Liet. mat. rink. LMD darbai, **50**, 2009, 420–425

Asymptotic distributions of the number of restricted cycles in a random permutation

Tatjana KARGINA
Vilnius University, Department of Mathematics and Informatics
Naugarduko 24, LT-03225 Vilnius, Lithuania
e-mail: tania.kargina@gmail.com

Abstract. The value distribution of additive functions defined on the symmetric group with respect to the Ewens probability is examined. For the number of cycles with restricted lengths, we establish necessary and sufficient conditions under which the distributions converge weakly to a limit law.

Keywords: symmetric group, weak convergence, factorial moments, convergence of moments, Ewens probability.

1. Introduction and result

We examine the distributions of additive functions defined on the symmetric group S_n with respect to the Ewens probability. The main result gives general conditions under which the distributions converge weakly to a discrete limit law.

Let $\sigma \in S_n$ be a permutation having $k_j(\sigma) \geqslant 0$ cycles of length j, $1 \leqslant j \leqslant n$. The structure vector is defined as $\bar{k}(\sigma) := (k_1(\sigma), \dots, k_n(\sigma))$ and $w(\sigma) = k_1(\sigma) + \dots + k_n(\sigma)$ is the number of cycles. The structure vector $\bar{k}(\sigma)$ satisfies the relation $1k_1(\sigma) + \dots + nk_n(\sigma) = n$.

Let us define a probabilistic measure on the symmetric group S_n which is characterized by a parameter $\Theta > 0$ setting

$$\nu_{n,\Theta}(\{\sigma\}) := \Theta^{w(\sigma)}/\Theta^{(n)}, \quad \forall \sigma \in S_n,$$

where $\Theta^{(n)} = \Theta(\Theta+1)\dots(\Theta+n-1)$. Then the probability of permutations with the structure vector $\bar{k} = (k_1, \dots, k_n)$ is expressed by the formula

$$\nu_{n,\Theta}(\bar{k}(\sigma) = \bar{k}) := \frac{n!}{\Theta^{(n)}} \prod_{j=1}^{n} \left(\frac{\Theta}{j}\right)^{k_j} \frac{1}{k_j!}, \tag{1}$$

where $k_j \geqslant 0$, $\Theta > 0$ and $1k_1 + \dots + nk_n = n$. See, for instance, [1].

As it has been proposed by W. Ewens (1972), the right-hand part of (1) can be ascribed as a probability to the vector set $\{\bar{k} \in \mathbf{Z}_+^n\colon 1k_1 + \dots + nk_n = n\}$. Since then, this probability, called the *Ewens Sampling Formula*, was widely applied in genetics (see [1] and the references therein).

We define a sequence of additive functions: h_n: $S_n \to \mathbf{R}$. Given a real sequence $\{h_{nj}(k)\}$, $1 \leqslant j \leqslant n$, $k \geqslant 0$, satisfying the condition $h_{nj}(0) \equiv 0$, we set

$$h_n(\sigma) = \sum_{j=1}^{n} h_{nj}\bigl(k_j(\sigma)\bigr).$$

If in addition, $h_{nj}(k) = a_{nj}k$, where $a_{nj} \in \mathbf{R}$ for all $k \geqslant 1$ and $1 \leqslant j \leqslant n$, then $h_n(\sigma)$ is called *completely additive function*.

The main problem is to establish general conditions under which the distribution functions

$$V_n\bigl(x; h_n, \alpha(n)\bigr) := \nu_{n,\Theta}\bigl(h_n(\sigma) - \alpha(n) < x\bigr),$$

where $\alpha(n) \in \mathbf{R}$ is a centralizing sequence, converge weakly to a limit distribution law. Here and afterwards we assume that $n \to \infty$. We can assume that $h_{nj}(k) = 0$ for all $jk \geqslant n$.

So far, general results were obtained in the case $\Theta = 1$ (see [4,7,8]). In particular, E. Manstavičius [7,8] established necessary and sufficient conditions under which, for a completely additive function with $a_{nj} \in \{0, 1\}$, the distribution function $\nu_{n,1}(h_n(\sigma) < x)$ converges weakly to a limit law. The partial case $\Theta > 1$ has been discussed in [3]. We generalize this result for $\Theta > 0$. It is worth to note that the very idea of such type results goes back to probabilistic number theory, in particular to J. Šiaulys' paper [9].

