X_1,X_2,Y}$ such that X_1, X_2 and Y have correspondingly I, J and K possible values and let $p_{ijk} = P(X_1 = x_i, X_2 = x_j, Y = y_k)$ with p_{ij} and p_i denoting corresponding marginals. Moreover, let

$$p^K_{ijk} = \frac{p_{ij}p_{ik}p_{jk}}{p_i p_j p_k}$$

be Kirkwood approximation of P. Interaction information is defined as

$$II(X_1; X_2; Y) := I[(X_1, X_2); Y] - I(X_1; Y) - I(X_2, Y), \tag{1}$$

where $I(X; Y)$ is mutual information of X and Y. II is a powerful tool to detect interactions used e.g. in Genome Wide Association Studies, see e.g. Moore et al. [2006]. Let $\widehat{II}(X_1; X_2; Y)$ be its plug-in estimator based on a i.i.d. sample of size n pertaining to P. The following result proved in Kubkowski and Mielniczuk [2018] will be discussed in the contribution.

Theorem 1 *We have*

$$n^{1/2}(\widehat{II} - II) \xrightarrow{d} N(0, \sigma^2_{II}), \tag{2}$$

where

$$\sigma^2_{II} = \sum_{i,j,k} p_{ijk} \ln^2 \left(\frac{p_{ijk}}{p^K_{ijk}}\right) - II^2(X_1, X_2, Y) = \mathrm{Var}\left(\ln \frac{p(X_1, X_2, Y)}{p_K(X_1, X_2, Y)}\right)$$

and σ^2_{II} equals 0 if and only if $P = P_K$.

In the opposite case the convergence is to the distribution of weighted centered chi-squared random variables. This case is of special importance as it roughly corresponds to interaction information being zero and the asymptotic distribution can be used for construction of formal tests for interaction detection. The result generalizes Han [1980] result for the case when all coordinate random variables are independent. For the case of $3 \times 3 \times 2$ contingency table corresponding to study of two interacting Single Nucleotide Polymorphisms (SNPs) for prediction of binary outcome, we provide complete description of the asymptotic law and construct approximate critical regions for testing of interactions when two SNPs are possibly dependent. The result is as follows.

Theorem 2 *Let $I = J = 3$, $K \geq 2$ and (X_1, X_2) be independent of Y. Then:*

$$2n\widehat{II} \xrightarrow{d} W,$$

where:

$$W = T_1 + \lambda_1(T_2 - T_3) + \lambda_2(T_4 - T_5),$$
$$T_1 \sim \chi^2_{4(K-1)},\ T_2, T_3, T_4, T_5 \sim \chi^2_{K-1},$$

T_1, T_2, T_3, T_4, T_5 are all independent,

$$\lambda_1^2 = \frac{H_1 + \sqrt{\Delta}}{2},\ \lambda_2^2 = \frac{H_1 - \sqrt{\Delta}}{2}, \tag{3}$$

where $\Delta = 2H_2 - H_1^2$ and

$$H_1 = \sum_{i,j} \frac{p^2_{ij}}{p_i p_j} - 1 = \sum_{i,j} \frac{(p_{ij} - p_i p_j)^2}{p_i p_j},$$
$$H_2 = \sum_{i,j,i',j'} \frac{p_{ij}p_{ij'}p_{i'j'}p_{i'j}}{p_i p_j p_{i'} p_{j'}} - 1.$$

If X_1, X_2 and Y are independent then $\lambda_1 = \lambda_2 = 0$.

In the case when X_1, X_2 and Y are independent and $K = 2$ the asymptotic distribution coincides with χ^2_4 distribution which is routinely used as a reference distribution for testing absence of interactions between two SNPs in predicting binary outcome. We show in numerical experiments that the test based on the derived asymptotic distribution is easy to implement and yields actual significance levels consistently closer to the nominal ones than the test based on chi-squared reference distribution.

References

T. S. Han. Multiple mutual informations and multiple interactions in frequency data. *Information and Control*, 46(1):26 – 45, 1980.

M. Kubkowski and J. Mielniczuk. Testing the significance of features interactions based on interaction information. *submitted*, 2018.

J. H. Moore, J. C. Gilbert, C. T. Tsai, F. T. Chiang, T. Holden, N. Barney, and B. C. White. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J. Theor. Biol.*, 241(2):256 – 261, 2006.

OCS 7
Association and Default
Harmony lnge
Tue
16:50-17:10

Rough Correlation and the Herd Behaviour (HIX) Index

Ashish Kumar[*,†], László Márkus[*]

[*]Eötvös Loránd University, Budapest, Hungary

email: [†]ashish.elte@gmail.com

In order to quantify the degree of synchronised movements in asset prices, dubbed the herd behavior in a financial market, the easy to calculate measure, the Herd Behavior Index (HIX) was introduced recently. When the mentioned asset price co-movements are modelled by stochastic correlations, the empirical copula of the prices can be inferred from simulations. Using diffusion processes for stochastic correlation in the fitted models to minutewise trading data, the obtained copula does not feature particularly strong tail dependence, and the fit is far from perfect. A seemingly better fit can be obtained when the stochastic correlation is described by a transformed Ornstein Uhlenbeck process driven by a fractional Brownian motion, with a moderately small Hurst index, around 0.4. The HIX index values calculated at various times can then be compared to the tail dependence indices of the empirical copulae at the same times, obtained from the simulated model. The subjects of our analysis are minutewise traded Apple and Microsoft stocks. The obtained results indicate that at certain time points the tail index of the corresponding copula may indicate stronger association of the prices than does the HIX index.

OCS 1
Large Random Graphs
Rm -1.64
Mon
15:50 - 16:10

W-random graphs

Dávid Kunszenti-Kovács

Rényi Alfréd Institute of Mathematics

email: daku@renyi.hu

The aim of this talk is to present how the theory of dense graph limits allows us to generate new models for large dense random graphs. This approach, starting from a limit object – essentially a function with two variables taking probability distributions as values – makes it possible to tune (with high probability) the densities of subgraphs arising in the generated large random graphs. The results presented are joint work with L. Lovász and B. Szegedy, and Á. Backhausz, and based on Kunszenti-Kovács et al. [2014] and Backhausz and Kunszenti-Kovács [2017].

References

Á. Backhausz and D. Kunszenti-Kovács. On the dense preferential attachment graph models and their graphon induced counterpart. 2017. arXiv: 1701.06760

D. Kunszenti-Kovács, L. Lovász, and B. Szegedy. Multigraph limits, unbounded kernels, and banach space decorated graphs. 2014. arXiv: 1406.7846

Asymptotic behaviour of randomised fractional volatility models

OCS 6
Rough Volatility
Harmony Inge
Thu
12:20 - 12:40

Blanka Horvath[*], Antoine Jacquier[*], Chloé Lacombe[*,†]

[*]Imperial College, London, U.K.
email: [†]chloe.lacombe14@imperial.ac.uk

We study the asymptotic behaviour of a class of small-noise diffusions driven by fractional Brownian motion, with random starting points. Different scalings allow for different asymptotic properties of the process (small-time and tail behaviours in particular). In order to do so, we extend some results on sample path large deviations for such diffusions. As an application, we show how these results characterise the small-time and tail estimates of the implied volatility for rough volatility models, recently proposed in mathematical finance.

Acknowledgement. B. Horvath acknowledges financial support from the SNSF Early Postdoc.Mobility grant 165248 and A. Jacquier from the EPSRC First Grant EP/M008436/1.

Bismut's way of the Malliavin Calculus for elliptic pseudodifferential operators on a Lie group

CS 18
Theoretical Probability
Rm -1.63
Mon
17:20-17:40

Rémi Léandre

Laboratoire de Mathématiques, Université de Bourgogne- Franche-Comté, 25030, Besançon, FRANCE
email: remi.leandre@univ-fcomte.fr

We show that there is an heat-kernel associated to an elliptic positive pseudo-differential operator right invariant on a Lie group which do not satisfy the maximum principle, such that there is no stochastic process associated, by using the technics of the Malliavin Calculus of Bismut type. Unlike the traditional Malliavin Calculus for Poisson process (jump processes are generated by pseudo-differential operator which satisfy the maximum principle), there is no limitation on the size of jumps.

References

Léandre, R. (2015) Stochastic analysis for a non-markovian generator: an introduction, *Russian Journal of Mathematical Physics*, 22 pp. 39–52.

Léandre, R. Perturbation of the Malliavin Calculus of Bismut type of large order. To appear in *XXXI. Int. Cong. Group. Methods in Physics* (J.P. Gazeau and al eds).

Léandre, R. Malliavin Calculus of Bismut type for an operator of order four on a Lie group. To appear *Journal of Pseudo-differential operators and applications*

Léandre, R. Bismut's way of the Malliavin Calculus of large order generators on a Lie group. To appear *6th Int. Eur. Conf. Math. Sciences and Applications* A.I.P. Proceedings (M. Tosun and al eds).

IS 34
Advances in Biostatistics
Rm -1.63
Mon
16:30 - 17:00

Distribution-Free Inference Methods for Threshold Regression

Meiling Ting Lee
University of Maryland
email: mltlee@umd.edu

In many medical and health-care contexts, a failure event (such as death, hospitalization or transplant) is triggered when a subject's deteriorating health first reaches a failure threshold. The failure process is well described as the sample path of a stochastic process hitting a boundary. The parameters and behaviors of such failure processes must often be inferred from data sets that include censored survival times and current health levels of survivors. A substantial input of expert experience with the health context is usually required to guide the data modeling. This paper describes a parsimonious model for the failure process that has only one distributional property, namely, stationary independent increments. As this property is frequently encountered in real applications, the stochastic model and its related statistical methodology have potential for general application in many fields. The mathematical underpinnings of the distribution-free methods for estimation and prediction are described as well as techniques for incorporating covariates. The methodology is essentially a distribution-free form of threshold regression. Computational aspects of the approach are straightforward. A case example is presented to demonstrate the methodology and its practical use. The methodology provides medical researchers and analysts with new and robust statistical tools for assessing failure risks, estimating effects of risk factors and treatments, and making inferences about residual lifetimes of survivors. The methodology can help to deepen scientific insights into the causes and nature of disease progression.

IS 52
Distribution Theory
Rm -1.64
Thu
12:00 - 12:30

Properties and Applications of Truncated T-X family of Distributions

Ayman Alzaatreh[*], Mohammad Aljarrah[†], Saman Hanif[‡], Muhammad Qaiser Shahbaz[‡], Felix Famoye[§], Carl Lee[§,¶]

[*]American University of Sharjah, UAE
[†]Tafila Technical University, Jordan
[‡]King Abdul Aziz University, Saudi Arabia
[§]Central Michigan University, USA
email: [¶]carl.lee@cmich.edu

The framework of the T-X family of distributions derived by Alzaatreh et al. [2013] has been used to developed many generalized distributions of existing distributions in the literature. In this article, we develop the truncated version of the T-X framework by introducing right-truncated and left-truncate T-X families. Some new generalized families of continuous distributions and properties are developed based on the truncated T-X framework. The relationship between right- and left-truncated T-X families are investigated. Two new left-truncated and right-truncated T-X distributions are applied to some real world data to demonstrate the flexibility of the truncated T-X family of distributions.

References

Ayman Alzaatreh, Carl Lee, and Felix Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013. doi: 10.1007/s40300-013-0007-y.

The Tilted Flashing Brownian Ratchet

CS 4
Diffusions and Random Walks
Rm -1.62
Thu
12:40 - 13:00

S. N. Ethier[*,‡], Jiyeon Lee[†,§]

[*]University of Utah, Salt Lake City, USA
[†]Yeungnam University, Gyeongsan, South Korea
email: [‡]ethier@math.utah.edu [§]leejy@yu.ac.kr

The flashing Brownian ratchet is a stochastic process that alternates between two regimes, a one-dimensional Brownian motion and a Brownian ratchet, the latter being a one-dimensional diffusion process that drifts towards a minimum of a periodic asymmetric sawtooth potential. The result is directed motion. In the presence of a static homogeneous force that acts in the direction opposite that of the directed motion, there is a reduction (or even a reversal) of the directed motion effect. Such a process may be called a tilted flashing Brownian ratchet. We show how one can study this process numerically using a random walk approximation.

Acknowledgement. The work of S. N. Ethier was partially supported by a grant from the Simons Foundation (429675).

References

Ajdari, A. and Prost, J. (1992) Drift induced by a spatially periodic potential of low symmetry: Pulsed dielectrophoresis. *C. R. Acad. Sci., Série 2*, **315**, p.1635-1639.

Allison, A. and Abbott, D. (2002) The physical basis for Parrondo's games. *Fluct. Noise Lett.*, **2**, L327-L341.

Astumian, R. D. (1997) Thermodynamics and kinetics of a Brownian motor. *Science*, **276**, p.917-922.

Astumian, R. D. and Hänggi, P. (2002) Brownian motors. *Phys. Today*, **55**, p.33-39.

Dinís, [Vizcaíno] L. I. (2006) Optimización y control de juegos de azar y motores brownianos colectivos. Doctoral Thesis, Universidad Complutense de Madrid, Spain.

Ethier, S. N and Lee, J. (2018) The flashing Brownian ratchet and Parrondo's paradox. *R. Soc. Open Sci.*, **5**, 171685.

Harmer, G. P. and Abbott, D. (1999) Parrondo's paradox. *Statist. Sci.*, **14**, p.206-213.

Harmer, G. P. and Abbott, D. (2002) A review of Parrondo's paradox. *Fluct. Noise Lett.*, **2**, R71-R107.

Harmer, G. P., Abbott, D. and Taylor, P. G. (2000) The paradox of Parrondo's games. *Proc. R. Soc. Lond. A*, **456**, p.247-259.

Harmer, G. P., Abbott, D., Taylor, P. G. and Parrondo, J. M. R. (2001) Brownian ratchets and Parrondo's games. *Chaos*, **11**, p.705-714.

Parrondo, J. M. R. and De Cisneros, B. J. (2002) Energetics of Brownian motors: a review. *Appl. Phys. A*, **75**, p.179-191.

Reimann, P. (2002) Brownian motors: noisy transport far from equilibrium. *Phys. Rep.*, **361**, p.57-265.

Reimann, P. and Hänggi, P. (2002) Introduction to the physics of Brownian notors. *Appl. Phys. A*, **75**, p.169-178.

Toral, R., Amengual, P. and Mangioni, S. (2003) Parrondo's games as a discrete ratchet. *Physica A*, **327**, p.105-110.

Toral, R., Amengual, P. and Mangioni, S. (2003) A Fokker–Planck description for Parrondo's games. *In Noise in Complex Systems and Stochastic Dynamics* (eds L. Schimansky-Geier, D. Abbott, A. Neiman, C. van den Broeck), Proc. SPIE Series, **5114**, p.309-317.

CS 17
Bayesian Approaches and MCMC
Rm 0.87
Wed
17:50-18:10

A polynomial sum-product algorithm for computing the derivatives of the likelihood in Bayesian Networks

Alexandra Lefebvre*, Gregory Nuel†

Stochastics and Biology Group, Probability and Statistics (LPSM, CNRS 8001), Sorbonne Université, Paris, France

email: *alexandra.lefebvre@sorbonne-universite.fr †nuel@math.cnrs.fr

We consider a Bayesian network over $(X_1, \dots, X_n) \in \mathbb{R}^n$ with a parameter $\theta \in \mathbb{R}^p$. It is well known that the probability of any *evidence* $\text{ev} = \cap_{i=1}^n \{X_i \in \mathcal{X}_i\}$ ($\mathcal{X}_i \subset \mathbb{R}$) can be computed through the following sum-product expression:

$$\mathbb{P}(\text{ev}|\theta) = \sum_{X_1,\dots,X_n} \prod_{i=1}^{n} K_i\left(X_i, X_{\text{pa}_i}|\theta\right) \tag{1}$$

where the pa_i are the *parent sets*, and where the $K_i\left(X_i, X_{\text{pa}_i}|\theta\right) = \mathbf{1}_{X_i \in \mathcal{X}_i} \mathbb{P}\left(X_i|X_{\text{pa}_i};\theta\right)$ are the *potentials* of each variable. From a statistical point of view, $L_n(\theta) = \mathbb{P}(\text{ev}|\theta)$ is the likelihood of θ given the evidence. Computing the derivatives of the likelihood function is of great interest, especially the first and second order derivatives from which one can derive the score and the observed Fisher information matrix. These quantities can not only help maximizing the likelihood function (*e.g.* through Newton-based algorithms) but also allow to obtain confidence intervals on parameters as well as performing hypothesis testing (likelihood ratio tests, score tests and Wald tests).

In probabilistic graphical models such as the Bayesian network, the sensitivity analysis allows to express $L_n(\theta)$ as a polynomial function of θ under the hypothesis that all potentials can be expressed as polynomials in θ [see Nielsen and Jensen 2009, pp 184–185]. But when the same parameter appears in many potentials, the resulting polynomial is usually of high order, and its computation cost prohibitive. Alternatively, in the particular context of the hidden Markov models (HMM), Cappé and Moulines [2005] suggest to use the Fisher and Louis identities to obtain the first and second likelihood derivatives through the so-called *smoothing recursions* allowing to compute efficiently the expectation of any additive functional. On the other hand, it is well known that polynomial versions of the sum-product algorithm can be very efficient for performing complex computation in probabilistic graphical models (*e.g.* order k moment of an additive functional in Bayesian networks [Cowell 1992, Nilsson 2001], moment/probability generating functions in pattern matching [Nuel 2008, 2010], etc.).

In the present work we want to generalize the smoothing recursions of Cappé and Moulines [2005] from the specific context of HMM to any Bayesian network, and we want to take advantage of polynomial arithmetic for simplified computations though a single sum-product recursion to compute both the likelihood function and all its derivatives. For a single parameter model ($\theta \in \mathbb{R}$), our method allows one to compute the derivatives up to the d^{th} order with a complexity of $C \times (d+1)(d+2)/2 = \mathcal{O}(C \times d^2)$ where C is the complexity for computing $L_n(\theta)$ through the original sum-product recursion. For a multi-parameter model ($\theta \in \mathbb{R}^p$) we obtain the likelihood, the gradient and the Hessian with a complexity of $\mathcal{O}(C \times p^2)$. The latter complexity is similar to the approach of Cappé and Moulines [2005], but the practical implementation is arguably simpler and easier to extend to higher order derivatives.

We illustrate our new method with the well known two-point linkage analysis model which is used in human genetics for localizing a disease susceptibility gene on the genome using *pedigrees* (family structures) where the disease segregates with a genetic factor.

Acknowledgement. This work was funded by the epidemiological doctoral program of the French *Ligue Nationale Contre le Cancer* (LNCC).

References

Olivier Cappé and Eric Moulines. Recursive computation of the score and observed information matrix in hidden markov models. In *Statistical Signal Processing, 2005 IEEE/SP 13th Workshop on*, pages 703–708. IEEE, 2005.

RG Cowell. Calculating moments of decomposable functions in bayesian networks. *Preprint*, 1992.

Thomas Dyhre Nielsen and Finn Verner Jensen. *Bayesian networks and decision graphs*. Springer Science & Business Media, 2009.

Dennis Nilsson. The computation of moments of decomposable functions in probabilistic expert systems. In *Proceedings of the Third International Symposium on Adaptive Systems*, pages 116–21, 2001.

Grégory Nuel. Pattern markov chains: optimal markov chain embedding through deterministic finite automata. *Journal of Applied Probability*, 45(1):226–243, 2008.

Grégory Nuel. On the first k moments of the random count of a pattern in a multistate sequence generated by a markov source. *Journal of Applied Probability*, 47(4):1105–1123, 2010.

On the Outcome of Epidemics with Detections

IS 26
Growth & Contagion Models
Rm 0.99
Mon
12:00 - 12:30

Claude Lefèvre[*,†,‡], Philippe Picard[†,§]

[*]Université Libre de Bruxelles, Bruxelles, Belgique
[†]ISFA, Université de Lyon 1, Lyon, France
email: [‡]clefevre@ulb.ac.be [§]philippe.picard69@free.fr

The classical SIR epidemic model is generalized to incorporate a detection process of infectives in the course of time. An approximating branching to this model was previously discussed by Trapman and Bootsma [2009]. Our purpose is to determine the distribution of the population state at the first detection instant and the next ones. An extension is also discussed that allows the parameters to change with the number of detected cases. The followed approach relies on simple martingale arguments and uses a special family of Abel-Gontcharoff polynomials. This material comes from Lefèvre and Picard [2017].

References

Lefèvre, C. and Picard, P. (2017). On the outcome of epidemics with detections. *Journal of Applied Probability* **54**, 890-904.

Trapman, P. and Bootsma, M.C.J. (2009). A useful relationship between epidemiology and queueing theory: The distribution of the number of infectives at the moment of the first detection. *Mathematical Biosciences* **219**, 15-22.

Modelling Preference Data with the Wallenius Distribution

CS 17
Bayesian Approaches and MCMC
Rm 0.87
Wed
17:30-17:50

Fabrizio Leisen

University of Kent
email: fabrizio.leisen@gmail.com

The Wallenius distribution is a generalisation of the Hypergeometric distribution where weights are assigned to balls of different colours. This naturally defines a model for ranking categories which can be used for classification purposes. Since, in general, the resulting likelihood is not analytically available, we adopt an approximate Bayesian computational (ABC) approach for estimating the importance of the categories. We illustrate the performance of the estimation procedure on simulated

datasets. Finally, we use the new model for analysing two datasets concerning movies ratings and Italian academic statisticians' journal preferences. The latter is a novel dataset collected by the authors.

Acknowledgement. This is a work in collaboration with Clara Grazian (University of Oxford) and Brunero Liseo (University of Rome la Sapienza)

IS 16
Fractional Stochastic Models
Rm 100/B
Wed
14:30 - 15:00

Fractional Poisson Fields and Martingales

Nikolai Leonenko
Cardiff University, UK
email: LeonenkoN@Cardiff.ac.uk

We present new properties for the Fractional Poisson process [1, 2, 3, 6, 9], Fractional non-homogeneous Poisson process [7, 8], Fractional Poisson fields on the plane [1, 5], and the fractional Skellam processes [4]. A martingale characterization for Fractional Poisson processes is given. We extend this result to Fractional Poisson fields, obtaining some other characterizations. The fractional differential equations are studied. The covariance structure is given. Finally, we give some simulations of the Fractional Poisson fields on the plane.

This is a joint work with G. Aletti (University of Milan, Italy) and E. Merzbach (Bar Ilan University, Israel).

References

[1] Giacomo Aletti, Nikolai Leonenko, and Ely Merzbach. Fractional Poisson Fields and Martingales. *J. Stat. Phys.*, 170(4):700–730, 2018. doi: 10.1007/s10955-018-1951-y.

[2] L. Beghin and E. Orsingher. Fractional Poisson processes and related planar random motions. *Electron. J. Probab.*, 14:no. 61, 1790–1827, 2009. doi: 10.1214/EJP.v14-675.

[3] L. Beghin and E. Orsingher. Poisson-type processes governed by fractional and higher-order recursive differential equations. *Electron. J. Probab.*, 15:no. 22, 684–709, 2010. doi: 10.1214/EJP.v15-762.

[4] Alexander Kerss, Nikolai N. Leonenko, and Alla Sikorskii. Fractional Skellam processes with applications to finance. *Fract. Calc. Appl. Anal.*, 17(2):532–551, 2014. doi: 10.2478/s13540-014-0184-2.

[5] Nikolai Leonenko and Ely Merzbach. Fractional Poisson fields. *Methodol. Comput. Appl. Probab.*, 17(1):155–168, 2015. doi: 10.1007/s11009-013-9354-7.

[6] Nikolai N. Leonenko, Mark M. Meerschaert, René L. Schilling, and Alla Sikorskii. Correlation structure of time-changed Lévy processes. *Commun. Appl. Ind. Math.*, 6(1):e–483, 22, 2014. doi: 10.1685/journal.caim.483.

[7] Nikolai Leonenko, Enrico Scalas, and Mailan Trinh. The fractional non-homogeneous Poisson process. *Statist. Probab. Lett.*, 120:147–156, 2017. doi: 10.1016/j.spl.2016.09.024.

[8] Nikolai Leonenko, Enrico Scalas, and Mailan Trinh. Limit theorems for the fractional non-homogeneous poisson process. arXiv: 1711.08768v1.

[9] Mark M. Meerschaert, Erkan Nane, and P. Vellaisamy. The fractional Poisson process and the inverse stable subordinator. *Electron. J. Probab.*, 16:no. 59, 1600–1620, 2011. doi: 10.1214/EJP.v16-920.

On Diffusion Approximation of Branching Processes in Random Envirionment

IS 25
Population Dynamics Studies
Rm -1.62
Mon
15:30 - 16:00

Nikolaos Limnios[*,‡], Elena Yarovaya[†,§]

[*]Université de Technologie de Compiègne, France
[†]Lomonosov Moscow State University, Russia
email: [‡]nlimnios@utc.fr [§]yarovaya@mech.math.msu.su

We consider near critical discrete- and continuous-time Markov branching processes. We present diffusion approximation results for discrete-time Bienaymé-Galton-Watson processes, and continuous-time Markov age-dependent branching processes. In a first part, Feller-Jirina theorem (Feller [1951], Jiřina [1969]) and Jagers theorem (Jagers [1971]) are revisited. The proofs of these theorems were obtained using the generating function technique. In contrast to this we suggest a new method to obtain diffusion approximations of such processes based on Markov generators convergence and semi-martingale relative compactness.

In a second part, we present also diffusion approximation of continuous-time branching processes in Markov random environment. The law of offsprings is supposed to be dependent of the state space of a jump Markov process with general state space.

Moreover, we prove that the near critical condition is a necessary and sufficient condition for a diffusion approximation of a Markov branching process to hold. An averaging result is also given.

References

Feller, W. (1951). *Diffusion Processes in genetics*. Proc. Second Berkley Symp. Math. Statist. Prob., 227–246. Univ. of California Press.
Jagers, P. (1971).*Diffusion Approximation of Branching Processes*. Annals Math. Statist., vol. 42, No 6, pp 2074–2078.
Jiřina, M. (1969). *On Feller's branching diffusion processes*. Časopsis. Pěst. Mat., vol. 94, pp 84–90.
Limnios, N. and Yarovaya, E. (2018). *A Note on the Diffusion Approximation of Branching Processes*, submitted.

On time-inconsistent stopping problems

CS 14
Stopping and First Exit
Rm 0.87
Wed
12:20 - 12:40

Kristoffer Lindensjö[*,‡], Sören Christensen[†]

[*]Stockholm University, Stockholm, Sweden
[†]University of Hamburg, Hamburg, Germany
email: [‡]kristoffer.lindensjo@math.su.se

Consider a diffusion X and the classical problem of choosing a stopping time τ that maximizes

$$\mathbb{E}_x(h(X_\tau)),$$

where h is a nice deterministic function. Recall that the solution to this problem is consistent in the sense that the optimal rule for stopping, i.e. 'stop the first time that X enters the stopping region', is independent of the initial state x. Now consider a non-linear deterministic function g and the problem of choosing a stopping time τ that maximizes

$$g(\mathbb{E}_x(h(X_\tau))).$$

The optimal stopping rule for this problem will, in contrast, typically depend on the initial state x. In the literature this is known as *time-inconsistency* and the problem described is an example of a *time-inconsistent stopping problem*.

Time-inconsistent problems are mainly studied using one of the following approaches:

1. The *game-theoretic approach*, which means formulating the problem as a game and look for equilibrium stopping times.

2. The *pre-commitment approach*, which means formulating the problem for a fixed initial state and allowing the corresponding optimal stopping rule to depend on that initial state.

Time-inconsistent problems were originally studied in financial economics where the inconsistency is due to: endogenous habit formation, non-exponential discounting, or mean-variance optimization/utility.

We have developed a general game-theoretic framework for time-inconsistent stopping problems, that can handle the economic problems mentioned above and that includes: i) a proper definition of equilibrium for time-inconsistent stopping problems, ii) an equilibrium characterization, iii) several other results with different necessary and sufficient conditions for equilibrium, e.g. a smooth fit result, iv) several examples, including mean-variance optimization and endogenous habit formation problems.

