Transactions of A. Razmadze
Mathematical Institute
Vol. **174** (2020), issue 2, 127–134

HIGHLIGHTS OF RESEARCH WORK OF ESTATE KHMALADZE

ROBERT M. MNATSAKANOV

The goodness of fit problem and scanning martingales. Here we briefly outline one of the central problems of mathematical statistics, the difficulties which remained open there from the mid 50s to the early 80s and the way they were overcome using very unexpected "martingale approach" developed by Khmaladze (1981), as well as its nontrivial extension to the case of multidimensional time.

Let F_n be an empirical distribution function of a sequence of n independent scalar random variables with distribution function F. The normalised difference $\sqrt{n}(F_n - F) = v_n$, one of the basic processes in statistical theory, is called the empirical process.

Certainly, both the distribution of the process v_n for finite n and its limiting distribution as $n \to \infty$ depend on F (the limiting process v is called F-Brownian bridge). However, the amazingly simple Kolmogorov's transformation (Kolmogorov (1933)), $u_n = v_n \circ F^{-1}$ with the condition that F is continuous, maps v_n into the so-called uniform empirical process u_n with the standard (and independent of F) distribution. This opens an extremely important possibility to use asymptotic theory for u_n *only* in asymptotic statistical inference concerning *any* continuous F. This "asymptotic distribution freeness" of u_n became one of the basic facts in nonparametric statistics and in the theory of the so called goodness of fit tests.

In the late 1950's and early 1960's it was discovered (see Kac, Kiefer and Wolfowitz (1955) or Gikhman (1953)) that in most practical cases, where $F = F_\theta$ is known only up to a finite dimensional parameter θ to be estimated from the data, the process $v_{n,\hat{\theta}} = \sqrt{n}(F_n - F_{\hat{\theta}})$ has the asymptotic distribution not only different from that of F-Brownian bridge, but also such that $v_{n,\hat{\theta}} \circ F_{\hat{\theta}}^{-1}$ is no longer asymptotically distribution free. The bibliography on this subject is huge; one review paper is Durbin (1973). Chibisov (1971) and Moore and Spruill (1975) demonstrated that the chi-square statistics with estimated parameter is, in general, also not asymptotically chi-square distributed. All developments in the late 1960's and throughout the 1970's persuaded statisticians that this was an unavoidable complication which needed to be lived with.

However, Khmaladze's paper (1981) changed this stereotype completely. It was shown that using a different point of view on v_n, a transformation of $v_{n,\hat{\theta}}$ can be found, $w_{n,\hat{\theta}} = \sqrt{n}(F_n - K_{\hat{\theta},n})$, which converges to F_θ-Brownian motion, and therefore $w_{n,\hat{\theta}} \circ F_{\hat{\theta}}^{-1}$, is asymptotically a standard Brownian motion on $[0,1]$, and thus asymptotically distribution free. The process $w_{n,\hat{\theta}}$ can be thought of as the innovation martingale of the process $v_{n,\hat{\theta}}$ with respect to the natural filtration of the later. In this way, the whole beauty and usefulness of asymptotically distribution free procedures were restored.

Further work of E. Khmaladze developed similar transformations in the difficult case of empirical processes based on multidimensional random variables. As is know, the existing theory of martingales in multidimensional time is complicated and involves restrictive conditions, not satisfied by many important processes with multidimensional time (cf., e.g., Wong and Zakai (1974), Cairoli and Walsh (1975), Hajek and Wong (1980), Gikhman (1982), Nualart (1983)). Therefore a new approach to the stochastic calculus of Gaussian random processes with multidimensional time was required. The notion of "scanning martingales", suggested by Khmaladze (1988a, 1993), provides such an approach and leads to an elegant and simple theory of innovation martingales in multidimensional spaces. In Khmaladze (1988a), the general goodness of fit problem for simple hypothesis was formulated for the first time, and Khmaladze (1993) gives the solution of this problem in a completely general setting and opens a way to distribution free "model testing" in multidimensional spaces, the possibility, previously nonexistent in the statistical theory.

