Theory of Stochastic Processes
Vol. 12 (28), no. 1–2, 2006, pp. 87–93

UDC 519.21

ALEXEY M. KULIK

MARKOV APPROXIMATION OF STABLE PROCESSES BY RANDOM WALKS

The notion of the Markov approximation is introduced. This notion is illustrated in the frameworks of the multidimensional functional CLT with a normal domain of attraction and the functional CLT with a stable domain of attraction.

Introduction

The famous A.V.Skorokhod' s *method of one probability space* is based on the following result: if a sequence of random elements $\{X^n\}$ in some Polish space $\mathcal{X}$ converges in distribution to an element X, then there exist a probability space $(\Omega, \mathcal{F}, P)$ and a two-component sequence (Y^n, Z^n) on it such that

$$Y^n \stackrel{d}{=} X^n, \quad Z^n \stackrel{d}{=} X, \quad \rho_{\mathcal{X}}(Y^n, Z^n) \stackrel{P}{\longrightarrow} 0, \quad n \to \infty. \tag{1}$$

If $\mathcal{X}$ is some functional space, say, $D(\Re^+)$ or some space of sequences, and the elements X^n, X are endowed by some Markov structure, it is natural to study the question how the method of one probability space interferes with this structure. For one possible result in this direction obtained in the context of the ergodic theorem for Markov chains, see [1]. In this paper, we study the question which is interesting by itself and also has non-trivial applications: suppose that the elements X^n, X, considered as processes on $\Re^+$, possess Markov properties w.r.t. their canonical σ-algebras at some points $\{t_n^k\} \subset \Re^+$. Is it then possible to construct the two-component processes (Y^n, Z^n) satisfying (1) in such a way that every of them possesses Markov property w.r.t. their canonical σ-algebras at these points? If it is so, then it is natural to say that the sequence $\{X_n\}$ provides *the Markov approximation* for the process X.

1. Main results

Let us formulate the main definition. The processes below are defined on $\Re^+$ and have the same locally compact phase space $(\mathcal{X}, \rho)$.

Definition 1. *The sequence of the processes $\{X_n\}$ provides the Markov approximation for the Markov process X, if, for every $\gamma > 0, T < +\infty$, there exist a constant $K(\gamma, T) \in \mathbb{N}$ and a sequence of two-component processes $\{\hat{Y}_n = (\hat{X}_n, \hat{X}^n)\}$, defined on another probability space, such that*

(i) $\hat{X}_n =^d X_n, \hat{X}^n =^d X$;

(ii) the process $\hat{Y}_n$ and the processes $\hat{X}_n, \hat{X}^n$ possess the Markov property at the points $\frac{iK(\gamma,T)}{n}, i \in \mathbb{N}$ w.r.t. the filtration $\{\hat{\mathcal{F}}_t^n = \sigma(\hat{Y}_n(s), s \le t)\}$;

2000 *AMS Mathematics Subject Classification.* Primary 60F17, 60J25.

Key words and phrases. Markov approximation, functional CLT, domain of attraction.

This article was supported (in part) by the Ministry of Education and Science of Ukraine, project N 01.01/103.

(iii) $\limsup_{n\to+\infty} P\left(\sup_{i\le \frac{Tn}{K(\gamma,T)}} \rho\left(\hat{X}_n\left(\frac{iK(\gamma,T)}{n}\right), \hat{X}^n\left(\frac{iK(\gamma,T)}{n}\right)\right) > \gamma\right) < \gamma.$

Remark 1. Condition (ii) implies that, for every $i \in \mathbb{N}, t > \frac{iK(\gamma,T)}{n}, (x,y) \in \mathcal{X}^2$ the marginal distributions of $P(\hat{Y}_n(t) \in \cdot|\hat{Y}_n(\frac{iK(\gamma,T)}{n}) = (x,y))$ coincide with $P(X_n(t) \in \cdot|X_n(\frac{iK(\gamma,T)}{n}) = x)$ and $P(X(t) \in \cdot|X(\frac{iK(\gamma,T)}{n}) = y)$.

