

**XVII-th International Summer Conference
on Probability and Statistics (ISCPS-2016)**

Seminar on Statistical Data Analysis

Workshop on Branching Processes and Applications

**Conference Proceedings
and Abstracts**

Edited by E. Stoimenova and M. Slavtchova-Bojkova

Pomorie, Bulgaria, 25 June – 1 July 2016



This volume contains a collection of abstracts and short contributed papers delivered by the participants at the XVII-th International Summer Conference on Probability and Statistics held in Pomorie, Bulgaria, from 25 June to 1 July 2016.

Co-organizers of this joint event were:

- Bulgarian Statistical Society
- Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences
- Faculty of Mathematics and Informatics of the University of Sofia “St. Kliment Ohridski”

The Conference took jointly with the Seminar on Statistical Data Analysis and Workshop on Branching Processes and Applications.

©Eugenia Stoimenova, Maroussia Slavtchova-Bojkova, Editors

Publisher: *Institute of Mathematics and Informatics*
Bulgarian Academy of Sciences
Acad. G.Bontchev str., bl.8, 1113 Sofia, BULGARIA

ISBN 978-954-8986-46-5

Contents

KEYNOTE INVITED SPEAKERS	6
GONZÁLEZ, M. ,MARTÍNEZ, R. ,MOTA, M. Extinction probability of some recessive alleles of X-linked genes in the context of two-sex branching processes	7
HAYNATZKI, G. Statistical model of Microenvironment-Dependent Carcinogenesis: estimation and applications	8
LOPEZ-MIMBELA, J.A. Perpetual Integral Functionals of Brownian Motion and Blowup of Semilinear Systems	9
GONZÁLEZ, M., MINUESA, C., del PUERTO, I. Bayesian estimation on controlled branching processes via disparities: robust estimators for the offspring parameters	9
GONZÁLEZ, M., MINUESA, C., del PUERTO, I. New research lines in controlled branching processes	10
SAVOV, M. Bernstein-Gamma functions and their applications in probability theory	11
VATUTIN, V. Family trees for decomposable branching processes	12
YANEV, G. Monotone empirical Bayes estimators for the reproduction parameter of Borel-Tanner distributions	13
SHORT COMMUNICATIONS	14
BARTOSZEK, K. A punctuated stochastic model of adaptation	15
KOPANOV, P., HRISTOVA, S. Gamma distribution, random impulses in differential equations and p-moment exponential stability	19
KHARIN, Y., ZHURAK, M. Statistical inferences on binomial conditional autoregressive model of spatio-temporal data ..	24

LERI, M. Forest fire model on configuration graphs with random node degree distribution	29
MITOV, K., YANEV, N. Markov branching process with immigration	33
SLAVTCHOVA-BOJKOVA, M., TRAYANOV, P. On the effect of lifespan distribution on mutation times in continuous-time branching models	38
YAKYMIV, A. Distribution of the Order in Some Classes of Random Mappings	42
ABSTRACTS	51
ANGELOV, A. Statistical analysis of interval-censored data in questionnaire surveys	52
ARANGELOV, V., MATEEV P. Animated Map of Mathematics	52
BALL, F., GONZÁLEZ, M., MARTÍNEZ, R., SLAVTCHOVA-BOJKOVA, M. Modelling of epidemics with vaccination: stochastic monotonicity and continuity properties	54
BENCHETTAH, A. Stochastic Navier-Stokes equation on the 2D- Riemannian manifold (Sphere)	55
CHOBANOV, G. Randomness uncertainty and ergodicity in nature and society: towards a sustainable society ethics in a non-ergodic world.	55
DONCHEV, D. A noisy drift forecast and trading policies for high-frequency markets	56
DRAGIEVA, V. On the successful and blocked calls in one finite queueing system with losses	57
GACOVSKA BARANDOVSKA, A. Randomly indexed central order statistics through examples and application	57
ILIJOSKI, B. A graph based methods for clustering a morphologically similar words of ruin	58
JANEVSKI, V., KUZMANOVSKI, A., POPESKA, Z., MIHOVA, M. Using data mining techniques for recognition of Macedonian national dances	58
JORDANOVA, P. Compound power series distribution with negative multinomial summands	59
KOLEVA, D., MILEV, M. Special application of Mellin transform to contingent claims pricing	60

KOLKOVSKA, E. Penalty functions for perturbed risk processes	60
KOSTADINOVA, K., MINKOVA, L. On the bivariate compound Pólya process	61
LAZAROVA, M., MINKOVA, L Noncentral Pólya-Aeppli process and ruin probability	61
MINKOVA, L Pólya-Aeppli of order K risk model	63
NIKOLOV, N. Two-sample tank tests based on Lee distance	63
PRODANOVA, K. Stochastic model of the relationship pharmacological effect - drug plasma concentration for new Bulgarian antihypertensive agent	65
RABHI, A. Asymptotic normality of conditional distribution estimation in the single index model	65
SONIN, I. Consensus algorithms and Decomposition-Separation Theorem	66
STANEVA, A., STOIMENOVA, V. Accelerated EM algorithm and statistical inference for multitype branching processes with multivariate power series offspring distributions	67
STOIMENOVA, E., PALEJEV, D. Comparison of partially ranked lists with application to RNA-seq data	68
TCHORBADJIEFF, A. Computational assessments of multidimensional branching processes derived from cosmic rays cascades	69
TZANINIS, S. The problem of the preservation of the distribution of compound mixed renewal processes under a change of measure	70
VELEVA, E. Generation of correlation matrices	71
MITOV, K., YANEV, N. Limiting Distributions for Alternating Regenerative Branching Processes	71
ZHELYAZKOVA, M., YORDANOVA, R., SLAVTCHOVA-BOJKOVA, M. A Bayesian spatial analysis of mumps data in Bulgaria	72

KEYNOTE INVITED SPEAKERS

Extinction probability of some recessive alleles of X-linked genes in the context of two-sex branching processes

M. GONZÁLEZ, C. GUTIÉRREZ, R. MARTÍNEZ, M. MOTA
*Department of Mathematics and Instituto de Computación Científica
Avanzada of the University of Extremadura
Badajoz, SPAIN*
e-mail: mvelasco@unex.es; cgutierrez@unex.es;
rmartinez@unex.es; mota@unex.es

Some defective alleles of certain genes can cause severe diseases or serious disorders in the organisms that carry them. Some of these genes, as could be the responsible for hemophilia, are linked to X chromosome. If the alleles causing the disorder are dominant, all the carriers are affected and most of them do not reach breeding age so they are rarely detected in a population. However, recessive pernicious alleles can survive since they only affect to carrier males and homozygous carrier females (the last ones must be daughters of a carrier male, so they rarely exist). Hence, heterozygous carrier females are not affected but can pass the allele onto offspring. In this work, we study a multitype two-sex branching process for describing the evolution of the number of individuals carrying the alleles, R and r , of a gene linked to X chromosome. The R allele is considered dominant and the r allele is supposed to be recessive and defective, responsible of a disorder. In this model we study the fate of the recessive allele in the population, i.e. under which conditions it eventually disappears. We also investigate conditions for the fixation of the dominant allele and the extinction of the population. Finally, conditions for the coexistence of both alleles are conjectured.

Keywords: two-sex branching processes, extinction probability, X-linked genes.

Acknowledgements: This research has been supported by the Ministerio de Economía y Competitividad (grant MTM2015-70522-P), the Junta de Extremadura (grant GR15105) and the FEDER.

Statistical model of Microenvironment-Dependent Carcinogenesis: estimation and applications

GLEB HAYNATZKI

Biostatistics Department
UNMC-College of Public Health
Omaha, NE, USA
E-mail: ghaynatzki@unmc.edu

One important aspect of carcinogenesis is the time to emergence of the first cancerous tumor after exposure to a carcinogen, also known as tumor latency time. The extreme-value theory (EVT) approach is particularly suitable for modeling of tumor latency time, and was first proposed by M.C. Pike in 1966. We have formulated in the past a statistical parametric model that extends the EVT approach, and have derived its survivor function. Our model was the first attempt to account for the interdependence of initiated cells while they are becoming cancerous within the microenvironment of the emerging tumor. Here we fit our model to experimental, observational, and simulated data of small and larger size, and compare several estimation procedures, including ones based on the maximum-likelihood (ML), the method of moments (MOM), and simulated annealing. Also, our model is compared with other EVT-based models.

Perpetual Integral Functionals of Brownian Motion and Blowup of Semilinear Systems

JOSE ALFREDO LOPEZ-MIMBELA

*Centro de Investigacion en Matematicas
Apartado Postal 402, 36000 Guanajuato, Mexico
E-mail: jalfredo@cimat.mx*

We investigate existence of nonglobal positive solutions of a system of nonlinear heat equations perturbed by a multiplicative Gaussian noise, where the noise factor is a Wiener process $W(t)$. We obtain a lower bound for the probability of explosion in finite time of such solutions, which is given in terms of a perpetual integral functional of $W(t)$.

Bayesian estimation on controlled branching processes via disparities: robust estimators for the offspring parameters

MIGUEL GONZÁLEZ, CARMEN MINUESA, INÉS M. DEL PUERTO

*Department of Mathematics.Faculty of Sciences.
University of Extremadura
Avda. de Elvas s/n 06006 Badajoz (SPAIN)
E-mail: cminuesaa@unex.es*

Branching processes are one of the most remarkable models for the description of the dynamic of populations. Within the class of these processes, our work is focused on the controlled branching process. This stochastic model is a generalization of the classical process that presents the novelty of incorporating a random control function which determines the number of progenitors in each generation.

The strong relation between the value of the parameters of this process and its behaviour has led the recent research in this field to focus on the inferential theory. The aim of this work is to develop a methodology which provides robust estimators against outliers for the offspring parameters in a Bayesian outlook. This work presents pioneer results due to the fact that robust procedures against outliers have been barely studied in this field and no result has been obtained in a Bayesian context. Assuming that the offspring distribution belongs to a parametric family, we consider the sample given by the whole family tree and we make use of disparity measures. The proposed procedure consists of replacing the log likelihood with an appropriately scaled disparity in the expression of the posterior density function. The resulting function is called D-posterior density function and Bayesian point estimators for the offspring parameters are obtained from it. The asymptotic properties and robustness measures of these estimators are analysed. Finally, the accuracy of the method is illustrated by the way of simulated examples developed with the statistical software R.

Acknowledgements: This research has been supported by the Ministerio de Educación, Cultura y Deporte (grant FPU13/03213), the Ministerio de Economía y Competitividad (grant MTM2015-70522-P), the Junta de Extremadura (grant GR15105) and the FEDER.

New research lines in controlled branching processes

MIGUEL GONZÁLEZ, CARMEN MINUESA, INÉS M. DEL PUERTO

Department of Mathematics
Faculty of Sciences
University of Extremadura, Spain
E-mail: idelpuerto@unex.es

Controlled branching processes (CBPs) are a class of discrete-time stochastic growth population models characterized by the existence of a

random control mechanism for determining in each generation (nonoverlapping generations) how many progenitors participate in the subsequent reproduction process. Once the number of progenitors is known, each one reproduces independently of the others according to the same probability law. The novelty of adding to the branching notion a mechanism that fixes the number of progenitors in each generation allows to model a great variety of random migratory movements. It is the aim of this work to extend this kind of model to new practical situations. CBPs can be also used to model the evolution of the number of individuals of a population in which the rate of growth not only depends on the current population size but also on the distance between this size and the carrying capacity of the environment, that is the maximum population size that the environment can admit in view of its resources. Another possible extension we consider is to introduce the continuous time, that is, to define branching processes allowing random migration according to some point process and by considering the lifetime of each individual is random. A first approach to deal with these new situations will be presented.

Acknowledgements: This research has been supported by the Ministerio de Educación, Cultura y Deporte (grant FPU13/03213), Ministerio de Economía y Competitividad (grant MTM2015-70522-P), the Junta de Extremadura (grant GR15105) and the FEDER.

Bernstein-Gamma functions and their applications in probability theory

MLADEN SAVOV

*Institute of Mathematics and Informatics, BAS
ul. "Akad. Georgi Bonchev" block 8; Sofia 1113, Bulgaria
E-mail: mladensavov@math.bas.bg*

Motivated by questions in spectral theory we introduce a general class of functions (Bernstein-Gamma functions) that naturally extend the Gamma function via its recurrent equation wherein the multiplier is taken

to be any Bernstein function. We analyze in detail the analytic properties of these functions and derive all the essential quantities that characterize them as meromorphic functions. Most notably we develop a general asymptotic representation for the Bernstein-Gamma functions which is reminiscent of the celebrated Stirling asymptotic for the Gamma function. Using this information we solve more general recurrent equations which are also satisfied by the moment transform of the exponential functionals of Levy processes. Then we proceed to obtain various new and general results for the law of these random variables. For example, as a by-product, we recover a number of results on the density of the supremum of a stable Levy process.

Family trees for decomposable branching processes

VLADIMIR VATUTIN

*Steklov Mathematical Institute
Gubkin street 8, 119991, Moscow, Russia
E-mail: vatutin@mi.ras.ru*

A critical Galton-Watson branching process $(Z(1; n), \dots, Z(N; n))$ with N types of particles labelled $1, 2, \dots, N$ is considered in which a type i parent may produce individuals of types $i, i + 1, \dots, N$ only. This model may be viewed as a stochastic model for the sizes of a geographically structured population occupying N islands, the location of a particle being considered as its type. The newborn particles of island $i < N$ either stay at the same island or migrate, just after their birth to the islands $i + 1, i + 2, \dots, N$. Particles of island N do not migrate. Let $Z(i; m, n)$ be the number of type i particles existing in the process at moment $m < n$ and having nonempty number of descendants at moment n . Assuming that the initial process is critical we study the structure of the process $(Z(i; m, n), i = 1, 2, \dots, N)$ given that $Z(1; n) + \dots + Z(N; n) > 0$.

Monotone empirical Bayes estimators for the reproduction parameter of Borel-Tanner distributions

GEORGE YANEV

*School of Mathematical and Statistical Sciences
University of Texas Rio Grande Valley
1201 West University Drive, Edinburg, TX 78539, USA
E-mail: george.yanev@utrgv.edu*

The Borel-Tanner (BT) distribution arises in stochastic models including branching processes and queuing theory. Originally it was derived as the distribution of the number of customers served in a busy period of a single server queuing process, assuming Poisson arrivals and constant service time. Later it appeared in the theory of branching processes. If the number of offspring of an individual is Poisson distributed with offspring mean $0 < m < 1$, then the total progeny of a Galton-Watson branching process is a BT random variable. More recently, the distribution has been used to model a variety of real-world phenomena including coalescence models, highway traffic flows, propagation of internet viruses, cascading failures of energy systems, and herd size in finance modeling.