Mathematically, the paper establishes new connections between the goodness of fit problem of statistics and empirical processes with functional time on one side, and the theory of stochastic differential equations for measure-value processes and Volterra decompositions of Hilbert-Schmidt operators on the other side.

The paper of Einmahl and Khmaladze (2001) gives a similar solution in $\mathbb{R}^d$ for another classical problem of statistics, the so- called two-sample problem in a multidimensional space.

Sequential ranks. The sequential rank S_k of a random variable X_k is its rank among random variables $X_1, X_2, \ldots, X_k$ as compared with the "ordinary" rank R_{kn} which is the rank of X_k amongst all n "available" random variables $X_1, X_2, \ldots, X_k, \ldots, X_n$, with $k \leq n$. Sequential ranks are practically very convenient when observations arrive one-by-one. However, their asymptotic theory meets with certain difficulties; it was not known how to study the efficiency of statistical procedures based on sequential ranks. Consequently, this theory fell into disuse, whereas the theory of "ordinary" ranks received considerable attention in the 1960's through to the 1980's. For example, in Sen (1978), although primarily devoted to sequential problems, it was necessary for the authors to work with "ordinary" ranks, which are inconvenient in this setting, because asymptotic methods were unavailable for sequential ranks.

On the other hand, if $X_1, X_2, \ldots, X_n$ are independent and identically distributed, the sequential ranks are of very nice behaviour: $S_1, S_2, \ldots, S_n$ are independent and each S_k has uniform distribution on integers $1, 2, \ldots, k$.

The difficulties connected with efficiency of the tests based on sequential ranks were overcome in the papers of Khmaladze and Parjanadze (1986), Pardzhanadze and Khmaladze (1986), which established an asymptotic theory of sequential ranks in the same basic framework as the existing theory for "ordinary" ranks. This was possible due to the development of asymptotic methods, not normally applied in the theory of rank statistics. Namely, it was shown that the partial sum processes based on (functions of) "ordinary" ranks and those based on sequential ranks may be asymptotically connected through a linear stochastic differential equation and hence the properties of one can be carried over into the properties of another.

In particular, it was shown that asymptotic distributions of linear statistics

$$\sum_{k=1}^{n} c_k a(R_{kn}/n)$$

from "ordinary" ranks and linear statistics

$$\sum_{k=1}^{n} \Bigg(c_k - \sum_{m \leq k} c_m/k \Bigg) a(S_k/k)$$

from sequential ranks have the same asymptotic distribution under *all* contiguous alternatives. Equivalently, statistics

$$\sum_{k=1}^{n} c_k a(S_k/k) \quad \text{and} \quad \sum_{k=1}^{n} \Bigg(c_k - \sum_{m \geq k} c_m/m \Bigg) a(R_{kn}/n)$$

have the same limit distributions under all contiguous alternatives. Thus, whenever one of them is optimal against some contiguous alternative, the other is also optimal for the same alternative.

Multinomial distributions of increasing dimension. The research in this field may be of interest to the colleagues in discrete mathematics.

The sequence of multinomial distributions $\mathcal{M}(\cdot, p_N, n)$, where $p_N = \{p_{in}\}_1^N$, $p_{in} > 0$ and $\sum\limits_1^N p_{in} = 1$, which have the number of different possible outcomes $N = N_n$ increasing with number of trials n, form a surprisingly rich class of distributions. They reflect and illustrate a very large number of interesting problems found in other parts of statistics, such as:

• the asymptotic behaviour and properties of statistics like the classical χ^2-statistic when $N_n \to \infty$ as $n \to \infty$ are sharply different from those which one can deduce when first letting $n \to \infty$ and then $N \to \infty$;

• statistical problems with increasing numbers of parameters, like the problem of estimating spectra of matrices of increasing dimension or problems with "fine" partitions, are very similar to what one meets in the asymptotic analysis of $\mathcal{M}(\cdot, p_N, n)$ — the normalised probabilities np_{in}, $i = 1, \ldots, N_n$, are these parameters;

• the class of asymptotic laws of which the famous Zipf - Mandelbrot's law is the most remarkable representative, are highly connected with the sequences $\mathcal{M}(\cdot, p_N, n)$ where $n \to \infty$. We comment on Zipf – Mandelbrot's law separately below.