The definition given above was initially motivated by the following result. Consider the sequence of the additive functionals of processes $\{X_n\}$ of the type

$$\varphi_s^{t,n} = \sum_{i:s\le \frac{i}{n}<t} F_n\Big(X_n(\frac{i}{n}),\dots,X_n(\frac{i+L}{n})\Big), \quad 0 \le s \le t.$$

For such functionals, we define their characteristics f^n by $f_s^{t,n}(x) = E[\varphi_s^{t,n}|X_n(0) = x]$.

Theorem. *Suppose that*

1) the sequence of the processes $\{X_n\}$ provides the Markov approximation for the Markov process X;

2) the sequence $\{f^n\}$ converges uniformly to a jointly continuous function f which is the characteristics of some W-functional φ of the process X;

3) $\|F_n\|_\infty \to 0, n \to +\infty$.

Then $\varphi_s^{t,n} \Rightarrow \varphi_s^t, 0 \le s \le t, n \to +\infty$.

We do not prove this theorem here. The detailed exposition of the proof of this theorem and its applications is given in the forthcoming paper [2]. Let us only mention that the scheme of the proof is very close to the one of the famous E.B.Dynkin's theorem on a convergence of W-functionals (see [3], Ch. 6.3), and thus the Markov property of the two-component processes $\hat{Y}_n$ is essential there.

This paper has two purposes. On the one hand, we illustrate the reasons why the main definition is given in such a complicated form. The main feature here is the following: even if X_n, X possess the Markov property at every point $\frac{i}{n}, i \in \mathbb{N}$, one typically can not take $K(\gamma,T) = 1$. In a "generic" situation $K(\gamma,T)$ should tend to $+\infty$ as $\gamma \to 0+$. This means that while the rate of approximation becomes better, the Markov properties of the pair of the processes should become worse. On the other hand, we show that Definition 1 is general enough to include the most typical limit theorems such as the multidimensional functional CLT with a normal domain of attraction and the functional CLT with a stable domain of attraction.

The main content of the paper is represented by two following theorems. In the first theorem, $\{\xi_k\}$ is a sequence of i.i.d. random vectors in $\Re^m$ with $E\|\xi_k\|^{2+\delta} < +\infty$ for some $\delta > 0$. We suppose ξ_k to have zero mean and identity covariation matrix. Define the processes X_n ("random broken lines") on $\Re^+$ by

$$(2) \qquad X_n(t) = \frac{S_{k-1}}{\sqrt{n}} + (nt-k+1)\left[\frac{S_k}{\sqrt{n}} - \frac{S_{k-1}}{\sqrt{n}}\right], \quad t \in [\frac{k-1}{n}, \frac{k}{n}), \quad k \in \mathbb{N},$$

where $S_n = \sum_{k=1}^n \xi_k$. Then the famous Donsker's invariance principle states that the distributions of the processes X^n in $C(\Re^+, \Re^m)$ converge to the distribution of the m-dimensional Wiener process X.

Theorem 1. I. *The random broken lines $\{X_n\}$ provide the Markov approximation for the process X.*

II. *Conditions (i)-(iii) hold true with* $\sup_{\gamma,T} K(\gamma,T) < +\infty$ *in the only trivial case of $\xi_k \sim \mathcal{N}(0, I_{\Re^m})$ ($I_{\Re^m}$ denotes the identity matrix in $\Re^m$).*

In the second theorem $\{\xi_k\}$ is a sequence of i.i.d. random vectors in $\Re^1$ that belongs to the normal domain of attraction of an α-stable law $\mathcal{L}$, $\alpha \in (0,2)$. This means, by definition, that

$$n^{-\frac{1}{\alpha}}[S_n - a_n] \Rightarrow \mathcal{L}, \quad a_n = \begin{cases} 0, & \alpha \in (0,1) \\ nE\xi_1, & \alpha \in (1,2) \\ n^2 E \sin \frac{\xi_1}{n}, & \alpha = 1 \end{cases}$$

(see [4], Chapter XVII.5). In order to shorten the notation, we suppose that $a_n \equiv 0$ and consider the processes X_n on $\Re^+$ defined by

$$(3)\quad X_n(t) = n^{-\frac{1}{\alpha}} S_{k-1} + (nt - k + 1)\left[n^{-\frac{1}{\alpha}} S_k - n^{-\frac{1}{\alpha}} S_{k-1}\right], \quad t \in [\frac{k-1}{n}, \frac{k}{n}), \quad k \in \mathbb{N}.$$

Denote, by X, the homogeneous process with independent increments such that $X(1) - X(0) =^d \mathcal{L}$. We call such a process α-stable.