Under the condition $a_{nj} \in \{0, 1\}$, the additive function $h_n(\sigma)$ is just the number of cycles with restricted lengths of a random permutation $\sigma \in S_n$. Upper and lower bounds for the distribution of this function were obtained in [5].

Apart from the one-dimensional limit laws, the weak convergence of the partial sum processes defined in terms of additive functions under the Ewens probability has been examined. In particular, the paper by G.J. Babu and E. Manstavičius [2] deals with the case of one normalized additive function while the paper [6] concerns the sequences of such functions but the probability measure is restricted to the case $\Theta = 1$. The sufficiency part of our theorem has some intersection with these results. Namely, the existence of a limit law for $V_n(x; h_n, 0)$ follows from the result in [6] but, in addition, we establish the convergence of moments.

We introduce some notation. For convenience, we add the asterisk over the sums to replace the condition $a_{nj} = 1$, $j = j_1, \ldots, j_m, \ldots$. Denote

$$\gamma_{nm,\Theta} = \Theta^m \sideset{}{^*}\sum_{j_1 \leqslant n} \frac{1}{j_1} \ldots \sideset{}{^*}\sum_{j_m \leqslant n} \frac{1}{j_m} \mathbf{1}\{j_1 + \cdots + j_m \leqslant n\} \frac{\Theta^{(n-j_1-\cdots-j_m)}}{(n - j_1 - \cdots - j_m)!} \qquad (2)$$

and

$$\hat{\gamma}_{nm,\Theta} = \frac{n!}{\Theta^{(n)}} \gamma_{nm,\Theta}.$$

THEOREM 1. *Let $h_n(\sigma)$ be a sequence of completely additive functions with $a_{nj} \in \{0, 1\}$ and $\Theta > 0$. The frequencies $V_n(x; h_n, 0) = \nu_{n,\Theta}(h_n(\sigma) < x)$ converge weakly a* *limit law if and only if there exist finite limits*

$$\lim_{n\to\infty} \hat{\gamma}_{nm,\Theta} =: \hat{\gamma}_{m,\Theta} \tag{3}$$

for all $m \in \mathbf{N}$. *Moreover, if* (3) *is satisfied, the characteristic function of the limit distribution is*

$$1 + \sum_{m=1}^{\infty} \frac{\hat{\gamma}_{m,\Theta}}{m!}\big(\mathrm{e}^{it} - 1\big)^m, \quad t \in \mathbf{R}.$$

The case $\Theta \geqslant 1$ has been examined in [3]. Some of the proof remains valid in the general case $\Theta > 0$, therefore in the present remark we present the complementing details for $0 < \Theta < 1$ only.

2. Proof of theorem

We start with two lemmas. Let $\mathbf{E}_{n,\Theta}$ be the mean value with respect to the probability $\nu_{n,\Theta}$. Denote $a_{(r)} := a(a-1)\dots(a-r+1)$.

LEMMA 1. *Let* $\Theta > 0, m \in \mathbf{N}, h_n(\sigma)$ *be a sequence of completely additive functions with* $a_{nj} \in \{0, 1\}$*, and let* $\hat{\gamma}_{nm,\Theta}$ *be defined above. Then*

$$\mathbf{E}_{n,\Theta} h_n(\sigma)_{(m)} = \widehat{\gamma}_{nm,\Theta} \tag{4}$$

for all $n \geqslant 1$.

Proof. See in [3].

The next lemma concerns the concentration function estimate. For a completely additive function $h_n(\sigma)$, we define

$$Q_{n,\Theta}(u) = \sup_{x\in\mathbf{R}} \nu_{n,\Theta}\big(|h_n(\sigma) - x| < u\big), \quad u \geqslant 0,$$

and $a_{nj}(\lambda) = a_{nj} - j\lambda$. Denote $a \wedge b = \min\{a, b\}$,

$$D_n(u; \lambda) = \sum_{j\leqslant n} \frac{u^2 \wedge a_{nj}(\lambda)^2}{j}, \qquad D_n(u) = \min_{\lambda\in\mathbf{R}} D_n(u; \lambda).$$

In what follows, the symbol $\ll$ is used in the sense of $O(\cdot)$, $a \asymp b$ stands for $a \ll b \ll a$, and $c, c_0, \dots, C, C_1, \dots$ denote some positive constants dependent on θ and other parameters which we shall indicate.