Our interest in estimating the parameter m stems from its role as the reproduction number of an epidemic infection modeled by a branching process. We construct a monotone version of an empirical Bayes estimator for the parameter m of the BT distribution. Properties of the estimator's regret risk are illustrated through simulations. Some open questions will be discussed.

SHORT COMMUNICATIONS

A punctuated stochastic model of adaptation

KRZYSZTOF BARTOSZEK

*Department of Mathematics, Uppsala University
Uppsala, Sweden*

e-mail: bartoszekkj@gmail.com

KB was supported by the Knut and Alice Wallenberg Foundation.

Abstract

Contemporary stochastic evolution models commonly assume gradual change for a phenotype. However the fossil record and biological theory suggests that development should be rather treated as undergoing punctuated change. In this work we present weak convergence results on for such an evolutionary model with jumps and illustrate it with simulations.

1 Introduction

Continuous trait evolution is currently modelled by stochastic differential equations (SDE) mainly in the framework of the Ornstein–Uhlenbeck (OU) process

$$dX(t) = -\alpha(X(t) - \theta(t))dt + \sigma_a dB_t, \quad (1)$$

where $\theta(t)$ can be piecewise linear. These models take into account the phylogenetic structure between the contemporary species. The trait follows the SDE along each branch of the tree (with possibly branch specific parameters). At speciation times this process divides into two independently evolving processes. However the fossil record indicates that change is not as gradual as Eq. (1) suggests. Rather that jumps occur and that the framework of Lévy processes is more appropriate.

The phylogeny is modelled by a pure birth process with speciation rate λ and conditioned on n contemporary tips. The key property of this branching process is that the time between speciation events k and $k + 1$ is exponential with parameter k . This is immediate from the memoryless property of the process and the distribution of the minimum of exponential random variables.

Regarding the phenotype I recently considered the Yule–Ornstein–Uhlenbeck with jumps (YOUj) model [3]. The trait follows an OU process but just after a branching point on the phylogeny, with a probability p_n , independently on each daughter lineage, a jump can occur. The jump is assumed to be normally distributed with mean 0 and variance $\sigma_{c,n}^2 < \infty$. Just after a speciation event at time t , independently for each daughter lineage, the trait value $X(t)$ will be

$$X^+(t) = (1 - Z)X(t^-) + Z(X(t^-) + Y). \quad (2)$$

In the above Eq. (2) $X(t^{-/+})$ means the value of $X(t)$ respectively just before and after time t , Z is a binary random variable with probability p_n of being 1 (i.e. jump occurs) and $Y \sim \mathcal{N}(0, \sigma_{c,n}^2)$.

2 Weak convergence results

Branching Ornstein–Uhlenbeck processes have been previously studied [e.g. 1, 4]. It was found that they have three asymptotic regimes depending on the relationship between the adaptation rate α and the branching rate λ . If $\alpha \geq \lambda/2$ then with appropriate scaling the contemporary sample average converges weakly to a normal distribution and behaves essentially as the mean of an i.i.d. sample. However if $0 \leq \alpha < \lambda/2$ then the process has “long memory” [“local correlations dominate over the OU’s ergodic properties”, cf. 1]. Denote $\kappa_n = 2p_n\sigma_{c,n}^2/(2p_n\sigma_{c,n}^2 + \sigma_a^2)$, $\gamma_n := (\sigma_a^2 + 2p_n\sigma_{c,n}^2)/2\alpha$, $\delta_n^* := (X_0 - \theta)/\sqrt{\gamma_n}$ and without loss of generality we take $\lambda = 1$ in what follows.

Theorem 2.1 *Let $\bar{Y}_n = (\bar{X}_n - \theta)/\sqrt{\gamma_n}$ be the normalized sample mean of the YOUj process with $\bar{Y}_0 = (\bar{X}_0 - \theta)/\sqrt{\gamma_n}$. If $\sigma_{c,n}^2\kappa_n \rightarrow 0$ then, depending on α , the process \bar{Y}_n has one of the following asymptotic (with n) types of behaviour.*

- (I) *If $0.5 < \alpha$ then $\sqrt{(n)}\bar{Y}_n$ is asymptotically normally distributed with mean 0 and variance $(2\alpha + 1)/(2\alpha - 1)$.*
- (II) *If $0.5 = \alpha$ then $\sqrt{(n/\ln n)}\bar{Y}_n$ is asymptotically normally distributed with mean 0 and variance 2.*
- (III) *If $0 < \alpha < 0.5$ then we can allow κ_n to take any value in $[0, 1]$ and $n^\alpha\bar{Y}_n$ converges almost surely and in L^2 to a random variable $Y_{\alpha, \gamma_\infty, \kappa_\infty}$ with first two moments*

$$\begin{aligned} \mathbb{E} \left[Y_{\alpha, \gamma_\infty, \kappa_\infty} \right] &= \gamma_\infty \Gamma(1 + \alpha), \\ \mathbb{E} \left[Y_{\alpha, \gamma_\infty, \kappa_\infty}^2 \right] &= \left(\gamma_\infty^2 + (1 - \kappa_\infty) \frac{4\alpha}{1 - 2\alpha} \right) \Gamma(1 + 2\alpha). \end{aligned}$$

The proof of the Central Limit Theorem 2.1, is based on a series of asymptotics, Lemma 2.1. Let \mathcal{Y}_n^* denote the σ -algebra that contains information on the Yule tree and jump pattern.

Lemma 2.1

$$\text{Var} \left[\mathbb{E} \left[e^{-2\alpha\tau^{(n)}} | \mathcal{Y}_n^* \right] \right] = \begin{cases} O(n^{-4\alpha}) & 0 < \alpha < 0.75, \\ O(n^{-4\alpha} \ln n) & \alpha = 0.75, \\ O(n^{-3}) & 0.75 < \alpha, \end{cases} \quad (3)$$

where $\tau^{(n)}$ is the time to coalescent of a randomly sampled pair of tip species [4]. In what follows C is a constant not scaling to 0 if $p \rightarrow 0$. Let J_i be a binary random variable indicating if a jump (p probability of jump) took place on the i -th (counting from the origin of the tree) speciation event of a randomly sampled lineage. For a fixed jump probability p we have

$$\text{Var} \left[\mathbb{E} \left[\sum_{i=2}^{\Upsilon^{(n)}+1} J_i e^{-2\alpha(t_{\Upsilon+1}+\dots+t_i)} | \mathcal{Y}_n^* \right] \right] \sim pC \begin{cases} n^{-4\alpha} & 0 < \alpha < 0.25 \\ n^{-1} \ln n & \alpha = 0.25 \\ n^{-1} & 0.25 < \alpha, \end{cases} \quad (4)$$

where $\Upsilon^{(n)}$ is the number of speciation events on a random lineage, t_{k+1} is the time between speciation events k and $k+1$ on a randomly chosen lineage,

$$\text{Var} \left[\mathbb{E} \left[\sum_{i=2}^{\mathfrak{v}^{(n)}+1} J_i e^{-2\alpha(\tau^{(n)}+t_{\mathfrak{v}}+\dots+t_{\mathfrak{v}+1})} | \mathcal{Y}_n^* \right] \right] \sim pC \begin{cases} n^{-4\alpha} & 0 < \alpha < 0.5, \\ n^{-2} \ln n & \alpha = 0.5, \\ n^{-2} & 0.5 < \alpha, \end{cases} \quad (5)$$

where $\mathfrak{v}^{(n)}$ is the number of common speciation events for a random pair of tips bar the splitting them event [2].

References

- [1] R. Adamczak and P. Miłoś. U–statistics of Ornstein–Uhlenbeck branching particle system. *J. Th. Probab.*, 27(4):1071–1111, 2014.
- [2] K. Bartoszek. Quantifying the effects of anagenetic and cladogenetic evolution. *Math. Biosci.*, 254:42–57, 2014.
- [3] K. Bartoszek. A central limit theorem for punctuated equilibrium. *ArXiv e-prints*, 2016.
- [4] K. Bartoszek and S. Sagitov. Phylogenetic confidence intervals for the optimal trait value. *J. App. Prob.*, 52(4) 1115–1132, 2015.

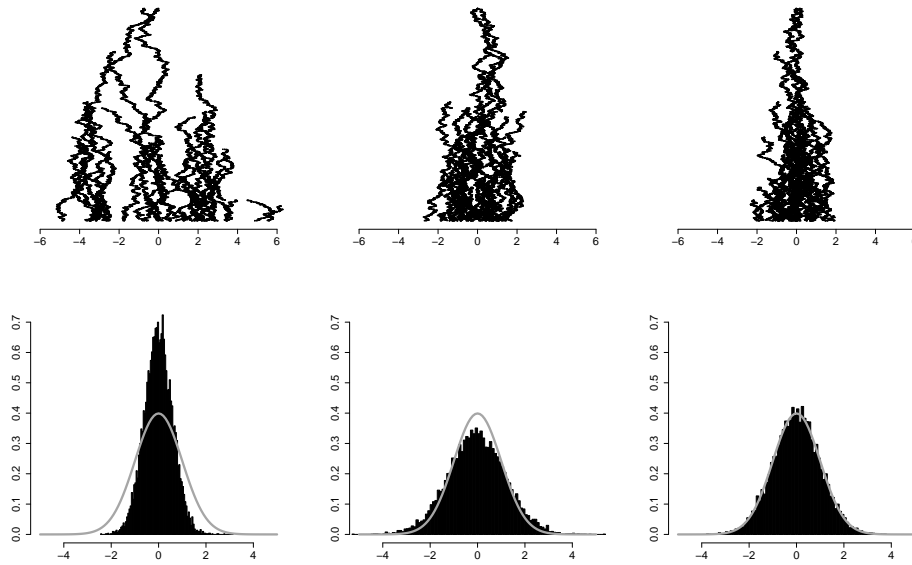


Figure 1: Left: $\alpha = 0.25$ centre: $\alpha = 0.5$ and right: $\alpha = 1$. Top row: examples of simulated YOUj process trajectories, bottom row: histograms of sample averages scaled to theoretical unit variance. In all three cases $p = 0.5$, $\sigma_c^2 = 1$, $\sigma_a^2 = 1$, $X_0 = \theta = 0$. The phylogenetic trees are pure birth trees with $\lambda = 1$ conditioned on number of tips, $n = 30$ for the trajectory plots and $n = 200$ for the histograms. The histograms are based on 10000 simulated trees. The gray curve painted on the histograms is the standard normal distribution. The simulations are done using TreeSim [7] and mvSLOUCH [5] R packages.

- [5] K. Bartoszek, J. Pienaar, P. Mostad, S. Andersson, and T. F. Hansen. A phylogenetic comparative method for studying multivariate adaptation. *J. Theor. Biol.*, 314:204–215, 2012.
- [6] F. Bokma. Time, species and separating their effects on trait variance in clades. *Syst. Biol.*, 59(5):602–607, 2010.
- [7] T. Stadler. Simulating trees with a fixed number of extant species. *Syst. Biol.*, 60(5):676–684, 2011.

Gamma distribution, random impulses in differential equations and p-moment exponential stability

P. KOPANOV*, S. HRISTOVA

Plovdiv University

Plovdiv, Bulgaria

* e-mail: pkopanov@yahoo.com

Abstract: Gamma distributed random moments of impulses in differential equation is studied. Some properties of the solutions being stochastic processes are studied and applied to be obtained sufficient conditions for p-moment exponential stability of the solutions.

Introduction

The modeling of real world phenomena in which the state of the investigated process changes instantaneously at uncertain moments lead to the application of an apparatus combining differential equations and probability theory. In the case the behavior of the studied variables are deterministically changed between two consecutive instantaneous changes at uncertain moments the adequate model is impulsive differential equations with random impulses. This type of equations is totally different than stochastic differential equation (see, for example, [3], [4]). Some considerations concerning deterministic differential equations with random impulses have been considered, for example, in [2], but there are some inaccuracies concerning properties of deterministic variables and random variables.

Random impulses in differential equations

Initially we will recall *impulsive differential equations* (IDE) with fixed points of impulses

$$\begin{aligned}x' &= f(t, x(t)) \quad \text{for } t \in (T_k, T_{k+1}], \quad k = 0, 1, 2, \dots, \\x(T_k + 0) &= I_k(x(T_k - 0)), \quad \text{for } k = 1, 2, \dots, \\x(T_0) &= x_0,\end{aligned}\tag{1}$$

where $x, x_0 \in \mathbb{R}^n$, $f : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $I_k : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and the increasing sequence of nonnegative points $\{T_k\}_{k=0}^\infty$ be given, $\lim_{k \rightarrow \infty} \{T_k\} = \infty$.

H1. $f(t, 0) = 0$ and $I_k(t, 0) = 0$ for $t \geq 0$, $k = 1, 2, \dots$.

The solution $x(t; T_0, x_0, \{T_k\})$ of IDE (1) depends not only on the initial condition (T_0, x_0) but on the moments of impulses T_k , $k = 1, 2, \dots$.

The main goal of the paper is studying stability properties of the solutions of impulsive differential equation in the case the waiting time between two consecutive impulses is Gamma distributed.

Let the probability space (Ω, \mathcal{F}, P) be given. Let $\{\tau_k\}_{k=1}^\infty$ be a sequence of random variables defined on the sample space Ω .

Assume $\sum_{k=1}^\infty \tau_k = \infty$ with probability 1.

H2. The random variables $\{\tau_k\}_{k=1}^\infty : \tau_k \in \Gamma(\alpha_k, \lambda)$ are independent where $\lambda > 0$ is the rate parameter and α_k is the shape parameter.

Define the sequence of random variables $\{\xi_k\}_{k=0}^\infty$ such that $\xi_0 = T_0$ and $\xi_k = T_0 + \sum_{i=1}^k \tau_i$, $k = 1, 2, \dots$ where $T_0 \geq 0$ is a fixed point. The random variables τ_k define the time between two consecutive impulses and ξ_n give the arrival time of n -th impulses.

Consider the initial value problem for impulsive differential equations with random moments of impulses (RIDE)

$$\begin{aligned} x'(t) &= f(t, x(t)) \quad \text{for } t \geq T_0, \quad \xi_k < t < \xi_{k+1}, \quad k = 0, 1, \dots, \\ x(\xi_k + 0) &= I_k(x(\xi_k - 0)), \quad \text{for } k = 1, 2, \dots, \\ x(T_0) &= x_0. \end{aligned} \tag{2}$$

Remark 1. The random variables $\Xi_n = \sum_{i=1}^n \tau_i \in \Gamma(\sum_{i=1}^n \alpha_i, \lambda)$.

Proposition 1. Let the condition (H1) be satisfied and the sequence of random variables $\{\Xi_k\}_{k=1}^\infty$ be such that $\Xi_n = \sum_{i=1}^n \tau_i = \xi_n - T_0$, $n = 1, 2, \dots$.

