Liet. matem. rink, **45**, *spec. nr.*, 2005, 51–56

The law of iterated logarithm for combinatorial multisets

Jolita NORKŪNIENĖ (VU, VIKO)
e-mail: jolita.norkuniene@mif.vu.lt

1. Introduction and results

In this paper we investigate the strong convergence of random variables given on sequences of probability spaces. We analyze mappings defined on a class of combinatorial structures $\mathcal{U}$, in some papers [1], [2] called multisets. Let σ be a combinatorial structure of size n, consisting of components of sizes $(k_1, k_2, ..., k_n)$, $k_j = k_j(\sigma) \geqslant 0$, $1 \leqslant j \leqslant n$ satisfying the condition

$$L_n(\bar{k}) := k_1 + 2k_2 + ... + nk_n = n. \tag{1}$$

The vector $\bar{k} = (k_1, k_2, ..., k_n)$ is called structure vector of σ. A component of size j may be taken with repetitions from some set having $1 \leqslant \pi(j) < \infty$ elements. Then the number of σ with the structure vector $\bar{k}$ is

$$N_n(\bar{k}) = \mathbf{1}(L_n(\bar{k}) = n) \prod_{j=1}^{n} \binom{\pi(j) + k_j - 1}{k_j}.$$

The number of structures of size n is $p(n) := \sum_{L_n(\bar{k})=n} N_n(\bar{k})$, where the summation is extended over vectors $\bar{k}$ satisfying condition (1).

The fundamental examples of the class of multisets $\mathcal{U}$ are integer partitions, polynomials over a finite field, additive arithmetical semigroups, forests of unlabeled trees, mapping patterns and others.

Let ν_n be the uniform probability measure on the set $\mathcal{U}_n \subset \mathcal{U}$ of multisets of size n. It is known [1], that, as $n \to \infty$, the asymptotic distribution of $k_j(\cdot)$ under ν_n for a fixed $j \geqslant 0$ is negative binomial with parameters $(\pi(j), q^{-j})$, $0 \leqslant j \leqslant n$, where $q > 1$ is a parameter depending on $\mathcal{U}$. We recall that this distribution is given by

$$P(\gamma_j = k) = \binom{\pi(j) + k - 1}{k} (1 - q^{-j})^{\pi(j)} q^{-jk}, \quad k \geqslant 0. \tag{2}$$

Let $\gamma_1, \gamma_2, ..., \gamma_n$ be independent negative binomial random variables (r.vs) γ_j with the parameters $(\pi(j), q^{-j})$, $1 \leqslant j \leqslant n$, and $q > 1$. The relation (1) makes $k_j(\cdot)$ dependent and satisfying the conditioning relation

$$\nu(k_1(\sigma) = k_1, ..., k_n(\sigma) = k_n) = P(\gamma_1 = k_1, ..., \gamma_n = k_n | \Theta_n = n) \tag{3}$$

for $\bar{k} = (k_1, ..., k_n) \in \mathbb{Z}^{+n}$. Here $\Theta_n = 1\gamma_1 + ... + n\gamma_n$. We assume that the class of multisets satisfies the following condition:

$$\pi(j) = \frac{\theta q^j}{j}\big(1 + \mathrm{O}(j^{-\beta})\big) \tag{4}$$

for some $\beta > 0$. This implies the logarithmic condition (see [1]).

Using the method going back to probabilistic number theory and proposed by E. Manstavičius [2] and generating functions analysis (see [4], [6]), we investigate the strong convergence of random variables defined via $k_j(\sigma)$, $1 \leqslant j < n$.