IS 45
Inference for Complex Data
Rm -1.64
Wed
12:00 - 12:30

Tolerance Tubes for Tracking Functional Data

Regina Y. Liu[*,‡], Yi Fan[†,§]

[*]Rutgers University, New Jersey, USA
[†]Amazon.com, Inc.
email: [‡]rliu@stat.rutgers.edu [§]fanyi0926@gmail.com

Tolerance intervals and tolerance regions are important tools for process monitoring or statistical quality control of univariate and multivariate data, respectively. We discuss their generalization to tolerance tubes in the infinite dimensional setting for functional data. In addition to the generalizations of the commonly accepted definitions of the tolerance level of β-content or β-expectation, we introduce the new definition of α-exempt β-expectation tolerance tube. The latter loosens the definition of β-expectation tolerance tube by allowing alpha (pre-set using domain knowledge) portion of each functional be exempt from the requirement. Those proposed tolerance tubes are completely nonparametric and broadly applicable. We discuss their general properties, and show that the α exempt beta-expectation tolerance tube is particularly useful in the setting where occasional short term aberrations of the functional data are deemed acceptable (or unpreventable) and they do not cause substantive deviation of the norm. This desirable property is elaborated further and illustrated with both simulations and real applications in continuous monitoring of blood glucose level in diabetes patients as well as of aviation risk patterns of aircraft landings.

Acknowledgement. The authors gratefully acknowledge the support from grants NSF-DMS 1107012 and DOT-FAA 09-G-017. The second author also acknowledges the generous graduate support from Rutgers University.

Products of random matrices: knowns and unknowns

IS 54
Random Matrix Theory
Rm -1.64
Tue
15:00 - 15:30

Dang-Zheng Liu

Key Laboratory of Wu Wen-Tsun Mathematics, CAS, School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, P.R. China
email: dzliu@ustc.edu.cn

The study of products of random matrices goes back to the pioneering work of Furstenberg and Kesten in 1960, in the context of dynamical systems and Lyapunov exponents. Later, applications were found in Schrödinger operator theory, statistical physics, wireless communication networks and combinatorics. Recently, exact solvability of singular values and eigenvalues for products of complex Gaussian random matrices and of truncations of random unitary matrices, due to Akemann and coworkers, has attracted much attention and eventually prompts a great revival. We give a brief survey of known and unknown results in this topic.

Quadratic variation of càdlàg semimartingales as a.s. limit of the normalized truncated variations

CS 3
SDEs and Semi-martin-gales
Rm 100/B
Mon
17:40-18:00

Rafał M. Łochowski

Warsaw School of Economics, Poland
email: rlocho@sgh.waw.pl

For a real càdlàg path x we define sequence of semi-explicit quantities, which do not depend on any partitions and such that whenever x is a path of a càdlàg semimartingale then these quantities tend a.s. to the continuous part of the quadratic variation of the semimartingale. Next, we derive several consequences of this result and propose a new approach to define Föllmer's pathwise integral.

References

Łochowski, R. M. (2017) Quadratic variation of càdlàg semimartingales as a.s. limit of the normalized truncated variations. arXiv: 1708.00732

Absorption time and absorption probabilities for a family of multidimensional gambler models

IS 14
Excursions of Lévy Processes
Rm 100/B
Mon
11:30 - 12:00

Paweł Lorek

Mathematical Institute, University of Wrocław, Poland
email: Pawel.Lorek@math.uni.wroc.pl

In Lorek, P. [2017] we presented some multidimensional generalization of a gambler's ruin problem. We explicitly gave formulas for winning/ruin probabilities (where one-step winning/losing probabilities can be arbitrary and can depend on the current fortune, model can be interpreted as a game of one player versus $d \geq 1$ players). In Lorek, P., Markowski, P. [2018] we provide similar result for a much wider family of multidimensional gambler models.

For a birth and death process on $\{0, 1, \ldots, d\}$, starting at 0, passage time to d is equal, in distribution, to the sum of d independent exponential random variables with parameters being the eigenvalues of the intensity matrix. The result is usually attributed to Keilson, the proof was purely analytical.

In Fill, J. A. [2009] a probabilistic proof for a discrete-time version was provided using some dualities between Markov chains. In Gong, Y., Mao, Y., Zhang, C. [2012] using similar method (and in Mao, Y., Zhang, C. [2016] using a different one) authors provided similar results for the case of two absorbing states when the chain starts at any $0 \leq i \leq d$. In Lorek, P., Markowski, P. [2018] we provide similar results for a wide family of multidimensional gambler models.

The talk will be mainly based on the results from Lorek, P., Markowski, P. [2018]. Among others, we heavily exploit Siegmund duality for Markov chains for partially ordered state spaces (we provided necessary tool in Lorek, P. [2017]) and spectral polynomials of stochastic matrices (see Fill, J. A. [2009]).

References

Fill, J. A. (2009) The Passage Time Distribution for a Birth-and-Death Chain: Strong Stationary Duality Gives a First Stochastic Proof. *Journal of Theoretical Probability*, **22**, 543–557.

Gong, Y., Mao, Y., Zhang, C. (2012) Hitting Time Distributions for Denumerable Birth and Death Processes *Journal of Theoretical Probability*, **25**, 950–980.

Lorek, P. (2017) Siegmund duality for Markov chains on partially ordered state spaces. *Probability in the Engineering and Informational Sciences*, 1–27.

Lorek, P. (2017) Generalized Gambler's Ruin Problem: explicit formulas via Siegmund duality. *Methodology and Computing in Applied Probability*, **19 (2)**, 603–613.

Lorek, P., Markowski, P. (2018) Absorption probabilities and hitting time distributions for family of multidimensional gambler models. *Submitted to: Advances in Applied Mathematics*, 2018.

Mao, Y. and Zhang, C. (2016) Hitting Time Distributions for Birth–Death Processes With Bilateral Absorbing Boundaries. *Probability in the Engineering and Informational Sciences*, 1–12.

IS 42
Scan Stats – Methods...
Rm 100/A
Thu
11:00 - 11:30

A Dimension-free Approach for Cluster Inference

Wendy Lou
University of Toronto, Canada
email: wendy.lou@utoronto.ca

Applications of scan statistics in the field of health science research often involve the identification of clusters, which are investigated for their associations with variables of interest; for example, the numbers of tuberculosis cases may be associated with the levels of environmental exposure to small atmospheric particulate matter. Quantifying such relationships statistically, while accounting for other factors including spatial correlations within and between clusters, is the focus of this presentation. A dimension-free approach with pivotal indices will be developed based on a real example, and compared to alternative approaches via numerical comparisons.

CS 13
Filtering
Rm -1.63
Tue
17:50-18:10

A Particle filter at low computational cost

Rodi Lykou*, George Tsaklidis
Aristotle University of Thessaloniki, Greece
email: *lykourodi@math.auth.gr

Interventions in the basic Particle Filter (PF) take place to create a PF at a low computational cost. For the case that outliers should not be rejected from our study, several approaches are proposed: First and foremost, when the PF is unable to predict the presence of an outlier, the nearest prediction to the measurement could be adopted for the state variable or the correction could be based mainly on

the measurement. Additionally, the distribution of whole data set could be approximated by sums of known distributions. Last but not least, phase type approximations are examined for their capability towards such processes.

References

Doucet, A. and Johansen, A. M. (2011) A tutorial on particle filtering and smoothing: fifteen years later, In D. Crisan and B. Rozovsky (Eds.), The Oxford Handbook of Nonlinear Filtering. Oxford University Press.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M. (1993) Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE PROC-F*, **140**, no. 2, pp. 107-113, doi: 10.1049/ip-f-2.1993.0015

Liu,B., (2017), Robust particle filter by dynamic averaging of multiple noise models, *ICASSP*, New Orleans, LA, pp. 4034-4038, doi: 10.1109/ICASSP.2017.7952914

Liu, B. (2018), ILAPF: Incremental Learning Assisted Particle Filtering, arXiv: 1710.10489

Pitt M. K. and Shephard N. (1999): Filtering via Simulation: Auxiliary Particle Filters, *J. Am. Stat. Assoc.*, **94**, no. 446,p. 590-599, doi: 10.2307/2670179

Random time-changes and asymptotic results for a class of continuous-time Markov chains on integers with alternating rates

IS 21
Renewal & semi-Markov...I.
Rm -1.62
Tue
11:30 - 12:00

Luisa Beghin[*,§], Claudio Macci[†,¶], Barbara Martinucci[‡,‖]

[*]Dipartimento di Scienze Statistiche, Sapienza Università di Roma
[†]Dipartimento di Matematica, Università di Roma Tor Vergata
[‡]Dipartimento di Matematica, Università di Salerno
email: [§]luisa.beghin@uniroma1.it [¶]macci@mat.uniroma2.it [‖]bmartinucci@unisa.it

We consider continuous-time Markov chains on integers which allow transitions to adjacent states only, with alternating rates. We give explicit formulas for probability generating functions, and also for means, variances and state probabilities of the random variables of the process. Moreover we study independent random time-changes with the inverse of the stable subordinator, the stable subordinator and the tempered stable subodinator. We also present some asymptotic results in the fashion of large deviations.

Local and global degree profiles of randomly grown self-similar hooking networks under uniform and preferential attachment

IS 50
Random Structures
Rm -1.64
Wed
17:30 - 18:00

Hosam Mahmoud

The George Washington University
email: hosam@gwu.edu

We investigate node degrees in a network grown from a seed by hooking self-similar components under two models of randomness: a uniform attachment model and a model based on preferential attachment. We study two degree profiles: a local profile tracking the evolution of the degree of a particular node over time, and a global profile concerned about counts of the number of nodes of a particular degree.

For the local profile, under uniform attachment growth, we have the exact mean, variance and probability distribution in terms of standard combinatorial numbers like generalized harmonic numbers and Stirling numbers of the first kind. Asymptotically, we observe phases: The early nodes have an asymptotically normal distribution, intermediate nodes have a Poisson distribution and the late

nodes have a degenerate distribution. In contrast, under preferential attachment, the moments of the degree of a node contain Stirling numbers of the second kind and (under appropriate scaling) has a gamma-type limit law.

As for the global profile, we use Pólya urns to derive strong laws. Four regimes arise according to the structure of the seed. Within these regimes, we identify a few degenerate cases. Barring these degenerate cases, we uncover an asymptotically normal joint multivarite distribution for nodes of very small degrees.

IS 49
Urn Models
Rm -1.64
Wed
14:30 - 16:00

Infinitely-many-colour Pólya urns and stochastic approximation

Cecile Mailler[*§], Jean-Francois Marckert[†], Denis Villemonais[‡]

[*]University of Bath, Bath, United-Kingdom
[†]Université de Bordeaux, Bordeaux, France
[‡]École des Mines de Nancy, Institut Élie Cartan, and Inria, Nancy, France
email: [§]c.mailler@bath.ac.uk

In this talk, I will present a joint work with Jean-François Marckert, and an ongoing work with Denis Villemonais. In Mailler and Marckert [2017], we introduced measure-valued Pólya processes (MVPPs) as a generalisation of Pólya urns to infinitely-many colours; this model was inspired by Bandyopadhyay and Thacker [2017], and the model of MVPPs was also introduced in Bandyopadhyay and Thacker [2016]. I will present Mailler and Marckert [2017] and explain, in particular, that each MVPP can be coupled with a Markov chain indexed by the random recursive tree and how this coupling allow to prove convergence in probability of a large class of MVPPs.

In an ongoing work with Denis Villemonais, we exploit the link between Pólya urns and quasi-stationary distributions (already exhibited in Aldous et al. [1988]) and use stochastic approximation techniques on a space of measures to prove almost sure convergence of a large class of MVPPs. We also apply our result to varied examples.

Acknowledgement. The presenter is grateful to EPSRC for funding through the fellowship EP/R022186/1.

References

David Aldous, Barry Flannery, and José Luis Palacios. Two applications of urn processes the fringe analysis of search trees and the simulation of quasi-stationary distributions of markov chains. *Probability in the engineering and informational sciences*, 2(3):293–307, 1988.

Antar Bandyopadhyay and Debleena Thacker. A new approach to Pólya urn schemes and its infinite color generalization. arXiv: 1606.05317, 2016.

Antar Bandyopadhyay and Debleena Thacker. Pólya urn schemes with infinitely many colors. *Bernoulli*, 23(4B): 3243–3267, 2017.

Cécile Mailler and Jean-François Marckert. Measure-valued Pólya processes. *Electronic Journal of Probability*, 2017. doi: 10.1214/17-EJP47.

Generalized Group Testing: Some Results and Open Problems

IS 29
Group Testing
Rm -1.63
Wed
16:30 - 17:00

Yaakov Malinovsky

University of Maryland, Baltimore County, United States
email: yaakovm@umbc.edu

Group testing has its origins in the identification of syphilis in the U.S. army during World War II. Consider a finite population of N items, where item i has a probability p_i to be defective independently from the other units. A group test is a binary test on an arbitrary group of units with two possible outcomes: all units are good, or at least one unit is defective. The goal is to identify all items through group testing with the minimum expected number of tests. This is the generalized group testing problem, or GGTP. The optimum procedure, with respect to the expected total number of tests, is unknown even in case when all p_i are equal. In the GGTP, an optimum regime is known only for one particular algorithm in the nested class of algorithms–the Dorfman procedure (F. Hwang (1975, 1981)). In this talk, we discuss how finding an optimal regime for additional algorithms in the nested class seems to be a computationally hard problem, and we present suboptimal solutions. In addition, some open problems and conjectures will be discussed.

Hitting times for Claim Number in Car Insurance Setting

IS 21
Renewal & semi-Markov...I.
Rm -1.62
Tue
10:30 - 11:00

Guglielmo D'Amico[*], Fulvio Gismondi[†], Jacques Janssen[‡], Raimondo Manca[§,**], Filippo Petroni[¶], Dmitrii Silvestrov[‖]

[*]University "G. d'Annunzio" of Chieti-Pescara, Italy
[†]University "Guglielmo Marconi", Rome Italy
[‡]Universitè Libre de Bruxelles, Belgium
[§]University of Rome "La Sapienza, Italy
[¶]University of Cagliari, Italy
[‖]University of Stockholm, Sweden
email: [**]raimondo.manca@uniroma1.it

In this paper, the phase space of homogeneous semi-Markov processes is constructed taking into account the number of claims that an insured will have will have during her/his driving life. The aim is the calculationfor a driver, of the mean time to report a given nymber of cliams. This problem can be solved constructing the probability distributionfunction of the first entry time for each state (number of claims) of the model. The age is considered as the homogeneous time variable where time 0 corresponds to the first age of driving. As well known, the age in car insurance contracts plays a fundamental relevance in the calculation of the behaviour of insured people. In this study, non-homogeneous semi-Markov models will be used for following the time evolution of the the claim number.

IS 1
Stochastic Methods in Finance
Harmony Inge
Tue
15:00 - 15:30

How to Choose Stochastic Correlations in Modelling Interdependence of Asset Prices

László Márkus*, Ashish Kumar

Eötvös Loránd University, Budapest, Hungary
email: *markus@cs.elte.hu

The interdependence of asset prices has an important impact in pricing derivatives dependent on multiple assets. It is also important in structural credit risk models, when joint defaults are of concern.

Usually the interdependence of assets is non-linear and has a temporal variation, hence Pearson correlation is insufficient to represent it properly. The use of stochastic correlations was suggested recently, and e.g. approximate pricing formula for quanto options was found for some specific models. However, the study of qualitative properties of the various models, such as tail dependence etc., allowing for checking goodness of fit are seemingly missing as yet. We assess this issue on historic asset data by the help of copulas.

We suppose that the individual asset prices follow one of the usual models of financial mathematics e.g. geometric Brownian motion or the Heston model. Eventually Variance Gamma or other subordinated Brownian motion model may also be considered within our framework. For illustrative purposes confining ourselves to two assets only, we build up *dependent* Wiener processes $W_i(t)$-s $i = 1, 2$. These will then drive the equations of asset prices.

$$dW_2(t) = \rho dW_1(t) + \sqrt{1-\rho^2}dV(t). \tag{1}$$

with another adapted Wiener process $V(t)$ *independent* of $W_1(t)$ A constant ρ here results in correlated Wiener processes and hence a two dimensional Wiener process. Instead, we change ρ to be an adapted process $\rho(t)$ called *stochastic correlation*. When $\rho(t)$ is adapted, Lévy's characterisation theorem guarantees that all $W_i(t)$-s are Wiener processes, however, the two dimensional process is no longer a Wiener one. As a result, the asset prices can still be generated from these marginal Wiener processes by the usual way.

The stochastic correlation $\rho(t)$ can either be created by a Jacobi process, or as a suitable transformation of an Ornstein-Uhlenbeck or other diffusion process. The copula of the increments is characterised by Kendall's K-function, which is computed from simulated processes. A simulated confidence domain is then created for the differences of K-functions, and the various possible choices of stochastic correlations can be tested against these. We illustrate our ideas on minutewise traded Microsoft and Apple stocks.

While the suggested model does not seem to have strong tail dependence, just as the Gaussian copula, its weak tail dependence is greater than that.

IS 44
Inference for Large Datasets
Rm 0.99
Wed
17:00 - 17:30

Real-time particle path interacting strategies for large datasets

Reinaldo Marques*†, Geir Storvik†

*University of Oslo, Norway & Federal University of Alfenas, Brazil
†University of Oslo & BigInsight Center, Norway
email: ‡ramarque13@gmail.com

The digital age exploits a number of applications which real-time inference is needed, including in the insurance and financial risk sectors. In particular, the prediction of latent states of general state

space models is a well-known challenging task, but particle algorithms [Del Moral 2004, 2013] have proven successful in a range of application. Prediction of functions of the whole sequence of latent variables, entitled path functionals [Del Moral et al. 2010, Del Moral and Penev 2017], may be of interest in itself, but are also crucial elements when estimation of static parameters is needed. Online inference of path functionals using particle filters have shown be a much harder computational task, mainly due to path degeneracy problems for large datasets. In this work, we present particle blocking approximations to mitigate the path degeneracy [Douc et al. 2014] when we have access to the use of recursive functionals after collecting a huge amount of data. We also provide illustrations to estimate path functionals and for online Bayesian parameter learning.

Acknowledgement. We gratefully acknowledge financial support from CAPES-Brazil and Statistics for Innovation Center, in Norway.

References

Pierre Del Moral. *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*. New York: Springer-Verlag, 2004.

Pierre Del Moral. *Mean field simulation for Monte Carlo integration*. CRC press, 2013.

Pierre Del Moral and Spiridon Penev. *Stochastic Processes: From Applications to Theory*. CRC Press, 2017.

Pierre Del Moral, Arnaud Doucet, and Sumeetpal S Singh. A backward particle interpretation of Feynman-Kac formulae. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(5):947–975, 2010.

Randal Douc, Eric Moulines, and David Stoffer. *Nonlinear time series: theory, methods and applications with R examples*. CRC Press, 2014.

Distributions of pattern statistics in sparse Markov models

IS 51
Probability models
Rm 0.87
Wed
10:30 - 11:00

Donald E.K. Martin

North Carolina State University
email: martin@stat.ncsu.edu

Markov models provide a good approximation to probabilities associated with many categorical time series, and thus they are applied extensively. However, a major drawback associated with them is that the number of model parameters grows exponentially in the order of the model, and thus only very low-order models are considered in applications. Another drawback is lack of flexibility, in that Markov models give relatively few choices for the number of model parameters. Sparse Markov models are a generalization of Markov models where transition probabilities are lumped into classes comprised of invariant probabilities. The contexts for conditioning may be either hierarchical (as in variable length Markov chains) or non-hierarchical. This supplies a model that takes care of the two problems given above that are associated with Markov models, and which thus give a better handling of the trade-off between bias associated with having too few model parameters and variance associated with having too many. In this work, methods for efficient computation of pattern distributions through Markov chains with minimal state spaces are extended to the sparse Markov framework.

CS 7
Actuarial Applications
Rm 0.99
Mon
17:00-17:20

Experience Rating by Means of Autoregressive Sequences

László Martinek[*,†,‡], Miklós Arató[*,§]

[*]Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest
[†]NN Group, The Hague
email: [‡]martinek@cs.elte.hu [§]arato@math.elte.hu

A posteriori ratemaking is widely applied in the premium calculation of property and casualty products, particularly in third party automobile insurance, where the premium adjustment is usually ruled by a bonus-malus system. The present talk suggests an alternative to common frameworks, which are designed as random walks on graphs of mostly finite states representing premium levels. The proposed premium calculation model is governed by the policyholder's claim history through a recursive equation Rosenblatt [1995]. This new autoregressive scheme structurally differs from the ones in use.

Instead of analysing the more optimal finite set of relativities Tan et al. [2015], Tzougas et al. [2014], we propose a fundamentally different set of transition rules from one premium class to another. Relevant metrics measuring the system's optimality are evaluated Coene and Doray [1996], Lemaire and Zi [1994], partially in analytical form. Through a comparison with existing models and parameterisation from real life data Gómez-Déniz [2016], the new model is put into perspective and its practical relevance is investigated. The purpose of the talk is to (a) introduce a new scheme which is structurally different from the ones in use, (b) to evaluate its metrics of higher relevance and (c) to put it into perspective through a comparison with existing models.

References

G. Coene and L. G. Doray. A financially balanced bonus-malus system. *ASTIN Bulletin*, 26(1):107–116, 1996.

E. Gómez-Déniz. Bivariate credibility bonus–malus premiums distinguishing between two types of claims. *Insurance: Mathematics and Economics*, 70:117–124, 2016.

J. Lemaire and H. Zi. A comparative analysis of 30 bonus-malus systems. *ASTIN Bulletin: The Journal of the IAA*, 24(2):287–309, 1994.

M. Rosenblatt. Prediction and non-gaussian autoregressive stationary sequences. *The Annals of Applied Probability*, 5(1):239–247, 1995.

C. I. Tan, J. Li, J. S-H. Li, and U. Balasooriya. Optimal relativities and transition rules of a bonus-malus system. *Insurance: Mathematics and Economics*, 61(C):255–263, 2015.

G. Tzougas, S. Vrontos, and N. Frangos. Optimal bonus-malus systems using finite mixture models. *ASTIN Bulletin*, 44(2):417–444, 2014. doi: 10.1017/asb.2013.31.

IS 27
Stochastic Comparisons
Rm -1.62
Wed
11:30 - 12:00

Sufficient conditions for some transform orders

Antonio Arriaza[*,‡], Félix Belzunce[†,§], Carolina Martínez-Riquelme[†,¶], Miguel A. Sordo[*,‖]

[*]Universidad de Cádiz, Facultad de Ciencias (Campus Universitario, Río San Pedro), Puerto Real, Cádiz (SPAIN)
[†]Universidad de Murcia, Facultad de Matemáticas (Campus de Espinardo), Murcia (Spain)
email: [‡]antoniojesus.arriaza@uca.es [§]belzunce@um.es [¶]carolina.martinez7@um.es
[‖]mangel.sordo@uca.es

In this talk we present several results that relate the unimodality of the ratio of two quantile density functions with some transform orders. In particular, we provide sufficient and, in some cases,

necessary conditions for the star-shaped, qmit and dmrl orders. These results are intended to be a tool for the comparison in the previous orders, when the convex order does not hold. Additional results for the relationship among the dmrl and star-shaped orders are also given, with implications in the context of ageing notions.

Acknowledgement. The authors want to acknowledge the support received by the Ministerio de Economía, Industria y Competitividad under grant MTM2016-79942-P (AEI/FEDER, UE).

On the squared telegraph process

IS 17
Stochastic Processes & Appl. I.
Rm -1.63
Tue
10:30 - 11:00

Barbara Martinucci*, Alessandra Meoli

Dipartimento di Matematica, Università degli Studi di Salerno, Fisciano (SA), Italy
email: *bmartinucci@unisa.it

The (integrated) telegraph process $X(t)$ describes a random motion with finite velocity on the line. This stochastic process is characterized by a probability law governed by a hyperbolic partial differential equation (the telegraph equation) widely used in mathematical physics, mostly in problems of electromagnetism. Moreover, such process deserves interest in many other applied fields, such as finance and mathematical biology. Among the first authors that studied the solution of the telegraph equation we recall Goldstein [1951] and Kac [1974]. Several aspects and generalization of the telegraph process have been provided in a quite large body of literature (see, for instance, Kolesnik and Ratanov [2013]).

In this talk we analyze the stochastic process $Y(t)$ defined as the square of the telegraph process. For this process we obtain a closed form expression of the moment generating function and study its limiting behavior as both the speed of the motion and the intensity of switchings tend to infinity. Specifically, under the Kac's scaling conditions, we show that, as expected, $Y(t)$ converges in distribution to the squared Brownian motion on $\mathbb{R}$.

Moreover, we consider two independent squared telegraph processes $Y_1(t)$ and $Y_2(t)$ both starting from the origin at the initial time $t = 0$ and developing on the line with finite constant velocity $c > 0$. Assuming that such processes are governed by two independent homogeneous Poisson processes characterized by the same rate, we obtain the probability law of the process $Y_1(t) + Y_2(t)$ and study its connection with the 2-dimensional squared-Bessel process.

Acknowledgement. The authors are member of the INdAM Research group GNCS.

References

Goldstein, S. (1951) On diffusion by discontinuous movements, and on the telegraph equation. *Quart J Mech Appl Math*, **4**, p. 129-156.

Kac, M. (1974) A stochastic model related to the telegrapher's equation. *Rocky Mountain J Math*, **4**, p. 497-509.

Kolesnik A.D. and Ratanov N. (2013) *Telegraph Processes and Option Pricing*, Springer Briefs in Statistics. Springer, Heidelberg.

IS 49
Urn Models
Rm -1.64
Wed
14:30 - 16:00

Urn Models with Irreducible, but Asymptotically Reducible, Replacement Matrices

KRISHANU MAULIK[*,‡], ROHAN SARKAR[†]
[*]Indian Statistical Institute, Kolkata, India
[†]Cornell University, Ithaca, USA
email: [‡]krishanu@isical.ac.in

The urn models with irreducible replacement matrices are well understood. In this presentation, we consider a sequence of replacement matrices which are irreducible at every stage, but converges to a matrix which is reducible. In particular, when there are two colors, the limiting replacement matrix can be taken as upper triangular. In this case, the count of the balls of the second color will continue to grow linearly. We provide a phase transition for the rate of growth of the count of the balls of the first color. We shall also discuss the models with higher numbers of colors.

Acknowledgement. K. Maulik's research was partially funded by an Unrestricted Research Grant from Microsoft Research India. Part of the research was done as Master's degree dissertation of R. Sarkar at Indian Statistical Institute, Kolkata, India under the supervision of the first author.

CS 15
State Space and Markov Models
Rm 0.99
Thu
10:50 - 11:10

A constrained Hidden Markov model for the Local Score of one sequence: an equivalence between HMM and the usual Local Score approach.

SABINE MERCIER[*,‡], GRÉGORY NUEL[†,§]
[*]Institut of Mathematics of Toulouse (IMT), University of Toulouse 2 Jean Jaurès, France
[†]Stochastics and Biology Group, Laboratory of Probability and Statistics (LPSM, CNRS 8001), Sorbonne Université, Paris, France
email: [‡]mercier@univ-tlse2.fr [§]nuel@math.cnrs.fr

Since the development of biological sequence databases in the 80's, the extraction of information from such an enormous amount of data has been the subject of great interest. Such works includes a large range of areas such as for example Markov models [Durbin et al 1998], segmentation [Luong et al 2013], and the critical problem of statistical significance that is omnipresent in biological sequence analysis [Karlin and Altschul 1990, Mercier and Daudin 2001].

We present here a new theoretical and practical approach for detecting atypical segments in a multi-state sequence $\mathbf{A} = A_1, \ldots, A_n \in \{1, \ldots, k\}^n$ using segmentation $\mathbf{S} = S_1, \ldots, S_n \in \{1, 2, 3\}^n$ through an underlying constrained Hidden Markov Model with the following unknown states: State 1 for components in a background region before an atypical segment and modeled using a given distribution q_0; state 2 for components in an atypical segment modeled by a distribution $q_1 \neq q_0$; state 3 for components in a background region after an atypical segment and modeled by q_0.

We provides exact estimates of posterior probabilities for the i-th component of the sequence $\mathbf{A}$ to be in an atypical segment, $\mathbf{P}(S_i = 2|\mathbf{A})$. Practical application is finally achievable following our adaptation of the standard Forward-Backward algorithm Durbin et al [1998] which provides exact estimates of posterior probabilities in a linear time. Additionally it can provide posterior probabilities on the segment length as well as its starting and ending indexes.