Then $\Xi_n \in \Gamma(\sum_{i=1}^n \alpha_i, \lambda)$ are continuous dependent random variables with cdf $F(t; \sum_{i=1}^n \alpha_i, \lambda) = P(\Xi_n < t) = \frac{\gamma(\sum_{i=1}^n \alpha_i, \lambda t)}{\Gamma(\sum_{i=1}^n \alpha_i)}$ for $t \geq 0$, respectively, where $\gamma(m, t) = \int_0^t x^{m-1} e^{-x} dx$, $m, t > 0$ is the lower incomplete gamma function.

Let the point t_k be an arbitrary value of τ_k , $k = 1, 2, \dots$. Define the increasing sequence of points $T_k = T_0 + \sum_{i=1}^k t_i$, $k = 1, 2, 3, \dots$ that are values of the random variables ξ_k and consider (1).

Definition 1. The solution $x(t; T_0, x_0, \{T_k\})$ of the IVP for the IDE with fixed points of impulses (1) is called a sample path solution of the IVP for the RIDE (2).

Definition 2. A stochastic process $x(t; T_0, x_0, \{\tau_k\})$ with an uncountable state space \mathbb{R}^n is said to be a solution of the IVP for the system of RIDE (2) if for any values t_k of the random variable $\tau_k, k = 1, 2, 3, \dots$ and $T_k = T_0 + \sum_{i=1}^k t_i, k = 1, 2, \dots$ the corresponding function $x(t; T_0, x_0, \{T_k\})$ is a sample path solution of the IVP for RIDE (2).

Gamma distributed moments of impulses

For any $t \geq T_0$ consider the events $S_k(t) = \{\omega \in \Omega : \xi_k(\omega) < t < \xi_{k+1}(\omega)\}$.

Lemma 0.0.1. Let the condition (H2) be satisfied. Then the probability that there will be exactly k impulses until time $t, t \geq T_0$ is given by

$$P(S_k(t)) = \frac{\gamma(\sum_{i=1}^k \alpha_i, (t - T_0)\lambda)}{\Gamma(\sum_{i=1}^k \alpha_i)} - \frac{\gamma(\sum_{i=1}^{k+1} \alpha_i, (t - T_0)\lambda)}{\Gamma(\sum_{i=1}^{k+1} \alpha_i)}.$$

Lemma 0.0.2. Let the condition (H1) be satisfied and $m_i + \lambda \geq m_k$ for all natural $i, k : i < k$. Then the solution of the IVP for the linear RIDE (2) is

$$u(t; T_0, u_0, \{\tau_k\}) = \begin{cases} u_0 e^{-m_0(t-T_0)} & \text{for } T_0 < t < \tau_1 \\ u_0 \left(\prod_{i=1}^k b_i \right) e^{-\sum_{i=1}^k m_{i-1} \tau_i} e^{-m_k(t-\xi_k)} & \text{for } \xi_k < t < \xi_{k+1} \end{cases}$$

and the expected value of the solution satisfies the inequality

$$\begin{aligned} E(|u(t; T_0, u_0, \{\tau_k\})|) &\leq |u_0| \left[e^{-m_0(t-T_0)} \left(1 - \frac{\gamma(\alpha_1, (t - T_0)\lambda)}{\Gamma(\alpha_1)} \right) \right. \\ &+ \sum_{k=1}^{\infty} e^{-m_k(t-T_0)} \prod_{i=1}^k \left(|b_i| \left(\frac{\lambda}{m_{i-1} - m_k + \lambda} \right)^{\alpha_i} \right) \times \\ &\left. \left(\frac{\gamma(\sum_{i=1}^k \alpha_i, (t - T_0)\lambda)}{\Gamma(\sum_{i=1}^k \alpha_i)} - \frac{\gamma(\sum_{i=1}^{k+1} \alpha_i, (t - T_0)\lambda)}{\Gamma(\sum_{i=1}^{k+1} \alpha_i)} \right) \right]. \end{aligned} \quad (3)$$

Remark 2. The estimate (3) is satisfied if $\{m_i\}_{i=0}^{\infty}$ is a nonincreasing sequence bounded from below by $M > 0$ and there exists a number $B \geq 1$ such that $\prod_{i=1}^k |b_i| \leq B$, $k = 1, 2, \dots$. Then from (3) it follows $E(|u(t; T_0, u_0, \{\tau_k\})|) \leq |u_0| B e^{-M(t-T_0)}$.

Remark 3. The estimate (3) is satisfied if $m_i + \lambda \geq m_k$ for all natural $i, k : i < k$, $m_k \geq M > 0$, $k = 0, 1, 2, \dots$ and $\prod_{i=1}^k |b_i| \left(\frac{\lambda}{m_{i-1} - m_k + \lambda}\right)^{\alpha_i} \leq B$ for $k = 1, 2, \dots$ where $B \geq 1$.

p-exponential stability of RIDE

Theorem 0.0.1. Let the following conditions be satisfied:

1. Conditions (H1), (H2) hold.
2. The function $V \in \Lambda(\mathbb{R}_+, \mathbb{R}^n)$ and there exist positive constants a, b such that

- (i) $a\|x\|^p \leq V(t, x) \leq b\|x\|^p$ for $t \in \mathbb{R}_+$, $x \in \mathbb{R}^n$;
- (ii) there exists a constant $m \geq 0$ such that $D_{(1)}^+ V(t, x) \leq -mV(t, x)$, $t \in \mathbb{R}_+$, $x \in \mathbb{R}^n$;
- (iii) for any $k = 1, 2, \dots$ there exist functions $w_k \in C(\mathbb{R}_+, \mathbb{R}_+)$ and constants $C_k > 0$, $w_k(t) \leq C_k$ for $t \geq 0$ such that $\sum_{k=0}^{\infty} \prod_{i=1}^k C_i = C < \infty$ and $V(t, I_k(x)) \leq w_k(t)V(t, x)$, $t \in \mathbb{R}_+$, $x \in \mathbb{R}^n$.

Then the trivial solution of the RIDE (2) is p -moment exponentially stable.

Remark 4. If $\alpha_k = 1$ for all k , i.e. $\tau_k \in \text{Exp}(\lambda)$, the p -moment exponential stability is studied in [1].

Acknowledgments. Research was partially supported by Fund Scientific Research MU15FMIIIT008, Plovdiv University.

References

- [1] Agarwal R., Hristova S., O'Regan D. (2013). Exponential stability for differential equations with random impulses at random times. *Adv. Difference Equ.* Vol. **2013**, pp. 303-307.
- [2] Anguraj A., Vinodkumar A. (2010). Existence, uniqueness and stability results of random impulsive semilinear differential systems. *Nonlinear Anal. Hybrid Syst.* Vol. **3**, pp. 475-483.

- [3] Wu H., Sun J. (2006). p-moment stability of stochastic differential equations with impulsive sump and Markovian switching. *Automatica*. Vol. **42**, pp. 1753-1759.
- [4] Yang J., Zhong S., Luo W. (2008). Mean square stability analysis of impulsive stochastic differential equations with delays, *J. Comput. Appl. Math.*. Vol. **216**,2, pp. 474-483.

Statistical inferences on binomial conditional autoregressive model of spatio-temporal data

YURIY KHARIN, MARYNA ZHURAK

Belarusian State University

Minsk, Belarus

E-mail: kharin@bsu.by; mzhurak@gmail.com

Abstract: Binomial conditional autoregressive model of spatio-temporal data is presented. Asymptotic properties of the maximum likelihood estimators of parameters for the binomial conditionally autoregressive model of spatio-temporal data are studied: asymptotic normality is proved and the asymptotic covariance matrix is found for the estimators. Results of computer experiments on simulated and real data are given.

Introduction

Studying the probabilistic models of spatio-temporal data is a new topical scientific direction. Statistical analysis and modeling of spatio-temporal data is a challenging task [1]-[5].

Models based on spatio-temporal data become widely used for solving practical problems in meteorology, ecology, economics, medicine and other fields. In [5] spatio-temporal model is used to analyse daily precipitation for 71 meteorological stations over 60 years in Austria. Bayesian spatio-temporal model is applied to predict cancer cases in [1].

Binomial conditional autoregressive model

Introduce the notation: (Ω, F, P) is the probability space; \mathbf{N} is set of positive integers; \mathbf{Z} is set of integers; $S = \{1, 2, \dots, n\}$ is the set of indexed spatial regions or space locations (let us agree to call them sites), into which the analyzed spatial area is partitioned; n is number of sites; $t \in \mathbf{Z}$ is discrete time; T is the length of observation period; $x_{s,t} \in A = \{0, \dots, N\}$ is a discrete random variable at time t at site s ;

$F_{<t} = \sigma\{x_{u,\tau} : u \in S, \tau \leq t-1\} \subset F$ is the σ -algebra generated by the indicated in braces random variables; $z_{j,t} \in R^1$, ($j = 1, \dots, m$) is an observed (known) level of the j -th exogenous factor at time t which influences $x_{s,t}$; $L(\xi)$ means the probability distribution law of random variable ξ ; $\mathbf{E}\{\cdot\}$, $\mathbf{D}\{\cdot\}$, $\mathbf{cov}\{\cdot\}$ are symbols of expectation, variance, covariance of random variables; $Bi(\cdot; N, p)$ is the binomial probability distribution law of random variable ξ with the parameters $N \in \mathbf{N}$, $0 \leq p \leq 1$:

$$P\{\xi = l\} = Bi(l; N, p) ::= C_N^l p^l (1-p)^{N-l}, l \in A; L\{\xi\} = Bi(\cdot; N; p). \quad (1)$$

We construct the binomial conditional autoregressive model for spatio-temporal data $\{x_{s,t}\}$, as in [3], [4]. Provided that prehistory $\{x_{s,\tau} : s \in S, \tau \leq t-1\}$ is fixed, random variables $x_{1,t}, \dots, x_{n,t}$ are assumed to be conditionally independent and

$$L\{x_{s,t} | F_{<t}\} = Bi(\cdot; N; p_{s,t}), \quad (2)$$

$$\ln \frac{p_{s,t}}{1-p_{s,t}} = \sum_{i=1}^n a_{s,i} x_{i,t-1} + \sum_{j=1}^m b_{s,j} z_{j,t}, t \in \mathbf{Z}, s \in S, \quad (3)$$

where $a_s = (a_{s,1}, \dots, a_{s,n})' \in R^n$, $b_s = (b_{s,1}, \dots, b_{s,m})' \in R^m$, $s \in S$, $\theta_s = (a'_s, b'_s)' \in R^{n+m}$, $\theta = (\theta'_1, \dots, \theta'_n)' \in R^{n(n+m)}$ is the composed vector of the parameters of the model; $p_{s,t}$ can be calculated as follows:

$$p_{s,t} = p_s(X_{t-1}, Z_t) ::= \exp\{\theta'_s Y_t\} \left(1 + \exp\{\theta'_s Y_t\}\right)^{-1}, s \in S, t \in \mathbf{Z}, \quad (4)$$

where $Z_t = (z_{1,t}, z_{2,t}, \dots, z_{m,t})' \in R^m$ is the column vector specifying exogenous factors at time t ; $X_t = (x_{1,t}, x_{2,t}, \dots, x_{n,t})' \in R^n$ is the column vector specifying the time slice of the process under consideration at time $t \in \mathbf{Z}$; $Y_t = (X'_{t-1}, Z'_t)' \in R^{n+m}$, $t \in \mathbf{Z}$.

Let $L = \{l_j = (l_{1,j}, \dots, l_{n,j})' \in A^n : j = 1, 2, \dots, (N+1)^n\}$ be the ordered set of all admissible values of the vector X_t ; $|L| = \nu = (N+1)^n$.

Theorem 1. For the model (1)-(3) the observed vector process X_t is the n -dimensional nonhomogeneous Markov chain with the finite state space L and the one-step transition probability matrix $Q(t) = (q_{I,J}(\theta, t))$, $I = (I_s), J = (J_s) \in L$:

$$q_{I,J}(\theta, t) = \prod_{s=1}^n \frac{C_N^{J_s} (\exp\{a'_s I + b'_s Z_{t-1}\})^{J_s}}{(1 + \exp\{a'_s I + b'_s Z_{t-1}\})^N}. \quad (5)$$

Corollary. Under conditions of Theorem 1, if vector of exogenous factors $Z_t = Z = (z_1, \dots, z_m)' \in R^m$ does not depend on t , then the one-step transition probability matrix does not depend on t , and Markov X_t chain is homogeneous :

$$Q = (q_{I,J}(\theta)) \in [0, 1]^{\nu \times \nu}, \quad I, J \in L, \quad (6)$$

$$q_{I,J}(\theta) = \prod_{s=1}^n C_N^{J_s} \left(\exp \left\{ a'_s I + b'_s Z \right\} \right)^{J_s} \left(1 + \exp \left\{ a'_s I + b'_s Z \right\} \right)^{-N}, \quad (7)$$

the n -dimensional Markov chain is ergodic and single stationary distribution $\pi = (\pi_I) \in [0, 1]^\nu$ exists:

$$Q' \pi = \pi, \quad \sum_{I \in A^n} \pi_I = 1.$$

Lemma 1. For the model (1)-(3) the observed vector process X_t is nondegenerate process for any finite coefficients values $\{\theta_s\}$ and finite $\{z_{i,t}\}$, that means in (4):

$$0 < p_{s,t} < 1, \quad s \in S, t \in \mathbf{Z}.$$

Lemma 2. For the model (1)-(3) in case of any finite coefficients values $\{\theta_s\}$ and finite $\{z_{i,t}\}$ the covariance matrix $\mathbf{cov}\{X_t, X_t\}$ is positively defined and takes the form:

$$\mathbf{cov}\{X_t, X_t\} = N \mathit{diag} \{p_i(X_{t-1}, Z_t)(1 - p_i(X_{t-1}, Z_t))\} + D \in R^{n \times n},$$

$$D = (d_{ij}) \in R^{n \times n}, d_{ij} = N^2 \mathbf{cov} \left\{ (1 + \exp(-\theta_i Y_t))^{-1}, (1 + \exp(-\theta_j Y_t))^{-1} \right\}.$$

Statistical estimation of parameters

Theorem 2. The log-likelihood function for the model (1)-(3) under the observed spatio-temporal data $\{X_t : t = 1, 2, \dots, T\}$ takes the additive form:

$$l(\theta) = \sum_{s=1}^n l_s(\theta_s), l_s(\theta_s) = \sum_{t=1}^T \left(x_{s,t} \theta'_s Y_t - N \ln \left(1 + \exp \left\{ \theta'_s Y_t \right\} \right) + \ln C_N^{x_{s,t}} \right). \quad (8)$$

To find the maximum likelihood estimators (MLE) $\{\hat{\theta}_s\}$ of the parameters of the model we need to maximize the log-likelihood function (8):

$$l(\theta) \rightarrow \max_{\theta \in R^{n(n+m)}} . \quad (9)$$

Theorem 3. In case of the model (1)-(3), if $m = 1$, $z_{1t} = z \neq 0$ does not depend on t and Markov chain $X_t \in L$ is stationary, then for any finite coefficients values $\{\theta_s\}$ and finite $z \in R^1$ the Fisher information matrix is nonsingular block-diagonal matrix ($Y_t = (X'_{t-1}, z)'$):

$$G = N \text{diag} \left\{ \mathbf{E} \left\{ Y_t Y_t' p_i(X_{t-1}, z) (1 - p_i(X_{t-1}, z)) \right\} \right\}, i = 1, \dots, n. \quad (10)$$

Theorem 4. Under Theorem 3 conditions, if $T \rightarrow +\infty$ the constructed by (9) maximum likelihood estimators $\{\hat{\theta}_s\}$ are consistent and asymptotically normally distributed:

$$L \left\{ \sqrt{T}(\hat{\theta} - \theta^0) \right\} \rightarrow N_{n(n+1)}(0, G^{-1}).$$

where G is determined by (10).