Let $\mathbb{G}$ be an additive abelian group. A map $h\colon \mathcal{U}_n \to \mathbb{G}$ is called an additive function if it satisfies the relation

$$h(\sigma) = \sum_{j=1}^{n} h_j\big(k_j(\sigma)\big)$$

for each $\sigma \in \mathcal{U}_n$, where $h_j(0) = 0$ and $h_j(k)$, $j \geqslant 1, k \geqslant 1$ is some double sequence in $\mathbb{G}$. Let $\mathbb{G} = \mathbb{R}$,

$$h(\sigma, m) := \sum_{j=1}^{m} h_j\big(k_j(\sigma)\big), \quad A(m) := \theta \sum_{j=1}^{m} \frac{a_j}{j}, \quad B(m) := \theta \sum_{j=1}^{m} \frac{a_j^2}{j}.$$

Here $a_j := h_j(1)$. As above, let γ_j, $1 \leqslant j \leqslant n$ be independent negative binomial r.vs, $\Xi_n = h_1(\gamma_1) + ...h_n(\gamma_n)$, and $S_n = a_1\gamma_1 + ...a_n\gamma_n$. As in [2], we compare distributions of $h(\sigma, n)$ with distribution of Ξ_n or S_n.

THEOREM 1. *Let $\alpha(m)$ and $\beta(m)$ be real sequences, $\beta(m) > 0, \beta(m) \uparrow \infty$, as $m \to \infty$. Then the following assertions are equivalent:*

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}}\, \nu_n \left(\max_{r \leqslant m \leqslant n} \beta(m)^{-1} \big| h(\sigma, m) - \alpha(m) \big| \geqslant \varepsilon \right) = 0, \tag{5}$$

for each $\varepsilon > 0$ and

$$\beta(n)^{-1}\big(S_n - \alpha(n)\big) \to 0 \quad P - a.s. \tag{6}$$

COROLLARY 2. *If $\beta(m) \to \infty$ and the series*

$$\sum_{j=1}^{\infty} \frac{|a_j|^p}{j\beta(j)^p}$$

converges for some $1 \leqslant p \leqslant 2$, then relation (5) holds with $\alpha(m) = A(m)$.

Let $Z_n := (S_n - A(n))/\beta(n)$. We write $Z_n \Rightarrow [-1, 1]$ if the sequence Z_n is relatively compact and the set of limit points is the interval [-1,1] with probability one. Denote $\beta(n) = (2B(n)LLB(n))^{1/2}$, where $Lu := \log\max(u, e)$, $u \in \mathbb{R}$, and $f(\sigma, m) = \big(h(\sigma, m) - A(m)\big)/\beta(m)$. We now state the law of iterated logarithm.

THEOREM 3. *Suppose $B(n) \to \infty$ and there exists a sequence $b = b(n) \to \infty$, $b = \mathrm{o}(n)$ such that $B(n) = \mathrm{o}(\beta^2(b))$. Then the following assertions are equivalent:*

$$Z_n \Rightarrow [-1, 1] \quad P - a.s.$$

and

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}}\, \nu_n\Big(\max_{r\leqslant m\leqslant n} \big|f(\sigma, m)\big| \geqslant 1+\delta \Big) = 0,$$

but

$$\lim_{r\to\infty} \underline{\lim}_{n\to\infty}\, \nu_n\Big(\min_{r\leqslant m\leqslant n} \big|f(\sigma, m) - b\big| < \delta \Big) = 1$$

for each $b \in [-1, 1]$ and $\delta > 0$.

THEOREM 4. *Let $j(\sigma, 1) < \cdots < j(\sigma, s)$ be the sizes of components of σ and $s = s(\sigma)$. Then*

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}}\, \nu_n\bigg(\max_{r\leqslant m\leqslant s} \frac{\big|\log j(\sigma, k) - k\big|}{(2kLLk)^{1/2}} \geqslant 1+\delta \bigg) = 0,$$

and

$$\lim_{r\to\infty} \underline{\lim}_{n\to\infty}\, \nu_n\bigg(\min_{r\leqslant m\leqslant s} \bigg|\frac{\log j(\sigma, k) - k}{(2kLLk)^{1/2}} - b\bigg| < \delta \bigg) = 1$$

for each $b \in [-1, 1]$ and $\delta > 0$.

2. Proofs

We will use the fundamental lemma and the tail probability estimates of conditional distributions.

LEMMA 5 ([1]). *If condition* (4) *is satisfied, then, in the above notation,*

$$\nu_n\Big(\big(k_1(\sigma), \ldots, k_b(\sigma)\big) \in A\Big) - P\big((\gamma_1, \ldots, \gamma_b) \in A\big) = \mathrm{O}(n^{-1}b)$$

uniformly in $A \subset \mathbb{Z}^{+b}$.