We also prove that the present segmentation approach is in fact equivalent to the local score approach [Karlin and Altschul 1990, Mercier and Daudin 2001] when the latter uses an appropriate rescaled scoring function. We illustrate this approach by rescaling the well-known Kyte and Doolitle

scoring function [Kyte and Doolittle 1982]. We finally explain how this approach might be used for learning a scoring function from a given dataset (*e.g.* a transmembrane protein data set).

Acknowledgement. S. Mercier and G. Nuel's research was supported by the CIMI project fund (International Center of Computer Science and Mathematics of Toulouse).

References

Durbin R, Eddy SR, Krogh A, Mitchison G (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge university press.

Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proceedings of the National Academy of Sciences*, **87(6)**, p.2264-2268

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology*, **157(1)**, p.105-132

Luong TM, Rozenholc Y, Nuel G (2013) Fast estimation of posterior probabilities in change-point analysis through a constrained hidden markov model. *Computational Statistics & Data Analysis*, **68**, p.129-140

Mercier S, Daudin JJ (2001) Exact distribution for the local score of one iid random sequence. *Journal of Computational Biology*, **8(4)**, p.373-380

An alternative common shock model and its applications to construct new copula families

CS 2
Copulas and independence
Rm 100/B
Thu
12:20 - 12:40

Mhamed Mesfioui

Université du Québec à Trois-Rivières
email: mhamed.mesfioui@uqtr.ca

The so-called trivariate reduction method is a popular approach widely used to construct multivariate distributions. It is well known that this method has two major drawbacks. On one hand, it can only model dependence positive; on the other hand, it cannot always span the full range of positive correlation. To remedy these drawbacks, Genest et al. [2018] introduced an alternative method which, contrary to the original, spans all possible degrees of dependence. This presentation will show that this novel idea can be used to construct a new class of copulas having an interesting stochastic representation. In particular, an extension of the Marshall–Olkin family of copulas will be presented. Some properties of this new family of copulas will be discussed.

Joint work with Christian Genest and Juliana Schulz.

References

C. Genest, M. Mesfioui, and J. Schulz. A new bivariate Poisson common shock model covering all possible degrees of dependence. *Statist. Probab. Lett.*, 2018. Under review.

POSTER
Mon
18:00-19:00

Moments and Mellin transform of the asset price in Stein and Stein model and option pricing

Zofia Michalik

Institute of Mathematics, University of Warsaw, Poland
email: z.michalik@mimuw.edu.pl

We consider the Stein and Stein stochastic volatility model. Under some constraints, we deliver closed formulas for moments and Mellin transform of the asset price. In the original paper of Stein and Stein [1991], the formulas for asset price distribution were obtained assuming that the noises driving the asset price and its volatility are uncorrelated. Here we avoid this restrictive assumption. We illustrate our results with some numerical examples of pricing financial derivatives such as self-quanto options and power options. We present different approaches based on computing either the moments or Mellin transform of the asset price. The presentation is based on Jakubowski et al. [2018].

References

Jakubowski, J., Michalik, Z., Wiśniewolski, M. (2018) Moments and Mellin transform of the asset price in Stein and Stein model and option pricing *Lithuanian Mathematical Journal*, doi: 10.1007/s10986-018-9380-9

Stein, E., Stein, J. (1991) Stock price distributions with stochastic volatility: an analytic approach *The Review of Financial Studies*, **4**, pp. 727-752, doi: 10.1093/rfs/4.4.727

IS 18
Stochastic Processes & Appl. II.
Rm -1.63
Tue
15:00 - 15:30

Stochastic Modelling and Statistical Analysis of Farmland Use Transitions in Lombardy Using a Markov Chains Approach

Giacomo Aletti[*,†,‡], Danilo Bertoni[*,§], Giulia Ferrandi[*,†,¶], Alessandra Micheletti[*,†,‖], Daniele Cavicchioli[*,**], Roberto Pretolani[*,††]

[*]Department of Environmental Science and Policy, Università degli Studi di Milano, Milano, Italy
[†]ADAMSS Center, Università degli Studi di Milano, Milano, Italy
email: [‡]alessandra.micheletti@unimi.it [§]giacomo.aletti@unimi.it
[¶]danilo.bertoni@unimi.it [‖]giulia.ferrandi1@studenti.unimi.it
[**]daniele.cavicchioli@unimi.it [††]roberto.pretolani@unimi.it

This work represents a preliminary attempt to evaluate ex-post impact of the common agricultural policy (CAP) *greening* payment on farmland use changes. Greening represents one of the main novelties of the current CAP programming period (entered in force in 2015), providing a horizontal payment for farmers, conditioned to the compliance with some "agricultural practices beneficial for the climate and the environment". These farm practices regard, and potentially influence, farmland allocation, particularly arable land and grassland.

On this ground, we pointed our attention on analysing at a very detailed (parcel) level the temporal and spatial dynamics of farmland use transitions before and after the introduction of greening commitments. We based our analysis on a huge dataset of about 2 millions of georeferenced land parcels in Lombardy over the period 2010-2016. Crop typologies have been aggregated into 23 different categories, in order to reduce the complexity of the analysis.

The system has been modelled as a Markov chain, where each land unit evolves, from one year to the other, into one of the 23 cultivation classes. Let us denote by $p_{ij}(t)$ the probability that a land unit evolves (i.e. is cultivated) from class i to class j, from year t to year $t+1$. Our aim was to check if any

statistically significant change in the transition probabilities $p_{ij}(t)$ and/or in the spatial distribution of the 23 cultivation categories, took place after the introduction of greening (that is between 2014 and 2015). Unfortunately a statistical test revealed a strong non stationarity in the $p_{ij}(t)$ for $t \leq 2014$, due to a possible correlation among data. This causes a problem in the statistical analysis, since the "physiological" variability registered before the new CAP must be filtered out for a correct comparison with the changes occurred from 2015 onwards. We solved the problem by introducing a new type of weighted χ^2 test, in which we determine the correct statistical unit that must be considered to accept the hypothesis of stationarity in a set of panel data. We applied this test to the complete set of available years (2010-2016) and we found evidence of change during 2015 in both the spatial distribution of the 23 cultivation classes and the transition probabilities of many relevant cultivations, like maize, maize for silage, wheat, soybean, etc. Furthermore we computed the Gini index to measure the heterogeneity of cultivations and the transition probabilities for the cultivation classes that resulted significant to the weighted χ^2 test. In this way we can visualise the zones of Lombardy which have mainly been affected by the greening policy.

Acknowledgement. This work is supported by the project CAPTION, funded by Fondazione Cariplo, and by the Universitá degli Studi di Milano grant project 2017 "Stochastic modelling, statistics and study of the invariance properties of stochastic processes with geometrical and space-time structure in applications".

Ruin probabilities for two collaborating insurance companies

IS 14
Excursions of Lévy Processes
Rm 100/B
Mon
12:00 - 12:30

Zbigniew Michna

Department of Mathematics and Cybernetics, Wrocław University of Economics, 53-345 Wrocław, Poland
email: zbigniew.michna@ue.wroc.pl

We find a formula for the supremum distribution of spectrally positive or negative Lévy processes with a broken linear drift. This gives formulas for ruin probabilities in the case when two insurance companies (or two branches of the same company) divide between them both claims and premia in some specified proportions. As an example we consider gamma Lévy process, α-stable Lévy process and Brownian motion. Moreover we obtain identities for Laplace transform of the distribution for the supremum of Lévy processes with randomly broken drift and on random intervals.

Individualized Fusion Learning (iFusion) with Applications to Personalized Inference

IS 45
Inference for Complex Data
Rm -1.64
Wed
11:00 - 11:30

Jieli Shen, Regina Liu, Minge Xie[*,†]

[*]Rutgers University, USA
email: [†]mxie@stat.rutgers.edu

Statistical inferences from multiple data sources can often be fused together to yield more effective inference than from individual source alone. Such fusion learning is of vital importance for big data where data are often assembled in various domains. This paper develops a fusion methodology called individualized fusion learning (iFusion), to enhancing inference for an individual via adaptive combination of confidence distributions obtained from its clique (i.e., peers of similar individuals). iFusion begins with obtaining inference for each individual, then adaptively forming a clique, and finally obtaining a combined inference from the clique. iFusion explores heterogeneity in the database to form a clique for each individual and, by drawing inference from the clique, it allows borrowing strength from similar peers to enhance the inference efficiency for each individual. Furthermore, iFusion can

be performed without using the entire data simultaneously and thus allow split-&-conquer to be implemented on individuals to substantially reduce the computational expense. We provide supporting theories for iFusion and also illustrate it using numerical examples.

Acknowledgement. The research is supported in part by grants: NSF-DMS 1513483 and NSF-DMS 1737857.

CS 5
Branching Processes
Rm -1.62
Tue
18:10-18:30

Branching processes in a varying environment with inhomogeneous immigration

Miguel González[*,†,§], Götz Kersting[‡,¶], Carmen Minuesa[*,‖], Inés del Puerto[*,†,**]

Department of Mathematics, University of Extremadura, Badajoz (Spain)[*]
Instituto de Computación Científica Avanzada, University of Extremadura, Badajoz (Spain) [†]
Institute of Mathematics, Goethe University Frankfurt, Frankfurt am Main (Germany)[‡]
email: [§]mvelasco@unex.es [¶]kersting@math.uni-frankfurt.de [‖]cminuesaa@unex.es
[**]idelpuerto@unex.es

In this talk, we present a modification of the standard branching model allowing immigration, the branching process in a varying environment with generation-dependent immigration. This model is characterised by having inhomogeneous reproduction and immigration, that is, the probability distributions governing the reproduction and immigration processes may change over the time. These features make the process more appealing from a practical point of view. On the one hand, the reproductive capacity of the individuals may be affected by factors that change over seasons and on the other hand, the influx of immigrants in the population is usually not constant over the generations. For this model, in relation to the extinction problem, a necessary and sufficient condition for the certain extinction of these populations is provided. Their limiting behaviour is studied for processes with critical offspring distributions according with the classification given in Kersting [2017]. Indeed, we establish the asymptotic distribution of the process -under a suitable normalization- when the immigration means stabilize to a positive value.

Acknowledgement. This research has been supported by the Ministerio de Educación, Cultura y Deporte (grants FPU13/03213 and EST16/00404), the Ministerio de Economía y Competitividad (grant MTM2015-70522-P), the Junta de Extremadura (grant IB16099) and the Fondo Europeo de Desarrollo Regional.

References

G. Kersting (2017) A unifying approach to branching processes in varying environment. arXiv: 1703.01960, p.1–23

PLENARY
Harmony Inge
Thu
9:00 - 10:00

Fractional Cox-Ross-Ingersoll processes and fractional stochastic volatility

Yuliya Mishura

Taras Shevchenko National University of Kyiv, Ukraine
email: myus@univ.kiev.ua

This is the common work with A. Yurchenko-Titarenko, V. Piterbarg and K. Ralchenko. We consider the fractional Cox-Ross-Ingersoll process with arbitrary Hurst index $H \in (0,1)$ that is the solution of the corresponding Stratonovich equation. Its properties are studied for the different values of the drift coefficients. Its behavior is different depending on whether $H < 1/2$ or, oppositely, $H > 1/2$. We consider some approximations of this process and study option pricing with stochastic volatility describing by the fractional Cox-Ross-Ingersoll process.

The steady state in the models of the population dynamics and their stability with respect to local or random perturbations

IS 25
Population Dynamics Studies
Rm -1.62
Mon
16:00 - 16:30

Stanislav Molchanov

UNC Charlotte, USA and HSE, Moscow, Russia
email: smolchan@uncc.edu

This is a review of the recent progress in the population dynamics based on the work of several groups involved in this research: E. Yarovaya, E. Chernousova (Moscow, Russia), Yu. Kondratiev, O. Kutotovyi (Bielefeld, Germany), O. Hryniv (Durham Great Britain), Y. Feng, D. Han, S. Molchanov (UNC Charlotte, USA).

There are several continuous and lattice models of the particle fields whose evolution includes the death and birth of processes , the random motion in the phase space (migration), immigration of the particles from outside and in some cases, the mean field type interaction. Under certain conditions, these models demonstrate the ergodicity, i.e. convergence to the statistical equilibrium (steady state). The proof of these results is based on the asymptotic analysis of the corresponding moments equations, i.e. the correlation functions. In the homogeneous medium the limiting correlation function (at least of order 1, 2, 3) can be explicitly calculated. But in the biological reality, the environment is not homogeneous. The study of the stability of the steady states with respect to different classes of perturbation (local, random etc.) contains as the main ingredient the spectral analysis of the non-local Schrödinger operators with the different assumptions on the potentials. Some of these results have the similarity with the classical case of the Hamiltonians in $\mathbb{R}^d$, some are essentially different. The highest level of stability is demonstrated by the models with immigration or mean field competition.

Probabilistic temperature forecasting using D-vine copula regression

IS 35
Environmental Models
Rm -1.63
Thu
12:00 - 12:30

Annette Möller[*,¶], Ludovica Spazzini[†], Daniel Kraus[‡], Claudia Czado[§]

[*]Institute of Applied Stochastics and Operations Research, Clausthal University of Technology, Clausthal-Zellerfeld, Germany
[†]Partners4Innovation, Milano, Italy
[‡]Bayrische Landesbank, Munich, Germany
[§]Department of Mathematics, Technical University of Munich, Munich, Germany
email: [¶]annette.moeller@tu-clausthal.de

To account for forecast uncertainty in numerical weather prediction (NWP) models it has become common practise to employ ensemble prediction systems that generate probabilistic forecast ensembles by multiple runs of the NWP model, each time with variations in the details of the numerical model and/or initial and boundary conditions.

However, forecast ensembles typically exhibit biases and dispersion errors as they are not able to fully represent uncertainty in the NWP models. Therefore, statistical postprocessing models have been developed to correct ensembles for biases and dispersion errors in conjunction with recently observed forecast errors. A popular and state-of-the art postprocessing method is the ensemble model output statistics (EMOS), which utilizes a regression framework, where a single parametric predictive distribution is obtained, with the parameters depending on the ensemble members in suitable ways.

We propose a novel postprocessing model for temperature forecasts based on D-vine copula quantile regression and compare its performance to the state-of-the-art EMOS model.

The D-vine copula regression model is a multivariate regression approach that predicts quantiles of a response (in our case temperature) based on a set of predictor variables, which will be the ensemble members. It allows to exploit the dependence structure of the observation and the predictors, accounting for non-gaussian dependencies in a flexible way. Furthermore, the approach is able to identify highly informative predictors within the procedure of selecting the vine structure. Therefore, the D-vine copula postprocessing method constitutes an alternative way to modelling of groups among members, namely by automatically selecting informative members and eventually discarding redundant (e.g. exchangeable) members from the model. The respective predictive distribution can be represented by sampling quantiles on a dense grid.

In a comparative study with temperature forecasts of different forecast horizons from the European Center for Medium Range Weather Forecast (ECMWF) the D-vine postprocessing approach shows to be highly competitive to the state-of-the-art EMOS model, clearly improving over standard EMOS for large forecast horizons.

As the case study indicates the D-vine postprocessing model provides an improvement specifically for settings with highly non-gaussian dependencies, a further extension is considered. The idea is to combine the advantages of both models in a regime-switching approach that monitors for example changes in the dispersion properties of the ensemble.

IS 41
Quickest Change Detection
Rm 100/A
Wed
10:30 - 11:00

Sequential Change-Detection for Markov Processes

George V. Moustakides

Rutgers University, New Brunswick, New Jersey, USA
University of Patras, Patras, Greece
email: gm463@rutgers.edu, moustaki@upatras.gr

We consider the problem of detecting an abrupt change in a discrete-time Markov process which is observed sequentially. A detector in this case is simply a stopping time (the time we stop and declare that a change took place) which is adapted to the observations. We assume that the pre- and post-change probability measures describing the statistical behavior of the observations are known, but the time of change is unknown. As performance measure for our detector we adopt the criterion proposed by Lorden [1971] which, for each fixed deterministic change-time, evaluates the average detection delay conditioned on the worst possible data before the change and then considers the worst possible deterministic change-time. Our goal is to minimize this worst-case scenario subject to the false alarm constraint that the average period of false alarms is no smaller than some minimal tolerable value. The solution of this problem when the observations are i.i.d. before and after the change with different pdfs, gives rise to the well-known CUSUM test (Lorden [1971], Moustakides [1986]). We extend this exact optimality result to cover data that are homogeneous Markov before and after the change with different conditional pdfs. We demonstrate that the optimum detection strategy is CUSUM-like containing two univariate functions that must be properly specified. For these two functions we provide equations that are sufficient for their determination but we also show that the same equations play a crucial role in the proof of optimality of the proposed scheme. Concluding, we should mention that this problem has been open for over thirty years.

Acknowledgement. This work was supported by the US National Science Foundation under Grant CIF 1513373, through Rutgers University.

References

Lorden, G. (1971) Procedures for reacting to a change in disrtibution. *Ann. Math. Stat.*, **42**, p.1897-1908.

Moustakides, G. V. (1986) Optimal stopping times for detecting changes in distributions. *Ann. Stat.*,**14**, p.1379-1387.

Functional Central Limit Theorems for Rough Volatility

OCS 6
Rough Volatility
Harmony Inge
Thu
12:00 - 12:20

Blanka Horvath[*,†], Antoine Jacquier[*,‡], Aitor Muguruza[*,§]

[*]Imperial College London
email: [†]b.horvath@imperial.ac.uk [‡]a.jacquier@imperial.ac.uk
[§]aitor.muguruza-gonzalez15@imperial.ac.uk

We extend Donsker's approximation of Brownian motion to fractional Brownian motion with Hurst exponent $H \in (0,1)$ and to Volterra-like processes. Some of the most relevant consequences of our 'rough Donsker (rDonsker) Theorem' are convergence results for discrete approximations of a large class of rough models. This justifies the validity of simple and easy-to-implement Monte-Carlo methods, for which we provide detailed numerical recipes. We test these against the current benchmark Hybrid scheme and find remarkable agreement (for a large range of values of H). This rDonsker Theorem further provides a weak convergence proof for the Hybrid scheme itself, and allows to construct binomial trees for rough volatility models, the first available scheme (in the rough volatility context) for early exercise options such as American or Bermudan.

Acknowledgement. The authors would like to thank Christian Bayer, Peter Friz, Paul Gassiat, Jim Gatheral, Mikko Pakkanen and Mathieu Rosenbaum for useful discussions. BH gratefully acknowledges financial support from the SNSF Early Postdoc.Mobility grant 165248, and AM is grateful to the Centre for Doctoral Training in Financial Computing & Analytics for financial support.

References

C. Bayer, P. Friz and J. Gatheral (2015).Pricing under rough volatility. *Quantitative Finance*: 1-18, 2015.

M. Bennedsen, A. Lunde and M. S. Pakkanen (2017). Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics 21(4)*, p.931-965.

J. Gatheral, T. Jaisson and M. Rosenbaum. Volatility is rough.arXiv: 1410.3394

J. Lamperti (1965). On convergence of stochastic processes. *Trans. Amer. Math. Soc. 104*, pp.430-435, 1965.

S. G. Samko, A. A. Kilbas and O. I. Marichev (1993). Fractional Integrals and Derivatives: Theory and Applications. *Gordon and Breach, Yverdon*, 1993

Equilibrium Asset Pricing with Transaction Costs

IS 7
Optimisatio & Machine Learning
Harmony Inge
Wed
11:00 - 11:30

Johannes Muhle-Karbe

Carnegie Mellon University, Pittsburgh
email: johannesmk@cmu.edu

We study the impact of quadratic transaction costs on risk-sharing economies. Using tools from stochastic analysis, we address the existence, uniqueness, and characterization of equilibria in this context. In particular, we discuss the effects trading costs have on equilibrium asset prices and their dynamics. (Based on joint works with Bruno Bouchard, Masaaki Fukasawa, Martin Herdegen, and Dylan Possamai.)

OCS 4
Markov & Semi-Markov Models
Rm -1.62
Tue
12:40 - 13:00

The Multivariate Wishart Distribution Based on Generalized Vandermonde Determinant

Asaph Keikara Muhumuza[*,†,§], Karl Lundengard[†], Jonas Osterberg[†], Sergei Silvestrov[†], John Mango Magero[‡], Godwin Kakuba[‡]

[*]Department of Mathematics, Busitema University, Box 236, Tororo Uganda.
[†]Division of Applied Mathematics, Malardalen University, Box 883, Hogskoleplan 1, SE 721 23 Vasteras, Sweden
[‡]Department of Mathematics, Makerere University, Box 7062, Kampala Uganda.
email: [§]asaph.keikara.muhumuza@mdh.se

A number of models from mathematics, physics, probability theory and statistics can be described in terms of orthogonal polynomials [3, 4] based on the Vandermonde determinant [1]. The most prominent example being the Laguerre ensembles of the spectrum of Wishart distribution. We aim to construct the distribution of the density of the eigenvalues of a random symmetric matrix using techniques for optimization of the generalized Vandermonde determinant over various surfaces [2].

References

[1] Abramowitz, M. and Stegun, I. (1964) Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. *Dover*, New York.
[2] Lundengård, K. Österberg, J. and Silvestrov, S. (2012) Optimization of the Determinant of the Vandermonde Matrix and Related Matrices. *Methodology and Computing in Applied Probability*, **19**, 4, p.1–12.
[3] Szegő G.(1939) Orthogonal Polynomials. *American Mathematical Society*.
[4] Vein, R. and Dale, P. (1999) Determinants and Their Applications in Mathematical Physics. *Applied Mathematical Sciences*, **134**, Springer, New York.

IS 36
Applied Prob. & Stat. Inference I.
Rm 100/A
Mon
12:00 - 12:30

Purely Sequential and Two-Stage Bounded-Length Confidence Interval Estimation Problems

Nitis Mukhopadhyay

University of Connecticut-Storrs, USA
email: nitis.mukhopadhyay@uconn.edu

Fisher's "Nile" example is a classic which involves responses from a bivariate random variable (X, Y) where X, Y are independent exponential random variables with respective means $1/b$ and b with b being a positive unknown parameter. We present bounded-length confidence interval estimation for the probability that X exceeds a given positive number a with some preassigned confidence coefficient using both purely sequential and two-stage methodologies. We show that both these methodologies enjoy asymptotic (i) first-order efficiency and asymptotic consistency properties; (ii) second-order efficiency properties. After summarizing relevant theory, we use simulations to empirically validate the theoretical properties. This is joint work with my PhD student, Yan Zhuang.

New Stochastic Comparisons Based on Tail Values at Risk

CS 19
Distributio
Rm 100/B
Tue
17:50-18:10

Julio Mulero[*,†,§], Félix Belzunce[†], Alba María Franco-Pereira[‡]

[*]University of Alicante, Alicante, Spain
[†]University of Murcia, Murcia, Spain
[‡]University Complutense, Madrid, Spain
email: [§]julio.mulero@ua.es

One of the most known and useful risk measures is the value at risk (VaR), but it has some limitations that became evident from the crisis that began in 2008. For this reason, in the Basel III agreements (2010) the use of the value-at-risk tail (TVaR), which offers more representative information on risks, was adopted. In the context of insurance theory, insurance companies go to reinsurance companies to protect their capital against possible large losses. In this sense, the claimed economic amount that happens to be assumed by the reinsurers can be studied from the residual life of said claims. In this work, we analyze the interest of comparing these quantities. In addition, we provide sufficient conditions for its study and illustrate its usefulness with some examples.

Acknowledgement. The research of Félix Belzunce and Julio Mulero is partially funded by the Ministerio de Economía y Competitividad (Spain) under grant **MTM2012-34023-FEDER**. Julio Mulero also wants to acknowledge the support received from the Conselleria d'Educació, Investigació, Cultura i Esport (Generalitat de la Comunitat Valenciana) under grant **GV/2017/015**. Alba M. Franco-Pereira acknowledges support received from the Ministerio de Economía y Competitividad (Spain) under grant **MTM2014-55966-P** and has also received financial support from the Xunta de Galicia (Centro Singular de Investigación de Galicia accreditation 2016-2019) and the European Union (European Regional Development Fund - ERDF).

Life Insurance Pricing with the Extended Marshall-Olkin Model

IS 10
Copula Functions
Rm 100/B
Wed
11:30 - 12:00

Fabio Gobbi[*], Nikolai Kolev[†], Sabrina Mulinacci[‡,§]

[*]Department of Statistics, University of Bologna, Italy
[†]Department of Statistics, University of São Paulo, São Paulo, Brazil
[‡]Department of Statistics, University of Bologna, Italy
email: [§]sabrina.mulinacci@unibo.it

Bivariate copula functions have been widely used to model the dependence structure between the residual lifetimes of the two individuals in a couple. However, considered copulas are absolutely continuous and do not allow for the case of a simultaneous death due to some catastrophic event. In order to include this case, we will consider the Extendend Marshall-Olkin distribution: this distribution is a generalization of the Marshall-Olkin one since marginal residual lifetimes are not constrained to be exponentially distributed and some dependence is allowed between the idiosyncratic residual lifetimes components. This model is applied to the same dataset provided by a large Canadian insurance company already considered in most of the existing studies in this topic.

IS 45
Inference for Complex Data
Rm -1.64
Wed
10:30 - 11:00

The Five Trolls under the Bridge: Principal Component Analysis with Asynchronous and Noisy High Frequency Data

Per A Mykland
University of Chicago
email: mykland@pascal.uchicago.edu

We develop a principal component analysis (PCA) for high frequency data. As in Northern fairly tales, there are trolls waiting for the explorer. The first three trolls are market microstructure noise, asynchronous sampling times, and edge effects in estimators. To get around these, a robust estimator of the spot covariance matrix is developed based on the Smoothed TSRV. The fourth troll is how to pass from estimated time-varying covariance matrix to PCA. Under finite dimensionality, we develop this methodology through the estimation of realized spectral functions. Rates of convergence and central limit theory, as well as an estimator of standard error, are established. The fifth troll is high dimension on top of high frequency, where we also develop PCA. With the help of a new identity concerning the spot principal orthogonal complement, the high-dimensional rates of convergence have been studied after eliminating several strong assumptions in classical PCA. As an application, we show that our first principal component (PC) closely matches but potentially outperforms the S&P 100 market index, while three of the next four PCs are cointegrated with two of the Fama-French non-market factors. From a statistical standpoint, the close match between the first PC and the market index also corroborates this PCA procedure and the underlying S-TSRV matrix, in the sense of Karl Popper. (Joint work with Dachuan Chen and Lan Zhang at University of Illinois at Chicago.)

IS 12
Efficient Monte Carlo Simulation
Rm 100/B
Thu
11:30 - 12:00

Effective methods for infinite variance in stochastic simulation

Leonardo Rojas Nandayapa
Institute of Financial and Actuarial Mathematics, University of Liverpool
email: leorojas@liverpool.ac.uk

In this talk I will discuss the challenges of implementing a Monte Carlo estimator having infinite variance, and illustrate with examples that such a setting is not so unusual in applications. Then I will discuss various effective methods to mitigate the effect of the infinite variance.

Acknowledgement. This is joint work with Thomas Taimre and Wangyue Xie

IS 27
Stochastic Comparisons
Rm -1.62
Wed
10:30 - 11:00

Stochastic comparisons of conditional distributions based on copula properties

Jorge Navarro[*,‡], Miguel A. Sordo[†]
[*]Universidad de Murcia, Spain
[†]Universidad de Cadiz, Spain
email: [‡]jorgenav@um.es

We show that different conditional distributions can be represented as distorted distributions. For example, in the bivariate case, we consider $(Y|X > x)$, $(Y|X \leq x)$, and $(Y|X = x)$. These

representations are used to obtain stochastic comparisons and bounds for them based on properties of the underlying copula and ordering properties for distorted distributions given in Navarro and del Águila [2017], Navarro et al. [2013, 2016] These properties can be used to explain the meaning of mathematical properties of copulas connecting them with dependence concepts. Some applications and illustrative examples will be provided. This talk is based on the paper Navarro and Sordo [2017].

Acknowledgement. J. Navarro research was supported by Ministerio de Economía, Industria y Competitividad of Spain under grant MTM2016-79943-P and M. A. Sordo by Ministerio de Economía, Industria y Competitividad of Spain under grant MTM2014-57559-P.

References

Navarro, J. and del Águila, Y. (2017) Stochastic comparisons of distorted distributions, coherent systems and mixtures with ordered components. *Metrika* **80**, p. 627–648.