The paper of Khmaladze (1983) completely modified the tools and approaches used in this field. Instead of considering sums, called "divisible statistics",

$$\sum_{i=1}^{N_n} g(\nu_{in}, np_{in})$$

and limit theorems for each sum, the paper studied partial sums

$$\sum_{i=1}^{k} g(\nu_{in}, np_{in}),\ k = 1, 2, \ldots, N_n,$$

and derived limit theorems for these processes. They are treated as semimartingales associated not with its natural filtration, but with richer filtration $\mathcal{F}_k = \sigma\{\nu_{1n}, \ldots, \nu_{kn}\}$, $k = 1, 2, \ldots, N_n$, based on underlying frequencies. The point of it is that the conditional distribution of ν_{in} given previous frequencies ν_{jn}, $j = 1, 2, \ldots, i-1$, is much simpler object, than conditional distribution of $g(\nu_{in}, np_{in})$ given previous summands $g(\nu_{jn}, np_{jn})$, $j = 1, 2, \ldots, i-1$.

It showed how new at a time limit theorems for semimartingales could be utilised and lead to general functional limit theorems for the basic statistics of the field – the so-called additively divisible statistics (statistics of increasing numbers of small, separate frequencies). The paper demolished an unnecessary partition between different parts of asymptotic statistics (for a better picture, see Ivchenko and Levin's review paper (1996)). It led to similar advances in the theory of general spacings (see Borovikov (1987)) and in the analysis of the so-called "very rare events" (see Mnatsakanov (1985) or Prakasa Rao (1987)).

Large number of rare events (LNRE) theory. The text of Dante's "Divina Comedia" is in length some 100,000 words. Approximately 13,000 of these words are different, that is, the vocabulary of "Divine Comedia" is only 13,000 words. It would be, however, very incorrect to suppose that each word was used by Dante approximately 8 times. There is certainly nothing like an even usage of words, few words were used hundreds of times, while about 6,000 words (half the vocabulary) were used only once and about 2,000 words were used only twice.

This is the typical situation in a surprisingly large number of applied statistical problems, not only in all sorts of large texts, but also in studies of the number of species in an environment, opinions in a survey, chemical analysis, income distributions, distribution of languages, etc.

According to Zipf's law, if $\mu_n(m)$ is the number of words (species, opinions, etc.) which occurred m times in a sample of size n and if μ_n is the number of all different words (species, opinions, etc.) in the same sample, then

$$\frac{\mu_n(m)}{\mu_n} \to \frac{1}{m(m+1)} \quad \text{as} \quad n \to \infty.$$

Its slightly modified form

$$\frac{\mu_n(m)}{\mu_n} \to \frac{1}{(a+bm)^q} \quad \text{as} \quad n \to \infty$$

is called Zipf - Mandelbrot's law. In the words of Mandelbrot (1953), "The form of Zipf's law is so striking and also so very different from any classical distribution of statistics that it is quite widely felt that it "should" have a basically simple reason, possibly as "weak" and general as the reasons which account for the role of Gaussian distribution. But, in fact, these laws turn out to be quite resistant to such an analysis. Thus, irrespective of any claim as to their practical importance, the "explanation" of their role has long been one of the best defined and most conspicuous challenges to those interested in statistical laws of nature".