Theorems 2-4 are used to construct statistical tests on the values of true unknown parameters $\{\theta_s^0\}$ and statistical forecasting statistics for future values $\{x_{t,s} : t = T + 1, \dots, T + \tau\}$.

Results of computer experiments on simulated and real data illustrate the theoretical results.

References

- [1] Kang S.Y., McGreed J., Baade P., Mengersen K. (2015). Case Study for Modelling Cancer Incidence Using Bayesian Spatio-Temporal Models *Australian and New Zealand Journal of Statistics.* Vol. **3**, pp. 325-345.
- [2] Kharin Yu.S., (2013). Robustness in Statistical Forecasting. *Springer: N.Y.* 356 P.
- [3] Kharin Yu.S., Zhurak M.K. (2015). Statistical Analysis of Spatio-Temporal Data Based on Poisson Conditional Autoregressive Model *INFORMATICA*. Vol. **26-1**, pp. 67-87.

- [4] Mariella L., Tarantino M. (2010). Spatial temporal conditional Auto-Regressive Model: A New Autoregressive Matrix *Austrian Journal of Statistics*. Vol. **39**, pp. 223-244.
- [5] Umlauf N., Mayr G., Messner J., Zeileis A. (2012). Why Does It Rain on Me? A Spatio-Temporal Analysis of Precipitation in Austria *Austrian Journal of Statistics*. Vol. **41**, pp. 81-92.

Forest fire model on configuration graphs with random node degree distribution

MARINA LERI

*Institute of Applied Mathematical Research
of the Karelian Research Centre RAS
Petrozavodsk, RUSSIA
e-mail: leri@krc.karelia.ru*

Abstract: We consider configuration graphs with node degrees being i.i.d. random variables following the power-law distribution. The distribution parameter is a random variable following the uniform distribution on a predefined interval.

We consider a destructive process which could be interpreted as a fire spreading over the graph links, and could be used for modeling forest fires as well as banking system defaults. This process is often referred to as a “forest fire model”. The probability of fire transfer over a graph link either possess a predefined value and is fixed for all the graph links or is a random variable following the standard uniform distribution. By computer simulation we study the robustness of such graphs from a viewpoint of node survival in the two cases of starting a fire propagation process: the “random ignition” and the “targeted lightning-up”. The results on finding the optimal interval of the node degree distribution parameter that would ensure maximum survival of trees in case of a fire are presented. A comparative analysis of various graph models in terms of their robustness to various fire propagation processes was performed.

Power-law random graph

We consider random graphs the number of nodes in which is equal to N . Node degrees $\xi_1, \xi_2, \dots, \xi_N$ are independent identically distributed random variables drawn from the following power-law distribution:

$$\mathbf{P}\{\xi \geq k\} = k^{-\tau}, \quad k = 1, 2, \dots, \tau > 1. \quad (1)$$

Node degrees form stubs, numbered in an arbitrary order. Then all stubs are joined one to another equiprobably to form links. If the sum of

node degrees appears to be odd, one stub is added to a randomly chosen node to form a lacking connection. In our work we consider power-law graphs with the parameter τ of distribution (1) following the uniform distribution on a predefined interval $[a, b]$. The studies of the power-law random graphs has been attracting consistent interest because of the wide use of these models for representing massive data networks (see e.g. [3, 4, 6]).

Forest fires

An interesting trend in the studies of random graphs is the aspect connected to the transmission of various destructive influences through graph links. In particular, we consider the process which could be interpreted as spreading of a fire over the graph links [2], and might have some other applications, for example, modeling of banking system defaults [1].

To study a forest fire model we view graph nodes as trees growing in a certain area of a real forest. Since we consider this area to be limited, the number of trees on it also has to be limited. For this purpose we proposed to use auxiliary square lattice graphs sized 100×100 (see [5]). If a real fire can be transferred from one tree to another, corresponding graph nodes are connected by a link. Thus, under a full packing every inner auxiliary graph node has 8 adjacent neighbours. Knowing that $i = \zeta(\tau)$ (where $\zeta(x)$ is the Riemann zeta function), we derived a regression relation between graph size $N \leq 10000$ and the parameter τ of node degree distribution (1):

$$N = 9256\tau^{-1.05}, \quad R^2 = 0.97. \quad (2)$$

The relation (2), where $\tau = \bar{\tau} = \frac{a+b}{2}$ confines the number of nodes in the graph N . Thus, we study a fire propagation process on power-law configuration graphs with N nodes, where N is calculated from (2) for various intervals $[a, b]$. The probability of fire transition p over the graph links is either a predefined value and is fixed for all the graph links or is a random variable following the standard uniform distribution. The purpose is to find the optimal interval for the node degree distribution parameter τ that would ensure maximum survival of graph nodes in case of a fire.

In our work we consider the three intervals $[a, b]$: $(1, 2]$, $[2, 3]$, $(1, 3]$ with the N values obtained from (2) being equal to 6046, 3536 and 4470, respectively. The two considered cases of starting a fire propagation process are the “targeted lightning-up” when the fire starts from the node with the highest degree and the “random ignition” starting from an equiprobably chosen node. For both cases we found the relations between the average number of trees surviving in a fire n and the probability of fire transition p . Figures 1 and 2 show how the percentage of nodes $\frac{n}{N}$ in % depends on the probability p for the two cases of starting a fire.

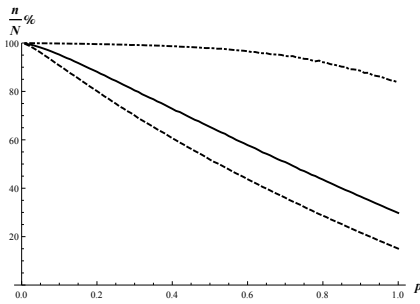


Fig.1 Relation between $\frac{n}{N}$ % and the probability of fire transition p (“targeted lightning-up”).

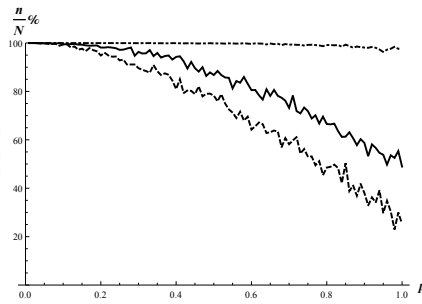


Fig.2 Relation between $\frac{n}{N}$ % and the probability of fire transition p (“random ignition”).

Obviously, the number of remaining nodes n as well as the proportion $\frac{n}{N}$ is decreasing with the increase of p . The results show that graphs with the parameter $\tau \in [2, 3]$ are more fire resistant, whereas graphs with $\tau \in (1, 2]$ are the most vulnerable to the fire in both fire start cases.

Furthermore, we study a case when the probability of fire transition p is a random variable drawn from the standard uniform distribution. The obtained results are presented in the following Table 1.

Table 1. An average proportion $\frac{n}{N}$ % of nodes surviving in a fire.

$\tau \in [a, b]$	“targeted lightning-up”	“random ignition”
$(1, 2]$	52.03	81.6
$[2, 3]$	97.96	99.92
$(1, 3]$	65.26	87.85

The results show a similar picture where in both cases of fire ignition graphs with the node degree distribution parameter $\tau \in [2, 3]$ will ensure a better survival of nodes in case of a fire.

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

References

- [1] Arinaminparty N., Kapadia S., May R. (2012). Size and complexity model financial systems. *Proceedings of the National Academy of Sciences of the USA*. Vol. **109**, pp. 18338-18343.
- [2] Bertoin J. (2012). Fires on trees. *Annales de l'Institut Henri Poincare Probabilites et Statistiques*. Vol. **48**, N **4**, pp. 909-921.
- [3] Durrett R. (2007). *Random Graph Dynamics*. Cambridge Univ. Press, Cambridge.
- [4] Faloutsos C., Faloutsos P., Faloutsos M. (1999). On power-law relationships of the Internet topology. *Computer Communications Rev.*. Vol. **29**, pp. 251-262.
- [5] Leri M., Pavlov Y. (2014). Power-law random graphs' robustness: link saving and forest fire model. *Austrian Journal of Statistics*. Vol. **43**, N **4**, pp. 229-236.
- [6] Reittu H., Norros I. (2004). On the power-law random graph model of massive data networks. *Performance Evaluation*, Vol. **55**, pp. 3-23.

Markov branching process with immigration

KOSTO V. MITOV, NIKOLAY M. YANEV

National Military University "Vasil Levski" - Aviation Faculty

D. Mitorpolia, Pleven, BULGARIA

e-mail: kmitov@yahoo.com

Institute of Mathematics and Informatics - BAS,

Sofia, BULGARIA

e-mail: yanev@math.bas.bg

Abstract: In the talk we consider critical Markov branching process with infinite variance and non-homogeneous Poisson immigration. The asymptotic of the probability for non extinction and limit theorems are proved in a particular case when the probability generating function of the process without immigration has an explicit representation.

Introduction and definitions

A Markov branching process can be described as follows: The particles of a given type evolve in a broad-media. Each particle, independently of the others, lives random time τ with exponential distribution function $G(t) = \mathbf{P}\{\tau \leq t\} = 1 - e^{-t/\mu}$, $t \geq 0$ and at the end of its life the particle produces random number $\xi \geq 0$ of new particles of the same type. Each of them evolves in the same way.

If we assume that this evolution started at time $t = 0$ with one new particle, then the number of particles at every moment $t \geq 0$, $Z(t)$ form the stochastic process, known as Markov branching process (see [1]).

Let us suppose that along the Markov branching process $Z(t)$ there is a sequence of random vectors (S_k, I_k) , $k = 0, 1, 2, \dots$, independent of $Z(t)$, where

$$0 = S_0 < S_1 < S_2 < S_3 < \dots$$

are the jump points of a non-homogeneous Poisson process $\nu(t)$ independent of $Z(t)$. Denote by $r(t)$ the intensity of $\nu(t)$ and by $R(t) = \int_0^t r(u)du$. Assume that at every jump-point S_k , I_k new particles immigrate into the

process $Z(t)$ and they participate in the evolution as the other particles. Let us denote the new process by $Y(t)$. It can be represent as follows

$$Y(t) = \begin{cases} \sum_{k=1}^{\nu(t)} Z_{(k)}(t - S_k), & \nu(t) > 0; \\ Y(t) = 0, & \nu(t) = 0, \end{cases}$$

where $Z_{(k)}(t)$ have the same branching mechanism as $Z(t)$ but they started with random number of ancestors at random time, i.e. $Z_{(k)}(t - S_k)$ starts with I_k ancestors at the moment S_k . The process $Y(t)$, $t \geq 0$ is a Markov branching process with non-homogeneous Poisson immigration (MBPNPI).

In the talk we assume that $I_k = 1$ almost surely.

Assumptions and results

We assume that the probability generating function of the offspring of one particle is $h(s) = Es^\xi = s + (1 - s)^{1+\gamma}$, $0 < \gamma < 1$ and $G(t) = 1 - e^{-t/\mu}$, $t \geq 0$. Then the probability generating function of the the process $Z(t)$, $t \geq 0$ has the following form

$$F(t; s) = 1 - \frac{1 - s}{(1 + ct(1 - s)^\gamma)^{1/\gamma}}, \quad t > 0, \quad s \in [0, 1]. \quad (1)$$

where $\gamma \in (0, 1)$ and $c = \frac{\gamma}{(1 + \gamma)\mu} \in (0, \infty)$. See Sevastyanov's book [1] or the recent paper of Sagitov and Lindo [2]).

Under these assumptions the process is *critical*, that is the number of the offspring of one particle is a random variable *with mean one and infinite variance*.

Such processes in general settings were first studied by V.M.Zolotarev [4].

The probability generating function of the process $Y(t)$ has the following form (see Yakovlev and Yanev[3]),

$$\Phi(t; s) := E(s^{Y(t)}) = \exp \left\{ - \int_0^t \frac{(1 - s)r(t - u)}{(1 + cu(1 - s)^\gamma)^{1/\gamma}} du \right\}. \quad (2)$$

For the intensity of the Poisson process we assume one of the following

conditions

$$r(t) \downarrow 0, r(t) = t^{-\theta} L_R(t), t \rightarrow \infty, \theta > 0, \quad (3)$$

$$r(t) \uparrow r > 0, \quad t \rightarrow \infty, \quad (4)$$

$$r(t) \uparrow \infty, \quad r(t) = t^\theta L_R(t), \quad t \rightarrow \infty, \theta > 0, \quad (5)$$

where $L_R(\cdot)$ is a slowly varying function at infinity.

From (1) we get

$$Q(t) = 1 - F(t; 0) = \frac{1}{(1 + ct)^{1/\gamma}}, \quad t \geq 0.$$

It is not difficult to calculate that $Q = \int_0^\infty Q(t) dt = \frac{\gamma}{c(1-\gamma)} \in (0, \infty)$. Therefore, $Q(t)/Q, t \geq 0$ is a density on $[0, \infty)$.

Depending on the rate of change of $r(t)$ we obtain the following results:

1. Assume that (1) and (3) with $\theta > 1$ hold.

In this case $R = \lim_{t \rightarrow \infty} R(t) = \int_0^\infty r(u) du \in (0, \infty)$ and $r(t)/R$ is also density on $[0, \infty)$. Then we have

$$P(Y(t) > 0) \sim \frac{\gamma}{c(1-\gamma)} r(t) + R \times (1 + ct)^{-1/\gamma}, \quad t \rightarrow \infty.$$

Depending on the asymptotic of $Q(t)/r(t)$ we have three different limit distributions.