Navarro, J., del Águila, Y., Sordo, M. A., Suárez-Llorens, A. (2013) Stochastic ordering properties for systems with dependent identically distributed components. *Applied Stochastic Models in Business and Industry* **29**, p. 264–278.

Navarro, J., del Águila, Y., Sordo, M. A., Suárez-Llorens, A. (2016) Preservation of stochastic orders under the formation of generalized distorted distributions. *Methodology and Computing in Applied Probability* **18**, p. 529–545.

Navarro, J. and Sordo, M. A. (2017) Stochastic comparisons and bounds for conditional distributions by using copula properties. *Submitted*.

Probabilistic Methods for Data Perturbation for Protecting Respondent's Privacy

IS 37
Applied Prob. & Stat. Inference II.
Rm 100/A
Mon
16:00 - 16:30

Tapan K. Nayak

George Washington University, Washington, DC, USA
email: tapan@gwu.edu

Various randomized response (RR) methods, which probabilistically perturb the true response values, have been proposed for using in statistical surveys to protect respondents' privacy. Interest in RR methods has grown significantly in recent years, for devising privacy-preserving data collection processes from on-line surveys, transactions, searches and postings. Although the RR topic has been investigated for over 50 years, how to set privacy protection goals and how to choose the randomization probabilities to achieve the stated privacy goals have received only modest attention and yielded inadequate guidance.

We consider a stringent view of privacy protection; namely, a privacy mechanism must guarantee that no intruder will gain much new information about any respondent from his/her response. For a categorical survey variable, we present a formal and general development of this idea and explore its implications and properties. In our formulation, the substantive outcomes and results depend largely on the specification of "much" in much new information. We bring out substantial guidance on setting privacy protection goals, and for any specified privacy requirement, we characterize all privacy satisfying RR procedures. By-products of this result show a direct connection to local differential privacy, and allow us to compare various privacy specifications, and privacy levels of different RR mechanisms. To compare data utility of all privacy preserving RR procedures, we use the sufficiency of experiments concept, which is agnostic to inferential goals and loss functions and thus very general, and obtain a complete characterization of the class of all admissible procedures. Moreover, we derive the optimum privacy preserving RR procedure under an intuitive and commonly accepted criterion.

IS 5
Rough Volatility & Market Impact
Harmony Inge
Thu
11:00 - 11:30

Incorporating Signals into Optimal Trading

Eyal Neuman
Imperial College London
email: e.neumann@imperial.ac.uk

Optimal trading is a recent field of research which was initiated by Almgren, Chriss, Bertsimas and Lo in the late 90's. Its main application is slicing large trading orders, in the interest of minimizing trading costs and potential perturbations of price dynamics due to liquidity shocks. The initial optimization frameworks were based on mean-variance minimization for the trading costs. In the past 15 years, finer modelling of price dynamics, more realistic control variables and different cost functionals were developed. The inclusion of signals (i.e. short term predictors of price dynamics) in optimal trading is a recent development and it is also the subject of this work.

We incorporate a Markovian signal in the optimal trading framework which was initially proposed by Gatheral, Schied, and Slynko (2012) and provide results on the existence and uniqueness of an optimal trading strategy. Moreover, we derive an explicit singular optimal strategy for the special case of an Ornstein-Uhlenbeck signal and an exponentially decaying transient market impact. The combination of a mean-reverting signal along with a market impact decay is of special interest, since they affect the short term price variations in opposite directions.

In order to support our models, we analyse nine months of tick by tick data on 13 European stocks from the NASDAQ OMX exchange. We show that orderbook imbalance is a predictor of the future price move and it has some mean-reverting properties. From this data we show that market participants, especially high frequency traders, use this signal in their trading strategies.

This is a joint work with Charles-Albert Lehalle.

OCS 2
Non-Marko Models in Finance
Harmony Inge
Tue
12:40 - 13:00

Log-optimal Portfolios with Memory Effect

Zsolt Nika[*,‡], Miklós Rásonyi[†]
[*]Pázmány Péter Catholic University, Budapest, Hungary
[†]Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary
email: [‡]zsolt.nika@gmail.com

I am going to talk about an investment problem with logarithmic utility function where we take into consideration the so-called 'memory effect' in the stock price dynamics. While general theories of log-optimal investment are well-elaborated, there is a lack of construction for the optimal strategy in parametric models. I will present an algorithm for a broad class of stock prices with two examples and then I will show how the parameters related to 'memory effect' affect the optimal solution.

The two examples for stock price dynamics will be presented in detail and also that the algorithm is computationally feasible.

Acknowledgement. The first author gratefully acknowledges support of the Széchenyi 2020 Program, of the Human Resource Development Operational Program, and of the Program of Integrated Territorial Investments in Central-Hungary (project numbers: EFOP- 3.6.2-16-2017-00013 and 3.6.3-VEKOP- 16-2017-00002), and of the European Structural and Investment Funds. The second author acknowledges support from the NKFIH (National Research, Development and Innovation Office, Hungary) grant KH 126505 and from the "Lendület" grant LP 2015-6 of the Hungarian Academy of Sciences

Conservative bounds in a certain first-passage-problem

IS 40
Sequential Methods
Rm 100/A
Wed
14:30 - 16:00

Igor Nikiforov

Université de Technologie de Troyes UTT/ICD/LM2S, UMR 6281, CNRS, 12, rue Marie Curie CS 42060, 10004, Troyes Cedex, France
email: Igor.Nikiforov@utt.fr

The goal is to calculate the probability that a discrete time autoregressive process (scalar or vector) $\{Q_n\}_{n\geq 1}$ leaves a given interval (domain) at least once during a certain period T. Let us define the stopping time $N = \inf\{n \geq 1 : |Q_n| \geq h\}$ ($N = \inf\{n \geq 1 : \|Q_n\|_2 \geq h\}$, respectively). We are interested in the calculation of the following probability $\mathbb{P}(N \leq T \mid |Q_0| < h)$ ($\mathbb{P}(N \leq T \mid \|Q_0\|_2 < h)$, respectively). The distributions of the innovation process and the initial state Q_0 are unknown but some special bounds for the cumulative distribution functions and/or for the probability density functions are available. Numerical methods to calculate the conservative bounds for the above-mentioned probability are considered. Such bounds can be useful in sequential change detection problem and also for risk estimation in some safety-critical applications.

Acknowledgement. The author gratefully acknowledges the research and financial support of this work from the Thales Alenia Space, France.

Convergence to the Mean Field Game Limit: A Case Study

IS 2
Optimal Transport
...
Rm 100/B
Mon
15:30 - 16:00

Marcel Nutz

Columbia University, New York
email: mnutz@columbia.edu

We study the convergence and multiplicity of equilibria in a tractable game of optimal stopping. If the mean field game has a unique equilibrium, any sequence of n-player equilibria converges to it as $n \to \infty$. Whereas in the case of non-uniqueness, it is shown that an additional stability condition is needed to ensure that a mean field equilibrium is the limit of n-player equilibria. (Joint work with Xiaowei Tan)

Computational Methods for Martingale Optimal Transport problems

IS 3
Optimal Transport
...
Harmony Inge
Wed
15:30 - 16:00

Jan Obłój

Mathematical Institute University of Oxford
email: jan.obloj@maths.ox.ac.uk

We develop numerical methods for solving the martingale optimal transport (MOT) problem. We prove that the MOT problem can be approximated through a sequence of linear programming (LP) problems which result from a discretisation of the marginal distributions combined with a suitable relaxation of the martingale constraint. Specialising to the one-step model in dimension one, we provide an estimation of the convergence rate. We adopt two computational algorithms to solve the LP problem that are related to a tailored discretisation of the marginals preserving the increasing convex order, based respectively on the iterative Bregman projection and stochastic averaged gradient method. Joint work with Gaoyue Guo.

POSTER
Mon
18:00-19:00

Bayesian Analysis of Multivariate Count Data Using Copula

Man-Suk Oh

Ewha Womans University, Seoul, Korea
email: msoh@ewha.ac.kr

There has been growing interest in modeling correlated multivariate count data over the past decade. Multivariate Poisson regression models, multivariate negative binomial regression models, and multivariate Poisson-Lognormal regression models have been used to incorporate correlations between multivariate count variables. We introduce a more general copula-based multivariate count regression model which incorporate the dependence among the multivariate counts by modeling multivariate random effects using copulas. By using copulas we can separate the modeling of the marginal distributions of the count varaibles from the modeling of the dependence stucture between the variables. Overdispersion and general correlation structures including both positive and negative correlations in multivariate counts can easily be accounted for by this approach. Our copular-based models can also encompass previously suggested multivariate negative binomial regression models and multivariate Poisson-Lognormal regression models. We employ a Bayesian method for estimating parameters in the proposed model using Markov chain Monte Carlo. The method is illustrated with the crash count data of five different severity levels collected from 451 three-leg unsignalized intersections in California.

IS 5
Rough Volatility & Market Impact
Harmony Inge
Thu
11:30 - 12:00

Modelling Limit Order Book Data by State-Dependent Hawkes Processes

Maxime Morariu-Patrichi[*], Mikko S. Pakkanen[*,†,‡]

[*]Imperial College London, London, UK
[†]CREATES, Aarhus University, Denmark
email: [‡]m.pakkanen@imperial.ac.uk

During the past ten years, self-exciting Hawkes processes have become a popular model of limit order book data in statistical finance, as they are able to capture the endogeneity and feedback effects in order flow data at high-frequency time scales. In an order flow model built on Hawkes processes, the arrival rate of new orders depends on the past order flow, but it cannot depend on any state variables of the limit order book, such as the current bid/ask price or queue imbalance. To address this limitation of Hawkes processes, we develop a state-dependent extension of a Hawkes process using the theory of hybrid marked point processes [Morariu-Patrichi and Pakkanen 2017]. Our new framework couples the Hawkes process to a state process that influences the arrival rate of new orders, whilst the arriving orders may, reciprocally, prompt the state process to move to a new state. In my talk, I will focus on estimation methodology for the new class of processes and present some preliminary empirical results using NASDAQ data.

Acknowledgement. M. S. Pakkanen acknowledges partial support from CREATES (DNRF78), funded by the Danish National Research Foundation, and from EPSRC through the Platform Grant EP/I019111/1.

References

Morariu-Patrichi, M. and Pakkanen, M. S. (2017) Hybrid marked point processes: characterisation, existence and uniqueness. arXiv: 1707.06970

Multivariate subordination of stable processes

CS 11
Lévy processes
Rm -1.62
Wed
12:00 - 12:20

Vladimir Panov

National Research University - Higher School of Economics, Moscow, Russia
email: vpanov@hse.ru

Stochastic time change is a well-used tool for construction of stochastic models which are able to represent the so-called stylized features of stock prices. From mathematical point of view, the main idea is to change the deterministic time t of a stochastic process $X(t)$ (usually - of a Lévy process) by another increasing process $T(s)$. As a result, one obtains a process $Y(s) = X(T(s))$, which is referred to as a time-changed process. The economical interpretation of this model is based on the idea that the "business" time $T(s)$ may run faster than the physical time in some periods, for instance, when the amount of transactions is high. Due to this interpretation, $Y(s)$ represents the log-returns of a stock price, and a natural candidate for $T(s)$ is a cumulative number of trades till time s.

The most popular choice of a process X is a Brownian motion with or without drift. This choice is mainly based on the Monroe theorem, which says that the class of time-changed Brownian motions in fact coincides with the class of all semimartingales.

In this research, we consider another case, when the class of stable processes is used for X. Empirically it turns out that the considered model is more appropriate than the subordinated Brownian motion for describing the stock returns. This can be explained by the observation that in our model rapid changes in log-returns are made not only due to jumps in number of trades (as in time-changed Brownian motion), but also due to stochastic factors, which are incorporated in X.

More precisely, we propose a multivariate time-changed model such that each component is a subordinated stable process and the dependence between subordinators is described via some Lévy copula. For this model, I will show a simulation method based on the series representation. Moreover, I will describe a method of statistical estimation of the parameters of copula and related distributions, and prove some properties of the considered estimates. The performance of the proposed method will be illustrated by the examples of asset prices.

References

Panov V. and Samarin, E. (2018) Multivariate subordination of stable processes. arXiv: 1802.02876

Panov, V. (2017). Series representations for multivariate time-changed Lévy models. *Methodology and computing in applied probability*, **19 (1)**, pp.97-119.

Belomestny, D. and Panov, V. (2013). Estimation of the activity of jumps in time-changed Lévy models. *Electronic journal of statistics*, **7**, pp. 2970-3003.

Exponential approximation for branching random walk via Stein's method

CS 15
State Space and Markov Models
Rm 0.99
Thu
11:30 - 11:50

Erol Pekoz

Boston University, Boston
email: pekoz@bu.edu

For the critical nearest-neighbor multidimensional branching random walk conditional on non-extinction, we show convergence to an exponential distribution for the number of sites with a given multiplicity of particles. We also get a rate of convergence using a version of Stein's method.

PLENARY
Harmony Inge
Mon
14:00 - 15:00

Modeling model uncertainty of pdf

Shige Peng

Shandong University
email: peng@sdu.edu.cn

We present a new method of using parameterized fully nonlinear parabolic PDE to calculate the uncertainty of probability distributions hidden behind real data. Typical examples are the nonlinear maximal distribution and the nonlinear normal distribution.

OCS 1
Large Random Graphs
Rm -1.64
Mon
16:10 - 16:30

Simulation results in a random network

István Fazekas[*], Attila Perecsényi[*,†], Bettina Porvázsnyik[*]

[*]University of Debrecen, Debrecen, Hungary
email: [†]perecsenyi.attila@inf.unideb.hu

Nowadays we can find many real-life networks in the society. To describe real-life networks such as the Internet, WWW, power grid and social networks (see the monograph of Barabási, A. L. [2016]), Barabási, A. L. and Albert, R. [1999] introduced a random graph model. They defined an evolving graph using the preferential attachment rule.

There exist several versions of preferential attachment models. For example Cooper, C., and Frieze, A. [2003] defined a mixed model. In that model, when a new vertex is born we can choose uniformly at random or by using the preferential attachment rule. This idea was followed by Backhausz, Á. and Móri, T. F. [2014] and Fazekas, I. and Porvázsnyik, B. [2016].

Recently Fazekas, I., Perecsényi, A. and Porvázsnyik, B. [2017] introduced a new random graph model, which is a generalization of the N-interaction model of Fazekas, I. and Porvázsnyik, B. [2016]. This is a growing, weighted, mixed model. The basic feature of the model is the cooperation (interaction) of N vertices. The choice of the cooperating individuals can be based either on the preferential attachment rule or can be made uniformly at random. In this model every step m new vertices are born, where m is a discrete random variable with values $0, 1, 2, \ldots, N-1$. Then the m new vertices interact with $(N-m)$ old nodes, so that they form a complete graph on N vertices.

In this presentation we will show numerical results of the above mentioned model. For example power-law distributions, clustering coefficients and diameters will be presented.

Acknowledgement. This work was supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project has been supported by the European Union, co-financed by the European Social Fund.

References

Backhausz, Á. and Móri, T. F. (2014) Weights and degrees in a random graph model based on 3-interactions. *Acta Math. Hungar.*, **143** (1), p.23-43.

Barabási, A. L. and Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, p.509-512.

Barabási, A. L. (2016) Network Science. *Cambridge University Press*.

Cooper, C., and Frieze, A. (2003) A general model of web graphs. *Random Structures Algorithms*, **22**, p.311-335.

Fazekas, I. and Porvázsnyik, B. (2016) Scale-free property for degrees and weights in an n-interactions random graph model. *Journal of Mathematical Sciences*, **214** (1), p.69-82.

Fazekas, I., Perecsényi, A. and Porvázsnyik, B. (2017) Numerical analysis of a network evolution model. *IEEE Computer Society 2017*, p.171-174.

From Hermitian processes to noncolliding particles to noncommutative processes: A review

PLENARY
Harmony Inge
Wed
9:00 - 10:00

Victor Pérez-Abreu

Center for Research in Mathematics, CIMAT, Guanajuato, Mexico
email: pabreu@cimat.mx

The Dyson-Brownian motion describes the evolution of a system of noncolliding processes corresponding to the process of eigenvalues of an Hermitian Brownian motion. It was first studied by the nuclear physicist Freeman Dyson as a time-varying model for energy levels of heavy nucleus. Its empirical measure valued process converges to the so called free Brownian motion, which is a premier example of a noncommutative process. We will present an overview of recent models arising as extensions of these matricial, free and noncommutative processes.

A probabilistic model of wind farm power generation via Copulas and indexed semi-Markov models

IS 22
Renewal & semi-Markov...II.
Rm -1.62
Tue
15:00 - 15:30

Guglielmo D'Amico[*,‡], Giovanni Masala[†,§], Filippo Petroni[†,¶]

[*]Department of Pharmacy, University "G. d'Annunzio" of Chieti-Pescara, Italy
[†]Department of Business and Economics, University of Cagliari, Italy
email: [‡]g.damico@unich.it [§]gb.masala@unica.it [¶]fpetroni@unica.it

In this paper we consider the problem of modelling the wind power production of a wind farm composed of a given number of wind turbines. As it is well known, the comprehension of stochastic properties of the total produced energy cannot be obtained simply considering the produced power of a single turbine with the total number of turbines. The reasons are mainly due to the so-called shear effect and to the morphological nature of the territory where the wind farm has been installed. For these reasons, it is interesting and crucial the development of a complete model that is able to correctly reproduce and forecast the power production of the whole wind farm. To this end we describe the stochastic production of energy of each turbine using an indexed semi-Markov chains (ISMC). This choice is motivated by recent research articles where the authors have demonstrated that ISMC model is able to reproduce the statistical properties of power production of a single wind turbine.. The ISMC model provides a very general approach that encompasses both semi-Markov processes and Markov chain based models. The superiority of the ISMC model resides in its accurate probabilistic description of the wind power evolution which accounts for the serial dependence of the wind power time series by incorporating past events (times and sizes of past power production) through an index process that increases the memory of the process. The modelling of the whole wind farm is executed by introducing a dependence structure among the considered ISMC that describe the wind power of each single wind turbine. The dependence structure is described by introducing copula functions. The advanced model is a modification of a previous paper where weighted-indexed semi-Markov chain model with copulas was applied to the description of financial performance of a portfolio of financial assets varying on a high frequency scale. A real application of the proposed multivariate model is performed on real data of energy produced have been analysed and compare with synthetic data obtained by implementing Monte Carlo simulation and the multivariate ISMC model.

IS 10
Copula Functions
Rm 100/B
Wed
12:30 - 13:00

On Copula Models derived by Convolution of Random Variables

Violetta E. Piperigou
University of Patras, Rio, Greece
email: vpiperig@math.upatras.gr

In the literature, the idea of creating a pair of dependant variables as convolution of other initially independent variables has often been used in the construction of either continuous, see Balakrishnan and Lai [2009], or discrete, see Papageorgiou and Piperigou [1997] and the references therein, bivariate distributions. However, the copula models that are constucted by this method have not been considered. These copulas are different from the fold product of copulas discussed by Longla [2015], yet some classes of them are special cases of a C-Convolution copula, Cherubini et al [2016]. Here several such copula models are derived using the Farlie-Gumbel-Morgenstern (FGM) family and the Fréchet-Hoeffding bounds with uniform marginals. Their dependence structure and relevant measures are examined in association with the initial variables, the conditional distributions are derived and are used for random variate generation. These models are also compared in practice with the generalized FGM family models presented in Piperigou [2009].

References

Balakrishnan, N. and Lai, C.-D. (2009) *Continuous bivariate distributions.* Second edition. Springer, Dordrecht.

Cherubini, U., Gobbi, F. and Mulinacci, S. (2016) *Convolution copula econometrics.* SpringerBriefs in Statistics. Springer.

Longla, M. (2015) On mixtures of copulas and mixing coefficients. *J. Multivariate Anal.* **139**, 259–265.

Papageorgiou, H. and Piperigou, V.E. (1997) On bivariate 'Short' and related distributions. *In: Advances in the Theory and Practice of Statistics-A Volume in Honor of Samuel Kotz* (eds N.L. Johnson, and N. Balakrishnan), pp. 397-413, New York: John Wiley & Sons.

Piperigou, V.E. (2009) Discrete distributions in the extended FGM family: the p.g.f. approach. *J. Statist. Plann. Inference* **139**, 3891-3899.

IS 17
Stochastic Processes & Appl. I.
Rm -1.63
Tue
11:00 - 11:30

Integrated Gauss-Markov processes and fractional integrated processes with applications

Mario Abundo[*], Enrica Pirozzi[†,‡]
[*]Dipartimento di Matematica, Università Tor Vergata, Roma, Italy
[†]Dipartimento di Matematica e Applicazioni, Università Federico II, Napoli, Italy
email: [‡]enrica.pirozzi@unina.it

Gauss-Markov (GM) processes and their time integral (Nobile et al. [2008]; Abundo and Pirozzi [2017]) have important applications in a variety of fields: in computational neuroscience, in finance mathematics, in queueing theory and other applied sciences (see e.g. Abundo [2015]; Taillefumier and Magnasco [2014]; Touboul and Faugeras [2008]). In the context of models for neural activity, the study of the neuronal dynamics can be improved by introducing the so-called colored noise in the neuronal stochastic model, in place of the classical white noise (see e.g. Bazzani et al. [2003]; Pirozzi [2017]). This kind of approach relies on the study of the integral over time of a Ornstein-Uhlenbeck process, or more generally, of a GM process and its first passage time. In Pirozzi [2017], a fractional stochastic model was also considered in order to preserve the memory of the neuronal membrane evolution; in such a description, the parameter that affects the firing activity is the fractional order of the involved derivative. Varying this parameter, it allows to put in evidence a number

of qualitative behaviors, and to compare them with those coming from real observations. Motivated by these considerations, we aim to study as replacing the ordinary Riemann integral with the fractional Riemann-Liouville (RL) integral of order $\alpha \in (0,1)$, affects the behavior, when varying α, of an integrated GM process, i.e. the process given by the fractional RL integral over time of a GM process. We provide some theoretical results and some comparisons by means of suitable simulations. We also discuss some specific applications in the context of neuronal modeling.

Acknowledgement. This work was partially supported by Gruppo Nazionale per il Calcolo Scientifico (GNCS-INdAM) .

References

Abundo, M., (2015) On the first-passage time of an integrated Gauss-Markov process. *Scientiae Mathematicae Japonicae Online* e-2015, **28**, 1–14.

Abundo,M., Pirozzi, E., (2017) Integrated stationary Ornstein–Uhlenbeck Process, and double integral processes. *Physica A*, doi: 10.1016/j.physa.2017.12.043

Bazzani, A., Bassi, G., Turchetti, G., (2003) Diffusion and memory effects for stochastic processes and fractional Langevin equations. *Physica A* **324**, 530–550.

Nobile, A.G., Pirozzi, E., Ricciardi, L.M., (2008) Asymptotics and evaluations of FPT densities through varying boundaries for Gauss-Markov processes. *Scientiae Mathematicae Japonicae* **67**, (2), 241–266.

Pirozzi, E., (2017) Colored noise and a stochastic fractional model for correlated inputs and adaptation in neuronal firing. *Biol Cybern* 1–15, doi: 10.1007/s00422-017-0731-0

Taillefumier T. and Magnasco M.O., (2014) A transition to sharp timing in stochastic leaky integrate-and-fire neurons driven by frozen noisy input, *Neural Computation*, **26** , 819 doi: 10.1162/NECO_a_00577

Touboul J. and Faugeras O. (2008) Characterization of the first hitting time of a double integral processes to curved boundaries. *Adv. Appl. Probab.*, **40**, pp. 501–528

A compound Poisson approximation for local sequence alignment

IS 23
Extremes and Bursts
Rm -1.62
Thu
11:00 - 11:30

Hrvoje Planinić*, Bojan Basrak

Department of Mathematics, Faculty of Science, University of Zagreb
email: *planinic@math.hr

We study ungapped local alignments of two independent i.i.d. sequences of letters from a finite alphabet. The motivation comes from measuring similarity between different biological sequences. The scores of all local alignments form a so–called score matrix in which, due to an interesting dependence structure, extreme scores appear in clusters along the diagonal lines. It is known that the number of clusters with extreme scores is approximately Poisson distributed. In particular, this determines the asymptotic distribution of the maximum local alignment score, see Dembo et al. [1994] and also Hansen [2006].

We show that it is possible to obtain a Poisson cluster limit for the point process of all local alignment scores together with their locations in the score matrix. Our approach is based on the theory of regularly varying random fields on the two-dimensional integer lattice and a point process version of the Poisson approximation result from Arratia et al. [1989]. This work is an extension of the previous joint work with Philippe Soulier (Université Paris Nanterre), see Basrak et al. [2018+].

References

R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: The Chen-Stein method. *Ann. Probab.*, **17**(1):9–25, 1989. doi: 10.1214/aop/1176991491.

Bojan Basrak, Hrvoje Planinić, and Philippe Soulier. An invariance principle for sums and record times of regularly varying stationary sequences. *Probab. Theory Relat. Fields*, Forthcoming 2018. doi: 10.1007/s00440-017-0822-9.

Amir Dembo, Samuel Karlin, and Ofer Zeitouni. Limit distribution of maximal non-aligned two-sequence segmental score. *Ann. Probab.*, **22**(4):2022–2039, 1994. doi: 10.1214/aop/1176988493.

Niels Richard Hansen. Local alignment of Markov chains. *Ann. Appl. Probab.*, **16**(3):1262–1296, 2006. doi: 10.1214/105051606000000321.

IS 16
Fractional Stochastic Models
Rm 100/B
Wed
15:30 - 16:00

Fractionality in discrete time: an application to Preferential attachment models

Federico Polito

Department of Mathematics "G. Peano", University of Torino, Italy
email: federico.polito@unito.it

In this talk we present a theory for fractional processes in discrete time. An example, namely a time-fractional preferential attachment model, is then described highlighting the differences with the classical non-fractional model.

Fractionality in continuous time is usually achieved by suitable time-changes and it is seen as a tool to extend Markov processes to models in which the presence of memory is taken into consideration. Interestingly enough these models are non-Markov, still they represent a class of processes retaining a certain mathematical treatibility. Even though the literature about continuous-time fractional processes is vast and growing, only few studies on their discrete-time counterparts have appeared so far.

We give a contribution in this sense outlining a possible theory based on discrete-time-changes and discrete infinite divisibility of random variables.

The talk collects joint works with Angelica Pachon and Costantino Ricciuti.

Acknowledgement. F. Polito has been partially supported by the projects *Memory in Evolving Graphs* (Compagnia di San Paolo/Università di Torino) and *Sviluppo e analisi di processi Markoviani e non Markoviani con applicazioni* (Università di Torino).

IS 31
Actuarial Risk Models II
Rm 0.99
Tue
15:00 - 15:30

Testing mixture component distribution

Denys Pommeret

Institut de Mathématique de Marseille, CNRS, Ecole Centrale de Marseille, Campus de Luminy, 13288 Marseille Cedex 9, France
email: denys.pommeret@univ-amu.fr

In this work we investigate a semiparametric testing approach to answer if the probability density function of an unknown component of a mixture model belongs to a given family. Based on a semiparametric estimation of the Euclidean parameters of the model, our method compares pairwise the nonparametric Fourier coefficients of the model estimated directly from the data with the ones obtained by plugging the estimated parameters. These comparisons are incorporated into a sum of square type statistic which order is controlled by a penalization rule. We prove under mild conditions that our test statistic is asymptotically $\chi^2(1)$-distributed and study its behavior under different types of alternatives, including contiguous nonparametric alternatives. Several level and power studies are numerically conducted. Finally we give various applications in Finance and Insurance models.

Open problems in contract theory

IS 7
Optimisatio & Machine Learning
Harmony lnge
Wed
10:30 - 11:00

Dylan Possamaï

Columbia University, New York
email: dp2917@columbia.edu

This talk will be the occasion to present recent progresses made on the treatment of continuous time contract theory, and more importantly to highlight several problems and areas where a general theory is still lacking. This includes notable adverse selection problems, time-inconsistent Agents, and general equilibrium issues between many Principals and many Agents. I will put forward the main difficulties encountered as well as possible approaches to tackle them.