The present interest in Zipf's law is, perhaps, characterised by the increase of interest in the more general concept of LNRE distributions, which was introduced and first systematically studied in an unpublished paper of Khmaladze (1988b), partly reproduced by Khmaladze and Chitashvili (1989) and called and treated as "fundamental" in the monograph of Baayen (2001).

"Chimeric" contiguous alternatives. The theory of contiguity of probability measures is a main tool in asymptotic statistics to study the efficiency and power of statistical procedures. Contiguous distributions or contiguous alternatives (to a given distribution) form a class of alternatives which are, heuristically speaking, most difficult to detect. It is well known that if $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are two distributions, then n-fold direct products $\mathbb{P}^{(n)}$ and $\tilde{\mathbb{P}}^{(n)}$ are either asymptotically singular as $n \to \infty$ or coincide (alternative of Kakutani). In order for $\tilde{\mathbb{P}}^{(n)}$ to be contiguous to $\mathbb{P}^{(n)}$, the distribution $\tilde{\mathbb{P}}$ must depend on n in such a way, basically, that

$$\left(\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right)^{1/2} = 1 + \frac{1}{\sqrt{n}} h_n \quad \text{with } \limsup_{n\to\infty} \|h_n\|_{L_2(\mathbb{P})} < \infty$$

(see, e.g., Oosterhof and van Zwet (1979)). Practically all papers which use contiguity theory replace the latter condition by $h_n \overset{L_2(\mathbb{P})}{\longrightarrow} h$ and for very good reasons. Nevertheless, the paper of Khmaladze (1998) studies such contiguous alternatives that $\|h_n\| \geq 1$, but $h_n \overset{\omega}{\to} 0$, that is, h_n has no limiting points in $L_2(\mathbb{P})$. It is clear that no classical goodness of fit test based on empirical process can detect any such "chimeric" alternative. Yet the paper of Khmaladze (1998) shows that new versions of empirical processes can be constructed and a goodness of fit theory can be developed which is no less rich than that which exists for converging contiguous alternatives.

The paper also shows that although they look exotic, "chimeric" alternatives can frequently be found in real problems. After all, the existence of our civilisation is itself an enormous "chimeric" alternative.

Change-set problem (spacial change-set problem). The idea of transferring the range of problems usually unified by the term "change-point problem" for real line to finite-dimensional Euclidean space was entertained and discussed by E. Khmaladze at the end of 80-ies, while still in Moscow, in particular, at the Moscow Seminar of Young Statisticians. But he started working himself only in 1996.

With the help of the concept of the local covering numbers, the papers of Khmaladze, Mnatsakanov and Toronjadze (2006a, 2006b) investigated the convergence of statistical estimators of the change-set and finally obtained the correct rate of n^{-1}. This is the rate of convergence of what is called "super-efficient" estimators in statistics. The smoothness of the boundaries were not required – only that the class of possible change-sets was locally compact.

Differentiation of set-valued functions. The research work of E. Khmaladze here has a story. Heuristically, the idea came from the work on the change-set problem. In this problem, the object of interest is a set, say $A \in \mathbb{R}^d$ as a hypothesis, and a sequence of sets $B_n \in \mathbb{R}^d$, converging in Hausdorff metric to A as a sequence of contiguous alternatives. What a statistician observes is a point process N_n in $\mathbb{R}^d$, and, as $n \to \infty$, the intensity of this point process increases, so that there appear more and more points, and symmetric difference $A\Delta B_n$ shrinks towards the boundary of A at the same time. Since the number of points increases, it is not necessary that their number in $A\Delta B_n$ decreases to zero. In the most interesting cases this number becomes a Poisson random variable. So, the corresponding points do not disappear; but where do they eventually "live"? The first intuition was that they must "live" on the boundary of A. But later the feeling grew that this should not be true. Some sort of "differentiation" seems to be lurking behind the scenes.

