Detailed summary of the results of the research project *PD106181 Asymptotics of stochastic processes*

Péter Kevei

My research plan is divided to three main parts: classical probability, stochastic geometry and branching processes. In the fourth part I gathered some other related results.

Classical probability: self-normalized sums and related results

The study of self-normalized sums began with the work of Breiman, where he investigated the large sample behavior of ratios of the form $T_n = \sum_{i=1}^n X_i Y_i / \sum_{i=1}^n Y_i$, where $X, X_1, X_2, \ldots$ are independent, identically distributed random variables, and $Y, Y_1, Y_2, \ldots$ are non-negative, independent, identically distributed random variables, independent of the X's. Breiman shows that whenever the ratio converges in distribution for any distribution of X with finite mean, and at least one of the limit distributions is non-degenerate, then the distribution of Y is necessarily in the domain of attraction of a stable law with index less than 1. In the end of his paper Breiman conjectures that the statement remains true if it is only assumed that the limit distribution exists and is non-degenerate for one such X. When it is assumed that X has finite absolute moment greater than two, the conjecture was proved by Mason and Zinn in 2005.

With David Mason, we considered the limit distribution of the numerator and denominator jointly as a random vector. With this method in [5] we gave a surprising characterization of the continuity of the subsequential limits of T_n. An important role in our study is played by those Y that are in the *centered Feller class*. A random variable Y is said to be in the *centered Feller class* if there exist sequence of norming constants $\{a_n\}_{n \geq 1}$ such that if $Y_1, Y_2, \ldots$ are iid Y then for every subsequence there exists a further subsequence $\{n'\}$ such that $\sum_{i=1}^{n'} Y_i / a_{n'} \to W$, where W is a non-degenerate random variable. We showed that each subsequential limit law of T_n is continuous for any non-degenerate X_1 with finite expectation, if and only if Y_1 is in the centered Feller class.

In the continuous time version of the problem (U_t, V_t) is a bivariate Lévy process, where V_t is a subordinator and U_t is a Lévy process formed by randomly weighting each jump of V_t by an independent random variable X_t having distribution function F. In fact, it is the limit of the properly normalized vector $(\sum_{i=1}^{[nt]} X_i Y_i, \sum_{i=1}^{[nt]} Y_i)$, where $[\cdot]$ stands for the integer part, and X and Y are as above. We investigated the asymptotic distribution of the self-normalized Lévy process U_t/V_t at 0 and at ∞. In [6] we showed the analog of the discrete time result, i.e. that all subsequential limits of this ratio at 0 (∞) are continuous for any nondegenerate F with finite expectation if and only if V_t belongs to the centered Feller class at 0 (∞). We also characterized when U_t/V_t has a non-degenerate limit distribution at 0 and ∞ in the case when F has a finite second moment.

The key step in the proof is a distributional representation of the bivariate Lévy process, where the jumps correspond to homogeneous Poisson process. With the same method it is possible to investigate the asymptotic behavior of the jumps of a subordinator.

Let V_t be a driftless subordinator, and let denote $m_t^{(1)} \geq m_t^{(2)} \geq \ldots$ its jump sequence on interval $[0,t]$. Put $V_t^{(k)} = V_t - m_t^{(1)} - \ldots - m_t^{(k)}$ for the k-trimmed subordinator. Let denote Λ the Lévy measure of the subordinator, and put $\overline{\Lambda}(x) = \Lambda((x,\infty))$. In [7] we showed that for any $k \geq 0$ the limit distribution of the ratio $V_t^{(k)}/m_t^{(k+1)}$ exists as $t \downarrow 0$ ($t \to \infty$) if and only if $\overline{\Lambda}(x)$ is regularly varying at 0 (at ∞). Moreover, we determined the Laplace transform of the limit. The corresponding result for iid random variables is due to Breiman and Darling. In the same paper we also proved similar results for the existence of the limit $m_t^{(k+1)}/m_t^{(k)}$ as $t \downarrow 0$ or $t \to \infty$. As a special case we have shown the following. In order that the limit $m_k^{(k+1)}/m_t^{(k)}$ exist as $t \downarrow 0$ (∞) it is necessary and sufficient that $\overline{\Lambda}(x)$ is regularly varying at 0 (∞) with index $-\alpha$. In this case the limit distribution is the beta$(k\alpha, 1)$ law.