Theorem 2. *The random broken lines $\{X_n\}$ provide the Markov approximation for the process X. Moreover, conditions (i)-(iii) hold true with $K(\gamma, T) = 1$ for every γ, T.*

Remark 2. Theorem 2 and statement II of Theorem 1 show that the approximation of an α-stable process by the random broken lines is much better (from the point of view of its Markov properties) than the one of the Wiener process.

2. Proof of Theorem 1

Proof of statement I. Due to CLT, $n^{-\frac{1}{2}} S_n \longrightarrow^d X(1)$. Since $E\|\xi_k\|^{2+\delta} < +\infty$, the family $\{\frac{S_n^2}{n}\}$ is uniformly integrable (one can verify this, by using the Burkholder and Hölder inequalities analogously to estimate (4) below). Thus, the Wasserstein–Kantorovich–Rubinstein distance between Law$(n^{-\frac{1}{2}} S_n)$ and Law$(X(1))$ tends to 0 as $n \to \infty$. Therefore, for every $\varepsilon > 0$, there exist n_ε and a random vector $(\eta_\varepsilon, \zeta_\varepsilon)$ such that

$$E\|\eta_\varepsilon - \zeta_\varepsilon\|^2_{\Re^m} < \varepsilon, \quad \eta_\varepsilon \stackrel{d}{=} \frac{S_{n_\varepsilon}}{\sqrt{n_\varepsilon}}, \quad \zeta_\varepsilon \stackrel{d}{=} X(1).$$

Now we construct the probability space $(\Omega^1, \mathcal{F}^1, P^1)$ in the following way:

$$\Omega^1 = (\Re^m)^{n_\varepsilon} \times C([0,1]), \qquad \mathcal{F}^1 = \mathcal{B}(\Omega^1).$$

Denote the coordinates of the point $\omega^1 \in \Omega^1$ by $\chi = (\chi_1, \dots, \chi_{n_\varepsilon}) \in (\Re^m)^{n_\varepsilon}, \varphi \in C([0,1])$. Define the following measures: $Q(du, dv)$ is the joint distribution of $(\eta_\varepsilon, \zeta_\varepsilon)$, $U_\varepsilon(d\chi, u)$ is the distribution of $\{\xi_1, \dots, \xi_{n_\varepsilon}\}$ conditioned by $\{\frac{S_{n_\varepsilon}}{\sqrt{n_\varepsilon}} = u\}$, and $V_\varepsilon(d\varphi, v)$ is the distribution of $X(\cdot)$ conditioned by $\{X(1) = v\}$. Put

$$P^1(A) = \int_{(\Re^m)^2} \left[\int_A U_\varepsilon(d\chi, u) V_\varepsilon(d\varphi, v)\right] Q(du, dv), \quad A \in \mathcal{F}^1.$$

Now we define the probability space $(\Omega, \mathcal{F}, P)$ as the infinite product of the copies of $(\Omega^1, \mathcal{F}^1, P^1)$. For $\omega = (\chi^1, \varphi^1, \chi^2, \varphi^2, \dots) \in \Omega$, we define the sequence $\{\hat{\xi}_k(\omega), k \geq 1\}$ by

$$\hat{\xi}_1(\omega) = \chi_1^1,\ \hat{\xi}_2(\omega) = \chi_2^1,\ \dots,\ \hat{\xi}_{n_\varepsilon}(\omega) = \chi^1_{n_\varepsilon},\ \xi_{n_\varepsilon+1}(\omega) = \chi_1^2,\ \dots.$$

The sequence $\{\hat{\xi}_k\}$ has the same distribution with $\{\xi_k\}$, and therefore the process $\hat{X}_n$, constructed from $\{\hat{\xi}_k\}$ by formula (2), has the same distribution with X_n. For $\omega = (\chi^1, \varphi^1, \chi^2, \varphi^2, \dots) \in \Omega$, we define

$$\hat{X}^n(t)(\omega) = \frac{1}{\sqrt{n}}\left[\sum_{k=1}^{[nt]} \varphi^k(1) + \varphi^{[nt]+1}\Big(\frac{t - [nt]}{n}\Big)\right], \quad t \geq 0,$$

and the process $\hat{X}^n$ has the same distribution with X.