All this were talks and guesses, and some reading for several years. The actual work started after 2004. In Karlsruhe, a very good colleague and very highly regarded geometer, Wolfgang Weil, came one day to the small temporary library room, where Estate was accommodated, and put down the famous book of J-P. Aubin and E. Frankowska (1990) on a set-valued analysis. The book contained two chapters on differentiation of set-valued functions. And this meant that secure, but not very original work, lied ahead. W. Weil was not too enthusiastic about this prospect. And in spite of joint paper of Khmaladze and Weil (2008), which was then the work in progress, Estate was on his own.

A year later, the draft of the paper on differentiation of set-valued functions was ready and in 2007 the paper Khmaladze (2007) was published. One short corollary of the new notion of a derivative is the following:

• if, as $t \to 0$, the symmetric difference $A\Delta B_t$ is differentiable at the boundary ∂A, then for any absolutely continuous distribution $\mathbb{P}$ in $\mathbb{R}^d$ there exists an absolutely continuous distribution $\mathbb{Q}$ on the normal cylinder of ∂A such that

$$\frac{d}{dt}\mathbb{P}(A\Delta B_t) = \mathbb{Q}\Big(\frac{d}{dt}B_t\Big).$$

Later, a review article with W. Weil was invited to the Annals of the Institute of Statistical Mathematics, Khmaladze and Weil (2018), where the derivative was given a name of "fold-up derivative" and was defined in more general class of situations. Before that, a paper with John Einmahl established CLT for the point process N_n on classes of sets in shrinking neighbourhoods of ∂A (see Einmahl and Khmaladze (2011)).

In a personal letter to E. Khmaladze, J.-P. Aubin calls the work of fold-up derivatives a "mathematical virtuosity".

Questionnaires – the problem of diversity in spaces of increasing dimension. A person is asked q binary questions: "yes" or "no". The person fills in this questionnaire and this is one "opinion". Altogether 2^q different opinions are possible. Lots of people, N, are asked to fill this questionnaire. So, there are lots of questionnaires with many possible opinions expressed in them. Exactly, how many different opinions will be found in the sample? How many opinions will be unique? These and all other similar questions have been answered by Khmaladze (2011). But the answers did not come without surprise also for the author.

Imagine, again, that we are in $[0,1]^q = [0,1] \times \cdots \times [0,1]$, and we divide the first interval $[0,1]$ in proportion $a_1 : 1 - a_1$, the second $[0,1]$ as $a_2 : 1 - a_2$, and so on. In this way one will obtain 2^q elementary cubes. In one-dimensional space, each of subintervals $[0, a_1]$ and $[a_1, 1]$ is divided as $a_2 : 1 - a_2$, then each of the resulting four are divided as $a_3 : 1 - a_3$, and so on, q times. The first impression was that this would be some other version of random partition of a "stick" into 2^q subintervals: if $0 < U_1 < U_2 < \cdots < U_{2^q-1} < 1$ are uniformly distributed random variables, arranged in increasing order, then the spacings $[U_i, U_{i+1}]$ are forming this random partition. If we now throw N random points on $[0,1]$, or in $[0,1]^q$, and count frequencies of these points in each subinterval, or small cubes, what will be their behaviour?, how many subintervals will remain empty?, how many will contain just one point?, etc. For a random partition, the answers are more or less known and initially Estate wanted an analogue of this.

However, the behaviour of these frequencies turn out to be very different. Very uneven. Behaviour of spacings is, strictly speaking, also "uneven", but not so far from being even. But sizes of intervals, or volumes of cubes, obtained through these a_i-s are sharply uneven. And behaviour of the frequencies of random points in them, consequently, is also uneven. First of all, the number of cells with some filling turns to be $o(N)$, i.e., much smaller than the number of points thrown; or the number of different opinions is much smaller, than the number of persons asked. The fraction of cells with one, two, and in general k points in them, relative to the number of all non-empty cells, follows some "law", which "almost" does not depend on the choice of a_i-s. Yet, this law is not the famous Zipf's law, which many of us could have heard about.