LEMMA 6. *Let $(G, +)$ be an additive abelian group, $A \subset G$, and $h_j(k)$ be the G-valued double sequence defining the additive function h: $\mathcal{U}_n \to \mathbb{G}$. If condition* (4) *is satisfied, then*

$$\begin{aligned} \nu_n\big(h(\sigma) \notin A + A - A\big) &= P(\Xi_n \notin A + A - A | \Theta_n = n) \\ &\leqslant C\big(P^{\theta\wedge 1}(\Xi_n \notin A) + n^{-\theta}\big). \end{aligned} \tag{7}$$

Proof of Lemma 6. The equality in (7) follows from (3). To prove the estimate, we use Lemma A of the paper [3]. We need to show that the negative binomial r.vs satisfy its assumptions. In our case condition (i) is obvious: $P(\gamma_j=0)=(1-q^{-j})^{\pi(j)}\geqslant c>0$.

Condition (ii) concerns the probabilities $P(\Omega_m) := P(\Theta_n = m)$ for $0 \leqslant m \leqslant n$. We have

$$P(\Omega_m) = \sum_{L_n(\bar{k})=m} \prod_{j=1}^{n} \binom{\pi(j)+k_j-1}{k_j} (1-q^{-j})^{\pi(j)} q^{-jk_j}$$
$$= \prod_{j\leqslant n} (1-q^{-j})^{\pi(j)} q^{-m} p(m).$$

Since the generating function of multisets is

$$Z(x) = 1 + \sum_{k\geqslant 1} p(k) z^k = \prod_{j=1}^{\infty} (1-x^j)^{-\pi(j)}, \quad |z| < q^{-1}.$$

using Proposition 3 from [4] we find the k-th Taylor coefficient of this generating function:

$$q^{-m} p(m) = K(\theta) n^{\theta-1} \big(1 + \mathrm{O}(n^{-\beta} \log n)\big),$$

here $K(\theta)$ is a constant depending on the class of multisets. Then $P(\Omega_n) \gg n^{-1}$ and

$$P(\Omega_m)/P(\Omega_n) \leqslant C\big(n/(m+1)\big)^{1-\theta}, \quad 0 \leqslant m \leqslant n-1.$$

These are the required estimates in conditions (ii) and (iii) of Lemma A. We omit easy technical estimates in the proof of

$$\sum_{jk=n} \binom{\pi(j)+k-1}{k} q^{-jk} = \mathrm{O}\Big(\frac{1}{n}\Big)$$

which is the remaining condition (iv) of this lemma. Lemma 5 is proved.

LEMMA 7. *Let $b_n \to 0$, $1 \leqslant s \leqslant n$ and $\varepsilon > 0$. Then*

$$\nu_n\bigg(\max_{s\leqslant m\leqslant n} b_m \Big| \sum_{j\leqslant m} a_j k_j(\sigma) - A(m) \Big| \geqslant \varepsilon\bigg)$$
$$\leqslant C_1(\varepsilon)\bigg(b_s^2 B(s) + \theta \sum_{s\leqslant j\leqslant n} \frac{b_j^2 a^2(j)}{j}\bigg)^{\theta\wedge 1} + C_2 n^{-\theta}.$$

Proof of Lemma 7. Use Lemma 6 and Theorem 3.3.15 of [5].

Proof of Theorem 1. As in [2], at first we notice, that it suffices to consider the linear function $\hat{h}(\sigma, m) := a_1 k_1(\sigma) + \cdots + a_m k_m(\sigma)$. Following E. Manstavičius and

G.J. Babu (see the proof of Theorem 1 in [3]), we have

$$\nu_n\Big(\max_{r\leqslant m\leqslant n}\beta(m)^{-1}\big|h(\sigma,m)-\hat h(\sigma,m)\big|\geqslant\varepsilon\Big)$$
$$\leqslant\nu_n\Big(\sum_{j\leqslant n}\big|h_j(k_j(\sigma))-a_jk_j(\sigma)\big|\geqslant\varepsilon\beta(r)\Big)=\mathrm{o}(1),$$