Maxima of Two Random Walks: Universal Statistics of Lead Changes

CS 4
Diffusions and Random Walks
Rm -1.62
Thu
12:20 - 12:40

Eli Ben-Naim[*], Paul Krapivsky[†,‡], Julien Randon-Furling[§,¶]

[*]Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico, USA
[†]Department of Physics, Boston University, Boston, Massachusetts, USA
[‡]Institut de Physique Théorique, Université Paris-Saclay and CNRS, Gif-sur-Yvette, France
[§]SAMM, Université Paris-1 Panthéon-Sorbonne, Paris, France
email: [¶]Julien.Randon-Furling@univ-paris1.fr

We investigate the statistics of lead changes of the maxima of two one-dimensional random walks. We show that the average number of lead changes grows as $\pi^{-1} \ln t$ in the long time limit. We present theoretical and numerical evidence of the universality of this asymptotic behavior, namely the independence on the jump distribution (as long as it is symmetric); e.g., the same asymptotic underlies the standard Brownian motion and symmetric Lévy flights with index $\mu \in (0, 2)$. We also show that the probability to have at most n lead changes behaves as $t^{-1/4}(\ln t)^n$ for Brownian motion and generally as $t^{-\beta(\mu)}(\ln t)^n$ for symmetric Lévy flights. The decay exponent $\beta \equiv \beta(\mu)$ varies continuously with the Lévy index μ for $0 < \mu < 2$.

Acknowledgement. Two of us (PLK and JRF) thank the Galileo Galilei Institute for Theoretical Physics for excellent working conditions during the program on "Statistical Mechanics, Integrability and Combinatorics" and the INFN for partial support. The work of EBN was supported through US-DOE grant DE-AC52-06NA25396.

References

Ben-Naim, E., Krapivsky, P.L. and Randon-Furling, J. (2016) Maxima of two random walks: universal statistics of lead changes. *J. Phys. A: Math. and Theor.*, **49** (20), 205003

CS 11
Lévy processes
Rm -1.62
Wed
12:20 - 12:40

Facets on the convex hull of d-dimensional Brownian and Lévy motion

Julien Randon-Furling[*,‡], Florian Wespi[†]

[*]SAMM, Université Paris-1 Panthéon-Sorbonne, Paris, France
[†]IMSV, Universität Bern, Bern, Switzerland
email: [‡]Julien.Randon-Furling@univ-paris1.fr

For stationary, homogeneous Markov processes (viz., Lévy processes, including Brownian motion) in dimension $d \geq 3$, we establish an exact formula for the average number of $(d-1)$-dimensional facets that can be defined by d points on the process's path. This formula defines a universality class in that it is independent of the increments' distribution, and it admits a closed form when $d = 3$, a case which is of particular interest for applications in biophysics, chemistry and polymer science.

We also show that the asymptotical average number of facets behaves as $\langle \mathcal{F}_T^{(d)} \rangle \sim 2 \left[\ln (T/\Delta t)\right]^{d-1}$, where T is the total duration of the motion and Δt is the minimum time lapse separating points that define a facet.

References

Randon-Furling, J. and Wespi, F. (2017) Facets on the convex hull of d-dimensional Brownian and Lévy motion. *Physical Review E*, **95** (3), 032129.

OCS 2
Non-Markov Models in Finance
Harmony Inge
Tue
12:00 - 12:20

Utility maximization without passing by the dual problem

Miklós Rásonyi

MTA Alfréd Rényi Institute of Mathematics, Budapest
email: rasonyi@renyi.mta.hu

We treat utility maximization from terminal wealth for an agent with utility function defined on the whole real line who dynamically invests in a continuous-time financial market and receives a possibly unbounded random endowment. We prove the existence of an optimal investment in a rather general framework without introducing the associated dual problem. Our results apply to non-smooth utilities and even strict concavity can be relaxed. We can handle certain random endowments with non-hedgeable risks, complementing earlier papers. Constraints on the terminal wealth can also be incorporated. As examples, we treat frictionless markets with finitely many assets and large financial markets.

Acknowledgement. The support received from the NKFIH (National Research, Development and Innovation Office, Hungary) grant KH 126505 and from the "Lendület" grant LP 2015-6 of the Hungarian Academy of Sciences is gratefully acknowledged.

Estimating the input of a Lévy driven queue by Poisson sampling of the workload process

CS 6
Queuing Theory
Rm 100/A
Tue
17:10-17:30

Liron Ravner[*,†,‡], Onno Boxma[†], Michel Mandjes[*]

[*]University of Amsterdam, The Netherlands
[†]Eindhoven University of Technology, The Netherlands
email: [‡]l.ravner@uva.nl

Inferring the statistical properties of queueing systems by observing the workload is a challenging task due to the dependence between transient samples. In particular, likelihood functions are typically intractable. We consider the problem of estimating the exponent function of the Laplace-Stieltjes Transform (LST) of the input process to a Lévy driven queue. The workload of the queue is observed at random times according to an independent Poisson process. We suggest a non-parametric estimation method that relies on the empirical Laplace transform of the workload. This is achieved by using a generalized method of moments approach on the conditional LST of the workload sampled after an exponential time. The consistency of the method requires an intermediate step of estimating a constant that is related to both the input distribution and the sampling rate. To this end, for the case of an M/G/1 queue, we construct a partial maximum likelihood estimator and show that it is consistent and asymptotically normal. For spectrally positive Lévy input we construct a biased estimator for the intermediate step by considering only high workload observations above some threshold. A bound on the bias is provided and we discuss the tradeoff between the bias and variance of the estimator with respect to the chosen threshold.

Asymptotic behavior of multivariate limit book prices under compound Hawkes processes

IS 8
Limit Order Books
Harmony Inge
Mon
12:00 - 12:30

Bruno Rémillard

HEC Montréal
email: bruno.remillard@hec.ca

We propose a simple stochastic model for the dynamics of a limit order book prices for several stocks when the number of price changes is modeled by a compound multivariate Hawkes process. The asymptotic behavior is shown to be a multivariate Brownian motion. Examples of application are given.

Application of time-dependent branching processes to an evolving random graph model

OCS 1
Large Random Graphs
Rm -1.64
Mon
16:30 - 16:50

Tamás F. Móri[*], Sándor Rokob[†]

Eötvös Loránd University, Budapest, Hungary
email: [*]mori@math.elte.hu [†]rodnasbokor@gmail.com

The theory of time-dependent branching processes (Crump-Mode-Jagers processes) is a popular tool used in the (mathematical) analysis of random trees. However, one can use the method of branching in such a way that the edges and not the vertices form the individuals of the population. In this scenario the construction gives back a random graph, which is mostly not a tree. In some cases the

resulting dynamics can be interpreted as the generalization of well-known random graph dynamics (e.g.: preferential attachment).

In our first approach, we dealt with an increasing group of individuals who are organized by pairwise collaborations, where a successful collaboration attracts newcomers, who start collaborating with one or both participants. Furthermore, the new connections can weaken and exhaust the attracting pair's collaboration, which eventually ceases. However, some properties can be stated in more general settings. Here we investigate the corresponding random graph processes in the framework mentioned above.

Acknowledgement. T. Móri was supported by the Hungarian National Research, Development and Innovation Office NKFIH, grant No.: K 125569. S. Rokob was supported by the ÚNKP-17-2 New National Excellence Program of the Ministry of Human Capacities.

References

Jagers, P. (1975) Branching Processes with Biological Applications. *Wiley Series in Probability and Mathematical Statistics*

Móri, T. F. and Rokob, S. (2017) A random graph model driven by time-dependent branching dynamics. *Annales Univ. Sci. Budapest., Sect. Comp.* **46**, 191-213.

Móri, T. F. and Rokob, S. (2018) Further properties of a random graph model driven by time-dependent branching dynamics. *Annales Univ. Sci. Budapest., Sect. Comp.*, submitted.

Nerman, O. (1981) On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeit.* **57**, 365-395.

IS 48
Stein's Method
Rm -1.64
Mon
11:00 - 11:30

Error bounds in local limit theorems using Stein's method

Adrian Röllin
National University of Singapore
email: adrian.roellin@nus.edu.sg

We provide a general result for bounding the difference between point probabilities of integer supported distributions and the translated Poisson distribution, a convenient alternative to the discretized normal. We illustrate our theorem in the context of the Hoeffding combinatorial central limit theorem with integer valued summands, of the number of isolated vertices in an Erdős–Rényi random graph, and of the Curie–Weiss model of magnetism, where we provide optimal or near optimal rates of convergence in the local limit metric. In the Hoeffding example, even the discrete normal approximation bounds seem to be new. The general result follows from Stein's method, and requires a new bound on the Stein solution for the Poisson distribution, which is of general interest.

This is joint work with A. B. Barbour and N. Ross; to appear in *Bernoulli*.

IS 10
Copula Functions
Rm 100/B
Wed
12:00 - 12:30

Measure-Invariance of Copula functions and statistical applications

Silvia Romagnoli
Department of Statistics, University of Bologna
email: silvia.romagnoli@unibo.it

Copulas which are invariant under margins' transforms induced by some change of measure, are investigated. It is emphasized that, this particular kind of transforms induced by some change of measure, largely used in pricing techniques, preserves the invariance of the aggregation operator and a sufficient condition to assure it is proved. The discussion is extended to the time-preserving

of measure-invariance; a characterization of its stability in time for multivariate stationary processes, based on the dynamic copula representation (see Cherubini et al., 2011), is provided. Finally a measure invariance-based statistical test for the absence of arbitrage opportunity assumption and its preservation in time is proposed and an empirical experiment based on quotes of $S\&P$ 500 futures and options traded on the Chicago Mercantile Exchange (CME), is discussed.

Market impact can only be power law and this implies diffusive prices with rough volatility

IS 5
Rough Volatility & Market Impact
Harmony Inge
Thu
10:30 - 11:00

Mathieu Rosenbaum

École Polytechnique

Market impact is the link between the volume of an order and the price move during and after the execution of this order. We show that under no-arbitrage, the market impact function can only be of power law-type. Furthermore we prove that this implies that the long term price is diffusive with rough volatility. Hence we simply explain the universal rough behavior of the volatility as a consequence of the no-arbitrage property. From a mathematical viewpoint, our study relies in particular on new results about hyper-rough stochastic Volterra equations. This is joint work with Paul Jusselin.

Random Knockout Tournaments

IS 55
Applied Probability
Rm 0.99
Wed
11:30 - 12:00

Sheldon M Ross

Department of Industrial and Systems Engineering, Viterbi School of Engineering, University of Southern California
email: smross@usc.edu

We consider random knockout tournaments with n players. We suppose that each player has a value, with player i having value v_i, $i = 1, \ldots, n$, that each game involves two players, and that a game played between i and j is won by i with probability $v_i/(v_i + v_j)$. The loser of a game leaves the tournament, whereas the winner moves on.

Our first model supposes that the number of games in each round is specified and that the games are played by randomly selected players from all remaining players. We determine bounds on layers' win probabilities and show that the probability that the player with the highest value, call it v^*, wins the tournament is at least $v^*/(v_1 + ... + v_n)$. We also consider a different type of random knockout tournament, which supposes that the players are randomly assigned to positions $1, \ldots, n$. The players in positions 1 and 2 play a game in the first round, and in each of the following rounds $r = 2, \ldots, n-1$, the winner of the previous round plays with the player in position $r + 1$. The winner of the game in round $n - 1$ is the winner of the tournament. Bounds on the probability that each player wins are given, and some conjectures are presented.

OCS 1
Large Random Graphs
Rm -1.64
Mon
16:50 - 17:10

Asymptotic degree distribution in preferential attachment graph models with multiple type edges

Bence Rozner

ELTE Eötvös Loránd University, Budapest
email: rbence@cs.elte.hu

Various types of random graphs with preferential attachment dynamics have been examined recently. The analysis of random graphs is motivated by large real networks, such as the internet and different kinds of biological and social networks. In some of the applications, the vertices and the edges may have different types. For example, in a social network, the vertices can be considered as males or females, and the edges can be considered as family or work relationships. This leads to different phenomena as if we assign weights to the vertices or to the edges.

In the current work, we extend the preferential attachment model by choosing the type of each edge randomly. We assume that there is a connection between the evolution of the structure of the graph and the types of the edges. For the sake of simplicity, we only consider the 2-type case, but the results and the proofs can be easily generalized for the n-type case. The types are represented by colours, thus every edge can be either blue or red. We prove the existence of an almost sure asymptotic degree distribution for a general family of coloured preferential attachment graphs. That is, we show that for every fixed k and l, the proportion of vertices with exactly k blue and l red edges tends to some random variable as the number of steps goes to infinity. We also provide recurrence equations for the asymptotic degree distribution. Finally, we show a possible way to generalize the scale-free property of random graphs for the multi-type case. This is a joint work with Ágnes Backhausz.

IS 2
Optimal Transport ...
Rm 100/B
Mon
16:00 - 16:30

Time consistency of the mean-risk problem

Birgit Rudloff[*,†], Gabriela Kováčová[*]

[*]Vienna University of Economics and Business, Institute for Statistics and Mathematics, Vienna, Austria
email: [†]birgit.rudloff@wu.ac.at

Multivariate risk measures appear naturally in markets with transaction costs or when measuring the systemic risk of a network of banks. Recent research suggests that time consistency of these multivariate risk measures can be interpreted as a set-valued Bellman principle. And a more general structure emerges that might also be important for other applications and is interesting in itself.

In this talk I will show that this set-valued Bellman principle holds also for the dynamic mean-risk portfolio optimization problem. In most of the literature, the Markowitz problem is scalarized and it is well known that this scalarized problem does not satisfy the (scalar) Bellman principle. However, when we do not scalarize the problem, but leave it in its original form as a vector optimization problem, the upper images, whose boundary is the efficient frontier, recurse backwards in time under very mild assumptions. I will present conditions under which this recursion can be exploited directly to compute a solution in the spirit of dynamic programming and will state some open problems and challenges for the general case.

Nested Kriging predictions for datasets with a large number of observations

IS 31
Actuarial Risk Models II
Rm 0.99
Tue
15:30 - 16:00

Didier Rullière[*,¶], Nicolas Durrande[†,‖], François Bachoc[‡,**], Clément Chevalier[§,††]

[*]ISFA, University Lyon 1, Lyon, France
[†]Prowler.io and Institut Fayol, Mines Saint-Étienne, Saint-Etienne, France
[‡]IMT, University of Toulouse, Toulouse, France
[§]Institute of Statistics, University of Neuchâtel, Neuchâtel, Switzerland
email: [¶]didier.rulliere@univ-lyon1.fr [‖]durrande@emse.fr
[**]francois.bachoc@math.univ-toulouse.fr [††]clement.chevalier@unine.ch

Actuarial studies often make use of expensive simulators, so that predicting the value of a simulator given some input values may ease some computations as Solvency II Solvency Capital Requirement and Nested Simulations problems. The presented work falls within the context of predicting the value of a real function at some input locations given observations of this function. The Kriging interpolation technique (or Gaussian process regression) is often considered to tackle such a problem, but the method suffers from its computational burden when the number of observation points is large. We introduce here nested Kriging predictors which are constructed by aggregating sub-models based on subsets of observation points. This approach is proven to have better theoretical properties than other aggregation methods that can be found in the literature. Contrarily to some other methods it can be shown that the proposed aggregation method is consistent.

Acknowledgement. Part of this research was conducted within the frame of the Chair in Applied Mathematics OQUAIDO, gathering partners in technological research (BRGM, CEA, IFPEN, IRSN, Safran, Storengy) and academia (Ecole Centrale de Lyon, Mines Saint-Etienne, University of Grenoble, University of Nice, University of Toulouse and CNRS) around advanced methods for Computer Experiments. The related paper has been finished during a stay of D. Rullière at Vietnam Institute for Advanced Study in Mathematics, the latter author thanks the VIASM institute and DAMI research chair (Data Analytics & Models for Insurance) for their support.

References

Rullière, D., Durrande, N., Bachoc, F. and Chevalier, C., (2017) Nested Kriging predictions for datasets with a large number of observations. *Statistics and Computing*, Online First, p.1-19. doi: 10.1007/s11222-017-9766-2.

Bachoc, F., Durrande, N., Rullière, D. and Chevalier, C. (2017) Some properties of nested Kriging predictors. arXiv: 1707.05708.

Subcritical branching walks with heavy tails

OCS 3
Branching Random Walks
Rm -1.62
Mon
17:40-18:00

Anastasiia Rytova

Lomonosov Moscow State University, Moscow, Russian Federation
email: rutovaa@gmail.com

Random walks are applicable in the studies of the processes with particle motion. For example, in [Descombes, X., Zhizhina, E. and Komech, S. [2015]] a random walk with heavy tails on the plane was used to model the growth of a nerve cell that has a mutation. We consider a more complex model of a continuous-time symmetric branching random walk (BRW) on a multidimensional lattice [see, e.g., Albeverio, S., Bogachev, L. V. and Yarovaya, E. B. [1998]]. It is assumed that the underlying random walk is homogeneous, irreducible and has heavy tails [see, e.g., Borovkov, A. and Borovkov,

K. [2008]], that implies the infinity of the variance of jumps [see Yarovaya, E. [2013]]. The random walks with finite variance of jumps are transient on lattices of dimension three and more, but in the case of heavy tails random walk may be transient even in dimension one [Yarovaya, E. [2013]]. As a consequence, the critical regime of the branching process is achieved at other values of the parameter that characterizes the branching intensity. The phase transitions also depends on the number of branching sources [Yarovaya, E. B. [2017]]. Here we focus on the asymptotic behavior of the moments of particle numbers for BRWs with heavy tails in the subcritical case.

Acknowledgement. This research is supported by the Russian Foundation for Basic Research, grant no. 17-01-00468.

References

Descombes, X., Zhizhina, E., Komech S. (2015) Modelling axon growing using CTRW. arXiv: 1512.02603.

Albeverio, S., Bogachev, L. V., Yarovaya, E. B. (1998) Asymptotics of branching symmetric random walk on the lattice with a single source. *C.R. Acad. Sci. Paris, Ser. I, Math.*, **326**, p.975-980.

Borovkov, A., Borovkov, K. (2008) Asymptotic Analysis of Random Walks. Heavy-Tailed Distributions. *Cambridge University Press*.

Yarovaya, E. (2013) Branching random walks with heavy tails. *Commun. Statist. Theory Methods.*, **42(16)**, p.3001-3010.

Yarovaya, E. B. (2017) Positive Discrete Spectrum of the Evolutionary Operator of Supercritical Branching Walks with Heavy Tails. *Methodology and Computing in Applied Probability, Springer*, **19(4)**, p.1151-1167, doi: 10.1007/s11009-016-9492-9.

POSTER
Mon
18:00-19:00

Parameter estimation of Heston model under the non-Gaussian regime

Dipanwita Saha

University of Houston, USA
email: dsaha@math.uh.edu

We investigate the behaviour of the Maximum Likelihood Estimates (MLEs) for the parameters of the Heston model in a non-Gaussian regime. We study a closed form formula for the maximum likelihood estimators by maximizing discretized log-likelihood of the N observation of true volatility available at times $T, 2T, .., NT$. In particular, we study the MLEs when the ratio of the parameters of the volatility equation of the Heston model, satisfying the classical Feller condition, falls under a certain critical value for fixed T and as $N \to \infty$. We demonstrate that the asymptotic behavior of parameters can be described as rational functions which involve variables following the stable distribution. We also develop a strategy to estimate parameters of the stable distribution and verify our approach using numerical simulations.

Theorem 1. *Consider the Heston volatility equation is given as given below*

$$dY_t = \kappa(\theta - Y_t)dt + \gamma \sqrt{Y_t} dB_t$$

It is parameterized by three parameters (κ, θ, γ). *If* $\zeta = \frac{\kappa\theta}{\gamma^2} \in (\frac{1}{2}, 1)$ *then there exists an explicit positive exponent* $q = q(\zeta)$ *such that, as* $N \to \infty$, *the parameters* $N^q(\kappa_N - \kappa_\infty)$, $N^q(\theta_N - \theta_\infty)$ *and* $N^q(\gamma_N - \gamma_\infty)$ *converges to heavy tailed stable distribution.*

Note, $(\kappa_N, \theta_N, \gamma_N)$ are the MLEs of (κ, θ, γ) obtained by maximizing the discrete likelihood function of N observation and $(\kappa_\infty, \theta_\infty, \gamma_\infty)$ are their corresponding limits as $N \to \infty$.

Acknowledgement. I would like to thank my advisors Dr. Robert Azencott and Dr. Ilya Timofeyev for accepting me as their student and guiding me in completing this paper and my PhD.

Optimal Risk Exchange of Correlated Risks

IS 30
Actuarial Risk Models I
Rm 0.99
Mon
16:00 - 16:30

Nabil Kazi-Tani[*], Yahia Salhi[†]

Univ Lyon, Université Lyon 1, ISFA, LSAF EA2429
email: [*]nabil.kazitani@gmail.com [†]yahia.salhi@univ-lyon1.fr

In this paper, we explore the optimal risk exchange between an insurer and a reinsurer exposed to two dependent risks. In the literature, various works have investigated the bilateral contract design problem to determine optimal reinsurance contracts for a given risk of the insurer. This problem is first formally analyzed by Borch (1960), who showed that if the stop-loss reinsurance treaty is the optimal strategy, see also Arrow (1963). In this work, we derive the optimal exchange contract for different assumptions on the agents measures of risk (entropic, average value-at-risk …etc.). The closed form formulas of the optimal contracts extend the work of Jouini, Schachermayer and Touzi (2008). This is a joint work with Nabil Kazi-Tani.

Extensions of the generalized Pólya process

IS 24
Reliability Analysis
Rm -1.62
Wed
17:00 - 17:30

Francisco Germán Badía[*,‡], Sophie Mercier[†,§], Carmen Sangüesa[*,¶]

[*]Department of Statistical Methods, University of Zaragoza, Zaragoza, Spain
[†]Univ Pau & Pays Adour/CNRS, IPRA-LMAP, Pau, France
email: [‡]gbadia@unizar.es [§]sophie.mercier@univ-pau.fr [¶]csangues@unizar.es

In this work we consider an extension of the generalized Pólya process, introduced in Cha [Advances in Applied Probability, 46, 1148-1171 (2014)]. This extension is appropriate for describing successive failures of a system subject to imperfect repairs. We study several probabilistic properties of this model (construction through a homogeneous pure-birth process, convergence or restarting properties, among others). We also find a connection between this model and Generalized Order Statistics, which are an extension of usual order statistics. We give several dependence properties both for the arrival points and increments of this counting process, and provide stochastic comparison results between the arrival points of two different processes with distinct parameters. We finally consider the classical maximum likelihood method in a parametric setting, which is tested on a few simulated data sets.

Variance Allocation and Shapley Value

IS 56
Stochastic Game Theory
Rm -1.64
Tue
12:00 - 12:30

Marco Scarsini

LUISS, Roma, Italy
email: marco.scarsini@luiss.it

Motivated by the problem of utility allocation in a portfolio under a Markowitz mean-variance choice paradigm, we propose an allocation criterion for the variance of the sum of n possibly dependent random variables. This criterion, the Shapley value, requires to translate the problem into a cooperative game. The Shapley value has nice properties, but, in general, is computationally demanding. The main result of this paper shows that in our particular case the Shapley value has a very simple form that can be easily computed. The same criterion is used also to allocate the standard

deviation of the sum of n random variables and a conjecture about the relation of the values in the two games is formulated.

IS 35
Environmental Models
Rm -1.63
Thu
10:30 - 11:00

Incorporating dependence structures when postprocessing ensemble weather forecasts

Roman Schefzik
German Cancer Research Center (DKFZ), Heidelberg, Germany
email: r.schefzik@dkfz.de

Nowadays, weather forecasts usually rest upon ensemble prediction systems, which consist of multiple runs of numerical weather prediction models varying with respect to the initial conditions and/or the parameterization of the atmosphere. Ensemble forecasts are typically biased and exhibit dispersion errors. Therefore, they need to be statistically postprocessed. Many established postprocessing approaches, such as the standard Bayesian model averaging (BMA) and the ensemble model output statistics (EMOS) methods, are univariate and apply to a single weather quantity at a single location and for a single prediction horizon only, failing to account for potentially crucial dependence structures. To address this drawback, several ensemble postprocessing methods that account for inter-variable, spatial, and/or temporal dependence structures have been designed in the last years and are reviewed and compared in this talk.

One strategy is to develop postprocessing models that yield truly multivariate predictive distributions. The mathematical background for many such multivariate approaches is provided by copulas and Sklar's theorem. There are essentially two classes of multivariate ensemble postprocessing methods. Parametric approaches, including Gaussian copula-based techniques, are typically tailored to specific inter-variable, spatial or temporal settings and perform well in low-dimensional scenarios. In contrast, non-parametric, empirical copula-based approaches, such as ensemble copula coupling and Schaake shuffle-based methods, are more general and can handle virtually any dimensionality. In such techniques, univariate postprocessed ensemble forecasts, for instance obtained using standard BMA or EMOS, are arranged according to the rank order structure of a specifically chosen dependence template. Finally, there also exist postprocessing methods that formally yield univariate predictive distributions, but account for dependencies through the design of the estimation procedure for the model parameters.

CS 1
Biological Statistics
Rm 100/A
Wed
12:40 - 13:00

Identifying differential distributions for single-cell RNA sequencing data comprising biological replicates

Roman Schefzik
German Cancer Research Center (DKFZ), Heidelberg, Germany
email: r.schefzik@dkfz.de

Major advances in technology in the current decade allow to sequence information from individual biological cells and thus offer a high resolution of cellular differences. Single cell sequencing facilitates fundamental insights into biology and was chosen as the method of the year 2013 by Nature Publishing Group. In particular, performing high-throughput ribonucleic acid (RNA) sequencing at the single-cell level (scRNA-seq) has enabled the quantification of cellular heterogeneity, and the design of appropriate statistical analysis methods for scRNA-seq data currently is of great interest and importance.

A typical task in the context of scRNA-seq data is to detect differences in gene expression distributions across conditions. For instance, a recently developed approach allows for a classification of genes with a differential distribution (DD) into categories that represent distinct DD patterns such as differential mean, differential modality or differential proportion of cells within each component.

Our aim is to develop a statistical method that identifies DDs in the scenario in which the scRNA-seq data comprises different biological replicates for each physiological condition. In such a case, the gene expression of each individual biological replicate for each condition separately is first represented using a discrete-continuous model including a point mass for zero expression and a logspline density estimate for non-zero (positive) expression that are fitted to the respective scRNA-seq data. Then, for each condition separately, the so obtained distributions for the corresponding multiple replicates are aggregated using a mixture distribution based on linear pooling. Finally, these condition-specific distributions are checked for significant differences applying the L^2 Wasserstein distance and a permutation test. Our procedure is illustrated and evaluated using data from a recent scRNA-seq experiment.

Optimal Alpha Spending for Sequential Analysis with Binomial Data and Variable Bernoulli Probabilities

IS 38
Appiled Prob. & Stat. Inference III.
Rm 100/A
Tue
14:30 - 15:00

Ivair R. Silva[*,§], Martin Kulldorff[†], Judith Maro[‡], Laura Hou[‡]

[*]Federal University of Ouro Preto, Ouro Preto, Brazil
[†]Harvard Medical School and Brigham and Women's Hospital, Boston, USA
[‡]Harvard Medical School and Harvard Pilgrim Health Care Institute.
email: [§]ivairest@gmail.com

Usage of alpha spending function is a comprehensive approach for sequential analysis hypothesis testing. Although many authors have tried to derive the optimal alpha spending shape for some classical performance measures, such as power and expected sample size, most of the solutions are based on near-optimal results or on normal distribution approximations. The present work introduces an analytical mechanism for finding optimal sequential analysis designs for binomial data type under time-variable Bernoulli probabilities. This is done for pre-specified overall alpha level and power. The solution works for a large class of objective functions, such as: expected loss functions, maximum sample size, expected sample size, and expected time to signal. Also, multiple constraints can be included in the optimization problem, which favors to find uniformly better designs.

Acknowledgement. This research was funded by the National Institute of General Medical Sciences, USA, grant #RO1GM108999. The first author has received additional support from Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Minas Gerais, Brasil (FAPEMIG).

On the Correspondence Between Frequentist and Bayesian Tests

IS 44
Inference for Large Datasets
Rm 0.99
Wed
16:30 - 17:00

Ivair R. Silva

Federal University of Ouro Preto, Ouro Preto, Brazil.
Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, USA.
email: ivairest@gmail.com

Hypothesis testing plays a fundamental role in the modern practice of statistical analysis. In statistical inference, a hypothesis is a statement about a population parameter $\theta \in \mathbb{R}$. The objective of a statistical hypothesis test is to decide which of two hypotheses is true. The general format of the two

hypotheses is $H_0 : \theta \in \Theta_0$ and $H_1 : \theta \in \Theta_1$, where Θ_i, $i = 0, 1$, is a subset of the parameter space, Θ, where $\Theta_0 \cap \Theta_1 = \emptyset$. H_0 and H_1 are called the null and the alternative hypothesis, respectively. The decision of taking H_i as true is based on a random sample, $X_1, \cdots, X_n$, from the population $F_X(x|\theta)$.