A complete description of the situation is given by Khmaladze (2011) and partly in the 18-years earlier paper of Khmaladze and Tsigroshvili (1993). This was a strong step forward within the theory of diversity and occupation problem. Division in more than two subintervals at each step is a fascinating problem for the future.

Unitary operators. The last several years a new development took place in the direction of distribution free testing theory. The main idea can be explained as follows.

The empirical process with estimated parameter $v_{n,\hat{\theta}}$, or estimated empirical process for short, is not just a process with different limit distribution from the empirical process v_n, it has the specific structure – its limit distribution is that of the projection of F_θ-Brownian bridge, orthogonal to the score function $\dot{f}_\theta/f_\theta$. Thus the distribution of the projected F_θ-Brownian bridge is dependent on this score function. This asymptotic phenomenon was first described by Khmaladze (1979). It implies that for any regular parametric family of distributions $G_\theta, \theta \in \Theta$, as a limit of the estimated empirical process, one will obtain again a projection of G_θ-Brownian bridge, orthogonal to a corresponding score function. However, if the dimension of parameters in both families is the same, than with the help of unitary mappings one projection can be mapped to another projection thus rendering the two testing problems equivalent, in the sense that one can be transformed into other and the other way around.

Convenient framework for application of operators on empirical processes is provided by the function-parametric version of empirical processes

$$v_{n,\hat{\theta}}(\phi) = \int \phi(x) v_{n,\hat{\theta}}(dx), \quad \phi \in L_2(F_\theta),$$

because then the operator U on $v_{n,\hat{\theta}}$ can be naturally defined as the adjoint operator U^* on $L_2(F_\theta)$:

$$(U v_{n,\hat{\theta}})(\phi) = v_{n,\hat{\theta}}(U^* \phi).$$

However, this notion of equivalence creates very wide classes of equivalence, and in each class one needs only one representative, for which the distributional work for test statistics should be carried through; this is no different to assuming that the sample came from uniform distribution while testing simple hypothesis.

The projections, as a result of estimation of parameters, are ubiquitous. They appear in situations where so far nobody considered testing problems. Estimation of parameters - yes, but not testing, in particular, not goodness of fit testing. From the families of discrete distributions, for which the goodness of fit testing theory appeared only in 2013 (see Khmaladze (2013)), to empirical processes in regression, and now testing models for point processes (see the article of Khmaladze (2020) in this issue), testing parametric hypothesis for Markov chains and for Markov diffusion processes, like the Ornstein – Uhlenbeck process, all are work in progress.

In this account of scientific contribution of Estate Khmaladze in statistics and stochastic models we do not comment on the other fields of his research such as

– kernel density estimators,
– asymptotic of non-crossing probabilities with moving boundaries,
– formulation of the strong law of large numbers for Voronoi tessalation,
– extreme value theory and record processes

and various others. One can find these, e.g., in Mnacakanov and Hmaladze (1981), Kotel'nikova and Khmaladze (1982), Khmaladze, Nadareishvili and Nikabadze (1997), Khmaladze and Shinjikashvili (2001), Khmaladze and Toronjadze (2001) (see also Schneider and Weil (2008)), Can, Einmahl, Khmaladze and Laeven (2015).