1.A. If $Q(t) = \frac{1}{(1 + ct)^{1/\gamma}} = o(r(t)), \quad t \rightarrow \infty$ then

$$\lim_{t \rightarrow \infty} P(Y(t) = k | Y(t) > 0) = p_k, \quad k = 1, 2, \dots,$$

$$\sum_{k=1}^{\infty} p_k s^k = 1 - (1 - s)^{1-\gamma}.$$

1.B. If $r(t) = o(Q(t)), \quad t \rightarrow \infty$ then

$$\lim_{t \rightarrow \infty} P(Y(t)Q(t) \leq x | Y(t) > 0) = D_\gamma(x), \quad x \geq 0.$$

where $D_\gamma(x)$ has Laplace transform $\hat{D}_\gamma(\lambda) = 1 - \frac{\lambda}{(1+\lambda^\gamma)^{1/\gamma}}, \quad \lambda > 0$.

1.C. If $r(t)/Q(t) \rightarrow d \in (0, \infty)$ then

$$\lim_{t \rightarrow \infty} P(Y(t) = k | Y(t) > 0) = p_k,$$

$$\sum_{k=1}^{\infty} p_k s^k = \frac{Q}{Q + R/d} (1 - (1 - s)^{1-\gamma}),$$

and

$$\lim_{t \rightarrow \infty} P(Y(t)Q(t) \leq x | Y(t) > 0) = \frac{Qd}{R + Qd} + \frac{R}{R + Qd} D_{\gamma}(x), \quad x \geq 0.$$

2. Assume that (1) and (3) with $\theta \in (0, 1)$ hold. Then

$$P(Y(t) > 0) \sim \frac{\gamma}{c(1 - \gamma)} r(t), \quad t \rightarrow \infty.$$

$$\lim_{t \rightarrow \infty} P(Y(t) = k | Y(t) > 0) = p_k, \quad k = 1, 2, \dots,$$

$$\sum_{k=1}^{\infty} p_k s^k = 1 - (1 - s)^{1-\gamma}.$$

3. Assume that (1) and (4) hold. Then

$$P(Y(t) > 0) \rightarrow 1 - e^{-\frac{r\gamma}{c(1-\gamma)}}, \quad t \rightarrow \infty.$$

$$\lim_{t \rightarrow \infty} P(Y(t) = k) = p_k, \quad k = 1, 2, \dots,$$

$$\sum_{k=1}^{\infty} p_k s^k = \exp\left(-r \frac{\gamma(1 - s)^{1-\gamma}}{c(1 - \gamma)}\right),$$

4. Assume that (1) and (5) hold. Then

$$P(Y(t) > 0) \rightarrow 1, \quad t \rightarrow \infty.$$

Acknowledgment. The research was supported by the NFSR of the Min. Edu. Sci. of Bulgaria, grant No. DFNI-I02/17

References

- [1] Sevastyanov B.A. (1971). *Branching processes*. Nauka, Moscow.
- [2] Sagitov S., Lindo A. (2015). A special family of Galton-Watson processes with explosions. *J. Appl. Prob.* Vol. **xx**, pp. P-P.
- [3] Yakovlev, A. Y., Yanev, N. M. (2006). Branching stochastic processes with immigration in analysis of renewing cell populations. *Math. Biosci.* Vol. **203**, pp. 37-63.
- [4] Zolotarev V.M. (1958). More exact statement of several limit theorems in the theory of branching processes. *Theor. Prob. Appl.* Vol. **2**, pp. 256-266.

On the effect of lifespan distribution on mutation times in continuous-time branching models

MAROUSSIA SLAVTCHOVA-BOJKOVA, PLAMEN TRAYANOV

Faculty of Mathematics and Informatics

St. Kl. Ohridski University of Sofia

Sofia, Bulgaria

e-mail: `bojkova@fmi.uni-sofia.bg`

Abstract: Recently branching stochastic models have been intensively used as models of mutations in cancer research (see [1]). On the other side, it turned out that the lifespan of a cell is not deterministic but random by nature (see [2, 3]). This is our motivation of considering the cellular life time as a continuous random variable. That is why we choose the so-called Bellman–Harris branching process (BHBP) or age–dependent BP as a model in which every cell has a continuously distributed lifespan and at the end of its life it reproduces or dies.

More precisely, we consider two–type BHBP in which subcritical type cells could mutate into supercritical cells. Our conjecture is that the time to appearance of a mutant depends strongly on the chosen lifespan distribution of the cells. We study this relationship in terms of BHBP and find out the most probable times for these mutants to appear in different model settings. Several types of lifespan distributions are considered, including non–exponential, and a possible application of the results in medicine is discussed.

Introduction

We consider two types of cells – type 1, which is subcritical, and type 0, which is supercritical. The type 0 cells produce cells of both types at the end of their life, but the cells of type 0 produce only cells of the same type. If the BP is subcritical, then we know the cells will extinct with probability 1, but if it is supercritical there is a positive probability that the cells will never extinct. We will call the supercritical type 0 cells - mutants. We will call “successful mutant” the first cell of type 0 that leads to exponential growth, i.e. it escapes extinction. The

waiting times for appearance of mutants and “successful mutants” are of interest in medicine where they model cancer cell development and possible appearance of metastasis. This paper studies the distribution of these waiting times and their dependence on the type and parameters of cellular lifespan distribution.

Preliminary theoretical results

The theorems presented in this section are the key results needed to calculate the distribution of the time to appearance of a “successful mutant” and the corresponding hazard function. Let the random variable T denotes the time to appearance of “successful mutant”, q_0 denotes the probability the mutant type to extinct, u denotes the probability for a mutation to occur in the offspring of type 1, $G_1(t)$ be the distribution of the cellular lifespan and f_1 - the probability generating function (p.g.f.) of type 1 cells.

Theorem 1. (see Slavtchova–Bojkova [4])

The distribution of T satisfies the following integral equation:

$$\mathbb{P}(T > t) = 1 - G_1(t) + \int_0^t f_1(uq_0 + (1 - u)\mathbb{P}(T > t - y)) dG_1(y). \quad (1)$$

Remark. Notice the solution of equation (1) neither depends on the life length distribution of mutant type cells, nor of the particular form of the offspring generating function of mutant type cells. It depends only on the probability of extinction of mutant cells - q_0 .

Theorem 2. *The joint probability that successful mutant has not been born yet and there are not cells of type 1 (with subcritical reproduction, $m_1 < 1$) satisfies the following integral equation:*

$$\mathbb{P}(T > t, Z^1(t) = 0) = \int_0^t f_1(uq_0 + (1 - u)\mathbb{P}(T > t - y, Z^1(t - y) = 0)) dG_1(y). \quad (2)$$

By definition the hazard function $g(t)$ is defined by $g(t)dt = \mathbb{P}(T \in [t, t + dt] | T > t, Z^1(t) > 0)$, which can be written also in the form

$$g(t) = \frac{F'_T(t)}{\mathbb{P}(T > t, Z^1(t) > 0)}, \quad (3)$$

where $F'_T(t)$ is the probability density function (PDF) of T . The denominator $\mathbb{P}(T > t, Z^1(t) > 0)$ in equation (3) satisfies

$$\begin{aligned} \mathbb{P}(T > t, Z^1(t) > 0) &= \mathbb{P}(T > t) - \mathbb{P}(T > t, Z^1(t) = 0) \\ &= F_T(t) - \mathbb{P}(T > t, Z^1(t) = 0). \end{aligned} \tag{4}$$

Using Theorem 1 and Theorem 2 we can calculate numerically $\mathbb{P}(T > t)$ and $\mathbb{P}(T > t, Z^1(t) = 0)$ and by substituting equation (4) in equation (3) we can also calculate the hazard function $g(t)$. The results from the calculations are presented in the next Section .

Results and applications

To illustrate the importance of the choice of lifespan distribution we consider several examples with normal, exponential, uniform, chi squared and beta distribution. The parameters of the distributions are chosen so that the life expectancy is equal to 10 time units in all of the cases. This makes the lifespan distributions comparable on a single plot and allows to study how the PDFs of the time to appearance of successful mutant depends on the “shape” of the distribution that we choose for the model. All other parameters of the model are kept the same in the simulations. We consider the probability for mutation to type 0 to be $u = 20\%$, the offspring p.g.f. for type 1 cells to be $f_1(s) = 0.625 + 0.375s^2$ and the extinction probability of BP, starting with a type 0 cell, to be $q_0 = 30\%$. On Figure 1 are shown the calculated PDFs of the waiting time T for appearance of “successful mutant” and the hazard function $g(t)$.

Conclusions

The simulation results reveal the importance of the distribution of the life length of type 1 cells as the time to the appearance of “successful mutant” depends on it. There are certain times at which the probability for appearance of a mutant is at its highest. In cancer treatment context this means we can target these moments to control the risk of metastasis and develop a better schedule for the therapy. We can also calculate the probability if a person has not yet developed mutant type of cancer, what are the chances he does not develop it in the future, namely $\mathbb{P}(T = \infty |$

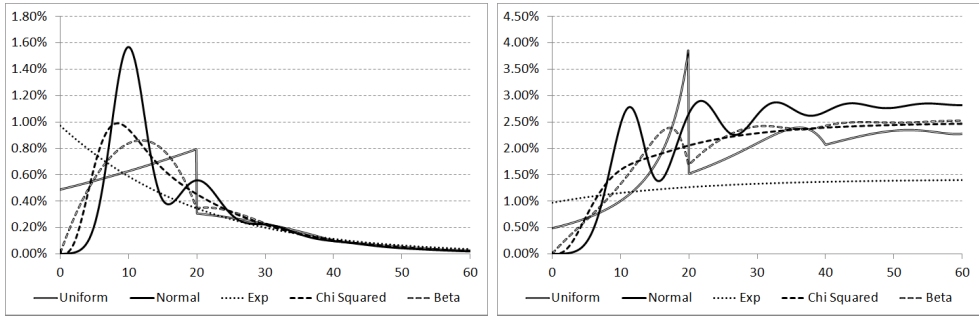


Figure 1: Probability density function of T (left) and the hazard function $g(t)$ (right).

$T > t$). The approach developed also allows to find the probability of a person who decided not to get treated again does not develop mutant type of cancer – $\mathbb{P}(T = \infty)$.

Acknowledgments. The research is partially supported by the National Fund for Scientific Research at the Ministry of Education and Science of Bulgaria, grant No DFNI-I02/17 and by the financial funds allocated to the Sofia University “St. Kl. Ohridski” (grant No 197/2016).

References

- [1] R. Durrett, 2014. *Branching Process Models of Cancer*, Springer, 2014.
- [2] K. Freise and J. Widness and R. Schmidt and P. Veng-Pedersen (2008). Modeling time variant distributions of cellular lifespans: increases in circulating reticulocyte lifespans following double phlebotomies in sheep, *J Pharmacokinet Pharmacodyn*, Vol **35(3)**, pp. 285-323.
- [3] J. Perez-Ruixo and A. Vermeulen. 2008. Basic pharmacodynamic models for agents that alter the lifespan distribution of natural cells, *J Pharmacokinet Pharmacodyn*, Vol **353**, pp.349-377.
- [4] M. Slavtchova-Bojkova, 2016, in press, On two-type decomposable branching processes in continuous time and time to escape extinction, *Proceeding WBPA2015, Lecture Notes in Statistics*, Springer

Distribution of the Order in Some Classes of Random Mappings

ARSEN YAKYMIV

*Steklov Mathematical Institute RAS
Gubkin St. 8, 119991 Moscow, Russia
E-mail: arsen@mi.ras.ru*

Abstract: This is the short review on limit theorems for the logarithm of the order in random mappings of different classes obtained earlier.

Key words: random mappings, random permutations, cyclic points, order of a mapping.

Description of a problem

Let \mathfrak{S}_n be a semigroup of all mappings of a finite set X with n elements into itself and a mapping σ belongs to \mathfrak{S}_n . The number of different elements in a sequence $\{\sigma, \sigma^2, \sigma^3, \dots\}$ of powers (i.e., iterations) of σ is called as an order of this mapping. Set an arbitrary probability distribution on \mathfrak{S}_n (in the majority of the cited articles it is uniform distribution on some subset $B_n \subseteq \mathfrak{S}_n$, having sufficiently large power). Let σ_n be a random mapping from \mathfrak{S}_n with this distribution and $M_n = M(\sigma_n)$ is it's order. The limit theorems received for $\ln M_n$ as $n \rightarrow \infty$ known to the author are given here.

A case of random permutations

In this section, we discuss the most investigated case when $P\{\sigma_n \in S_n\} = 1$, where S_n is a group of one-to-one mappings from \mathfrak{S}_n , i.e., various permutations of a set X . Note that for the first time such problem (together with some other problems of the statistical theory of groups) was studied by P. Erdős and P. Turán 1965 - 1969 [6] with uniform distribution at $B_n = S_n$. In their researches, in particular, it was shown

that, for any fixed $x \in R$

$$\mathbb{P} \left\{ \ln M_n \leq \frac{\ln^2 n}{2} + x \sqrt{\frac{\ln^3 n}{3}} \right\} \rightarrow \Phi(x)$$

as $n \rightarrow \infty$ where $\Phi(x)$ is the standard normal distribution function. It is curious to note that, according to E. Landau [12] (1909)

$$\max_{\sigma_n \in S_n} \ln M(\sigma_n) \sim \sqrt{n \ln n}.$$

Thus, we see that "typical" permutation from S_n has much a smaller order, than that at which the maximum in the last expression is reached. The proof of this result of P. Erdős and P. Turán was rather difficult. Probably, therefore the present limit theorem drew to itself close attention of many researchers: it was proved again many times (see, for instance, the articles of Best [3] (1970), Kolchin [11] (1977), Pavlov [16] (1980), Bovey [4] (1980), DeLaurentis and Pittel [13] (1985), Arratia and Tavaré [1] (1992)). Proofs became simpler step by step, but as a result of them nevertheless it was impossible to call simple. In 1985, Nicolas [15] showed that the error in Erdős-Turán law with a mean shift of $-\ln n \ln \ln n$ is of order at most $O(\ln^{1/2} n \ln \ln n)$. Barbour and Tavaré [2] obtained in 1994 the estimation of correct order in this law. Note that they considered more general model of random permutation $\sigma_n(\varrho)$, having so called Ewens sampling distribution with parameter $\varrho > 0$ [7], i.e., for any $\sigma \in S_n$ with k cycles $\mathbb{P}\{\sigma_n(\varrho) = \sigma\} = \varrho^k / \varrho(n)$ where $\varrho(n) = \varrho(\varrho + 1) \cdots (\varrho + n - 1)$ (the uniform distribution on S_n turns out for $\varrho = 1$). In particular, Barbour and Tavaré showed in [2] that

$$\sup_{x \in R} \left| \mathbb{P} \left\{ \frac{\ln M(\sigma_n(\varrho)) - \frac{\varrho \ln^2 n}{2} - \varrho \ln n \ln \ln n}{\sqrt{\frac{\varrho \ln^3 n}{3}}} \leq x \right\} - \Phi(x) \right| = O\left(\frac{1}{\sqrt{\ln n}}\right).$$