LEMMA 2. *We have*

$$Q_{n,\Theta}(u) \ll u D_n(u)^{-1/2} \tag{5}$$

with the constant in "$\ll$*" depending at most on* Θ.

Proof. See in [3].

Proof of Theorem. Sufficiency. Condition (3) of Theorem implies $\hat{\gamma}_{n1,\Theta} \leqslant C < \infty$ for all $n \geqslant 1$. We further use (3), (4), and the expansion

$$\mathbf{E}_{n,\Theta} \mathrm{e}^{ith(\sigma)} = 1 + \sum_{m=1}^{L} \frac{\hat{\gamma}_{nm,\Theta}}{m!} (\mathrm{e}^{it} - 1)^m + O\left(\frac{\hat{\gamma}_{n,L+1,\Theta}}{(L+1)!} |\mathrm{e}^{it} - 1|^{L+1} \right). \qquad (6)$$

To estimate the reminder, we have to prove that

$$\hat{\gamma}_{nm,\Theta} \leqslant C_1 \hat{\gamma}_{n,m-1,\Theta}, \qquad (7)$$

where C_1 does not depend on $m \geqslant 1$. The argument is based upon the following well known estimate

$$c \frac{\Theta^{(n)}}{n!} \leqslant (1+n)^{\Theta-1} \leqslant C_2 \frac{\Theta^{(n)}}{n!} \qquad (8)$$

for all $n \geqslant 1$.

We now examine expression (2) of $\gamma_{n,m,\theta}$. For brevity, we denote $J = j_1 + \cdots + j_{m-1}$ and majorise the most inner sum over $j_m =: j$. Using (8) repeatedly, we obtain

$$\begin{aligned}
&\Theta \sum_{j \leqslant n-J}^{*} \frac{1}{j} \frac{\Theta^{(n-J-j)}}{(n-J-j)!} \\
&= \frac{\Theta n!}{\Theta^{(n)}} \sum_{j \leqslant n-J}^{*} \frac{\Theta^{(n-j)}}{j(n-j)!} \left(\frac{\Theta^{(n)}}{n!} \frac{(n-j)!}{\Theta^{(n-j)}} \frac{\Theta^{(n-J-j)}}{(n-J-j)!} \right) \\
&\leqslant \frac{C_2}{c^2} \frac{\Theta n!}{\Theta^{(n)}} \sum_{j \leqslant n-J}^{*} \frac{\Theta^{(n-j)}}{j(n-j)!} \left(1 - \frac{j}{n+1} \right)^{1-\Theta} (1+n-J-j)^{\Theta-1} \\
&\leqslant \frac{C_2}{c^2} \hat{\gamma}_{n1,\Theta} \leqslant \frac{C_2 C}{c^2} =: C_1.
\end{aligned}$$

This and (2) prove inequality (7). By induction we now obtain $\hat{\gamma}_{nm,\Theta} \leqslant C_3^m$, where $C_3 = \max\{C, C_1\}$. Hence using (3) and (6) we have

$$\mathbf{E}_{n,\Theta} \mathrm{e}^{ith(\sigma)} = 1 + \sum_{m=1}^{L} \frac{\hat{\gamma}_{m,\Theta}}{m!} \left(\mathrm{e}^{it} - 1\right)^m + O\left(\frac{{C_3}^L}{(L+1)!} \right) + o_L(1),$$

where either of the estimates is uniform in $t \in \mathbf{R}$ and the second one depends on $L \geqslant 1$. Taking now $n \to \infty$ and later $L \to \infty$, we complete the proof of convergence of the characteristic function and find the claimed formula of its limit. This implies the weak convergence of $V_n(x; h_n, 0)$.

Necessity. Let $V_n(x; h_n, 0)$ converges weakly to a limit distribution $P(\xi < x)$, where ξ is a random variable taking values in the set $\mathbf{Z}_+$. Hence for the concentration function $Q_{n,\Theta}(1)$ we have

$$Q_{n,\Theta}(1) \gg \max_{m\in\mathbf{Z}_+} P(\xi = m) \geqslant c_1 > 0,$$

where the constant c_1 depends at most on ξ provided that n is sufficiently large. In what follows we disregard such dependence. Consequently, by (5) in Lemma 2, we obtain $D_n(1; \lambda) \ll 1$ with some $\lambda = \lambda_n \in \mathbf{R}$. By virtue of $a_{nj} \in \{0, 1\}$, this implies