The processes $\hat{Y}_n = (\hat{X}_n, \hat{X}^n)$ satisfy conditions (i), (ii) of Definition 2 with $K(\gamma, T) = n_\varepsilon$. On the other hand, the values of the difference $\hat{X}_n - \hat{X}^n$ at the points $\frac{iK(\gamma,T)}{n}, i \in \mathbb{N}$, are equal to the sums of i.i.d. random vectors. Thus, it follows from the Kolmogorov's maximal inequality that

$$P\Big(\sup_{i \le \frac{Tn}{K(\gamma,T)}} \left\| \hat{X}_n(\frac{iK(\gamma,T)}{n}) - \hat{X}^n(\frac{iK(\gamma,T)}{n}) \right\|_{\Re^m} > \gamma\Big) \le$$

$$\le \frac{C_m}{\gamma^2} E \left\| \hat{X}_n \left(\Big[\frac{Tn}{K(\gamma,T)}\Big] \frac{K(\gamma,T)}{n} \right) - \hat{X}^n \left(\Big[\frac{Tn}{K(\gamma,T)}\Big] \frac{K(\gamma,T)}{n} \right) \right\|^2_{\Re^m} \le \frac{C_m}{\gamma^2} \cdot T\varepsilon,$$

where C_m is some constant depending only on the dimension of the phase space. If ε was taken less than $\frac{\gamma^3}{C_m T}$ at the beginning of the construction, then condition (iii) of Definition 1 also holds true, which completes the proof.

Proof of statement II. In order to shorten the notation, we consider only the case $m = 1$, the general case is completely analogous.

Suppose that conditions (i)-(iii) hold true with $\sup_{\gamma,T} K(\gamma, T) < +\infty$. Then there exists such a constant $K \in \mathbb{N}$ and a sequence $\gamma_k \to 0+$ that $K(\gamma_k, 1) = K$. For every k, consider the corresponding processes $\hat{X}_{n,k}, \hat{X}^{n,k}$.

The sequence $\{\hat{X}_{n,k}(iK) - \hat{X}^{n,k}(iK), i \in \mathbb{N}\}$ is a martingale. Applying the Burkholder inequality with $p = 2 + \delta$ and then the Hölder inequality with $p = 1 + \frac{\delta}{2}$, we have

$$E \left| \hat{X}_{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) - \hat{X}^{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) \right|^{2+\delta} \le B_{2+\delta} n^{-1-\frac{\delta}{2}} E \left\{ \sum_{i=1}^{[\frac{n}{K}]} \Delta_i^2 \right\}^{1+\frac{\delta}{2}} \le \frac{B_{2+\delta}}{n} E \sum_{k=1}^{n} \Delta_k^{2+\delta}, \tag{4}$$

where $\Delta_i = \left| (\hat{\xi}_{iK-K+1} + \cdots + \hat{\xi}_{iK}) - \sqrt{n} \left(\hat{X}^{n,k}(\frac{iK}{N}) - \hat{X}^{n,k}(\frac{(i-1)K}{N}) \right) \right|$. All the variables $\hat{\xi}_i$ have the same finite moment of the order of $2 + \delta$. The variable

$$\sqrt{n} \left(\hat{X}^{n,k}(\frac{iK}{N}) - \hat{X}^{n,k}(\frac{(i-1)K}{N}) \right)$$

has the distribution $\mathcal{N}(0, K)$, and its moment of the order of $2 + \delta$ is some constant. This, together with (4), gives that

$$E \left| \hat{X}_{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) - \hat{X}^{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) \right|^{2+\delta} \le C, \quad n, k \in \mathbb{N}. \tag{5}$$