A.I. Pavlov [17] (1981) considered a random permutatin $\sigma_n = \sigma_{nk}$, having uniform distribution on the set $S_n^{(k)} = \{a : a = x^k | x \in S_n\}$ and proved the next limit theorem for the logarithm of the order ζ_{nk} of σ_{nk} :

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{\zeta_{nk} - \mathbb{E}\zeta_{nk}}{\sqrt{D\zeta_{nk}}} \leq x \right\} = \Phi(x),$$

where $k \in N$ and $x \in R$ are fixed. Also he showed that $E\zeta_{nk} \sim \frac{\varphi(k)}{2k} \ln^2(n)$ and $D\zeta_{nk} \sim \frac{\varphi(k)}{3k} \ln^3(n)$ as $n \rightarrow \infty$ and k is fixed, where $\varphi(\cdot)$ is the Euler's function. Note also curious fact (though it is beyond of the present review), that a weak limit for a number cycles $\xi_{sk}(n)$ of fixed length s in this model of random permutation, have not, as usual, Poisson distribution, though the set of $S_n^{(k)}$ has quite large power: $|S_n^{(k)}| \sim An!n^{-1+\varphi(k)/k}$, nevertheless

$$\mathbf{P}\{\xi_{sk}(n) = \nu\} \rightarrow \frac{1}{\nu!s^\nu} h_d^{-1} \left(\frac{1}{s} \right)$$

as $n \rightarrow \infty$ for any fixed k , where

$$\bar{k} = \prod_{p|k} p, \bar{d} = (s, k), d = \prod_{p|\bar{d}} p^{\alpha_p}, p^{\alpha_p} | k, p^{\alpha_p+1} \nmid k, h_d(z) = \sum_{m=0, d|m}^{\infty} \frac{z^m}{m!},$$

p is a prime number and A depends only of k . The similar effect for some class of random mappings has found L. Mutafchiev [14](1978). In 2004, Zacharovas [24] obtained a rate of convergence in Pavlov's limit theorem:

$$\sup_{x \in R} \left| \mathbf{P} \left\{ \frac{\zeta_{nk} - E\zeta_{nk}}{\sqrt{\frac{\varphi(k)}{3k} \ln^3(n)}} \leq x \right\} - \Phi(x) - r_k(x) \frac{e^{-x^2/2}}{\sqrt{\ln n}} \right| = O \left(\frac{\ln \ln n}{\ln n} \right)^{2/3}$$

as $n \rightarrow \infty$, where $k \in N$ is fixed and

$$r_k(x) = \sqrt{\frac{3}{\beta}} \frac{(1 - 8c - x^2)}{8\sqrt{2\pi}}, \quad c = \beta \int_0^1 \frac{(1-y)^{\beta-1} - 1}{y} dy, \quad \beta = \frac{\varphi(k)}{k}.$$

As we see, Zacharovas allocated the member having order $\ln^{-1/2} n$ (see the result of Barbour and Tavaré [2] specified earlier) and estimated an order of the corresponding difference. In the same work, Zacharovas specified mentioned above Pavlov's result concerning an asymptotics of mean ζ_{nk} :

$$E\zeta_{nk} = \frac{\beta}{2} \ln^2 n - \beta \ln n (\ln \ln n + C(k)) + \sum_{\varrho} \Gamma(-\varrho) (\beta \ln n)^\varrho + O((\ln \ln n)^2),$$

where

$$C(k) = \ln \beta - 1 - \int_0^1 \frac{(1-y)^{\beta-1} - 1}{y} dy - \sum_{p|k} \frac{\ln p}{p-1}$$

and \sum_{ρ} means the sum on nontrivial zeros of Riemann's zeta-function.

Let $S_n^{(k)}$ be a set of permutations from S_n , having no more than k cycles of identical length, and $\sigma_n^{(k)}$ be a random permutation uniformly distributed on $S_n^{(k)}$. In the paper [18] (1996), A.I. Pavlov obtained the next limit theorem:

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \ln M(\sigma_n^{(k)}) - \frac{\ln^2 n}{2} \leq x \sqrt{\frac{\ln^3 n}{3}} \right\} = \Phi(x).$$

where $x \in \mathbb{R}$ and $k \in \mathbb{N}$ are fixed. As it is well known, every permutation decomposes into the cycles. Fix some set $A \subseteq \mathbb{N}$. Permutation $\sigma \in S_n$ is called as A -permutation if all lengths of its cycles belong to a set A [20]. By $T_n = T_n(A)$ we designate the set of all A -permutations of n -th degree. Let $\tau_n = \tau_n(A)$ be a random permutation uniformly distributed on $T_n(A)$. In 2010, it was shown by the author [23] that a sequence of random variables

$$\frac{\ln M(\tau_n(A)) - \sum_{i \in A, i \leq n} \ln(i)/i}{\sqrt{\frac{\rho}{3} \ln^3 n}}$$

converges weakly to the standard normal r.v. if the next relations are valid:

$$\frac{|k : k \leq n, k \in A|}{n} \rightarrow \rho > 0, \quad \frac{|k : k \leq n, k \in A, m - k \in A|}{n} \rightarrow \rho^2 \quad (1)$$

uniformly by $m \in [n, Cn]$ for an arbitrary constant $C \in (1, \infty)$ as $n \rightarrow \infty$. Since it is not evident, which sets A satisfy assumptions (1), then some examples are given in the following section. More detailed analysis of examples is available in the book [22]. Here we note only that for $A = \mathbb{N}$ Erdős-Turán result mentioned at the beginning of this section follows from here.

In the recent work [8] (2012), Gnedin, Iksanov and Marynych considered the next model of random permutation. Let a random variable W has a distribution D on $(0, 1)$ and let W_1, W_2, \dots be its independent copies. Set

$$Q_0 = 1, \quad Q_j = \prod_{i=1}^j W_i, \quad j \in \mathbb{N}.$$

Suppose that U_1, \dots, U_n be independent random variables uniformly distributed of $[0, 1]$ and not depending on W_1, W_2, \dots . We assume that permutation $\Pi_n \in S_n$ has the cycle $(i_1 \rightarrow \dots \rightarrow i_l \rightarrow i_1)$ if $U_{i_1} < \dots < U_{i_l}$ and the points U_{i_1}, \dots, U_{i_l} and only they from U_1, \dots, U_n fall into interval $(Q_j, Q_{j+1}]$. It is clear that a permutation Π_n identically defines by it's cycles. Let a distribution D is absolutely continuous with a density f and, for some $\delta_1, \delta_2 \geq 0$, a function $f(x)$ is nondecreasing on $(0, \delta_1)$, is bounded on $[\delta_1, 1 - \delta_2]$ and is nonincreasing on $(1 - \delta_2, 1)$. In the article [8], it is shown that if r.v. $|\ln W|$ is in domain of attraction of stable distribution with index $\alpha \in (1, 2]$, then, for some sequences a_n and b_n (which are written out explicitly) distribution of r.v. $(\ln M(\Pi_n) - a_n)/b_n$, weakly converges to a stable law with the same parameter α . In particular, if $f(x) = \varrho x^{\varrho-1}$, then the limit law is normal ($\alpha = 2$), and permutation Π_n has Ewens sampling distribution with parameter ϱ , and, if $\varrho = 1$, we have the uniform distribution on S_n .

A case of random mappings

Let, as earlier, \mathfrak{S}_n be the set of all mappings of a finite set X with n elements into itself. For the first time, we consider the situation when a random mapping σ_n has uniform distribution on \mathfrak{S}_n . In the work [10] (1973), B. Harris proved, that, for an arbitrary fixed real x ,

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \ln M(\sigma_n) - \frac{\ln^2 n}{8} \leq x \sqrt{\frac{\ln^3 n}{24}} \right\} = \Phi(x).$$

The mapping $\sigma \in \mathfrak{S}_n$ is called A -mapping, if all cycles of σ belong to the set $A \subseteq N$. For more details on the A -mappings, see Sachkov's book [20]. Let $\sigma_n(A)$ be a random mapping uniformly distributed on the set $V_n(A)$ of all A -mappings of the n th degree. Suppose that the relations (1) are valid (see the previous section). Recently the author proved (unpublished), that, under these assumptions, for an arbitrary fixed real x

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \frac{\ln M(\sigma_n(A)) - \sum_{i \in A, i \leq \sqrt{n}} \ln(i)/i}{\sqrt{\frac{\varrho}{24} \ln^3 n}} \leq x \right\} = \Phi(x). \quad (2)$$

Now we give some examples of sets A for which the relations (1) hold and therefore, the equality (2) takes place.

Example 1. *Let a set A has an asymptotic density 1 in the set of natural numbers. Then the relations (1) and hence (2) are valid with $\varrho = 1$. In particular, for $A = N$ the just mentioned Harris result follows from here.*

Example 2. *Let a set A be random and random variables $\eta_1 = \chi\{1 \in A\}, \eta_2 = \chi\{2 \in A\}, \dots$ are independent and*

$$\mathbf{P}\{\eta_n = 1\} \rightarrow \varrho > 0.$$

as $n \rightarrow \infty$. Then the relations (1) hold for a set A almost shurely (a.s.) [22, section 3.6]. Thus, for an arbitrary fixed real x , as $n \rightarrow \infty$

$$\mathbf{P} \left\{ \frac{\ln M(\sigma_n(A)) - \sum_{i \in A, i \leq \sqrt{n}} \ln(i)/i}{\sqrt{\frac{\varrho}{24} \ln^3 n}} \leq x \middle| A \right\} \rightarrow \Phi(x) \text{ a.s.}$$

Example 3. *Let an arbitrary finite union of intervals Δ from $[0, 1]$ and a function $g(t) = t^\alpha l(t), \forall t > 0$ are given, where funcrion $l(t)$ slowly varies at infinity and number α is positive and noninteger. We include a number $n \in N$ into a set A if and only if $\{g(n)\} \in \Delta$. It is shown in the book [22, theorem 3.5.2, conclusion 3.5.1] that, if for an arbitrary $m = 1, 2, \dots, [\alpha] + 3$*

$$\frac{d^m}{dt^m} l(t) = o(t^{-m} l(t)) \quad (t \rightarrow \infty),$$

then the relations (1) are valid, therefore the equality (2) holds for A .

And one more new result. Let τ be some permutation. Set $\Theta(\tau) = (\xi_1, \xi_2, \dots)$, where $\xi_i = \xi_i(\tau)$ is a number of cycles in permutation τ of lenghts $i \in N$. Suppose that, for some sequence of random permutations $\tau_n \in S_n$ and for positive sequences $a(n)$ and $b(n)$, the next relation

$$\frac{\ln M(\tau_n) - a(n)}{b(n)} \Rightarrow \eta \quad (n \rightarrow \infty)$$

holds, where η is a r.v. with continuous distriburion function. Let $\sigma_n \in \mathfrak{S}_n$ be an arbitrary sequence of random mappings such that

$$D(\Theta(\sigma_n^*) | \lambda_n = k) = D(\Theta(\tau_k)),$$

for any $k, n \in N$, $k \leq n$ where $D(\cdot)$ denotes a distribution, λ_n is a number of cyclic points in a random mapping σ_n , and σ_n^* is a reduction of σ_n into it's cyclic points. Suppose that $\lambda_n/r(n) \Rightarrow \xi > 0$ for some positive sequence $r(n) \rightarrow \infty$ such that $r(n) \geq n^\varepsilon$ for any $n \in N$ and some $\varepsilon > 0$. Assume that the inequality $a(n) \geq c \ln n$ holds for some constant $c > 0$ and all $n \in N$ and $a(n)/b(n) \rightarrow \infty$ as $n \rightarrow \infty$. Let there exists a positive function $d(x) \rightarrow \infty$ such that

$$\sup_{y \in (d^{-1}(x), d(x))} \frac{a(x) - a(xy)}{b(x)} \rightarrow 0 \quad (x \rightarrow \infty)$$

(we set $a(x) = a([x])$ and $b(x) = b([x])$ for $x > 0$). Then

$$\frac{\ln M(\sigma_n) - a(r(n))}{b(r(n))} \Rightarrow \eta \quad (n \rightarrow \infty).$$

Using the sequences of random permutations from section 2, we can obtain from here the assertions for corresponding classes of random mappings σ_n .

Nonuniform distributions on \mathfrak{S}_n were considered by different authors, for instance, by Cheplyukova [5], Hansen and Jaworski [9], Mutafchiev [14], Pavlov and Myllari [19], Timashëv [21] and others.

References

- [1] Arratia R.A., Tavaré S. (1992). Limit theorems for combinatorial structures via discrete process approximations. *Rand. Struct. Alg.* Vol. **3**, pp. 321-345.
- [2] Barbour A.D., Tavaré S. (1994). A rate for Erdős-Turán law. *Combinatorics, Probability and Computing*. Vol. **3**, pp. 167-176.
- [3] Best M.R. (1970). The distribution of some variables on symmetric group. *Proc. Kon. Ned. Akad. Wet. Ser. A*, Vol. **73**, pp. 385-402.
- [4] Bovey J.D. (1980). An approximate probability distribution for the order of elements of the simmetric group. *Bull. London Math. Soc.* Vol. **12**, no 1, pp. 41-46.

- [5] Cheplyukova I.A. (2004). A case of the limit distribution of the number of cyclic vertices in a random mapping. *Discrete Math. Appl.* Vol. **14**, no. 4, pp. 343-352.
- [6] Erdős P., Turán P. (1965-1969). On some problems of a statistical group theory I-IV. *Acta Math. Acad. Sci. Hung.*
- [7] Ewens W.J. (1972). The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* Vol. **3**, pp. 87-112.
- [8] Gnedin Alexander, Iksanov Alexander, Marynych Alexander (2012). A generalization of the Erdős-Turán law for the order of random permutation. *Comb. Probab. Comput* Vol. **21**, No. 5, pp. 715-733.
- [9] Hansen, Jennie C., Jaworski Jerzy (2013). Random mappings with Ewens cycle structure. *Ars Comb.* Vol. **112**, pp. 307-322.
- [10] Harris B. (1973). The asymptotic distribution of the order of elements in symmetric semigroups. *Journal of Combinatorial Theory (A)*. Vol. **15**, pp. 66-74.
- [11] Kolchin V.F. (1977). New proof of asymptotic lognormality of the order in random permutation.- Proceedings "Combinatorial and asymptotic analysis", Published by Krasnoyarsk State University, pp. 82-93.
- [12] Landau E. (1909). Handbuch der Lehre von der Verteilung der Primzahlen. Bd. I.
- [13] DeLaurentis J.M., Pittel B.G. (1985). Random permutations and Brownian motion. *Pacific Journal of Mathematics*. Vol. **119**, no. 2, pp. 287-301.
- [14] Mutafchiev L.R. (1978). Random mappings with restrictions on the corresponding graphs. (Russian). *Serdica*. Vol. **4**, pp. 105-110.
- [15] Nicolas J.-L. (1985). Distribution statistique de l'ordre d'un element du groupe symetrique. *Ada Math. Hungar.* Vol. **45**, pp. 69-84.
- [16] Pavlov A.I. (1980). On the Erdős-Turán theorem. *Problems of Cybernetics (Moscow)*. Vol. **64**, pp. 57-66. (in Russian).