$$1 \gg \sum_{2/|\lambda|<j\leqslant n} \frac{1 \wedge (a_{nj} - \lambda j)^2}{j} = \sum_{2/|\lambda|\leqslant j\leqslant n} \frac{1}{j} \geqslant c_0 \log(n|\lambda|)$$

if $\lambda \neq 0$, where $c_0 > 0$ and n is sufficiently large. Hence $|\lambda| \leqslant C_4/n$. This and the relation $(x + y)^2 \leqslant 2(x^2 + y^2)$ yield

$$\begin{aligned} D_n(1; 0) &= \sum_{j\leqslant n}^{*} \frac{1}{j} \ll D_n(1; \lambda) + \sum_{j\leqslant n} \frac{1 \wedge (C_4 j/n)^2}{j} \\ &\ll 1 + \frac{1}{n^2} \sum_{j\leqslant n/C_4} j + \sum_{n/C_4<j\leqslant n} \frac{1}{j} \leqslant C_5 < \infty \end{aligned}$$

for all $n \geqslant 1$. Hence and from (8) we obtain

$$\begin{aligned} \hat{\gamma}_{n1,\Theta} &\ll n^{1-\Theta}\Biggl(\sum_{j\leqslant n/2}^{*} + \sum_{n/2<j\leqslant n}\Biggr)\frac{1}{j}(1 + n - j)^{\Theta-1} \\ &\leqslant 2C_5 + \frac{1}{n^{\Theta}} \sum_{n/2<j\leqslant n} (1 + n - j)^{\Theta-1} \ll 1 \end{aligned}$$

for $n \geqslant 1$. As in the sufficiency part, the induction argument leads to $\hat{\gamma}_{nm,\Theta} \ll {C_6}^m$ and, by Lemma 1, to the inequality

$$\sup_n \mathbf{E}_{n,\Theta} h_n(\sigma)_{(m)} \leqslant C_6^m$$

for every $m \geqslant 1$. Consequently, from the weak convergence of distribution functions frequencies $V_n(x; h_n, 0)$ we obtain convergence of factorial moments.

Theorem is proved.

References

1. R. Arratia, A.D. Barbour, S. Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach*, EMS Monographs in Mathematics. EMS Publishing House, Zürich, 1992.
2. G.J. Babu, E. Manstavičius. Processes with independent increments for the Ewens sampling formula. *Ann. Inst. Statist. Math.*, 54(3):607–620, 2002.
3. T. Kargina. Additive functions on permutations and the Ewens probability. In: *Šiauliai Math. Semin.*, Vol. 2(10), 33–41, 2007.

4. E. Manstavičius. Additive and multiplicative functions on random permutations. *Lith. Math. J.*, 36(4):400–408, 1996.
5. E. Manstavičius. On the probability of combinatorial structures without some components. In: M.A. Bennett *et al.* (Eds.), *Number Theory for the Millennium*, Vol. II. A.K.Peters, Boston, 387–401, 2001.
6. E. Manstavičius. Functional limit theorem for sequences of mappings on the symmetric group. In: A. Dubickas *et al.* (Eds.), *Anal. Probab. Methods Number Theory*. TEV, Vilnius, 175–187, 2002.
7. E. Manstavičius. The Poisson distribution for the linear statistics on random permutations. *Lith. Math. J.*, 45(4):434–446, 2005.
8. E. Manstavičius. Discrete limit laws for additive functions on the symmetric group. *Acta Math. Universitatis Ostraviensis*, 13:47–55, 2005.
9. J. Šiaulys. Factorial moments method for distributions of additive functions. *Lith. Math. J.*, 40(4):508–525, 2000.

REZIUMĖ

T. Kargina. Atsitiktinio keitinio apribotų ciklų skaičiaus asimptotiniai skirstiniai

Nagrinėjamas adityviųjų funkcijų, apibrėžtų simetrinėje grupėje, asimptotinių skirstinių egzistavimas. Keitiniai imami su Evenso tikimybe. Rastos būtinos ir pakankamos sąlygos, kai adityvioji funkcija išreiškia keitinio ciklų su bet kokiais apribojimais skaičių.

Raktiniai žodžiai: simetrinė grupė, silpnasis konvergavimas, faktorialiniai momentai, momentų konvergavimas, Evenso tikimybė.