In the joint paper with Gábor Fukker and László Györfi [3] we investigated the behavior of generalized St. Petersburg sums conditioned on its maxima. We say that X is St. Petersburg(α, p) random variable if $\mathbf{P}(X = r^{k/\alpha}) = q^{k-1}p$ for $k \geq 1$, where $\alpha > 0$, $p \in (0,1)$, $q = 1-p$ and $r = 1/q$. Let $X, X_1, X_2, \ldots$ be iid St. Petersburg(α, p) random variables, put S_n and M_n for the partial sum and maximum, respectively. We determined the joint limiting behavior of the properly normalized version of (S_n, M_n). It was already known that $M_n/n^{1/\alpha}$ has limiting distribution along subsequences. We showed that conditioning on small maximum the limit distribution of the sum is Gaussian, for large maximum the limit is governed by the maximum, while for typical values the well-known semistable behavior appears. With this approach we obtained an infinite series representation for the limiting semistable distribution, where each of the summands have finite exponential moment. As another application of the approach we proved the 'almost subexponentiality' of the St. Petersburg distribution, i.e. for fixed n we determined the tail behavior of the sum S_n. It was already noted by Goldie that the St. Petersburg distribution is not subexponential. One of the many characterizing properties of the subexponential distributions is that $P(X_1 + \ldots + X_n > x) \sim nP(X_1 > x)$, where $X_1, \ldots, X_n$ are iid. We proved that this asymptotics almost holds, in fact fails only on small intervals, that is in a certain sense the distribution is almost subexponential.

Stochastic geometry

The following stochastic geometry problem was formulated by Bálint Tóth with K being the unit disc in the plane. Let $K = K_0$ be a convex body in $\mathbb{R}^d$ that contains the origin, and define the diminishing process (K_n, p_n), $n \geq 1$, as follows: let p_{n+1} be a uniform random point in K_n, and set $K_{n+1} = K_n \cap (p_{n+1} + K)$. Clearly, (K_n), $n \geq 1$, is a nested sequence of convex bodies which converges to a non-empty limit object, which is again a convex body in $\mathbb{R}^d$. What can we say about the distribution of this limit body, and how fast this happens? If K is the unit disc, the limiting object is almost surely a convex disc of constant width 1. It is an indication of the difficulty of the problem that almost nothing is known about these questions even in the plane.

In one dimension the process was handled by Ambrus, Kevei and Vígh. In [10] with Viktor Vígh we investigated this process for K being a regular simplex or a cube in any dimension, or a regular convex polygon in the plane with an odd number of vertices. Let K be a regular d-dimensional simplex with centroid $(0, 0, \ldots, 0)$ and vertices $(e_0, e_1, \ldots, e_d)$, such that $e_0 = (1, 0, \ldots, 0)$. Let denote $\rho_d = 1/d$ the radius of the inscribed sphere of K. We showed that in this case the limit object is a regular simplex and its center expressed in barycentric coordinates has a Dirichlet distribution, which is the generalization of the arcsine law. Moreover, we showed that $n^{1/d}(m_n - \rho_d)$ converges in distribution to a Weibull distribution, that is the speed of the process is $n^{-1/d}$. The cube process is just d independent copies of the interval process. In [10] we also investigated the regular polygon process in the plane. We showed the surprising result, that for square the speed of the process is n^{-1}, while for any odd-sided regular polygon the speed is $n^{-1/2}$. We also derived some new results in one dimension for non-uniform distributions.

Another closely related problem is the following. Consider the convex hull of n uniformly chosen points in a given convex body. The asymptotics of the expected number of vertices was determined by Rényi and Sulanke already in 1963.

In [2] with Ferenc Fodor and Viktor Vígh we generalized some of the classical results of Rényi and Sulanke in the context of spindle convexity. A planar convex disc S is R-spindle convex if it is the intersection of congruent closed circular discs with radius R. The intersection of finitely many congruent closed circular discs is called a disc-polygon. We prove asymptotic formulas for the expectation of the number of vertices, missed area and perimeter difference of uniform random disc-polygons contained in a sufficiently smooth spindle convex disc. In particular, we have proved that $Ef_0(S_n) \sim n^{1/3}\, C(S)$, where $f_0(S_n)$ is the number of vertices of the random disc-polygon, and the constant $C(S)$ depends only on S. We also showed that from our results the (linear) convex case can be deduced as $R \to \infty$. An interesting new feature in our results is the case when S is a disc. In this case the expected number of vertices converges to $\pi^2/2$ as $n \to \infty$. Roughly speaking this means that after choosing many random points from a circle, the spindle convex hull has few vertices. This phenomena has no counterpart in the (linear) convex case.

Branching processes

Consider an inhomogeneous Galton–Watson branching process with immigration, and assume that the offspring means are less than 1, yet tend to 1, that is the process is nearly critical. This model was first investigated by Györfi, Ispány, Pap and Varga in 2007 in the case when the offsprings have a Bernoulli distribution. The main idea is that the extinction effect coming from the subcritical branching mechanism can be balanced with the immigration in order to obtain a non-trivial limit distribution. Kevei in 2011 showed that these results can be extended to general offspring distributions.