References

1. J. -P. Aubin, H. Frankowska, *Set-valued Analysis*. Birkhäuser Boston, MA, 1990.
2. R. H. Baayen, *Word Frequency Distributions*. Kluwer Academic Publishers, Dordrecht, 2001.
3. V. P. Borovikov, Some limit theorems for statistics which are partial sums of functions of spacings. *Theory Probab. Appl.* **32** (1987), no. 1, 86–97.
4. R. Cairoli, J. B. Walsh, Stochastic integrals in the plane. *Acta Math.* **134** (1975), 111–183.
5. S. U. Can, J. H. J. Einmahl, E. V. Khmaladze, R. J. A. Laeven, Asymptotically distribution-free goodness-of-fit testing for tail copulas. *Ann. Statist.* **43** (2015), no. 2, 878–902.
6. D. M. Chibisov, Certain chi-square type tests for continuous distributions. *Theory Probab. Appl.* **16** (1971), no.1, 1-22.
7. J. Durbin, *Distribution Theory for Tests Based on the Sample Distribution Function.* Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, no. 9. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973.
8. J. H. J. Einmahl, E. Khmaladze, The two-sample problem in $\mathbb{R}^m$ and measure-valued martingales. *IMS Lecture Notes Monogr. Ser.* **36** (2001), 434-463.
9. J. H. J. Einmahl, E. Khmaladze, Central limit theorems for local empirical processes near boundaries of sets. *Bernoulli* **17** (2011), no. 2, 545–561.

10. J. I. Gikhman, Some remarks on A. N. Kolmogorov's criterion of fit. (Russian) *Doklady Akad. Nauk SSSR* **91** (1953), 715–718.
11. J. I. Gikhman, Biparametric martingales. (Russian) *Uspekhi Mat. Nauk* **37** (1982), 215–237.
12. B. Hajek, E. Wong, Representation and transformation of two-parameter martingales under a change of measure. *Z. Wahrsch. Verw. Gebiete* **54** (1980), no. 3, 313–330.
13. G. I. Ivchenko, V. V. Levin, Processes in discrete schemes (A survey, Part I and II). *Math. Methods Statist.* **2** (1993), no. 1, 72–84; no. 2, 115–129.
14. M. Kac, J. Kiefer, J. Wolfowitz, On tests of normality and other tests of goodness of fit based on distance methods. *Ann. Math. Statist.* **26** (1955), no. 2, 189–211.
15. E. V. Khmaladze, The use of ω^2-tests for testing parametric hypotheses. *Theory Probab. Appl.* **24** (1979), no. 2, 283–301.
16. E. V. Khmaladze, A martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26** (1981) no. 2, 240–257.
17. E. V. Khmaladze, Martingale limit theorems for divisible statistics, *Theory Probab. Appl.* **28** (1983), no. 3, 530–549.
18. E. V. Khmaladze, A. M. Parjanadze, Functional limit theorems for linear statistics from sequential ranks. *Probab. Theory Related Fields* **73** (1986), no. 4, 585–595.
19. E. V. Khmaladze, An innovation approach to goodness-of-fit tests in $\mathbf{R}^m$. *Ann. Statist.* **16** (1988a), no. 4, 1503–1516.
20. E.V. Khmaladze, The statistical analysis of a large number of rare events. *Technical Report MS-R8804, CWI, Amsterdam*, 1988b.
21. E. V. Khmaladze, R. Ya. Chitashvili, Statistical analysis of a large number of rare events and related problems. (Russian) Probability theory and mathematical statistics. (Russian) *Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR* **92** (1989), 196–245.
22. E. V. Khmaladze, Goodness of fit problem and scanning innovation martingales. *Ann. Statist.* **21** (1993), no. 2, 798–829.
23. E. V. Khmaladze, Z. P. Tsigroshvili, On polynomial distributions with a large number of rare events. *Math. Methods Statist.* **2** (1993), no. 3, 240–247.
24. E. Khmaladze, M. Nadareishvili, A. Nikabadze, Asymptotic behaviour of a number of repeated records. *Statist. Probab. Letters* **35** (1997), no.1, 49-58.
25. E. V. Khmaladze, Goodness of fit tests for "himeric" alternatives. *Statist. Neerlandica* **52** (1998), no. 1, 90–111.