In practice, an arbitrary real-valued function of the sample, a measure of evidence in favor of H_i, $i = 0, 1$, plays the role of test statistic and hence is used to define the decision rule. Small observed values of the measure of evidence in favor of H_i suggest rejection of H_i. The reasoning used to construct a measure of evidence is the main point of divergence between the Bayesian and the frequentist approaches.

In the frequentist approach, the well-known measure of evidence is the p-value, and this is so because the focus under the frequentist point of view is on controlling the Type I error probability under a alpha level. In contrast, for the Bayesian approach the Type I error probability is not considered a critical performance measure. But the Bayesian construction is seriously concerned on how to use the empirical information in order to update the analyst's prior uncertainty about the plausibility of each hypothesis. This is possible through the so called posterior distribution about the veracity of each hypothesis, which in turn is used to construct a Bayesian measure of evidence against H_0. Two common Bayesian measures of evidence are the posterior expected loss and the Bayes Factor.

Currently, the efforts to accommodate both approaches under the same decision rule are not applicable for the general case of any hypothesis testing problem. This presentation aims to explore the main results of the recently published work of Silva (2017). In that paper, Silva offers a theoretical framework for unification of frequentist and Bayesian approaches by the same decision rule.

Acknowledgement. The author is very grateful to Martin Kulldorff for important advices on statistical fundamentals, to Thiago Morais Pinto for the rich suggestions on writing style.

References

Silva, I.R. (2017) On the Correspondence Between Frequentist and Bayesian Tests. *Communications in Statistics - Theory and Methods*, doi: 10.1080/03610926.2017.1359296

CS 15
State Space and Markov Models
Rm 0.99
Thu
11:10 - 11:30

Testing Stochastic Trend in Space State Models for the Location-scale Family

I. R. Silva[*], D. Ernesto[†], F. L. P. Oliveira[*,‡]

[*]Universidade Federal de Ouro Preto, (Ouro Preto), Brasil
[†]Universidade Federal de Viçosa, (Viçosa), Brasil
email: [‡]fernandoluizest@gmail.com

Space state models are commonly used to describe the statistical behavior of time series. For non-stationary processes, a key point is the decision between modeling the trend through a deterministic or a stochastic term. The present work introduces a hypothesis testing procedure to guide in such a decision. The method works for any time series distribution belonging to the location-scale family, such as the Gaussian and the Logistic distributions.

Acknowledgement. This work was funded by: Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq $(300825/2016 - 1)$, Fundação de Amparo a Pesquisa do Estado de Minas Gerais - FAPEMIG $(CEX - APQ - 02322 - 14 - 525111)$, Project Prevenção da Fadiga, and Instituto de Bolsas de Moçambique, IBE.

Semi-Markov Reward Algorithms and their Applications

IS 21
Renewal & semi-Markov...I.
Rm -1.62
Tue
11:00 - 11:30

Dmitrii Silvestrov[*,‡], Raimondo Manca[†]

[*]Department of Mathematics Stockholm University, Stockholm, Sweden
[†]Department of Methods and Models for Economics, Territory and Finance, University "La Sapienza" Rome, Italy
email: [‡]silvestrov@math.su.se

A survey of results on Markov and semi-Markov reward algorithms is presented. Effective doubly recurrent algorithms for computing moments of rewards accumulated on trajectories of semi-Markov processes up to some stopping times are described. These algorithms are based on methods of sequential phase space reduction, recurrent relations connecting moments of different orders for accumulated rewards and special time-space truncation procedures. Insurance, finance, credit rating dynamics and other applications are also discussed. The survey is based on results presented in papers [1]-[7] and some latest new results in the area.

References

[1] Stenberg, F., Manca, R., Silvestrov, D (2006) Semi-Markov reward models for disability insurance. *Theory Stoch. Process*, **12(28)**, no. 3-4, 239-254.
[2] Stenberg, F., Manca, R., Silvestrov, D (2007) An algorithmic approach to discrete time non-homogeneous backward semi-Markov reward process with an application to disability insurance. *Metodol. Comput. Appl. Probab.*, **9**, 497-519.
[3] Biffi, E., D'amigo, G., Di Biase, G., Janssen, J., Manca, R. , Silvestrov, D. (2008) Monte Carlo semi-Markov methods for credit risk migration and Basel II rules. I and 2. *J. Numer. Appl. Math.*, **1(96)**, Part 1: 28-58, Part 2: 59-86.
[4] Silvestrov, D., Silvestrova, E., Manca, R. (2008) Stochastically ordered models for credit rating dynamics. *J. Numer. Appl. Math.*, **1(96)**, 206-215.
[5] Silvestrov, D., Manca, R., Silvestrova, E. (2014) Computational algorithms for moments of accumulated Markov and semi-Markov rewards. *Comm. Statist. Theory, Methods*, **43**, no. 7, 1453-1469.
[6] Silvestrov, D., Manca, R. (2017) Reward algorithms for semi-Markov processes. *Metodol. Comput. Appl. Probab.*, **19**, no. 4, 1191-1209.
[7] Silvestrov, D., Manca, R. (2017) Recurrent algorithms for mixed power-exponential moments of hitting times for semi-Markov processes. 17th ASMDA Conference Proceedings, London, UK, CMSIM, 735-753.

Individual Ergodic Theorems for Perturbed Alternating Regenerative Processes

OCS 4
Markov & Semi-Markov Models
Rm -1.62
Tue
12:00 - 12:20

Dmitrii Silvestrov

Stockholm University, Stockholm, Sweden
email: silvestrov@math.su.se

Results of complete analysis and classification of ergodic theorems for perturbed alternating regenerative processes with semi-Markov modulation, based on quasi-ergodic theorems for perturbed regenerative processes, are presented. New short, long and super-long time individual ergodic theorems for regularly, singularly and super-singularly perturbed alternating regenerative processes are given. Application to queuing, reliability and other types of perturbed stochastic system and processes are also discussed.

IS 8
Limit Order Books
Harmony Inge
Mon
11:30 - 12:00

Decomposition of an L^2 space with nonlinear stochastic integrals

Clarence Simard
Université du Québec à Montréal
email: simard.clarence@uqam.ca

We characterize the solution to the optimization problem of minimizing the distance between a nonlinear stochastic integral and a square-integrable random variable respect to the L^2-norm. A by-product of this solution is a generalization of the Kunita-Watanabe decomposition with nonlinear stochastic integrals.

IS 26
Growth & Contagion Models
Rm 0.99
Mon
11:30 - 12:00

Cross-infection in epidemics spread by carriers

Matthieu Simon
University of Melbourne
email: matthieus@unimelb.edu.au

In this talk, we present a simple and flexible Markov model to describe some epidemics spreading within two linked populations, of susceptibles and carriers. The population of carriers represents the external source of infection. It may be constituted by individuals without disease symptoms but who are able to transmit the infection. It may also cover more general sources of infection such as pollution streams and animal vectors. The susceptibles usually constitute the population of main interest. Any infected susceptible is immediately detected and do not intervene in the contagion process, so the carriers are the only source of infection.

In real life, the spread of the disease among the susceptibles can also have some influence on the carrier process. Indeed, as more infections are detected, it may be expected that some procedures will be implemented to speed up their removal. On the other hand, for certain diseases, the infected cases transmit the infection to vectors (e.g. mosquitoes), which increases the population of carriers. For these reasons, we propose a model that accounts for a crossed influence between carriers and susceptibles.

Our main purpose is to determine the state of the process at the end of the epidemic, that is, when there are no more possibility of infection in the population. We model the epidemic by a block-structured transient Markov process and use standard matrix-analytic methods to determine the final distribution of the number of susceptibles and carriers, as well as the distribution of any integral path for the susceptible process. We also present an extended model that accounts for the recovery of infected individuals. Finally, we illustrate the results on some examples.

This is a joint work with Claude Lefèvre (Université libre de Bruxelles).

Entropy, Extropy, and Information, in the Courtroom

IS 55
Applied Probability
Rm 0.99
Wed
11:00 - 11:30

Nozer D. Singpurwalla

The City University of Hong Kong, Hong Kong
email: nsingpur@cityu.edu.hk

The information-theoretic ideas of Hartley, Shannon, and Rényi can be nicely exposited via the courtroom scenario of a witness being cross examined by a prosecutor for pining down the guilty suspect among n possible suspects. In this expository style talk we question several key features of Shannon's theory (indeed Barnard and Kolmogorov have preceded us) and then make the claim that Shannon's entropy *overestimates* both the "guessing" (or hacking) entropy, as well as the "partitioning" entropy. We close this talk by presenting Frank Ladd's notion of "extropy", as the dual of Shannon's entropy to conclude that a more complete picture of the uncertainty in a probability distribution is provided by a tandem use of both entropy and extropy.

Models and Analysis of Mental Health Data Using Monotone Single Index Model

IS 46
High-dim. Bayesian Inference
Rm 0.99
Wed
15:30 - 16:00

Debajyoti Sinha[*,†], Kumaresh Dhara[*]

[*]Florida State University, USA
email: [†]sinhad@stat.fsu.edu

Single index model is a practical tool for dimension reduction and prediction in presence of a complex relationship between the response and the predictors. We introduce monotone single index model using Bernstein polynomial basis for the link function to adevelop clinically useful models for various mental health studies. For the frequentist approach, we use an iterated profiling based algorithm to estimate the index vector and the unknown link function. For the Bayesian inference, we use a first order Taylor expansion of the link function to facilitate the Metropolis-Hastings step for sampling of the index vector because the conditional posterior distribution of the index vector does not have a closed form expression. This allows us to have an efficient and easily implementable Markov Chain Monte Carlo tool. We show the novelty of our methods in the simulation study and later apply the method to the study of Dysphoria among the adolescent girls.

Acknowledgement. Dr.Sinha's research was supported by NIH grant and Pfeiffer Foundation.

The Doob-Meyer decomposition in continuous-time as a limit of its discrete-time analogue

IS 1
Stochastic Methods in Finance
Harmony Inge
Tue
15:30 - 16:00

Pietro Siorpaes

Imperial College London, London, United Kingdom
email: p.siorpaes@imperial.ac.uk

As showed in Rao, M. [1969], one can give a constructive proof of the existence of the Doob-Meyer decomposition $S = M + A$ of a submartingale $(S_t)_{t\in[0,1]}$ of class D by taking limits of the Doob decomposition $(M^n_t + A^n_t)_{t\in D_n}$ of the sampled process $(S_t)_{t\in D_n}$, where $(D_n)_n$ are refining partitions of $[0,1]$. However, A^n in general only converge in the $\sigma(L^1, L^\infty)$-topology, and A is showed to be natural instead of predictable.

As showed in Jakubowski, A. [2005], there exist forward convex combinations B^n of the A^n (extended to $[0,1]$) such that $\limsup_n B^n_t = A_t$ a.s. for all t, and this implies that A is predictable. Here we show that in fact $B^n_t \to A_t$ a.s. (and in L^1) for all t. In particular, this allows us to show that an increasing process is predictable iff it is natural by passing to the limit the analogous (and trivial) discrete time statement. Also, it enables us to prove that a predictable stopping time can be approximated 'from below' by predictable stopping times which take finitely many values.

References

Jakubowski, A. (2005), An almost sure approximation for the predictable process in the Doob-Meyer decomposition theorem. *Sem. de Probab. XXXVIII*, p.158-164

Rao, M. (1969), On Decomposition Theorems of Meyer. *Math. Scandinavica*, **24**, p.66-78.

CS 14
Stopping and First Exit
Rm 0.87
Wed
12:00 - 12:20

Life time estimation of complicated machines: a first exit time theory approach

Christos H Skiadas

ManLab, Technical University of Crete, Chania, Greece
email: skiadas@cmsim.net

We assume that the functionality level of a complicated mechanical system follows a stochastic process during time. The end of the functionality of the system comes when the functionality function reaches a zero level. We develop a first exit time methodology to model the life time process of the complicated machineAfter solving several technical details including the Fokker-Planck equation for the appropriate boundary conditions we estimate the transition probability density function and then the first exit time probability density of the functionality of the system reaching a barrier during time. The formula we arrive is essential for complicated mechanical forms as for several machines. A simpler case has the form called as Inverse Gaussian and was proposed to express the probability density of a simple first exit time process hitting a linear barrier. We propose a more general form with the Inverse Gaussian as a special case. A geometric derivation of the general model resulting from the Inverse Gaussian is also presented. Applications to the functionality life time of cars are done.

IS 47
Entropy Estimates & Appl.
Rm -1.62
Mon
12:00 - 12:30

Limit theorems for random sum of random variables

Nikolai Slepov

Lomonosov Moscow State University, Moscow, Russia
email: naslepov@mail.ru

The Stein method is a powerful tool permitting to establish limit theorems with convergence rates for appropriately normalized sums of random variables. An important role here is played by the zero bias transformation (see, e.g., [1]) and its analogues. The interesting results were obtained for approximations of the mentioned sums distributions by exponential, geometrical and the Laplace laws (see, e.g., [2], [3], [4]).

We develop the Stein method to investigate the distributions of random sums of random variables. In particular an analogue of the famous Renyi theorem is proved for a sequence of m-dependent random variables. Approximations by the Laplace law are discussed as well. We also tackle some properties of the Renyi entropy estimates.

References

[1] Goldstein, L., Reinert, G. (1997) Stein's method and the zero bias transformation with application to simple random sampling. *The Annals of Applied Probability*, **7**, p. 935-952. doi: 10.1214/aoap/1043862419
[2] Pekoz, E., Rollin, A. (2011) New rates for exponential approximation and the theorems of Renyi and Yaglom. *Annals of Probability*, **39**, p. 587-608. doi: 10.1214/10-AOP559
[3] Pekoz, E., Rollin, A., Ross, N. (2013) Total variation error bounds for geometric approximation. *Bernoulli*, **19**, p. 610–632. doi: 10.3150/11-BEJ406
[4] Pike, J., Ren, H. (2014) Stein's method and the Laplace distribution. *ALEA*, **11** p. 571-587.

Coupled random self-similar sets for contractions type operators

CS 18
Theoretical Probability
Rm -1.63
Mon
17:40-18:00

A. Petrusel*, A. Soos

Babes-Bolyai University, Cluj, Romania
email: *asoos@math.ubbcluj.ro

The aim of this paper is to present random fixed set theorems, collage type and anti-collage type results for single-valued operators $T : X \times X \to X$ in the framework of a complete metric space X. Based on the coupled fixed point theory, existence of random fixed sets, collage type and anti-collage type results for iterated function systems are also presented. The results are related to random self-similar sets theory and the mathematics of fractals. Several examples of random coupled fractals illustrate our results.

References

Petruşel, A., Petruşel, G.: Nonlinear dynamics, fixed points and coupled fixed points in generalized gauge spaces withapplications to a system of integral equations, *D*iscrete Dynamics in Nature and Society, **2015**:1–10, Article ID 143510, 2015.
Petruşel, A.,Soós, A.: Coupled fractals in complete metric spaces, *N*onlinear Anal. Modelling and Control, 23(2018), No. 2, 141–158.

Stochastic precedence and usual stochastic ordering for vectors of dependent lifetimes: a comparison based on the multivariate conditional hazard rates

IS 27
Stochastic Comparisons
Rm -1.62
Wed
11:00 - 11:30

Fabio L. Spizzichino

University of Rome Sapienza, Department of Mathematics. Piazzale Aldo Moro, 5, I-00185, Rome, Italy
email: fabio.spizzichino@uniroma1.it

We consider n non-negative random variables $X_1, ..., X_n$ with an absolutely continuous joint distribution and set, for $i = 1, ..., n$,

$$\alpha_i = \mathbb{P}\left(\min_{1\le j\le n} X_j = X_i\right).$$

We say that $X_1, ..., X_n$ are ordered in a *stochastic-precedence* sense when

$$\alpha_1 \le \alpha_2 \le ... \le \alpha_n.$$

In the talk, several different properties of this notion will be pointed out. In particular, we will focus on some apparently paradoxical aspects related with *non-transitivity* behaviour of stochastic-precedence. Furthermore, we will establish a comparison with corresponding aspects of stochastic orderings in the *usual sense*. In this respect, we will start from observing that the condition $X_1 \leq_{st} X_2$ does not necessarily imply

$$\mathbb{P}(X_1 \leq X_2) \geq \frac{1}{2},$$

even if such an implication does obviously hold true when X_1, X_2 are stochastically independent.

The assumption of absolute continuity allows us to describe the joint probability distribution of $(X_1, ..., X_n)$ in terms of the multivariate conditional hazard rates. This description is equivalent, from a theoretic viewpoint, to the one based on the joint density function. However, from a practical point of view, it is generally not easy at all to pass from one to the other description.

As a main purpose of the talk, we aim to point out that the multivariate conditional hazard rates provide the appropriate tool for the purpose of analyzing the issues mentioned above and for understanding the related heuristic meaning. Our arguments may also suggest some more general meditation about the implications of joint absolute continuity, when dealing with random variables having the meaning of lifetimes of units that start work simultaneously.

Simple and helpful examples can be shown by considering the special case of *load-sharing* models.

This is joint work with Emilio De Santis and Yaakov Malinovsky.

IS 9
Hedging, Model Fitting, Estimation
Rm 100/B
Tue
10:30 - 11:00

Nonparametric Bayesian volatility estimation

Shota Gugushvili[*,¶], Frank van der Meulen[†,‖], Moritz Schauer[*,**], Peter Spreij[‡,§,††]

[*]Mathematical Institute, Leiden University
[†]Delft Institute of Applied Mathematics, Delft University of Technology
[‡]Korteweg-de Vries Institute for Mathematics, University of Amsterdam
[§]Institute for Mathematics, Astrophysics and Particle Physics, Radboud University
email: [¶]shota.gugushvili@math.leidenuniv.nl [‖]f.h.vandermeulen@tudelft.nl
[**]m.r.schauer@math.leidenuniv.nl [††]spreij@uva.nl

Given discrete time observations over a fixed time interval, we study a nonparametric Bayesian approach to estimation of the volatility coefficient of a stochastic differential equation. We postulate a histogram-type prior on the volatility with piecewise constant realisations on bins forming a partition of the time interval. The values on the bins are assigned an inverse Gamma Markov chain (IGMC) prior. Posterior inference is straightforward to implement via Gibbs sampling, as the full conditional distributions are available explicitly and turn out to be inverse Gamma. We also discuss in detail the hyperparameter selection for our method. Our nonparametric Bayesian approach leads to good practical results in representative simulation examples. Finally, we apply it on a classical data set in change-point analysis: weekly closings of the Dow-Jones industrial averages.

Change Detection and Inference for High-Dimensional Covariance Matrices

IS 36
Applied Prob. & Stat. Inference I.
Rm 100/A
Mon
11:00 - 11:30

Ansgar Steland*, Rainer von Sachs†

*RWTH Aachen University, Germany
†UC Louvain, Belgium

New results about inference and change point analysis of high dimensional vector time series are discussed. The results deal with change-point procedures that can be based on an increasing number of bilinear forms of the sample variance-covariance matrix as arising, for instance, when studying change-in-variance problems for projection statistics and shrinkage covariance matrix estimation.

Contrary to many known results, e.g. from random matrix theory, the results hold true without any constraint on the dimension, the sample size or their ratio, provided the weighting vectors are uniformly ℓ_1-bounded. Extensions to ℓ_2-bounded projections are also discussed. The large sample approximations are in terms of (strong resp. weak) approximations by Gaussian processes for partial sum and CUSUM type processes, which imply (functional) central limit theorems under certain conditions. It turns out that the approximations by Gaussian processes hold not only without any constraint on the dimension, the sample size or their ratios, but even without any such constraint with respect to the number of bilinear form. For the unknown variances and covariances of these bilinear forms nonparametric estimators are proposed and shown to be uniformly consistent.

We present related change-point procedures for the variance of projection statistics as naturally arising in principal component analyses, aggregated data from sensor arrays and dictionary learning, amongst others. Further, we discuss how the theoretical results lead to novel distributional approximations and sequential methods for shrinkage covariance matrix estimators in the spirit of Ledoit and Wolf.

Acknowledgement. Part of the work of the presenting author has been supported by a grant from Deutsche Forschungsgemeinschaft (DFG), grant STE 1034/11, which he gratefully acknowledges.

References

Ansgar Steland and Rainer von Sachs. Large-sample approximations for variance-covariance matrices of high-dimensional time series. *Bernoulli*, 23(4A):2299–2329, 2017. doi: 10.3150/16-BEJ811.

Ansgar Steland and Rainer von Sachs. Asymptotics for high-dimensional covariance matrices and quadratic forms with applications to the trace functional and shrinkage. *Stochastic Process. Appl.*, 2018, arXiv: 1711.01835. In press.

Stein's Method for Dynamical Systems

IS 48
Stein's Method
Rm -1.64
Mon
12:00 - 12:30

Mikko Stenlund

University of Helsinki, Finland
email: mikko.stenlund@helsinki.fi

We present a version of Stein's method of multivariate normal approximation that can be implemented in the setting of deterministic dynamical systems. This method skips auxiliary randomization and is direct in nature. We also discuss subsequent extensions to certain non-autonomous systems, which in particular model non-equilibrium physical processes. This is joint work with Olli Hella and Juho Leppänen.

Acknowledgement. The author is grateful to Emil Aaltosen Säätiö, and the Jane and Aatos Erkko Foundation for their generous support.

OCS 6
Rough Volatility
Harmony Inge
Thu
12:40 - 13:00

Pathwise large deviations for the rough Bergomi model

Henry Stone
Imperial College London
email: henrymstone@hotmail.co.uk

We prove a pathwise large deviations principle for a small-noise version of the rough Bergomi model, introduced by Bayer, Friz and Gatheral in 'Pricing under rough volatility'.

IS 23
Extremes and Bursts
Rm -1.62
Thu
11:30 - 12:00

Inference for Continuous Time Random Maxima with Heavy-Tailed Waiting Times

Peter Straka[*,§], Katharina Hees[†], Smarak Nayak[‡]
[*]UNSW Sydney, Australia
[†]TU Dortmund, Germany
[‡]National Australia Bank
email: [§]p.straka@unsw.edu.au

In many complex systems of interest, inter-arrival times between events such as earthquakes, trades and neuron voltages have a heavy-tailed distribution. The set of event times is fractal-like, being dense in some time windows and empty in others, a phenomenon dubbed "bursty" in the physics literature. Renewal processes with heavy-tailed waiting times reproduce this bursty behaviour. We develop an inference method for "Continuous Time Random Maxima" (also called "Max-renewal processes"), which assume i.i.d. magnitudes at the renewal events and model the largest cumulative magnitude. For high thresholds and infinite-mean waiting times, we show that the times between threshold crossings are Mittag-Leffler distributed, i.e. form a fractional Poisson Process. Exceedances of thresholds are known to be Generalized Pareto distributed, according to the Peaks Over Threshold approach. We model threshold crossing times and threshold exceedances jointly and provide graphical means of estimating model parameters. We show that these methods yield meaningful insights on real-world datasets.

Acknowledgement. P. Straka was supported by the Discovery Early Career Research Award DE160101147 on the Project "Predicting Extremes when Events Occur in Bursts" by the Australian Research Council. K. Hees was supported by the DAAD co-financed by the German Federal Ministry of Education and Research (BMBF)

References

Katharina Hees, Smarak Nayak, and Peter Straka. Inference for continuous time random maxima with heavy-tailed waiting times. arXiv: 1802.05218.

Pseudo Marginal Approach for Metropolis Hastings in Smooth Transition Auto-Regressive Models

CS 17
Bayesian Approaches and MCMC
Rm 0.87
Wed
16:30-16:50

Osnat Stramer*, Chuyi Wang

University of Iowa, Iowa, USA
email: *osnat-stramer@uiowa.edu

We study a relative new MCMC algorithm called pseudo marginal approach introduced in Andrieu and Roberts [2009] and compare it to other approaches such as Monte Carlo within Metropolis. We apply the different MCMC algorithms on Smooth Transition Auto-regressive models with missing values. In particular we examine the performance of pseudo marginal approach under different settings and measure the efficiency of pseudo marginal approach.

References

Christophe Andrieu and Gareth O. Roberts. The pseudo-marginal approach for efficient Monte Carlo computations. *Ann. Statist.*, 37(2):697–725, 2009. doi: 10.1214/07-AOS574.

Osnat Stramer and Matthew Bognar. Bayesian inference for irreducible diffusion processes using the pseudo-marginal approach. *Bayesian Anal.*, 6(2):231–258, 2011. doi: 10.1214/11-BA608.

Timo Teräsvirta. Specification, estimation, and evaluation of smooth transition autoregressive models. *J. Amer. Statist. Assoc.*, 89(425):208–218, 1994. doi: 10.1080/01621459.1994.10476462.

General Compound Hawkes Processes in Limit Order Books

IS 8
Limit Order Books
Harmony lnge
Mon
11:00 - 11:30

Anatoliy Swishchuk

University of Calgary
email: aswish@ucalgary.ca

In this talk, we present two new Hawkes processes, namely, so-called general compound and regime-switching general compound Hawkes processes, to model the price processes in the limit order books. We prove Law of Large Numbers (LLN) and Functional Central Limit Theorems (FCLT) for these processes. The latter two FCLTs are applied to limit order books where we use these asymptotic methods to study the link between price volatility and order flow in our two models by investigating the diffusion limits of these price processes. The volatilities of price changes are expressed in terms of parameters describing the arrival rates and price changes. A numerical example will be presented as well.

Statistical post-processing of dual resolution ensemble forecasts

POSTER
Mon
18:00-19:00

Sándor Baran*, Zied Ben-Bouallègue†, Martin Leutbecher†, Marianna Szabó*,‡

*University of Debrecen, Hungary †European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom
email: ‡szabo.marica@gmail.com

Recently, all major weather prediction centres provide forecast ensembles of different weather quantities, which are obtained from multiple runs of numerical weather prediction models with various initial conditions and model parametrizations. The European Centre for Medium-Range Weather

Forecasts (ECMWF) issues global ensemble forecasts since 1992. According to its strategic plans till 2025, ECMWF wants to improve the resolution of ensemble forecasts from 18 km to a 5 km grid which requires a substantial increase of computation resources. In connection with this strategy, experiments are made at the ECMWF on mixing high- and low resolution ensemble forecasts having a fixed total computation cost, in order to determine the optimal combination having the best predictive performance. In the present work we investigate the behavior of post-processed dual resolution global ensemble forecast for 2 m temperature, where for statistical calibration the ensemble model output statistics [EMOS; Gneiting *et al.* 2005] approach is applied. As high resolution we consider the 50-member operational TCo639 ensemble of the ECMWF (18km resolution) together with the 200-member TCo399 ensemble having a horizontal resolution of 29 km. Tests with local and semi-local [Lerch and Baran 2017] EMOS post-processing support the existence of a superior combination of high- and low resolution forecasts, however, statistical calibration reduces the differences in verification scores.

Acknowledgement. Sándor Baran and Marianna Szabó were supported by the EFOP-3.6.3-VEKOP-16-2017-00002 project. The project was co-financed by the Hungarian Government and the European Social Fund. Sándor Baran also acknowledges the support of the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the EFOP-3.6.1-16-2016-00022 project. The project is co-financed by the European Union and the European Social Fund.

References

Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea. Rev.* **133**, 1098–1118.

Lerch, S., Baran, S. (2017) Similarity-based semi-local estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Stat.* **66**, 29–51.