In [4] with László Györfi, Márton Ispány and Gyula Pap we investigated the multitype case. In a multi-type homogeneous Galton–Watson process (without immigration) the main data of the process is the spectral radius ϱ of the mean matrix, which is its largest eigenvalue. By classical results, a positively regular, non-singular multi-type Galton–Watson process dies out almost surely if and only if $\varrho \leq 1$. The process is called subcritical, critical or supercritical if $\varrho < 1$, $= 1$ or > 1, respectively. In the multi-type setup we consider nearly-critical processes, that is we assume that the sequence of offspring mean matrices converge to a critical limit matrix. However, contrary to the one-dimensional case, there are a lot of critical matrices, and thus a lot of nearly-critical processes. Under general conditions we obtain limit distribution for the process, that is it is possible to balance the extinction effect with immigration. One of the interesting features of our limit theorem is that no scaling is necessary, the process converges itself. On the other hand the limit distribution can be rather complicated, the coordinates of the limit vector are not necessarily independent. The proof relies on quite explicit calculation with multivariate probability generating functions.

In [1] with Attila Dénes, Hiroshi Nishiura and Gergely Röst we investigate the probability of epidemic outbreaks. The European Centre for Disease Prevention and Control called the attention in March 2012 to the risk of measles in Ukraine among visitors to the 2012 UEFA European Football Championship. Large populations of supporters traveled to various locations in Poland and Ukraine, depending on the schedule of Euro 2012 and the outcome of the games, possibly carrying the disease from one location to another. In [1], we proposed a novel two-phase multitype branching process model with immigration to describe the risk of a major epidemic in connection with large-scale sports-related mass gathering events. By analytic means, we calculated the expected number and the variance of imported cases and the probability of a major epidemic caused by the imported cases in their home country. Applying our model to the case study of Euro 2012 we demonstrated that the results of the football games can be highly influential to the risk of measles outbreaks in the home countries of supporters. To prevent imported epidemics, it should be emphasized that vaccinating travelers would most efficiently reduce the risk of epidemic, while requiring the minimum doses of vaccines as compared to other vaccination strategies. Our theoretical framework can be applied to other future sport tournaments too.

Others

In [8] we show somewhat unexpectedly that whenever a general Bernstein-type maximal inequality holds for partial sums of a sequence of random variables then a maximal form of the inequality is also valid.

Let $X_1, X_2, \ldots$ be any sequence of random variables, and denote $S(k,l) = \sum_{i=k}^{l} X_i$ the partial sum and $M(k,l) = \max\{|S(k,k)|, \ldots, |S(k,l)|\}$ the partial maximum. It turns out that under a variety of assumptions the partial sums $S(k,l)$ satisfy a generalized Bernstein-type inequality of the following form: for suitable constants $A > 0$, $a > 0$, and for suitable g_n for all $m \geq 0$, $n \geq 1$ and $t \geq 0$,

$$\mathbf{P}\{|S(m+1, m+n)| > t\} \leq A \exp\left\{-\frac{at^2}{n + g_n(t)}\right\}.$$

We showed that in this case the same inequality holds with $M(m+1, m+n)$ instead of $S(m+1, m+n)$, possibly with different constants A and a.

In [9] we prove strong approximation results to time dependent empirical and quantile processes. We define a time dependent empirical process based on n iid fractional Brownian motions and establish Gaussian couplings and strong approximations to it by Gaussian processes. They lead to functional laws of the iterated logarithm for this process.

References

[1] A. Dénes, P. Kevei, H. Nishiura, G. Röst: Risk of infectious disease outbreaks by imported cases with application to the European Football Championship 2012, *International Journal of Stochastic Analysis*. Volume 2013 (2013), Article ID 576381, 9 pages.

[2] F. Fodor, P. Kevei, V. Vígh: On random disc-polygons in smooth convex discs. *Advances in Applied Probability*, **46**, 2014.

[3] G. Fukker, L. Györfi, P. Kevei: Asymptotic behavior of the St. Petersburg sum conditioned on its maximum. Submitted.

[4] L. Györfi, M. Ispány, P. Kevei, Gy. Pap: Asymptotic behavior of a multitype nearly critical Galton–Watson processes with immigration. *Teoriya Veroyatnostei i ee Primeneniya*, to appear.

[5] P. Kevei, D.M. Mason: The Asymptotic Distribution of Randomly Weighted Sums and Self-normalized Sums. *Electronic Journal of Probability*, **17**, no. 46, 2012, 1–21.

[6] P. Kevei, D.M. Mason: Randomly Weighted Self-normalized Lévy Processes. *Stochastic Processes and their Applications*, **123** (2) 2013, 490–522.

[7] P. Kevei, D.M. Mason: The limit distribution of ratios of jumps and sums of jumps of subordinators. Submitted.

[8] P. Kevei, D.M. Mason: A More General Maximal Bernstein-type Inequality, In: *High dimensional probability VI: The Banff Volume*. Progress in Probability, Vol. 66, Houdré, C.; Mason, D.M.; Rosiński, J.; Wellner, J.A. (Eds.), Springer, pp. 55–62, 2013.

[9] P. Kevei, D.M. Mason: Strong Approximations to Time Dependent Empirical and Quantile Processes Based on Independent Fractional Brownian Motions. Submitted.

[10] P. Kevei, V. Vígh: On the diminishing process of B. Tóth. Submitted.