- [17] Pavlov A.I. (1982). On the limit distribution of the number of cycles and the logarithm of the order of a class of permutations. *Math. USSR, Sb.* Vol. **42**, pp. 539-567.
- [18] Pavlov A.I. (1996). On the Erdős-Turán theorem on the logarithm of the order of a random permutation. *Dokl. Math.* Vol. **54**, No.2, pp. 688-691.
- [19] Pavlov Yu.L., Myllari T.B. (2012). Limit distributions of the number of vertices of given degree in the forest of a random mapping with a given number of cycles. *Discrete Math. Appl.* Vol. **22**, No. 2, pp. 225-234.
- [20] Sachkov V.N. (1996). *Probabilistic Methods in Combinatorial Analysis*. Cambridge Univ. Press, Cambridge.
- [21] Timashëv A.N. (2011). *Asymptotic expansions in probabilistic combinatorics*. Publishing House "TVP", Moscow.
- [22] Yakymiv A.L. (2005). *Probabilistic applications of Tauberian theorems*. Modern Probability and Statistics, VSP, Leiden, Boston.
- [23] Yakymiv A.L. (2010). A limit theorem for the logarithm of the order of a random A-permutation. *Discrete Math. Appl.* Vol. **20**, no 3, pp. 247-275.
- [24] Zacharovas V. (2004). Distribution of the logarithm of the order of a random permutation.- *Lithuanian Mathematical Journal*. Vol. **44**, No. 3, pp. 296-327.

ABSTRACTS

Statistical analysis of interval-censored data in questionnaire surveys

ANGEL ANGELOV

Umea University

901 87 Umea, Sweden

E-mail: agangelov@fmi.uni-sofia.bg

Interval-censored data arise in questionnaire surveys when, instead of being asked to provide an exact value, respondents are free to answer with any interval. Such a format is particularly suitable when asking questions that are difficult to answer with an exact amount (e.g. recall questions). In this context, the assumption of noninformative censoring is violated, and thus standard methods for interval-censored data can not be applied. We suggest a nonparametric maximum likelihood estimator of the underlying distribution function, based on a certain data collection scheme. The consistency of the estimator is established under general conditions and an iterative procedure for finding the estimate is proposed. The performance of the method is evaluated in a simulation study.

Animated Map of Mathematics

VELICHKO ARANGELOV, PLAMEN MATEEV

Sofia University "St. Kliment Ohridski"

Sofia, BULGARIA

e-mail: velichko@gmail.com pmateev@gmail.com

Abstract: A graphical representation of connections between different areas of mathematics through the last 25 years of XX century.

Dave Rusin created his MathAtlas at the end of last age [1]. We present a replication of his idea. We made a graphical representation of connections between different areas of mathematics using MDS algorithm (realised in R [2]). The data (downloaded from [3]), contains the classifications of Mathematical Reviews according Mathematical Subject Classification through the last 25 years of XX century. Graphics are made by moving bubbles chart, a powerful visualization tool in Google spreadsheets.

Acknowledgment

This work is partially supported by the financial funds allocated to the Sofia University "St.Kl.Ohridski", grant No 197/2016.

References

- [1] Rusin D. (1997). *The Mathematical Atlas*, <http://www.mathatlass.org> (accessed April 7, 2015).
- [2] Everitt B.S.,Hothorn T. (2006). *A Handbook of Statistical Analyses Using R*. Chapman & Hall/CRC, Boca Raton, Florida.
- [3] Mathematics on the Web. (2011) <http://www.mathontheweb.org/> (accessed January 19, 2016).

Modelling of epidemics with vaccination: stochastic monotonicity and continuity properties

FRANK BALL^a, M. GONZÁLEZ^b, R. MARTÍNEZ^b, MAROUSSIA
SLAVTCHOVA-BOJKOVA^{c,d},

^a *School of Mathematical Science, the University of Nottingham, UK*

^b *Department of Mathematics and Instituto de Computación Científica
Avanzada of the University of Extremadura, Badajoz, SPAIN*

^c *Sofia University, Faculty of Mathematics and Informatics,*

^d *Institute of Mathematics and Informatics, BAS, Sofia*

e-mail: `bojkova@fmi.uni-sofia.bg`

Vaccination is one of the key tools in controlling the spread of an epidemic and there have been numerous mathematical studies of its effect on disease dynamics. Recently, Ball et al. [1] have developed a framework for analyzing time-dependent vaccination policies for epidemics which are modelled by a Crump–Mode–Jagers branching process. Stochastic monotonicity and continuity results for a wide class of functions (e.g., extinction time and total number of births over all time) defined on such a branching process are proved, leading to optimal vaccination schemes to control corresponding functions of epidemic outbreaks. In this study we focus on expanding these properties in the framework of Sellke construction for SIR epidemic models [2].

Acknowledgements This work is partially supported by the financial funds allocated to the Sofia University “St. Kl. Ohridski” (grant 197/2016), the Ministerio de Economía y Competitividad (grant MTM2015–70522–P), the Junta de Extremadura (grant GR15105) and the FEDER, and the National Fund for Scientific Research at the Ministry of Education and Science of Bulgaria (grant DFNI–I02/17).

References

[1] Ball, F., González, M., Martínez, R. and Slavtchova–Bojkova, M. (2014). Stochastic monotonicity and continuity properties of functions defined on Crump–Mode–Jagers branching processes, with application to vaccination in epidemic modelling. *Bernoulli*, 20(4), 2076–2101.

[2] Sellke, T. (1983) On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Prob* 20 (20), 390–394.

Stochastic Navier-Stokes equation on the 2D- Riemannian manifold (Sphere)

AZZEDINE BENCHETTAH

LANOS Laboratory, Mathematical Department

Badji Mokhtar University

BP12, Annaba, Algeria

E-mail: abenchettah@hotmail.com

We consider the stochastic Navier-Stokes equation on the unit sphere considered as a Riemannian manifold. Using the method of Galerkin approximations and the stochastic calculus, we prove the existence of the solution and the invariant measure for the discretized equation (finite dimensional case). We prove also the convergence of the discretized equation solutions to the stochastic Navier stokes solution on the sphere.

References:

[1] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1981.

[2] V. Barbu and G. Da Prato, The Kolmogorov equation for a 2D-Navier-Stokes Stokes Flow in a Channel, Nonlinear Analysis 69 (2008), 940-949.

[3] H. Breckner, Approximation of the Solution of the Stochastic Navier-Stokes Equation, Optimization 49 (2001),15-38.

Randomness uncertainty and ergodicity in nature and society: towards a sustainable society ethics in a non-ergodic world

GEORGI CHOBANOV

Faculty of Economics and Business Administration

Sofia University "St. Kliment Ohridski"

Tsarigradsko chaussee blvd. 125, block 3, Sofia, Bulgaria

E-mail: georgech@feb.uni-sofia.bg

People live and take decisions under uncertainty in nature and society environment. Since ancient times they try to reduce randomness in nature by recognizing laws of nature as causal relationships among different events or processes. In society, laws and regulations are created by people's mind according people's value system beliefs. Randomness in society is incorporated in uncertainty of human's behavior. Humans create institutions according a value system beliefs in order to reduce uncertainty and risk and to achieve stability and sustainability and they can make essential changes in society establishing a new value system of beliefs. It is difficult to predict occurrence of essential changes in society since processes in society are path-dependent and non-ergodic. This paper is concerned with problems of a transformation process towards a sustainable society value system of beliefs or secular ethics under the conditions of a non-ergodic world hypothesis.

A noisy drift forecast and trading policies for high-frequency markets

DONCHO DONCHEV

*Faculty of Mathematics and Informatics
Sofia University
5, J. Bourchier blvd, 1164 Sofia, Bulgaria
E-mail: doncho@fmi.uni-sofia.bg*

We study the impact of the noise in predicting the near-future behavior of high-frequency markets. In contrast to the existing methods that smooth or neglect the noise, we assume that it contains an important information about the hidden news that drive the price dynamics. Thus, in case of high-frequency markets, we construct a trading policy by analysing not the whole time series, but only its noisy part. Making use of suitable neural networks and more than three year historical data for USD/JPY exchange rates, we show that market's near-future behavior is partially predictable. A definition of a trading edge as a criterion for policy's performance is also given.

On the successful and blocked calls in one finite queueing system with losses

VELIKA DRAGIEVA

Sofia University of Forestry
10 Kliment Ohridsky street ,Sofia, Bulgaria
E-mail: dragievav@yahoo.com

The paper deals with two important descriptors of any queueing system with losses and/or blocked calls: the number of successful and the number of lost (blocked) calls, made during a time interval $(0,t)$. The system under consideration has one server that serves finite number of customers. The customer, arriving at a moment of a busy server is not allowed either to queue, or to repeat his/her attempts. He/she is blocked for an exponentially distributed time interval in the orbit of blocked customers. Formulas for computing the main characteristics of the system, including the number of successful and lost calls are obtained in previous papers. Objective of the present paper is to extend and generalize them, and on the basis of numerical results to investigate the influence of the system parameters on the behavior of these descriptors.

Randomly indexed central order statistics through examples and application

ANETA GACOVSKA BARANDOVSKA

Faculty of Natural Sciences and Mathematics
Institute of Mathematics, University Ss Cyril and Methodius
Skopje, Republic of Macedonia
E-mail: agacovska@gmail.com

In our paper from 2013, we have investigated the asymptotic behavior of central order statistics with random size, using regular time-space changes. Here, an example using the results of randomly indexed central order statistics is given and a survey on the possible applications in insurance is made.

A graph based methods for clustering a morphologically similar words

BOJAN ILIJOSKI

*Faculty of Computer Science and Engineering
University Ss Cyril and Methodius
"Rugjer Boshkovikj" 16, P.O. Box 393,1000 Skopje, Republic of Macedonia
E-mail: bojan.ilijoski@finki.ukim.mk*

In this paper we present a method for grouping a morphologically similar words. At first we use some metrics for measuring the distance between elements in nonempty finite sets. We use those metrics in order to create a similarity matrix between words. The distance between two words in our case depends only on their morphological similarity. This allow us to make word clustering without a dictionary and also help us to handle incorrectly written words (typos).

After this process we create a weight graph which represent the similarity between words. Words are represented as vertices and there is an edge between two of them if and only if the similarity between words are above some threshold. The weight of the edge presents similarity between words. On this graph we use some methods to extract only the most morphologically similar words. With those methods we practically do a graph clustering i.e. we make a grouping of similar words in a clusters. This approach should increase efficiency for words clustering and should give us a better way of clustering words which can be further used for making better bag of words, finding similar words etc.

Using data mining techniques for recognition of Macedonian national dances

VIKTOR JANEVSKI, ALEKSANDAR KUZMANOSKI, ZANETA POPESKA,
MARIJA MIHOVA

*Faculty of Computer Science and Engineering
University Ss Cyril and Methodius
"Rugjer Boshkovikj" 16, P.O. Box 393,1000 Skopje, Republic of Macedonia*

E-mail: zaneta.popeska@finki.ukim.mk

Nowadays there are numerous sensors and hardware devices that can be creatively used to solve a vast variety of problems. In our research we used the Microsoft Kinect sensor in order to collect data about the movement of different parts of the human body from individuals who perform Macedonian national dances. The main goal is to employ supervised learning techniques to create models of several traditional dance choreographies based on the built dataset. We discovered that, even by using simple data mining techniques and algorithms such as Naive Bayes, kNN, decision trees and classification rules, we obtained satisfactory results with accuracy values up to 95% for kNN and 93% for Naive Bayes on a randomly chosen test set. The potential applications of the proposed solution range from analyzing the constitutive elements of a particular traditional dance choreography, through implementing a dancing tutor, to digitizing and preserving the Macedonian cultural heritage.

Compound power series distribution with negative multinomial summands

PAVLINA JORDANOVA

*Faculty of Mathematics and Informatics
Shumen University*

115 Universitetska Street, Shumen 9712, Bulgaria.

E-mail: pavlina.kj@abv.bg

The class of Power Series Distributions is quite general. It contains the most important discrete distributions, e.g. Poisson, Binomial, Negative binomial, Logarithmic series etc. The paper considers a multivariate distribution which coordinates are dependent Compound Power Series Distributed random variables with Negative Multinomial summands. We show the relation between these distributions and Multinomial distribution. In the most important particular cases we found appropriate equivalent presentations. For example in geometric case the resulting vector is just stochastically equivalent to a product of independent Bernoulli random variable and Multivariate geometrically distributed random vector. Using these relations we obtain some additional properties of Compound Power Series Distribution with Negative Multinomial Summands and its particular cases.

Special application of Mellin transform to contingent claims pricing

DESSLAVA KOLEVA, MARIYAN MILEV

*Faculty of Mathematics and Informatics
Sofia University "St.Kl.Ohridski", Sofia, BULGARIA
e-mail: koleva_dn@yahoo.com
University of Food Technologies-Plovdiv
bul. Maritza 26, Plovdiv 4002, Bulgaria
e-mail: marianmilev2002@gmail.com*

Obtaining the fair value of a contingent claim is a broadly explored area in financial mathematics. Based on the assumption of a Geometric Brownian Motion process for the underlying asset we need to solve a parabolic equation with the relevant special conditions depending on the contract structure. Mellin transform, under certain conditions, can be a useful tool for obtaining a price estimate. Moreover, it turns out that after applying an inverse technique based on maximum entropy principle, the solution remains smooth and positive. This effect is best illustrated when pricing path-dependent contracts.

Penalty functions for perturbed risk processes

EKATERINA KOLKOVSKA

*Centro de Investigacion en Matematicas
Apartado Postal 402 36000 Guanajuato, Mexico
E-mail: todorova@cimat.mx*

We study several penalty functionals of classical risk processes with two-sided jumps, perturbed by an alpha-stable process. We obtain a formula for the Laplace transform of such functionals, as well as a representation of the functionals as infinite series of convolutions of given known functions.