26. E. Khmaladze, E. Shinjikashvili, Calculation of noncrossing probabilities for Poisson processes and its corollaries. *Adv. in Appl. Probab.* **33** (2001), no. 3, 702–716.
27. E. Khmaladze, N. Toronjadze, On almost sure coverage property of Voronoi tessellation: The $\mathbf{R}^1$ case. *Adv. Appl. Probab.* **33** (2001), no. 4, 756–764.
28. E. Khmaladze, R. Mnatsakanov, N. Toronjadze, The change-set problem for Vapnik-Červonenkis classes. *Math. Methods Statist.* **15** (2006a), no. 2, 224–231.
29. E. Khmaladze, R. Mnatsakanov, N. Toronjadze, The change set problem and local covering numbers. *Math. Methods Statist.* **15** (2006b), no. 3, 289–308.
30. E. V. Khmaladze, Differentiation of sets in measure. *J. Math. Anal. Appl.* **334** (2007), no. 2, 1055–1072.
31. E. V. Khmaladze, Wolfgang Weil, Local empirical processes near boundaries of convex bodies. *Ann. Inst. Statist. Math.* **60** (2008), no. 4, 813–842.
32. E. V. Khmaladze, Convergence properties in certain occupancy problems including the Karlin-Rouault law. *J. Appl. Probab.* **48** (2011), no. 4, 1095–1113.
33. E. V. Khmaladze, Note on distribution free testing for discrete distributions. *Ann. Statist.* **41** (2013), no. 6, 2979–2993.
34. E. Khmaladze, W. Weil, Fold-up derivatives of set-valued functions and the change-set problem: a survey. *Ann. Inst. Statist. Math.* **70** (2018), no. 1, 1–38.
35. E. V. Khmaladze, Distribution free testing for linear regression. Extension to general parametric regression. *Ann. Inst. Statist. Math.* (2020) (under review).
36. A. N. Kolmogorov, Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn.* **4** (1933), 83–91.
37. V. F. Kotel'nikova, E. V. Khmaladze, Calculation of the probability of an empirical process not crossing a curvilinear boundary. (Russian) *Teor. Veroyatnost. i Primenen.* **27** (1982), no. 3, 599–607.
38. B. Mandelbrot, An informational theory of the statistical structure of language. *Communication theory* **84** (1953), 486–502.
39. R. M. Mnacakanov, E. V. Hmaladze, On L_1-convergence of statistical kernel estimators of distribution densities. *Soviet Math. Dokl.* **23** (1981), no. 3, 633-636.
40. R. M. Mnatsakanov, A functional limit theorem for additively separable statistics in the case of very rare events. *Theory Probab. Appl.* **30** (1986), no. 3, 622–626.
41. D. S. Moore, M. C. Spruill, Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.* **3** (1975), 599–616.
42. D. Nualart, Two-parameter diffusion processes and martingales. *Stochastic Process. Appl.* **15** (1983), no. 1, 31–57.
43. J. Oosterhoff, W. van Zwet, A Note on Contiguity and Hellinger Distance. In: *Contributions to Statistics*, pp. 157–166, Reidel, Dordrecht-Boston, Mass.-London, 1979.

44. A. M. Pardzhanadze, E. V. Khmaladze, On the asymptotic theory of statistics of sequential ranks. *Theory Probab. Appl.* **31** (1986), no. 4, 669-682.
45. B. L. S. Rao Prakasa, *Asymptotic Theory of Statistical Inference.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1987.
46. R. Schneider, W. Weil, *Stochastic and Integral Geometry.* Springer, New York, 2008.
47. P. K. Sen, Invariance principles for linear rank statistics revisited. *Sankhya Ser. A* **40** (1978), no. 3, 215–236.
48. E. Wong, M. Zakai, Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrsch. Verw. Gebiete* **29** (1974), 109–122.

Department of Mathematics, West Virginia University, P.O. Box 6310, Morgantown, WV 26506, USA

A. Razmadze Mathematical Institute of I. Javakhishvili Tbilisi State University, 6 Tamarashvili Str., Tbilisi 0177, Georgia

E-mail address: Robert.Mnatsakanov@mail.wvu.edu