On the other hand, if the distribution of $\hat{\xi}_1$ is not equal $\mathcal{N}(0, 1)$, then the distribution of $\hat{\xi}_1 + \dots \hat{\xi}_K$ is not equal $\mathcal{N}(0, K)$. Then the Wasserstein—Kantorovich—Rubinstein distance between Law$(\hat{\xi}_1 + \cdots + \hat{\xi}_K)$ and $\mathcal{N}(0, K)$ is equal to some $d > 0$, which means that $E(\eta - \zeta)^2 \ge d$ for every random vector (η, ζ) with $\eta =^d \hat{\xi}_1 + \cdots + \hat{\xi}_K$, $\zeta =^d \sqrt{n} \left(\hat{X}^{n,k}(\frac{iK}{N}) - \hat{X}^{n,k}(\frac{(i-1)K}{N}) \right) \sim \mathcal{N}(0, K)$. Therefore, $E\Delta_i^2 \ge d, i \ge 1$ and

$$E \left(\hat{X}_{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) - \hat{X}^{n,k} \left(\frac{K}{n} \Big[\frac{n}{K}\Big] \right) \right)^2 = n^{-1} \sum_{i=1}^{[\frac{n}{K}]} E\Delta_i^2 \ge \frac{d}{K} \cdot \frac{n-K}{n}. \tag{6}$$

Now condition (iii) and inequalities (5),(6), together with the elementary inequality

$$E\xi^2 \le \gamma^2 + E\xi^2 \mathbb{I}_{|\xi|>\gamma} \le \gamma^2 + [E\xi^{2+\delta}]^{\frac{2}{2+\delta}} [P(|\xi| > \gamma)]^{\frac{\delta}{2+\delta}},$$

give contradiction. The theorem is proved.

3. Proof of Theorem 2

Denote, by F and U, the distribution functions of ξ_1 and $X(1)$, respectively. One has (see [4], Chapter XVII, §6)

$$\exists C_-, C_+ : \quad x^\alpha[1 - G(x)] \to C_+, \quad x^\alpha G(-x) \to C_-, \quad x \to +\infty, \tag{7}$$

where G denotes either F or U. Put $F_{-1}(z) = \sup\{y|F(y) \le z\}, z \in [0,1]$ and $\Phi(x) = F_{-1}(U(x)), x \in \Re$. Since $U(\cdot)$ is continuous, $U(X(1))$ is uniformly distributed on $[0,1]$ and $\Phi(X(1)) =^d \xi_1$. We take the stable process $\{X(t), t \ge 0\}$ with $X(0) = 0, X(1) =^d \mathcal{L}$ on some probability space $(\Omega, \mathcal{F}, P)$. We put

$$\hat{\xi}_k = \Phi(X(k) - X(k-1)), \quad k \ge 1$$

and construct $\hat{X}_n$ from $\{\hat{\xi}_k\}$ by formula (3). We also put $\hat{X}^n(t) = n^{-\frac{1}{\alpha}} X(tn), t \ge 0$.

Conditions (i),(ii) of Definition 1 obviously hold true with $K(\gamma, T) = 1$. Now let us proceed with the condition (iii). We will prove that, for every $\gamma > 0$,

$$\lim_{n\to+\infty} P\Big(\sup_{i\le T} \Big|\hat{X}_n\left(\frac{i}{n}\right) - \hat{X}^n\left(\frac{i}{n}\right)\Big| > \gamma\Big) = 0. \tag{8}$$

It follows from (7) that, for every $\varepsilon > 0$, there exists $x_\varepsilon > 0$ such that

$$\Phi(x) \in ((1-\varepsilon)x, (1+\varepsilon)x), \quad x > x_\varepsilon, \qquad \text{and} \qquad \Phi(x) \in ((1+\varepsilon)x, (1-\varepsilon)x), \quad x < -x_\varepsilon,$$

and consequently

$$|x - \Phi(x)| \le \varepsilon|x|, \quad |x| > x_\varepsilon. \tag{9}$$

Denote, by D, the distribution function of the variables $\Delta_k \equiv (X(k) - X(k-1)) - \hat{\xi}_k$. It follows from (7),(9) that