On the bivariate compound Pólya process

KRASIMIRA Y. KOSTADINOVA¹, LEDA D. MINKOVA²

¹ *Faculty of Mathematics and Informatics
Shumen University "K. Preslavsky"
Shumen, BULGARIA*

e-mail: kostadinova@shu-bg.net

² *Faculty of Mathematics and Informatics
Sofia University "St. Kl. Ohridski"
Sofia, BULGARIA*

e-mail: leda@fmi.uni-sofia.bg

Key words: Compound distributions, bivariate geometric distribution, Pólya process.

In this paper we define a bivariate counting process as a compound negative binomial process with bivariate geometric compounding distribution. The resulting process is called a Bivariate compound Pólya process. Then we consider a risk model in which the claim counting process is a Bivariate compound Pólya process and call it a Bivariate compound Pólya risk model. We also consider two types of ruin probability for this risk model and find the corresponding Laplace transforms. We discuss in detail the particular case of exponentially distributed claims.

Noncentral Pólya-Aeppli process and ruin probability

MEGLENA LAZAROVA, LEDA MINKOVA

*Faculty of Mathematics and Informatics
Sofia University "St.Kl.Ohridski", Sofia, BULGARIA*

e-mail: meglena.laz@fmi.uni-sofia.bg

e-mail: leda@fmi.uni-sofia.bg

In this paper we introduce a stochastic process which is a sum of homogeneous Poisson process and Pólya-Aeppli process and call it a non-central Pólya-Aeppli process (NPAP). The probability mass function, recursion formulas and some properties are derived. As application we consider a risk model with NPAP counting process. The joint distribution of the time to ruin and deficit at the time of ruin is defined. The differential equation of the ruin probability is given. As example we consider the case of exponentially distributed claims.

Noncentral Pólya-Aeppli of order K risk model

LEDA MINKOVA

*Faculty of Mathematics and Informatics
Sofia University "St.Kl.Ohridski", Sofia, BULGARIA
e-mail: leda@fmi.uni-sofia.bg*

In this study we define the Polya-Aeppli process of order K as a compound Poisson process with truncated geometric compounding distribution with success probability $1 - \rho > 0$ and investigate some of its basic properties. Using simulation we provide a comparison between the sample paths of the Polya - Aeppli process and the Poisson process. Also, we consider a risk model in which the claim counting process $N(t)$ is Polya-Aeppli process of order K, and call it a Polya-Aeppli of order K risk model. For the Polya-Aeppli of order K risk model we derive the joint distribution of the time to ruin and the deficit at ruin as well as the ruin probability. We discuss in detail the particular case of exponentially distributed claims and provide simulation results for more general cases

Two-sample tank tests based on Lee distance

NIKOLAY NIKOLOV

*Institute of Mathematics and Informatics - BAS
Acad. G.Bontchev str., bl.8, 1113 Sofia, BULGARIA
e-mail: n.nikolov@math.bas.bg*

Let X_1, X_2, \dots, X_m and Y_1, Y_2, \dots, Y_n be two independent random samples with continuous distribution functions $F(x)$ and $G(x)$, respectively. We consider rank tests for the two-sample location problem of testing the null hypothesis H_0 against the alternative H_1

$$\begin{aligned} H_0 &: F(x) \equiv G(x) \\ H_1 &: F(x) \geq G(x), \end{aligned}$$

with strict inequality for some x .

For nonparametric testing problem there are various procedures to constructing rank tests via metrics on the permutation group. Critchlow proposed a unified approach based on distance defined by the minimum interpoint distance between appropriate sets of permutations. The test statistic for the above problem is derived by the distance between the class of equivalence $[\alpha]$ and the extremal set E

$$d([\alpha], E) = \min_{\pi \in [\alpha], \sigma \in E} d(\pi, \sigma), \quad (1)$$

where d is an arbitrary metric on the set S_N of all permutation of N integers. The set $[\alpha]$ consists of all permutations in S_N which are equivalent to the observed permutation α . The set E of extremal permutations consists of all permutations in S_N which are least in agreement with H_0 and most in agreement with H_1 . The test rejects the null hypothesis for small values of $d([\alpha], E)$.

For two permutations α and β from the set S_N , Lee distance is defined by

$$L(\alpha, \beta) = \sum_{i=1}^N \min(|\alpha(i) - \beta(i)|, N - |\alpha(i) - \beta(i)|) .$$

The goal of this research is to derive the test statistic (1) induced by the Lee distance function and study the distribution under null hypothesis.

Stochastic model of the relationship pharmacological effect - drug plasma concentration for new Bulgarian antihypertensive agent

KRASIMIRA PRODANOVA

Technical University of Sofia
blv "Kl. Fhridski" 8, 1000 Sofia, Bulgaria
E-mail: kprod@tu-sofia.bg

In this work the mathematical models of pharmacokinetics are used for description of the pharmacological effect of the medicine as a function of experimental plasma concentration. Pharmacokinetics is a branch of medico-pharmaceutical sciences and its aim is to examine the qualitative changes of the medicine when it enters the body. Base on the experimental data the parameters of two stochastic models for the new Bulgarian antihypertensive agent are estimated. For the description of the relationships plasma concentration - time the compartment models are used. For the relationship pharmacological effect - plasma concentration, the model of enzyme kinetics Michaelis - Menten is applied.

Asymptotic normality of conditional distribution estimation in the single index model

ABBES RABHI

Department of Mathematics
University of Sidi Bel Abbes
P.O.Box 89, Sidi Bel Abbes 22000, Algeria
E-mail: rabhi_abbes@yahoo.fr

This work deals with the estimation of conditional distribution function based on the single-index model for functional time series data. The asymptotic normality of the conditional distribution estimator is established. Moreover, as an application, the asymptotic $(1 - \gamma)$ confidence interval of the conditional distribution function is given for $0 < \gamma < 1$.

Consensus algorithms and Decomposition-Separation Theorem

ISAAK SONIN

*Department of Mathematics and Statistics
University of North Carolina at Charlotte
9201 University City Blvd,
Charlotte, NC 28223-0001, USA
E-mail: imsonin@uncc.edu*

A classical dictionary definition describes consensus as general agreement about something: an idea or opinion that is shared by all the people in a group. Nowadays this definition is outdated. Numerous studies of the last 20-30 years have demonstrated that many particle (agents) systems exhibit cooperative behavior that can lead to a consensus, or to multiple consensuses. Examples of this kind can be found in biology, robotics and control, sociology, political science, physics, mechanics, computer science etc. Consensus algorithms are algorithms leading to consensus. The most studied of these algorithms are linear (averaging) defined by stochastic matrix transformation. This leads to the study of nonhomogeneous Markov chains in reversed time. The Decomposition-Separation theorem proved in a final form by I. Sonin about 30 years ago and started in the thirties by A. Kolmogorov, and continued by D. Blackwell and H. Cohn, is a general theorem describing the behavior of such arbitrary Markov chain. One of the current publications states that "D-S Theorem relates to and generalizes most of the existing results in the literature on convergence properties of linear consensus algorithms". I will give a brief survey of consensus algorithms and their relationship to the DS Theorem.

References:

- [1] Dimov, I., Karaivanova, A., Overconvergent Monte Carlo Methods for Density-function Modelling Using B-Splines, Advances in Numerical Methods and Appl., World Scientific, pp. 85-93, 1994
- [2] Hammersley, J.M, Handscomb, D.C., Monte Carlo Methods, Jonh Wiley & Sons, inc., New York, London, Methuen, 1964

[3] Karaivanova, A., Dimov, I., Error Analysis of an Adaptive Monte Carlo Method for Numerical Integration, Mathematics and Computers in Simulation 47, Elsevier, pp.201-213, 1998

[4] Sobol,I, Monte Carlo Numerical Methods, Nauka, Moscow, 1973.

Accelerated EM algorithm and statistical inference for multitype branching processes with multivariate power series offspring distributions

ANA STANEVA, VESELA STOIMENOVA

*Faculty of Mathematics and Informatics
Sofia University*

5 James Bourchier blvd., Bl.25A, 1164, Sofia, Bulgaria

E-mail: anais@uni-sofia.bg

In our work we use the Accelerated EM algorithm to estimate the individual parameters of the class of discrete-time multitype branching processes with power series offspring distributions.

We consider an observation scheme in which only the population sizes of each generation are known and there are unobservable data about the quantity of the children.

A possible iterative procedure for maximum likelihood estimation in this model is the regular EM algorithm. Previous work has shown that this algorithm is effective, but time consuming in case of statistical inference for multitype branching processes.

We propose an R-implementation of the Accelerated EM algorithm, simulations and computational results.

Acknowledgements: This work is partially supported by the financial funds allocated to the Sofia University "St. Kl. Ohridski", grant No 197/2016, and by the National Fund for Scientific Research at the Ministry of Education and Science of Bulgaria, grant No DFNI-I02/17

Comparison of partially ranked lists with application to RNA-seq data

EUGENIA STOIMENOVA, DEAN PALEJEV

*Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G.Bontchev str., bl.8, 1113 Sofia, BULGARIA
E-mail: jeni@math.bas.bg*

In many situations, there are different methods for analyzing the same data. For example, several methods exist for finding differentially expressed genes using RNA-seq data. These methods tend to produce similar, but not identical significant genes and partial rankings of the gene list.

Here we consider partial rankings of all genes of a given species. The genes are split into several groups, so that there is a ranking between the groups and not necessarily within each group. An important special case is a setup in which the most significant k genes are ranked, with group $k+1$ consisting of the remaining genes.

When comparing different methods applied to the same data, we are interested in how close are their outputs. Prior studies have compared several RNA-seq differential expression methods using empirical results based on real and simulated data. Our approach is based on the statistical distribution of the distance on the set of partial rankings. This allows us not only to measure the relative similarities between different methods, but also to infer whether two methods are significantly different. The outputs of several RNA-seq differential expression methods are analyzed using this approach.

Computational assessments of multidimensional branching processes derived from cosmic rays cascades

ASSEN TCHORBADJIEFF

Institute of Mathematics and Informatics - BAS
Acad. G.Bontchev str., bl.8, 1113 Sofia, BULGARIA
e-mail: `atchorbadjieff@math.bas.bg`

The cosmic rays cascades are one of the most famous examples of branching processes in physics. They consist of many different types of secondary high energy particles which are offsprings of leading primary one, usually high energy nucleon or gamma, after collision with atmosphere. The rate of expansion of cascade depends on multiple different conditional probabilities as the chance of survival in atmosphere without interactions, particle lifetime, the number of daughter particles and etc. The presentation demonstrates quantitative results of simulated in R procedures for parallel estimations of energy spectra distribution of multiple types of daughter particles. The models are based on simplified cascades with only most distributed particles. The probability dependencies on particle's energy and free path in atmosphere are not constrained only to available experimental results. Moreover, for simplicity the scattering angles are not considered in this version of software.

The problem of the preservation of the distribution of compound mixed renewal processes under a change of measure

SPYRIDON TZANINIS

*Department of Statistics and Insurance Science
University of Piraeus, Greece
E-mail: stzaninis@unipi.gr*

Let (Ω, Σ, P) be a probability space, let N be a mixed renewal process with mixing parameter a random vector Θ , and let X a family of independent and identically distributed random variables (i.i.d. for short), such that X and N as well as X and Θ are independent. Let S be the compound mixed renewal process with respect to the probability measure P induced by N and X . We characterize all probability measures Q on Σ such that S remains a compound mixed renewal process with respect to Q , and Q and P are progressively equivalent.

The fact that the interarrival process W is not i.i.d. but conditionally i.i.d. with respect to Θ makes the problem more complicated and difficult than the one in the Classical Risk Theory which was posed and solved by F. Delbaen & J. Haezendock [1].

We provide a characterization of probability measures Q as above in terms of martingales and disintegrations. As a consequence we obtain as special cases well known results from F. Delbaen & J. Haezendock [1] and P. Embrechts & S. Meister [2]. Finally, we provide concrete examples of the calculation of the Radon-Nikodym derivative of Q with respect to P .

References:

[1] Delbaen, F. and Haezendock, J.: A martingale approach to premium calculation principles in an arbitrage free market, *Insurance: Mathematics and Economics* Volume 8, Issue 4, 269-277. (1989)

[2] Embrechts, P. and Meister, S.: Pricing insurance derivatives, the case of CAT-futures, *Proceedings of the 1995 Bowles Symposium on Securitization of Risk*. (1997)

[3] Schmidt, K.D.: *Lectures on Risk Theory*, B.G. Teubner, Stuttgart (1996).

Generation of correlation matrices

EVELINA VELEVA

*Faculty of Mathematics and Informatics
Sofia University
5, J. Bourchier blvd, 1164 Sofia, Bulgaria
E-mail: eveleva@abv.bg*

Correlation and covariance matrices representing the strength of the dependence for each pair of factors involved in a statistical experiment are always positive definite matrices. The article discusses ways to generate such matrices with or without restrictions on their elements and laws of distribution. Special attention is paid to the case when the resulting correlation matrix presents a weak correlation between pairs of factors involved in the experiment.

Limiting Distributions for Alternating Regenerative Branching Processes

KOSTO V. MITOV, NIKOLAY M. YANEV

*National Military University "Vasil Levski" - Aviation Faculty
D. Mitorpolia, Pleven, BULGARIA
e-mail: kmitov@yahoo.com
Institute of Mathematics and Informatics - BAS,
Sofia, BULGARIA
e-mail: yanev@math.bas.bg*

Three new classes of alternating regenerative branching processes (ARBP) are proposed and investigated applying renewal and regenerative methods developed in Mitov and Yanev [J. Appl. Prob. 2001], [Adv. Appl. Prob. 2001]. In this way new limiting distributions are obtained for AR Sevastyanov processes, for AR Sevastyanov processes with non-homogeneous Poisson immigration and for AR randomly indexed branching processes.

A Bayesian spatial analysis of mumps data in Bulgaria

MAYA ZHELYAZKOVA¹, ROUMYANA YORDANOVA², MAROUSSIA
SLAVTCHOVA-BOJKOVA^{1,2}

¹ *Sofia University, Faculty of Mathematics and Informatics,
Department of Probability, Operations Research and Statistics, Sofia,
Bulgaria*

² *Institute of Mathematics and Informatics, BAS, Sofia
e-mail: zhelyazkova@fmi.uni-sofia.bg*

Bayesian spatial methods have been widely applied in different scientific areas such as epidemiological studies, image processing, fMRI data analysis and many others. We will apply Bayesian hierarchical model with conditionally autoregressive (CAR) prior to a collection of weekly mumps data from 2000–2008 in Bulgaria. We will generate a disease mapping of the crude standardized incidence ratio (SIR) across all regional centers. Similar mappings will also be produced for the smoothed relative risk. The combination of methods for estimates of the relative risk is a powerful tool to identify high risk regions and may be used to inform local authorities and programs.

Acknowledgements: This work is partially supported by the financial funds allocated to the Sofia University “St. Kl. Ohridski”, grant No 197/2016.