IS 11
Monte Carlo, Importance Sampling...
Rm 100/B
Tue
14:30 - 15:00

Optimal importance sampling for Lévy and affine processes

Aurélien Alfonsi[*], Adrien Génin[†], Zorana Grbac[‡], David Krief[‡], Peter Tankov[§,¶]

[*]Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée, France
[†]Société Générale, Paris, France
[‡]Paris-Diderot University, Paris, France
[§]ENSAE ParisTech, Palaiseau, France
email: [¶]peter.tankov@ensae.fr

We present an overview of several ongoing works devoted to importance sampling methods for Monte Carlo pricing of European and path-dependent options in models driven by Lévy and affine processes. These methods are based on the theory of large deviations and extend earlier works in the context of Gaussian vectors and the Black-Scholes model [Glasserman et al. 1999, Guasoni and Robertson 2008]. For Lévy processes, we use the pathwise large deviations principle of Léonard [Leonard 2000], and for affine processes we prove a new pathwise LDP which was absent in the literature. Using Varadhan's lemma, we then compute an explicit asymptotic approximation for the variance of the pay-off under an time-dependent Esscher-style change of measure. Minimizing this asymptotic variance using convex duality, we then obtain an importance sampling estimator of the option price. We show that our estimator is logarithmically optimal among all importance sampling estimators. Numerical tests in the variance gamma model, Heston model, and the Wishart multidimensional model show consistent variance reduction with a small computational overhead.

References

P. Glasserman, P. Heidelberger, and P. Shahabuddin, *Asymptotically optimal importance sampling and stratification for pricing path-dependent options*, Mathematical finance, 9 (1999), pp. 117–152.

P. Guasoni and S. Robertson, *Optimal importance sampling with explicit formulas in continuous time*, Finance and Stochastics, 12 (2008), pp. 1–19.

C. Léonard, *Large deviations for Poisson random measures and processes with independent increments*, Stochastic processes and their applications, 85 (2000), pp. 93–121.

Asymptotic Bayesian Theory of Quickest Change Detection for Hidden Markov Models

IS 41
Quickest Change Detection
Rm 100/A
Wed
11:00 - 11:30

Alexander G. Tartakovsky[*,†,§], Cheng-Der Fuh[‡]

[*]Moscow Institute of Physics and Technology, Moscow, Russia
[†]AGT StatConsult, Los Angeles, CA, USA
[‡]National Central University, Taoyuan, Taiwan;
email: [§]agt@phystech.edu

Sequential changepoint detection problems deal with detecting changes in a state of a random process via observations obtained sequentially. If the state is normal, one wants to continue obtaining observations. If the state changes and becomes abnormal, one wants to detect this change as rapidly as possible. In such a problem, there is always a tradeoff between reducing false alarms and quickening detection. The two goals must be balanced. A conventional criterion is to minimize the expected delay to detection while controlling a risk associated with false detections. An optimality criterion and a solution depend heavily on what is known about the models for the observations and for the change point. In this talk, we are interested in a Bayesian setup when one attempts to find a detection rule that minimizes the average delay to detection, or more generally, higher moments of the delay to detection, for a given weighted probability of false alarm, assuming that the prior distribution of the change point is given.

In the early 2000s, Tartakovsky and Veeravalli (2004) developed an asymptotic Bayesian theory of changepoint detection for general stochastic models assuming a certain stability of the log-likelihood ratio process, which was expressed in terms of r-quick convergence. Recently, Tartakovsky (2017) refined this theory relaxing r-quick convergence to r-complete convergence. While several examples related to Markov and hidden Markov models were considered, these are only very particular cases where the main condition on the r-quick convergence of the normalized log-likelihood ratio was verified. Moreover, even these particular examples show that verifying this condition typically represents a hard task. At the same time, there is a class of very important stochastic models – hidden Markov models – that find extraordinary applications in a wide variety of fields such as speech recognition; handwritten recognition; computational molecular biology and bioinformatics; human activity recognition; object detection and tracking; and modeling, rapid detection, and tracking of malicious activity of terrorist groups, to name a few. In this work, we investigate the performance of the Bayesian Shiryaev changepoint detection rule for hidden Markov models. We propose a set of regularity conditions under which the Shiryaev procedure is first-order asymptotically optimal in a Bayesian context, minimizing moments of the detection delay up to certain order asymptotically as the probability of false alarm goes to zero. The developed theory for hidden Markov models is based on Markov chain representation for the likelihood ratio by Fuh (2003) and r-quick convergence for Markov random walks. In addition, applying Markov nonlinear renewal theory, we present a high-order asymptotic approximation for the expected delay to detection and the probability of false alarm of the Shiryaev detection rule. We also study asymptotic properties of another popular change

detection rule, the Shiryaev–Roberts rule. The efficiency of both detection rules is evaluated not only asymptotically but also with Monte Carlo simulations.

Acknowledgement. The research of A.G. Tartakovsky in this direction was supported in part by the U.S. Air Force Office of Scientific Research under MURI grant FA9550-10-1-0569, by the U.S. Defense Advanced Research Projects Agency under grant W911NF-12-1-0034 and by the U.S. Army Research Office under grant W911NF-14-1-0246 during 2008–2013 at the University of Southern California and during 2013–2014 at UConn as well as by the RF Ministry of Education and Science through the Russian Academic Excellence Project 5-100 and the Arctic project at the Moscow Institute of Physics and Technology in 2017–2018. The research of Cheng-Der Fuh was supported in part by MOST 105-2410-H-008-025-MY2 and MOST 106-2118-M-008-002-MY2.

IS 7
Optimisatio & Machine Learning
Harmony lnge
Wed
12:00 - 12:30

Deep Hedging

Josef Teichmann[*,‡], Hans Buehler[†], Lukas Gonon[*], Ben Wood[†]

[*]Eidgenössische Technische Hochschule Zürich
[†]J.P. Morgan, London
email: [‡]josef.teichmann@math.ethz.ch

We present a framework for hedging a portfolio of derivatives in the presence of market frictions such as transaction costs, market impact, liquidity constraints or risk limits using modern deep reinforcement machine learning methods.

We discuss how standard reinforcement learning methods can be applied to non-linear reward structures, i.e. in our case convex risk measures. As a general contribution to the use of deep learning for stochastic processes, we also show that the set of constrained trading strategies used by our algorithm is large enough to ϵ-approximate any optimal solution.

Our algorithm can be implemented efficiently even in high-dimensional situations using modern machine learning tools. Its structure does not depend on specific market dynamics, and generalizes across hedging instruments including the use of liquid derivatives. Its computational performance is largely invariant in the size of the portfolio as it depends mainly on the number of hedging instruments available.

We illustrate our approach by showing the effect on hedging under transaction costs in a synthetic market driven by the Heston model, where we outperform the standard solution.

References

Hans Buehler, Lukas Gonon, Josef Teichmann, and Ben Wood. Deep hedging. 2018. doi: 10.2139/ssrn.3120710.

CS 16
Spatial structures
Rm -1.63
Wed
12:20 - 12:40

Estimating the covariance function of isotopic fields on the sphere

György Terdik

University of Debrecen, Faculty of Informatics, Hungary
email: Terdik.Gyorgy@inf.unideb.hu

Estimating the covariance function of a Gaussian isotropic random field $T(x)$ on the unit sphere $\mathbb{S}_2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ of the Euclidean space $\mathbb{R}^3$ has primary importance. We assume that an observation $\widehat{T}(x)$, $x \in \mathbb{S}_2$ is given and estimate the covariance function by

$$\widehat{\mathcal{C}}(\cos\gamma) = \int_{\mathbb{S}_2} \int_{C(x,\gamma)} \widehat{T}(x)\, \widehat{T}(x_\gamma(\varphi, x)) \frac{d\varphi}{2\pi} \frac{\Omega(dx)}{4\pi}. \tag{1}$$

where $\Omega(dx) = \sin\vartheta d\vartheta d\varphi$ is the Lebesgue element of the surface area on $\mathbb{S}_2$ and $C(x,\gamma)$ is the set of unit vectors $x_\gamma(\varphi, x)$ such that the central angle between x and $x_\gamma(\varphi, x)$ is γ. We show that $\widehat{\mathcal{C}}(\cos\gamma)$ follows a Rosenblatt type distribution. Following results of Veillette and Taqqu [2012] the asymptotic distribution of the truncated version of the estimator $\widehat{\mathcal{C}}(\cos\gamma)$ is also given. One can substitute the usual estimator of the spectrum, see Marinucci and Peccati [2011], by estimating the covariance function first then using the Gauss-Legendre quadrature for estimation of the spectrum, see Szapudi et al. [2001] as well. The problem of cosmic variance Weinberg [2008] is also considered in accordance with simulations and real data. In practice the observations are given on a high resolution discretized sphere, for instance the Cosmic Microwave Background anisotropies (CMB) data are given on a pixel structure called HEALPix Collaboration et al. [2015], hence the estimator (1) is approximated in the high accuracy.

This talk is based on the work Leonenko et al. [2017].

Acknowledgement. This work is supported by the EFOP-3.6.2-16-2017-00015 project. The project has been supported by the European Union, co-financed by the European Social Fund.

References

Planck Collaboration, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, N. Bartolo, E. Battaner, and et al. (214 more). Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *Astronomy & Astrophysics*, 2015.

N. N. Leonenko, M. S. Taqqu, and Gy. Terdik. Estimation of the covariance function of Gaussian isotropic random fields on spheres and related Rosenblatt distributions. Technical report, Draft, ArXiv e-prints, 2017.

D. Marinucci and G. Peccati. *Random Fields on the Sphere*, volume 389 of *London Mathematical Society, Lecture Notes Series*. Cambridge University Press, Cambridge, 2011.

I. Szapudi, S. Prunet, D. Pogosyan, A. S Szalay, and J R. Bond. Fast cosmic microwave background analyses via correlation functions. *The Astrophysical Journal Letters*, 548(2):L115, 2001.

M. S. Veillette and M. S. Taqqu. Berry–Esseen and Edgeworth approximations for the normalized tail of an infinite sum of independent weighted gamma random variables. *Stochastic Processes and their Applications*, 122(3):885–909, 2012.

S. Weinberg. *Cosmology*. Oxford University Press, 2008.

Border Aggregation Model

IS 50
Random Structures
Rm -1.64
Wed
17:00 - 17:30

Debleena Thacker[*,‡], Stanislav Volkov[†]

[*]Uppsala University, Uppsala
[†]Lund University, Lund
email: [‡]thackerdebleena@gmail.com

Start with a graph with a subset of vertices called *the border*. A particle released from the origin performs a random walk on the graph until it comes to the immediate neighbourhood of the border, at which point it joins this subset thus increasing the border by one point. Then a new particle is released from the origin and the process repeats until the origin becomes a part of the border itself. We are interested in the total number ξ of particles to be released by this final moment.

Interestingly, this model can be considered as a generalization of the OK Corral model as studied in Kingman and Volkov [2003]. We obtain distributions and bounds for ξ in cases where the graph is star graph, regular tree, and a $d-$dimensional lattice.

Levine and Peres [2007] observed that the border aggregation model on $d-$dimensional lattice can be considered as an "inversion" of the classical diffusion-limited-aggregation model (DLA). We strengthen bounds obtained in Kesten [1987] for DLA model to obtain a lower bound on ξ.

References

Harry Kesten. How long are the arms in DLA? *J. Phys. A*, 20(1):L29–L33, 1987. doi: 10.1088/0305-4470/20/1/007.

J. F. C. Kingman and S. E. Volkov. Solution to the OK Corral model via decoupling of Friedman's urn. *J. Theoret. Probab.*, 16(1):267–276, 2003. doi: 10.1023/A:1022294908268.

Lionel Levine and Yuval Peres. Internal erosion and the exponent $\frac{3}{4}$. 2007. http://www.math.cornell.edu/~levine/erosion.pdf.

CS 13
Filtering
Rm -1.63
Tue
16:50-17:10

A method for constrained Kalman filtering

Ourania Theodosiadou*, George Tsaklidis†

Department of Mathematics, Aristotle University, Thessaloniki, Greece.
email: *outheod@math.auth.gr †tsaklidi@math.auth.gr

A method for estimating the hidden states of a standard Kalman filter is provided when the states are subject to certain one side constraints. The equations of the prediction and updating stages are derived, and comparison with known constrained filter methodologies is given. Finally, an application concerning a financial time series is presented.

IS 33
Neuron Models, Neural Fields
Rm -1.63
Mon
11:00 - 11:30

Simulation of the spiking times of a neuron using Piecewise Deterministic Markov Processes.

Michèle Thieullen

Sorbonne Université - Campus Jussieu (UPMC), LPSM, 4 Place Jussieu, 75252 Paris Cedex 05, France
email: michele.thieullen@upmc.fr

The excitability of a neuron (or of an excitable cell) results from the gating of ion channels located in its membrane. I will explain why Piecewise Deterministic Markov Processes (PDMP) are good models to study channel noise. In particular, on PDMP stochastic Hodgkin-Huxley models, I will consider the issue of the exact simulation of the spiking times. Our results are general and apply to a large class of PDMP when one is interested in the exact simulation of the successive jump times. Part of this talk is based on joint work with my colleagues Vincent Lemaire and Nicolas Thomas at LPSM, Sorbonne Université-Campus Pierre et Marie Curie (UPMC).

IS 13
Lévy Processes & Appl.
Harmony Inge
Mon
15:30 - 16:00

The Gaussian distribution is freely selfdecomposable

Steen Thorbjørnsen

Department of Mathematics, University of Aarhus

The class of selfdecomposable distributions in free probability theory was introduced by Barndorff-Nielsen and the speaker. It constitutes a fairly large subclass of the freely infinitely divisible distributions, but so far specific examples have been limited to Wigner's semicircle distributions, the free stable distributions, two kinds of free gamma distributions and a few other examples.

In 2011 it was proved by S. Belinschi, M. Bozejko, F. Lehner and R. Speicher that the Gaussian distributions are freely infinitely divisible. In recent work T. Hasebe, N. Sakuma and the speaker established that the Gaussian distributions are in fact freely selfdecomposable. More generally it was established that the Askey-Wimp-Kerov distribution μ_c is freely selfdecomposable for any c in $[-1, 0]$.

The talk will provide a summary of the theory of free selfdecomposability and an outline of the proof of the result described above.

Fractional type equations, semi-Markov processes and anomalous diffusion/aggregation phenomena

IS 18
Stochastic Processes & Appl. II.
Rm -1.63
Tue
15:30 - 16:00

Bruno Toaldo

Dipartimento di Matematica e Applicazioni "Renato Caccioppoli", Università degli Studi di Napoli "Federico II"
email: bruno.toaldo@unina.it

A process $X(t)$, $t \geq 0$, is said to be a semi-Markov process in the sense of Gihman and Skorohod Gihman and Skorohod [1975] if the couple $(X(t), \gamma(t))$, where $\gamma(t)$ is the sojourn time in the current position, is a strong (homogeneous) Markov process. The Kolmogorov's equations of such processes are investigated. An evolutionary form (integro-differential) for such equations is derived by using the construction of a semi-Markov processes as time-changed Markov processes Meerschaert and Toaldo [2016], Orsingher et al., [2016, 2018], Toaldo [2015a,b]. The time-fractional equation is an interesting particular case. The long time behaviour of processes is investigated and related to anomalous diffusion/aggregation phenomena.

References

Gihman, I.I. and Skorohod A.V. 1975 *The theory of stochastic processes II*. Springer-Verlag (1975).

Meerschaert M.M. and Toaldo B. 2016 *Relaxation patterns and semi-Markov dynamics*. In revision (2018).

Orsingher E., Ricciuti C. and Toaldo B. *Time-inhompgeneous jump processes and variable order operators*. Potential Analysis, 45(3), 435 - 461 (2016)

Orsingher E., Ricciuti C. and Toaldo B. 2016 *On semi-Markov processes and their Kolmogorov's integro-differential equations*, Journal of Functional Analysis, in press (2018).

Toaldo B. *Convolution-type derivatives, hitting-times of subordinators and time-changed C_0-semigroups*. Potential Analysis, 42(1), 115 – 140 (2015)

Toaldo B. *Lévy mixing related to distributed order calculus, subordinators and slow diffusions*. Journal of Mathematical Analysis and Applications, 430(2), 1009 – 1036 (2015)

Routeing properties in a Gibbsian model for highly dense multihop networks

CS 17
Bayesian Approaches and MCMC
Rm 0.87
Wed
16:50-17:10

Wolfgang König[*], András Tóbiás[†,‡]

[*]WIAS Berlin and TU Berlin, Germany
[†]TU Berlin, Germany
email: [‡]tobias@math.tu-berlin.de

In my talk, I will present our recent paper König, W. and Tóbiás, A. [2017/2], which investigates a probabilistic model for routeing in a multihop ad-hoc communication network. Messages travel in random hops via the other users. Their trajectories are chosen at random according to a Gibbs distribution that favours trajectories with low interference and little total congestion. This model was

introduced in our earlier paper König, W. and Tóbiás, A. [2017/1], where we expressed, in the high-density limit, optimal trajectories as the minimizer of a characteristic variational formula.

In König, W. and Tóbiás, A. [2017/2], we derive qualitative properties of the minimizer, in the special case in which congestion is not penalized. We quantify emerging typical pictures in analytic terms in three extreme regimes. We analyze the typical number of hops and the typical length of a hop, and the deviation of the trajectory from the straight line in two regimes, (1) in the limit of a large communication area and large distances, and (2) in the limit of a strong interference weight. The typical trajectory turns out to quickly approach a straight line. Surprisingly, in regime (1), the typical hop length diverges logarithmically in the distance. We further analyze the local and global repulsive effect of (3) a densely populated area on the trajectories.

Acknowledgement. The research of A. Tóbiás was supported by the Phase II scholarship of the Berlin Mathematical School.

References

W. König and A. Tóbiás (2017), A Gibbsian model for message routing in highly dense multi-hop networks, arXiv: 1704.03499.

W. König and A. Tóbiás (2017), Routeing properties in a Gibbsian model for highly dense multihop networks. Available here.

CS 13
Filtering
Rm -1.63
Tue
17:10-17:30

Kalman Filtering with censored states and measurements

Kostas Loumponias, George Tsaklidis*

Aristotle University of Thessaloniki, Greece
email: *tsaklidi@math.auth.gr

Recently, progress is made in estimating hidden states via state space modeling when censored measurements have to be taken into account. The Tobit Kalman filter (TKF) and the Adaptive Tobit Kalman filter (ATKF) are proposed in order to estimate the vector of hidden states, by partially known (censored) measurements, presented by the measurement vector. In the standard TKF and ATKF filtering, only measurements are considered to be censored, while the hidden states evolve in an unsencored environment. In our study censoring concerning the hidden states as well as the measurements has been taken into account, and the equations for the associated KF are derived.

References

Allik, B., Miller, C., Piovoso, M. J. and Zurakowski, R. (2016), The Tobit Kalman filter: An estimator for censored measurements, *IEEE Trans. Control Syst. Technol.*, **24**, no. 1, pp. 365–371.

Tobin, J. (1958), Estimation of relationships for limited dependent variables, *Econometrica*, **26**, no. 1, pp. 24–36

POSTER
Mon
18:00-19:00

Distribution of suprema for generalized risk processes

Ivana Geček Tuđen

University of Zagreb, Croatia
email: igecek@math.hr

The basic risk model, known as the Cramér-Lundberg model, has been revisited many times in the risk theory and generalized in a few ways. We study the generalized risk process $X(t) = Y(t) - C(t)$, $t \in [0, \tau]$, where Y is a Lévy process, C an independent subordinator and τ an independent exponential time. This allows us to observe the process in the context of the fluctuation theory for Lévy

processes. In this surrounding, we derive a Pollaczek-Khinchine type formula for the supremum of the dual process $\widehat{X} = -X$ on $[0, \tau]$ which generalizes the previously known results. We also drop the standard assumptions on the finite expectations of the processes Y and C and the net profit condition, and discuss which assumptions are necessary for deriving our results. At the end we revisit and explain the assumptions and obtained results from the point of view of the ladder process.

Approximate and exact simulation methods for the maximum of a stable process

IS 13
Lévy Processes & Appl.
Harmony lnge
Mon
16:00 - 16:30

Jorge González Cázares[*], Aleksandar Mijatović[*,†], Gerónimo Uribe Bravo[‡,§]

[*]King's College, London, UK
[†]The Alan Turing Institute, London, UK
[‡]Universidad Nacional Autónoma de México, México City, México
email: [§]geronimo@matem.unam.mx

We present two algorithms to sample the maximum of a stable process on any finite interval. Both algorithms are based on exploiting the scaling properties of stable processes to deduce an unexpected perpetuity equation which highlights the self-decomposing character of the stable maximum. The first algorithm provides is approximate, (comparatively) fast, and is accompanied by explicit error bounds on the γ-Wasserstein distance to the target distribution. The second algorithm is exact and based on perfect simulation through dominated coupling from the past. Though slower than the first, we will argue that this exact algorithm is still of practical use.

Acknowledgement. GUB's research supported by CoNaCyT grant FC-2016-1946 and UNAM-DGAPA-PAPIIT grant IN115217.

Fractional Weibull and exponential distributions

CS 19
Distributio
Rm 100/B
Tue
16:30-16:50

Lotfi Boudabsa[*], Pierre Vallois

Université de Lorraine, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F54506, France
email: [*]pierre.vallois@univ-lorraine.fr

We define the fractional Weibull distribution function as the unique solution of $D^\alpha F_{\alpha,\rho,\lambda}(x) = \lambda x^{\rho-\alpha}\big(1 - F_{\alpha,\rho,\lambda}(x)\big)$ in $[0, +\infty[$, with initial condition $F_{\alpha,\rho,\lambda}(0) = 0$, where $\lambda, \rho > 0, \alpha \in]0, 1]$ and D^α is the fractional operator of order α. When $\alpha = 1$, $F_{\alpha,\rho,\lambda}$ coincides with the classical Weibull cumulative distribution function (CDF). We prove that $F_{\alpha,\rho,\lambda}$ is actually a CDF, is the sum of an explicit series and $1 - F_{\alpha,\rho,\lambda}$ is completely monotonic when $\rho \in]0, 1]$. We determine the associated probability measure whose Laplace transform equals $1 - F_{\alpha,\rho,\lambda}$. In the case where $\rho = \alpha$, $F_{\alpha,\rho,\lambda}$ is the fractional exponential CDF and can be expressed via stable symmetric laws with parameter α.

CS 11
Lévy processes
Rm -1.62
Wed
12:40 - 13:00

Maximum Loss of Spectrally Negative Lévy Processes

Ceren Vardar-Acar[*,‡], Mine Çağlar[†]
[*]Department of Statistics, Middle East Technical University, Ankara, Turkey
[†]Department of Mathematics, Koç University, Istanbul, Turkey
email: [‡]cvardar@metu.edu.tr

The maximum loss, or maximum drawdown of a process X is the supremum of X reflected at its running supremum. The motivation comes from mathematical finance as it is useful to quantify the risk associated with the performance of a stock.

The maximum loss at time $t > 0$ is formally defined by

$$M_t^- := \sup_{0 \le u \le v \le t} (X_u - X_v)$$

which is equivalent to $\sup_{0\le v\le t}(\sup_{0\le u\le v}(X_u - X_v))$ and $\sup_{0\le v\le t}(S_v - X_v)$, that is, the supremum of the reflected process $S - X$, or the so-called loss process, where S denotes the running supremum.

The loss process has been studied for Brownian motion [Salminen and Vallois 2007, Vardar-Acar et al. 2013], and some Lévy processes [Mijatović and Pistorius 2012]. A spectrally negative Lévy process X is a Lévy process with no positive jumps, that is, its Lévy measure is concentrated on $(-\infty, 0)$.

In this study, the joint distribution of the maximum loss and the maximum gain is obtained for a spectrally negative Lévy process until the passage time of a given level. Their marginal distributions up to an independent exponential time are also provided. The existing formulas for Brownian motion with drift are recovered using the particular scale functions.

References

Aleksandar Mijatović and Martijn R. Pistorius. On the drawdown of completely asymmetric Lévy processes. *Stochastic Process. Appl.*, 122(11):3812–3836, 2012. doi: 10.1016/j.spa.2012.06.012.

Paavo Salminen and Pierre Vallois. On maximum increase and decrease of Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.*, 43(6):655–676, 2007.

Ceren Vardar-Acar, Craig L. Zirbel, and Gábor J. Székely. On the correlation of the supremum and the infimum and of maximum gain and maximum loss of Brownian motion with drift. *J. Comput. Appl. Math.*, 248:61–75, 2013. doi: 10.1016/j.cam.2013.01.010.

CS 9
Stability of financial systems
Harmony Inge
Tue
17:30-17:50

Estimating short and long term effects of capital requirements – evidence from Hungary

Tibor Szendrei, Katalin Varga[*]
Magyar Nemzeti Bank
email: [*]vargaka@mnb.hu

Estimating long term benefits of capital regulation is troublesome highlighted by the contradictory results of several studies in the field. The linkage between financial system and the real economy is state-dependent, as suggested by standard models with financial frictions; hence, the transmission of policies varies with the state of the economy. A Threshold Bayesian Vector Autoregression (TBVAR) framework is used where both the threshold value and the delay parameter are determined endogenously. The variable splitting the time series into two regimes is a financial stress index: the FISS. The TBVAR completes the FISS as this modelling framework is capable of yielding a threshold

value that is informative for policymakers while giving insights about the underlying interactions in the economy. The effect of a regulatory capital shock on the variable describing the real economy has the same direction across the two regimes with the extent of the impacted shrinkage being different. When distinguishing between loans to households and non-financial corporations there is evidence that banks try to reduce their household loan activity before shrinking their corporate loan portfolio. In fact, in the normal regime there is no significant reaction of the corporate loan segment to a regulatory capital increase. Evidence of heterogeneity is found in the effects of the regulation across the banking sector. When distinguishing between banks with foreign ownership and domestic ownership there is a significant difference in the reaction to a regulatory capital shock. Although, this is a crude way to introduce differences it clearly highlights the possibility of further improving the precision of the estimates by tackling the cross sectional dimension of the banking system.

Generalised block bootstrap in temperature data modelling

IS 35
Environmen
tal Models
Rm -1.63
Thu
11:30 - 12:00

László Varga*, András Zempléni

Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest
email: *vargal4@cs.elte.hu

In an earlier article (Rakonczai et al. [2014]) we emphasized the importance of investigating the effective sample size in case of autocorrelated data. The simulations were based on the block bootstrap methodology. However, the discreteness of the usual block size did not allow for exact calculations.

In the paper Varga and Zempléni [2017] a new generalisation is proposed of the block bootstrap methodology, which allows for any positive real number as expected block size. We relate it to the existing optimisation procedures and apply it to a temperature data set. Our other focus is on statistical tests, where quite often the actual sample size plays an important role, even in case of relatively large samples. This is especially the case for copulas. These are used for investigating the dependencies among data sets. As in quite a few real applications the time dependence cannot be neglected, we investigated the effect of this phenomenon on the used test statistic. The critical value can be computed by the proposed new block bootstrap simulation, where the block size is determined by fitting a vector autoregression (VAR) model to the observations.

The results are illustrated on models for the used temperature data. We modelled the dependence structure of 5 site pairs in the Carpathian Basin of the gridded temperature database of E-OBS. Our main goal was to investigate the change of dependence structure with testing the homogeneity of two copulas (first and second half of the samples). Our main meteorological conclusion is that we have found some significant changes in the dependence structure between the standardised temperature values, which is more obvious at grid point pairs lying farther from each other.

Acknowledgement. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES and the data providers in the ECA&D project (http://www.ecad.eu).

References

Rakonczai, P. and Varga, L. and Zempléni, A. (2014) Copula fitting to autocorrelated data with applications to wind speed modelling. *Advances in Statistical Climatology, Meteorology and Oceanography*, **3(1)**, p.55-66

Varga, L. and Zempléni, A. (2017) Generalised block bootstrap and its use in meteorology. *Ann. Univ. Sci. Budapest. Sect. Comput.*, **43**, p.3-20

CS 8
Financial Risk and Valuation
Rm 100/B
Tue
12:00 - 12:20

Dynamic Probability Scoring Rules: Comparing the Quality of Time Evolving Probabilistic Predictions

Jan Vecer
Charles University, Prague, Czech Republic
email: vecer@karlin.mff.cuni.cz

Measuring the quality of a probabilistic estimate of an outcome is done by a scoring rule. Different types of scoring have been extensively studied in the previous literature. A class of proper scoring rules gives the highest expected reward to the true probability distribution. Examples of proper scoring rules include the logarithmic scoring rule and the Brier (quadratic) score. The scoring rules apply to one probability estimate and one time observation. However, it is not clear how to measure the quality of the probabilistic estimate if such an estimate evolves in time. Examples of such probabilistic estimates are election predictions, weather predictions, or probabilities of some market events that appear in hedging of financial products, such as probabilities that a price of an asset will end up over or below a certain level.