$$x^\alpha[1 - D(x)] \to 0, \; x^\alpha D(-x) \to 0, \quad x \to +\infty, \tag{10}$$

and this is a key point: the tails of the distribution of the "error" Δ_k are less valuable than the tails of the initial or limiting distributions. Therefore, these errors will be overwhelmed by the normalizing coefficient $n^{-\frac{1}{\alpha}}$. This idea is not a new one. It was used before as a key point in the proofs of an *invariance principle* (either in its weak or strong form) for the sums of independent summands with heavy-tailed distributions (see, for instance, [5]). In our exposition, we adapt this idea to our needs.

Denote $Z_n(t) = \hat{X}_n(t) - \hat{X}^n(t), t \in [0,T]$. In order to prove (8), it is sufficient to prove that $Z_n(\cdot) \to 0$ weakly in $\mathbb{D}([0,T])$. Let us show first that $Z_n(t) \Rightarrow 0$ for every given $t \in [0,T]$. We consider only $t = 1$, the general case is analogous.

Consider the sequence of the canonical measures (see [4], Chapter XVII) $\mu_n(dx) = nx^2 D(n^{\frac{1}{\alpha}} dx)$, and let us show that μ_n converge to the zero measure properly, i.e. that

$$\mu_n((x,y]) \to 0, \quad x, y \in \Re, \quad \int_{-\infty}^{-x} \frac{1}{z^2}\mu_n(dz) \to 0, \int_x^{+\infty} \frac{1}{z^2}\mu_n(dz) \to 0, \quad x > 0. \tag{11}$$

The convergence of two integrals in (11) is just relation (10) written in another form. The convergence of $\mu_n([x,y])$ to 0 also follows from (10). Indeed, for $x>0$, we have

(12)
$$n\int_{(0,x]} z^2 D(n^{\frac{1}{\alpha}}dz) = -n^{1-\frac{2}{\alpha}}\int_{(0,xn^{\frac{1}{\alpha}}]} u^2 d(1-D(u))$$
$$= n^{1-\frac{2}{\alpha}}\Big[z^2(1-D(z))\Big]\Big|_{xn^{\frac{1}{\alpha}}}^{0} + 2n^{1-\frac{2}{\alpha}}\int_0^{xn^{\frac{1}{\alpha}}} z[1-D(z)]dz,$$

and due to (10)

$$n^{1-\frac{2}{\alpha}}\Big[z^2(1-D(z))\Big]\Big|_{xn^{\frac{1}{\alpha}}}^{0} = n^{1-\frac{2}{\alpha}}\cdot O(n^{\frac{2}{\alpha}})\cdot o((n^{\frac{1}{\alpha}})^{-\alpha}) = o(1), \quad n\to+\infty,$$

$$n^{1-\frac{2}{\alpha}}\int_0^{xn^{\frac{1}{\alpha}}} z[1-D(z)]dz = n^{1-\frac{2}{\alpha}}\cdot o((n^{\frac{1}{\alpha}})^{2-\alpha}) = o(1), \quad n\to+\infty.$$

Here, we used the relation $x^{2-\alpha}\int_0^x x[1-D(x)]\,dx\to 0, x\to+\infty$, that follows immediately from (10). Thus, $n\int_{(0,x]} z^2 D(n^{\frac{1}{\alpha}}dz)\to 0, n\to+\infty$ for every $x>0$. Analogously, $n\int_{(x,0]} z^2 D(n^{\frac{1}{\alpha}}dz)\to 0, n\to+\infty$ for every $x<0$, which gives the needed convergence for μ_n.

Now using Theorem 2 [4], Chapter XVII, §2, we deduce that there exists such a sequence $\{b_n\}\subset\Re$ that $Z_n(1)-b_n\Rightarrow 0$. Recall that $Z_n(1)=\hat{X}_n(1)-\hat{X}^n(1)$, $\hat{X}^n(1)=^d\mathcal{L}$, $\hat{X}_n(1)\Rightarrow\mathcal{L}$. This implies that b_n is bounded. Moreover, if we suppose that $\{b_n\}$ has some partial limit $b_*\neq 0$, then we obtain that $\mathcal{L}=^d\mathcal{L}+b_*$, which is evidently impossible. This means that $b_n\to 0$ and $Z_n(1)\Rightarrow 0, n\to+\infty$.