From the Lemma 9.2.5 [5], one can see that relation (6) is equivalent to

$$P\Big(\sup_{m\geqslant r}\beta(m)^{-1}\big|S(m)-\alpha(m)\big|\geqslant\varepsilon\Big)$$
$$=\lim_{n\to\infty}P\Big(\sup_{r\leqslant m\leqslant n}\beta(m)^{-1}\big|S(m)-\alpha(m)\geqslant\varepsilon\big|\Big)=\mathrm{o}(1),$$

for each $\varepsilon>0$ and $r\to\infty$. From the conditioning relation (3) and Lemma 6 we have

$$\nu_n\Big(\max_{r\leqslant m\leqslant n}\beta(m)^{-1}\big|h(\sigma,m)-\alpha(m)\big|\geqslant\varepsilon\Big)$$
$$\ll P^{1\wedge\theta}\Big(\sup_{r\leqslant m\leqslant n}\beta(m)^{-1}\big|S(m)-\alpha(m)\big|\geqslant\varepsilon/3\Big)+n^{-\theta}.$$

Thus from (6) we obtain (5). Now using Lemma 5, we have

$$\nu_n\Big(\max_{r\leqslant m\leqslant b}\beta(m)^{-1}\big|h(\sigma,m)-\alpha(m))\big|\geqslant\varepsilon\Big)$$
$$=P\Big(\sup_{r\leqslant m\leqslant b}\beta(m)^{-1}\big|S_m-\alpha(m)\big|\geqslant\varepsilon\Big)+\mathrm{o}(1))$$

for each $b=b(n)$, $b\to\infty$, $b=\mathrm{o}(n)$ and $r\leqslant b$. Taking limits with respect to n, later with respect to r, from (5) we deduce (6). Theorem is proved.

Proof of Theorem 2. The desired assertion can be obtained using the arguments from the proof of Theorem 1, Lemma 5 and Lemma 6. See the proof of Theorem 2 in [2] for details.

Proof of Theorem 3. At the beginning, we apply Theorem 2 to the additive function $h(\sigma,m)=s(\sigma,m):=k_1^0(\sigma)+\cdots+k_m^0(\sigma)$, which is the count of all different cycle lengths in decomposition (1), $1\leqslant m\leqslant n$, $0^0:=0$. In this case we have

$$\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\,\nu_n\Big(\max_{r\leqslant m\leqslant n}\frac{|s(\sigma,m)-\log m|}{(2\log mLLLm)^{1/2}}\geqslant1+\delta\Big)=0$$

and

$$\lim_{r\to\infty}\underline{\lim_{n\to\infty}}\,\nu_n\Big(\min_{r\leqslant m\leqslant n}\Big|\frac{s(\sigma,m)-\log m}{(2\log mLLLm)^{1/2}}-t\Big|<\delta\Big)=1$$

for each $t \in [-1, 1]$ and $\delta > 0$. If $s(\sigma, m) = k$, then from the relation $k = s(\sigma, j(\sigma, k))$ it follows that last two assertions are also satisfied for $j(\sigma, k)$. The proof of Theorem 3 is completed.

References

1. R. Arratia, A.D. Barbour, S. Tavaré, *Logarithmic Combinatorial Structures: a Probabilistic Approach*, EMS Monographs in Math., The EMS Publishing House, Zürich (2003).
2. E. Manstavičius, The law of iterated logarithm for random permutations, *Lith. Math. J.*, **38**, 160–171 (1999).
3. G.J.Babu, E. Manstavičius, Brownian motion for random permutations, *Sankhya Ser. A*, **61**(3), 312–327 (1999).
4. H.-K. Hwang, Asymptotic of Poisson approximation of random discrete distributions, *Adv. Appl. Prob.*, **31**, 448–491 (1999).
5. V.V. Petrov, *Sums of Independent Random Variables*, Nauka, Moscow (1972) (in Russian).
6. E. Manstavičius, On the frequency of multisets without some components, *Liet. matem. rink.*, **40**(spec. issue), 55–60 (2000).

REZIUMĖ

J. Norkūnienė. Kartotinio logaritmo dėsnis kombinatorinėms multiaibėms

Nagrinėjamas adityvių funkcijų, apibrėžtų multiaibėse, sekų stiprusis konvergavimas.