In our research, we present a novel approach to measuring a quality of probabilistic estimates. It is based on a comparison of two probabilistic series rather than on a fit of a single estimate with the outcome. The basic idea of our approach is that if we have two different quotes of two probability estimates, one can use this discrepancy for setting a bet of these two values against each other. For instance, estimates of 0.5 and 0.6 in two time series give a possibility to set a bet in the interval [0.5, 0.6]. The exact value of the probability set by the bet and the corresponding volume is determined by optimization of some utility function that describes the hypothetical behavior of these two bettors. Such optimization procedure finds an equilibrium, where the supply and demand functions of the two agents meet. This creates a sequence of trades that matches every discrepancy that was not reflected in the past trades.

We show that the expected profit loss of the true probability series is positive against any other probability sequence. This is an analogous concept to the proper scoring rule. We illustrate the performance of these dynamic probabilistic scoring rules on examples of simulated evolutions of probabilities of a price asset ending above a certain level using various parameters and demonstrate that the paths corresponding to the correct parameters are indeed performing best.

Acknowledgement. Jan Vecer's research was supported in part by the Grant Agency of the Czech Republic, Grant GACR 16-21216S.

References

Gneiting, T., & Raftery, A. E (2007) Strictly proper scoring rules, prediction, and estimation. *Journal of American Statistical Association*, 102(477), 359-378.

Winkler, R. L. (1969) Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association*, 64(327), 1073-1078.

Quickest Detection of Dynamically Evolving Events in Networks

IS 41
Quickest Change Detection
Rm 100/A
Wed
11:30 - 12:00

VENUGOPAL V. VEERAVALLI[*], SHAOFENG ZOU[†]

University of Illinois at Urbana-Champaign, USA
email: [*]vvv@illinois.edu [†]szou3@illinois.edu

We consider the problem of quickest detection of dynamically evolving events in networks. After an event occurs, a number of nodes in the network are affected and undergo a change in the statistics of their observations. We assume that the event is dynamic and can propagate with time, i.e., different nodes perceive the event at different times. The goal is to design a sequential algorithm that can detect when the event has affected η or more nodes as quickly as possible, subject to false alarm constraints. We design a computationally efficient algorithm called the Spartan-CuSum test for this problem. We show that the Spartan-CuSum test is adaptive to unknown propagation dynamics, and demonstrate its asymptotic optimality as the false alarm rate goes to zero. We also provide numerical simulations to validate our theoretical results.

Acknowledgement. This research was supported in part by the US National Science Foundation under grant CCF 1618658, by the US Army Research Laboratory under cooperative agreement W911NF-17-2-0196, through the University of Illinois at Urbana-Champaign.

Markov Decision Processes with long duration

IS 56
Stochastic Game Theory
Rm -1.64
Tue
11:30 - 12:00

XAVIER VENEL[*,‡], BRUNO ZILIOTTO[†]

[*]PSE-Université Paris 1, CES
[†]CEREMADE - CNRS
email: [‡]xavier.venel@univ-paris1.fr

The standard model of Markov Decision Process (or Controlled Markov chain) was introduced by Bellman [1957] and has been extensively studied since then. In this model, at the beginning of every stage, a decision-maker perfectly observes the current state, and chooses an action accordingly, possibly randomly. The current state and the selected action determine a stage payoff and the law of the next state. There are two standard ways to aggregate the stream of payoffs. Given a strictly positive integer n, in the n-stage MDP, the total payoff is the Cesaro mean $n^{-1}\sum_{m=1}^{n} g_m$, where g_m is the payoff at stage m. Given $\lambda \in (0,1]$, in the λ-discounted MDP, the total payoff is the λ-discounted sum $\lambda\sum_{m\geq 1}(1-\lambda)^{m-1}g_m$. The maximum expected payoff that the decision-maker can obtain in the n-stage problem (resp. λ-discounted problem) is denoted by v_n (resp. v_λ).

A huge part of the literature investigates *long-term* MDPs, that is, MDPs which are repeated a large number of times. It can be done following several approaches. The first approach is to determine whether (v_n) and (v_λ) converge when n goes to infinity and λ goes to 0, and whether the two limits coincide. When this is the case, the MDP is said to have an *asymptotic value*. A second approach is to define the payoff in the infinite problem as the inferior limit of the expectation of $n^{-1}\sum_{m=1}^{n} g_m$. In the literature, this is referred as the *long-run average payoff criterion* (AP criterion, see Araposthathis et al. [1993] for a review of the subject). When the asymptotic value exists and coincides with the value in behavior (resp. pure) strategies of the infinite problem, the MDP is said to have a *uniform value* in behavior (resp. pure) strategies. A third approach is to define the payoff as being the expectation of $\liminf_{n\to+\infty} n^{-1}\sum_{m=1}^{n} g_m$ as studied in Gillette [1957]. The decision maker is particularly pessimistic when he aggregates the payoff. Opposite to this case, one can look at the MDP where the payoff

in the infinite problem is the expectation of $\limsup_{n\to+\infty} n^{-1}\sum_{m=1}^{n} g_m$ and more generally study evaluations which depends on the strategy of the decision maker.

Renault [2011] showed the existence of the uniform value for MDPs with compact set of states under some regularity assumptions on the transitions. Renault and Venel [2017] showed that the uniform value in this case has more properties since the same strategy is not only good for Cesaro Means and Abel means but for other mean evaluations. In a previous work, Venel and Ziliotto [2016] proved that the value for the inferior limit also coincides with the uniform value when the set of states and the set of actions are compact and with the same regularity assumptions. In this paper, we extend the result to more general evaluation and in particular to the limsup evaluation proving that all the notions coincide and yield the uniform value. In particular, even if the decision maker is particularly optimist in the way he is aggregating the payoff, he can not guarantee more than the uniform value.

Acknowledgement. X. Venel's research was supported by the French National Agency under grant ANR CIGNE ANR-15-CE38-0007-01.

References

Arapostathis A., Borkar V., Fernandez-Gaucherand E., Gosh M. and Marcus S. (1993) Discrete-time controlled Markov processes with the average cost criterion: a survey. *SIAM Journal on Control and Optimization*, **31**, p.282-344

Bellman R. (1957) A Markovian decision process. Technical report *DTIC Document*.

Gillette D. (1957) Stochastic games with zero stop probabilities. *Contributions to the Theory of Games*, **3**, p.179-187

Renault J. (2011) Uniform value in dynamic programming. *Journal of European Mathematical Society*, **13**, p.309-330.

Renault J. and Venel X. (2017) Long-Term Values in Markov Decision Processes and Repeated Games, and a New Distance for Probability Spaces. *Mathematics of Operations Research*, **42**, p.349-376.

Venel X. and Ziliotto B. (2016) Strong Uniform Value in Gambling Houses and Partially Observable Markov Decision Processes *SIAM Journal on Control and Optimization*, **54**, p.1983-2008.

IS 13 Lévy Processes & Appl.
Harmony Inge
Mon
16:30 - 17:00

A universal approach to estimate the conditional variance in semimartingale limit theorems

Mathias Vetter

Christian-Albrechts-Universität zu Kiel
email: vetter@math.uni-kiel.de

The typical central limit theorems in high-frequency asymptotics for semimartingales are results on stable convergence to a mixed normal limit with an unknown conditional variance. Estimating this conditional variance usually is a hard task, in particular when the underlying process contains jumps. For this reason, several authors (Mykland and Zhang (2017) doi: 10.3982/ECTA12501, Christensen et al. (2017) SSRN: 2665345) have recently discussed methods to automatically estimate the conditional variance, i.e. they build a consistent estimator from the original statistics, but computed at various different time scales. Their methods work in several situations, but are essentially restricted to the case of continuous paths always. The aim of this work is to present a new method to consistently estimate the conditional variance which works regardless of whether the underlying process is continuous has jumps. We will discuss the case of power variations in detail and give insight to the heuristics behind the approach.

On model fitting and estimation of stationary processes with applications to finance

IS 9
Hedging, Model Fitting, Estimation
Rm 100/B
Tue
11:30 - 12:00

Lauri Viitasaari

University of Helsinki, Helsinki
email: lauri.viitasaari@iki.fi

Stationary processes form an important class of stochastic processes that has been extensively studied in the literature. Their applications include modelling and forecasting numerous real life phenomenon including natural disasters, sustainable energy sources, sales and market movements.

One of the most essential families of stationary processes is the ARMA family. When modelling existing data with ARMA process, the first step is to fix the orders of the model. After that, one can estimate the related parameters by using standard methods such as maximum likelihood (ML) or least squares (LS) estimators. The final step is to conduct various diagnostic tests in order to determine the quality of the model.

In this talk we present a novel way of fitting a model to a data that is assumed to be a realization from a discrete time stationary process. Our approach is based on a recently proved AR(1) characterisation of stationary processes, where the noise is not assumed to be white. As a result, we obtain more general and easier way to fit a model into a stationary time series, thus outperforming traditional ARMA approaches. In particular, we obtain closed form consistent estimators of various model parameters and their asymptotic normality under general conditions. The results are then applied to the ARCH model with a memory effect. ARCH models can be employed, e.g. in modeling time-varying volatility. We also discuss continuous time extensions.

FCFS Parallel Service Systems and Matching Models

CS 6
Queuing Theory
Rm 100/A
Tue
18:10-18:30

Ivo Adan[*], Rhonda Righter[†], Gideon Weiss[‡,§]

[*]Technical University of Eindhoven, Netherlands
[†]University of California at Berkeley, USA
[‡]The University of Haifa, Israel
email: [§]gweiss@stat.haifa.ac.il

We consider three parallel service models in which customers of several types are served by several types of servers subject to a bipartite compatibility graph, and the service policy is first come first served. Two of the models have a fixed set of servers. The first is a queueing model in which arriving customers are assigned to the longest idling compatible server if available, or else queue up in a single queue, and servers that become available pick the longest waiting compatible customer, as studied by Adan and Weiss [2014]. The second is a redundancy service model where arriving customers split into copies that queue up at all the compatible servers, and are served in each queue on FCFS basis, and leave the system when the first copy completes service, as studied by Gardner et al. [2016]. The third model is a matching queueing model with a random stream of arriving servers. Arriving customers queue in a single queue and arriving servers match with the first compatible customer and leave the system at the moment of arrival, or they leave without a customer. The last model is relevant to organ transplants, to housing assignments, to adoptions and many other situations.

We study the relations between these models, and show that they are closely related to the FCFS infinite bipartite matching model, in which two infinite sequences of customers and servers of several types are matched FCFS according to a bipartite compatibility graph, as studied by Adan et al. [2017].

Acknowledgement. The work of Gideon Weiss is supported in part by Israel Science Foundation Grant 286.13

References

Adan, I., Weiss, G. (2014) A queue with skill based service under FCFS-ALIS: steady state, overloaded system, and behavior under abandonments. *Stochastic Systems*, **4**(1):250-299.

Gardner, K., Zbarsky, S., Doroudi, S., Harchol-Balter, M., Hyytia, E., Scheller-Wolf, A. (2016). Queueing with redundant requests: exact analysis. *Queueing Systems*, 83(3-4), 227-259.

Adan, I., Busic, A., Mairesse, J., Weiss, G. (2017). Reversibility and further properties of FCFS infinite bipartite matching. *Mathematics of Operations Research* published online.

CS 8
Financial Risk and Valuation
Rm 100/B
Tue
12:40 - 13:00

Asian Option Pricing with Orthogonal Polynomials

Sander Willems

EPFL and Swiss Finance Institute, Lausanne, Switzerland
email: sander.willems@epfl.ch

In this paper we derive a series expansion for the price of a continuously sampled arithmetic Asian option in the Black-Scholes setting. The expansion is based on polynomials that are orthogonal with respect to the log-normal distribution. All terms in the series are fully explicit and no numerical integration nor any special functions are involved. We provide sufficient conditions to guarantee convergence of the series. We address the moment indeterminacy of the log-normal distribution and numerically investigate its impact on the asymptotic behavior of the series.

IS 54
Random Matrix Theory
Rm -1.64
Tue
15:30 - 16:00

Gaussian multiplicative chaos in random matrix theory

Nathanael Berestycki[*,‡], Christian Webb[†], Mo Dick Wong[*]

[*]University of Cambridge, Cambridge, United Kingdom
[†]University of Aalto, Helsinki, Finland
email: [‡]N.Berestycki@statslab.cam.ac.uk

Gaussian multiplicative chaos (GMC) is a random multifractal measure that has appeared in many different areas of mathematics, from Kolmogorov-Obukhov's model of turbulence, Liouville's theory of quantum gravity to probabilistic number theory. In this talk I shall first give an introduction to the subject, and then discuss an ongoing programme in establishing GMC as a universal object in random matrix theory. In particular I shall explain that the characteristic polynomials of a large class of random Hermitian matrices, when suitably normalized, behave like GMC as the size of the matrix goes to infinity. This is based on a joint work with Nathanael Berestycki and Christian Webb.

Acknowledgement. N. Berestycki's work was supported by EPSRC grants EP/L018896/1 and EP/I03372X/1. M.D. Wong is a PhD student at the Cambridge Centre for Analysis, supported by EPSRC grant EP/L016516/1 and a Croucher Foundation Scholarship. C. Webb was supported by the Academy of Finland grants 288318 and 308123.

References

Berestycki, N., Webb, C. and Wong, M.D. (2018) Random Hermitian matrices and Gaussian multiplicative chaos. *Probab. Theory Relat. Fields*. doi: 10.1007/s00440-017-0806-9

Webb, C. and Wong, M.D. (2017) On the moments of the characteristic polynomial of a Ginibre random matrix. arXiv: 1704.04102

On a multivariate renewal-reward process involving time delays and discounting: Applications to IBNR process and infinite server queues.

IS 55
Applied Probability
Rm 0.99
Wed
10:30 - 11:00

Jae Kyung Woo

School of Risk and Actuarial Studies, UNSW, Sydney, Australia.
email: j.k.woo@unsw.edu.au

In this talk, we consider a particular renewal-reward process with multivariate discounted rewards (inputs) where the arrival epochs are adjusted by adding some random delays. Then this accumulated reward can be regarded as multivariate discounted Incurred But Not Reported (IBNR) claims in actuarial science and some important quantities studied in queueing theory such as the number of customers in $G/G/\infty$ queues with correlated batch arrivals. We study the long term behavior of this process as well as its moments. Asymptotic expressions and bounds for the quantities of our interest are studied. Next, assuming exponentially distributed delays, we derive some explicit and numerically feasible expressions for the limiting joint moments. In such case, for an infinite server queues with renewal arrival process, we obtain limiting results on the expectation of the workload, and the covariance of queue size and workload. Finally, some numerical examples are given. This is joint work with Landy Rabehasaina.

Sequential Low-Rank Change-Point Detection

IS 40
Sequential Methods
Rm 100/A
Wed
14:30 - 16:00

Yao Xie

Georgia Institute of Technology
email: yao.xie@isye.gatech.edu

We study the problem of detecting an abrupt change to the covariance matrix using a sequence of observations. In particular, the covariance changes from a "white" identity matrix to an unknown spiked covariance matrix. We propose a CUSUM type of procedure based subspace tracking without having to specify all the parameters, and the drift parameters are estimated from data. We show that the procedure can perform better than adapting the largest eigenvalue statistics over a sliding window, which would be a default option if one extends the low-rank hypothesis test.

Acknowledgement. This is joint work with George Moustakides and Liyan Xie. This work is partially supported by NSF CAREER Award CCF-1650913.

Systems with weighted components with application to wind energy

IS 53
Engineering Systems
Rm 0.99
Tue
17:00 - 17:30

Femin Yalcin[*,‡], Serkan Eryilmaz[†]

[*]Izmir Katip Celebi University, Izmir, Turkey
[†]Atilim University, Ankara, Turkey
email: [‡]femin.yalcin@ikc.edu.tr

In a system with weighted components, the system's components contribute differently to the capacity of the system. Systems with weighted components are useful to model various capacity-based engineering systems such as oil transportation systems, power generation systems, and production systems. Consider a system consisting of n components, each with its own positive weight. The weight of a component might be assumed to be its performance rate. In this study, systems with binary and multi-state weighted components are used to model and evaluate wind power systems.

References

Eryilmaz, S. (2018) Reliability analysis of multi-state system with three-state components and its application to wind energy. *Reliab. Eng. Syst. Safe.*, **172**, p.58-63.

IS 25
Population Dynamics Studies
Rm -1.62
Mon
16:30 - 17:00

Branching Random Walks and their Applications to Population Dynamics

Elena Yarovaya

Lomonosov Moscow State University, Moscow, Russian Federation
email: yarovaya@mech.math.msu.su

The modelling of population dynamics in biology, demography, medicine and genetics is a significant stimulus for the development of stochastic processes. We consider stochastic processes with generation and transport of particles that essentially determined by properties of a particle motion and dimension of the space in which the particles evolve. Such processes on discrete structures are usually called *branching random walks*. The main object of study is a continuous-time symmetric branching random walk on a multidimensional lattice with a finite set of the particle generation centers, named *branching sources*. The description of an underlying random walk in terms of Green's function allows us to offer a general approach for investigating random walks with finite, as well as with infinite, variance of jump. Such branching random walks can be applied for modeling complex stochastic systems with different spatial dynamics, implying the existence of heavy-tailed distributions of random walk jumps. The behavior of branching random walks is mainly defined by properties of the evolutionary operator for the mean number of particles both at an arbitrary point and on the entire lattice. The main attention is paid to the investigation of how the geometric configuration of the branching sources in a branching random walk affects the behavior of the process. In particular, we obtain limit theorems for the case when pairwise distance between branching sources of different intensities tends to infinity.

Acknowledgement. This research is supported by the Russian Foundation for Basic Research, grant no. 17-01-00468.

IS 36
Applied Prob. & Stat. Inference I.
Rm 100/A
Mon
11:30 - 12:00

Inference on High-Dimensional Mean Vectors Under the Strongly Spiked Eigenvalue Model

Kazuyoshi Yata[*,‡], Makoto Aoshima[*], Aki Ishii[†]

[*]Institute of Mathematics, University of Tsukuba, Ibaraki, Japan
[†]Department of Information Sciences, Tokyo University of Science, Chiba, Japan
email: [‡]yata@math.tsukuba.ac.jp

Aoshima, M. and Yata, K. [2018] provided two disjoint models: the strongly spiked eigenvalue (SSE) model and the non-SSE (NSSE) model, for high-dimensional data. In this talk, we consider asymptotic normality for inference on high-dimensional mean vectors under the SSE model. We note that, under the SSE model, the asymptotic normality of statistics is not valid because it is heavily influenced by strongly spiked eigenvalues. In order to give a treatment of the SSE models, we develop data transformation techniques that transform the SSE models to the non-SSE models and create new statistics under the SSE model. We verify that the proposed statistics are asymptotically distributed as a normal distribution under the SSE model. With the help of the asymptotic normality, we consider inference on high-dimensional mean vectors under the SSE model.

Acknowledgement. The research of the first author was partially supported by Grant-in-Aid for Young Scientists (B), Japan Society for the Promotion of Science (JSPS), under Contract Number 26800078. The research of the second author was partially supported by Grants-in-Aid for Scientific Research (A) and Challenging Research (Exploratory), JSPS, under Contract Numbers 15H01678 and 17K19956.

References

Aoshima, M. and Yata, K., (2018) Two-sample tests for high-dimension, strongly spiked eigenvalue models. *Stat. Sinica* **28**, 43-62.

Some Monotonicity Properties of Parametric and Nonparametric Bayesian Bandits

CS 17
Bayesian Approaches and MCMC
Rm 0.87
Wed
17:10-17:30

Yaming Yu

Department of Statistics, University of California at Irvine
email: yamingy@uci.edu

One of two independent stochastic processes (arms) is to be selected at each of n stages. The selection is sequential and depends on past observations as well as the prior information. The objective is to maximize the expected future-discounted sum of the n observations. We study structural properties of this classical bandit problem, in particular how the maximum expected payoff and the optimal strategy vary with the priors, in two settings: (a) observations from each arm have an exponential family distribution and different arms are assigned independent conjugate priors; (b) observations from each arm have a nonparametric distribution and different arms are assigned independent Dirichlet process priors. In both settings, we derive results of the following type: (i) for a particular arm and a fixed prior weight, the maximum expected payoff increases as the prior mean yield increases; (ii) for a fixed prior mean yield, the maximum expected payoff increases as the prior weight decreases. Specializing to the one-armed bandit, the second result captures the intuition that, given the same immediate payoff, the less one knows about an arm, the more desirable it becomes because there remains more information to be gained when selecting that arm. A key tool in the derivation is stochastic orders.

On the Methods of Bounding the Rate of Convergence for Inhomogeneous Continuous-time Markov Chains

IS 17
Stochastic Processes & Appl. I.
Rm -1.63
Tue
11:30 - 12:00

Alexander Zeifman[*,‡], Galina Shilova[†], Anastasia Kryukova[†], Ksenia Kiseleva[†]

[*]Vologda State University, Vologda, Russia; IPI FRC CSC RAS; ISEDT RAS, Russia
[†]Vologda State University, Vologda, Russia
email: [‡]a_zeifman@mail.ru

In our previous papers we presented an approach to finding sharp upper bounds in natural metrics via essential positivity of the reduced intensity matrix of a Markov chain. These bounds are sharp for nonnegative difference of the initial conditions of two probability distributions of a Markov chain. However, in a general situation the assumption of nonnegativity of this difference does not hold. Here we suggest a new approach for obtaining sharp bounds on the rate of convergence to the limiting characteristics for a general class of (inhomogeneous) finite continuous-time Markov chains. Some queueing examples are considered.

References

Zeifman, A. I., Korolev, V. Y. (2015) Two-sided Bounds on the Rate of Convergence for Continuous-time Finite Inhomogeneous Markov Chains. *Statistics & Probability Letters*, **103**, p. 30–36.

Zeifman, A., Sipin, A., Korolev, V., Shilova, G., Kiseleva, K., Korotysheva, A., Satin, Y. (2018) On Sharp Bounds on the Rate of Convergence for Finite Continuous-Time Markovian Queueing Models. *Lecture Notes in Computer Science*, **10672**, p. 20–28.

Zeifman, A. I., Korolev, V. Y., Satin, Y. A., Kiseleva, K. M. (2018) Lower Bounds for the Rate of Convergence for Continuous-time Inhomogeneous Markov Chains with a Finite State Space. *Statistics & Probability Letters*, **137**, p. 84–90.

Zeifman, A., Razumchik, R., Satin, Y., Kiseleva, K., Korotysheva, A., Korolev, V. (2018) Bounds on the Rate of Convergence for One Class of Inhomogeneous Markovian Queueing Models with Possible Batch Arrivals and Services. *Int. J. Appl. Math. Comp. Sci.*, **28**.

IS 50
Random Structures
Rm -1.64
Wed
16:30 - 17:00

Characterizations of dynamic preferential attachment networks

Panpan Zhang[*,‡], Hosam M. Mahmoud[†]

[*]University of Connecticut, Storrs, CT, U.S.A.
[†]The George Washington University, Washington, DC, U.S.A.
email: [‡]panpan.zhang@uconn.edu

We investigate a popular random network model which receives a lot of attention in the last two decades—the preferential attachment (PA) network proposed by Barabási, A.-L. and Albert, R. [1999]. Since it was released, mathematicians, statisticians and applied scientists have extended this model by considering and adding various dynamical flavors. Several recent representative research articles include but not limited to Peköz et al. [2013, 2017], Resnick, S. I. and Samorodnitsky, G. [2016], Wang, T. and Resnick, S. I. [2017]. In the first part of the presentation, we briefly review the models proposed in the above research articles and the results therein. Next, we specifically look into the PA network model proposed in Zhang, P. [2016]. We uncover two typical properties for this class of PA networks: the joint distribution of the number of nodes of small degrees and the degree profile of a node as network evolves. For the former property, we develop a Gaussian law via multivariate martingale techniques; while for the latter one, we exploit Pólya urn model to determine the first two moments of the random variable of interest, where we discover a phase transition in the asymptotic mean. Lastly, we propose some directions in our future study of PA network models. The preprint of this research can be found in Zhang, P. and Mahmoud, H. M. [2016].

References

Barabási, A.-L. and Albert, R. (1999) Emergence of scaling in random networks. *Nature*, **286**, p.509-512

Pekö, A. E., Röllin, A. and Ross, N. (2013) Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.*, **23**, p.1188-1218

Pekö, A. E., Röllin, A. and Ross, N. (2013) Joint degree distributions of preferential attachment random graphs. *Adv. Appl. Probab.*, **49**, p.368-387

Resnick, S. I. and Samorodnitsky, G. (2016) Asymptotic normality of degree counts in a preferential attachment model. *Adv. in Appl. Probab.*, **48**, p.283-299.

Wang, T. and Resnick, S. I. (2017) Asymptotic normality of in- and out-degree counts in a preferential attachment model. *Stoch. Models*, **33**, p. 229-255.

Zhang, P. (2016) On terminal nodes and the degree profile of preferential dynamic attachment circuits. In *Proceedings of the Thirteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 80-92, Arlington, VA.

Zhang, P. and Mahmoud, H. M. (2016) On nodes of small degrees and degree profile in preferential dynamic attachment circuits. arXiv: 1610.04675.

On Heterogeneity in the Individual Model with both Dependent Claim Occurrences and Severities

CS 7
Actuarial Applications
Rm 0.99
Mon
17:20-17:40

Yiying Zhang

Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong
email: zhangyiying@outlook.com

It is a common belief for actuaries that the heterogeneity of claim severities in a given insurance portfolio tends to increase its dangerousness, which results in requiring more capital for covering claims. In this talk, we aim to investigate the effects of orderings and heterogeneity among scale parameters on the aggregate claim amount when both claim occurrence probabilities and claim severities are dependent. Under the assumption that the claim occurrence probabilities are LWSAI, the actuaries' belief is examined from two directions, i.e., claim severities are comonotonic or RWSAI. An application in assets allocation is addressed as well. This is a joint work with Xiaohu Li and Ka Chun Cheung.

Acknowledgement. Yiying Zhang thanks the Hong Kong Ph.D. Fellowship Scheme (PF14-13413) supported by Hong Kong Research Grants Council. Ka Chun Cheung acknowledges the financial support by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 17324516).

Dynamic Pricing in Insurance

POSTER
Mon
18:00-19:00

Yuqing Zhang*, Neil Walton†, Kees van Schaik‡

School of Mathematics, The University of Manchester, Manchester, UK.
email: *yuqing.zhang@manchester.ac.uk †neil.walton@manchester.ac.uk
‡kees.vanschaik@manchester.ac.uk

We develop a pricing policy that enables an insurance company to find the optimal price and maximize the expected profit. The insurance company sells a single product and adopts prices to learn its customers' responses. The pricing policy is determined by the selling price and other model parameters. The parameters of the underlying model are initially unknown to the insurance company, so each price decision involves a trade-off between learning and earning. Maximum quasi-likelihood estimation (MQLE) is used to estimate the unknown parameters in the model. We build an algorithm that guarantees that MQLE parameter estimates eventually exist and converge to the correct values, which implies that the sequence of chosen prices also converge to the optimal price. The performance of the pricing policy is measured in terms of the *regret*: the expected revenue loss caused by not using the optimal price. Upper bounds on the regret can be achieved by our pricing algorithm. The advantages of this new pricing policy are clear. Firstly, it formulates a learning-and-earning problem, where price is used as a learning tool to explore the demand and claims response to different prices. Secondly, only very limited assumptions are made on the model. Finally, although similar results exist in other domains, this is among the first to consider dynamic pricing problems with model uncertainty in the field of insurance.

Acknowledgement. Y. Zhang's research is supported by the Chinese Scholarship Council. Moreover, this is a collaborative effort of Arnoud den Boer, and special thanks to him.

References

den Boer, A. (2003) Dynamic Pricing and Learning. PhD thesis, VU University of Amsterdam.

IS 56
Stochastic Game Theory
Rm -1.64
Tue
11:00 - 11:30

Constant payoff in zero-sum stochastic games

Bruno Ziliotto
Paris Dauphine University
email: brunoziliotto01@gmail.com

In a zero-sum stochastic game, at each stage, two adversary players take decisions and receive a stage payoff determined by these actions and by a random variable called state of nature. The total payoff is the discounted sum of the stage payoffs. Assume that players are very patient and use optimal strategies. We then prove that at any point in the game, players get essentially the same payoff: the payoff is constant. The proof builds on a recent work by Freidlin and Wentzell about Markov chains with slow transitions.

Author index

Session index