Since all the finite-dimensional distributions of $Z_n(\cdot)$ converge to the distributions of the process that is equal to zero with probability 1, in order to prove (8), it is enough to show that the distributions of $Z_n(\cdot)$ are weakly compact in $D([0,T])$. The simplest case here is $\alpha>\frac{1}{2}$, then the needed compactness follows from the general sufficient conditions given in [6], Chapter 3. The case of general α is more delicate, we use here the Theorem 2 from [7]. This theorem gives a non-trivial generalization of Theorem 2 [4], Chapter XVII, §2 and provides the weak convergence in $D([0,T])$ of the random broken lines generated by the triangular array of the independent random variables. The main condition there is the proper convergence of the sequence of canonical measures defined on $\Re^+\times\Re$. Let us formulate the statement which we refer to, it is a very partial corollary of Theorem 2 [7] and Lemma 2 [8].

Proposition 1. *Let $Z_n(\cdot)$ be the random broken lines with the vertices in $\{\frac{k}{n},k\in\mathbb{N}\}$ and*

$$Z_n\Big(\frac{k}{n}\Big)=\sum_{i=1}^k \eta_{n,i}, \quad k\in\mathbb{N},$$

where $\{\eta_{n,i},i\in\mathbb{N}\}$ are the i.i.d. random variables with the distribution function Φ_n. Define the measures M_n on $\Re^+\times\Re$ by

$$M_n((a,b]\times(x,y]) = \#\left\{i\,\middle|\,\frac{i}{n}\in(a,b]\right\}\cdot\int_{(x,y]} z^2\Phi_n(dz).$$

Suppose that, for every a,b, the measures $\mu_n^{(a,b]}(\cdot)\equiv M_n((a,b]\times\cdot)$ converge to the zero measure properly (see (11)). Then there exists a sequence $\{c_n\}\subset\Re$ such that, for every $T>0$, the processes $\{Z_n(t)-c_nt, t\in[0,T]\}$ converge weakly in $D([0,T])$ to 0.

We have already proved that the condition of Proposition 1 holds true. The considerations analogous to those made above show that $c_n \to 0, n \to +\infty$. This implies that $Z_n(\cdot)$ converge weakly in $D([0,T])$ to 0, and (8) holds true. Theorem 2 is proved.

Bibliography

1. A.A.Dorogovtsev, *A variant of Skorokhod theorem for Markov sequences*, Skorokhod ideas in probability theory, Institute of Math. Ukrainian Acad. Sci., 2000, pp. 147 – 151.
2. Iu.N. Kartashov, A.M. Kulik, *Invariance principle for additive functionals of random walks.* (to appear)
3. E.B. Dynkin, *Markov processes*, Moscow: FIZMATGIZ, 1963. (Russian)
4. W. Feller, *An Introduction to Probability Theory and Its Applications. Vol II*, Mir, Moscow, 1984, pp. 752. (Russian, translated from W. Feller, An Introduction to Probability Theory and Its Applications, John Wiley & Sons, New York, 1971.)
5. G. Simons, W. Stout, *A weak invariance principle with applications to domains of attraction*, Ann. Probability **6** (1978), no. 2, 294 – 315.
6. P. Billingsley, *Convergence of Probability Measures*, Nauka, Moscow, 1977, pp. 352. (Russian, translated from P. Billingsley, Convergence of Probability Measures, John Wiley & Sons, New York, 1971.)
7. P. Major, *Limit theorems and infitely divisible distributions. Part III, Functional limit theorems*, available electronically at http://www.renyi.hu/˜major/probability/divisible.html.
8. P. Major, *Limit theorems and infitely divisible distributions. Part II*, available electronically at http://www.renyi.hu/˜major/probability/divisible.html.

3, Tereshchenkivska Str., Kyiv 01601, Institute of Mathematics, Ukrainian National Academy of Sciences

E-mail: kulik@imath.kiev.ua