Lithuanian Mathematical Journal, Vol. 42, *No.* 3, 2002

INFINITELY DIVISIBLE LIMIT PROCESSES FOR THE EWENS SAMPLING FORMULA

G. J. Babu[1]
Department of Statistics, 319 Thomas Building, The Pennsylvania State University, University Park, PA 16802, USA
(e-mail: babu@stat.psu.edu)

E. Manstavičius[2]
Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, LT-2600 Vilnius, Lithuania
(e-mail: eugenijus.manstavicius@maf.vu.lt)

Abstract. The *Ewens sampling formula* in population genetics can be viewed as a probability measure on the group of permutations of a finite set of integers. Functional limit theory for processes defined through partial sums of dependent variables with respect to the Ewens sampling formula is developed. Using techniques from probabilistic number theory, it is shown that, under very general conditions, a partial sum process weakly converges in a function space if and only if the corresponding process defined through sums of independent random variables weakly converges. As a consequence of this result, necessary and sufficient conditions for weak convergence to a stable process are established. A counterexample showing that these conditions are not necessary for the one-dimensional convergence is presented. Very few results on the necessity part are known in the literature.

Keywords: random partition, population genetics, allelic partition, permutation, probabilistic number theory, Skorokhod topology, functional limit theorem, stable process.

1. INTRODUCTION

Ewens [9] derived a formula to describe the distribution of a sample of n genes from a population that was evolved over many generations, by a partly heuristic argument. In several genetic models, it is an exact formula and, in others, it is a close approximation. The formula is derived under the null hypothesis that there are no selection effects. In this case, the *allelic partition*, $\bar{k} = (k_1, \ldots, k_n)$, contains all the information available in a sample of n genes, where k_j denotes the number of alleles represented j times in the sample, $j = 1, \ldots, n$.

The Ewens sampling formula [9] is given by

$$\nu_{n,\theta}(\bar{k}) := \nu_{n,\theta}\big(k_1, \ldots, k_n\big) := \frac{n!}{\theta_{(n)}} \prod_{j=1}^{n} \left(\frac{\theta}{j}\right)^{k_j} \frac{1}{k_j!}, \tag{1.1}$$

where $\theta > 0$, $\theta_{(n)} = \theta(\theta+1)\cdots(\theta+n-1)$, $k_j \geqslant 0$, and

$$1k_1 + \cdots + nk_n = n. \tag{1.2}$$

[1]Research supported in part by NSA grant MDA904-97-1-0023, NSF grant DMS-9626189, and by National Research Council's 1997-99 Twinning fellowship.
[2]Research supported by National Research Council's 1997-99 Twinning fellowship and by a Lithuanian State Stipend.

Translated from Lietuvos Matematikos Rinkinys, Vol. 42, No. 3, pp. 294–307, July–September, 2002. Original article submitted March 7, 2001.

0363–1672/02/4203–0232$27.00

The vector $\bar{k}$ represents a partition of an integer n.

For each $\theta > 0$, the Ewens sampling formula describes a probability measure on the symmetric group $\mathbb{S}_n$ of permutations on $\{1, \ldots, n\}$. Let $\sigma \in \mathbb{S}_n$ and let

$$\sigma = \varkappa_1 \cdots \varkappa_w \tag{1.3}$$

be its unique representation (up to the order) by the product of independent cycles $\varkappa$, where $w = w(\sigma)$ denotes the number of cycles. Recall that $\sigma, \sigma' \in \mathbb{S}_n$ are conjugate if $\sigma = \tau\sigma'\tau^{-1}$ with some $\tau \in \mathbb{S}_n$. So $\mathbb{S}_n$ splits into the union of classes of conjugate elements. There is a bijective correspondence between the set of classes and the set of vectors $\bar{k} := (k_1, \ldots, k_n)$ satisfying (1.2) such that the elements in the conjugate class corresponding to $\bar{k}$ will all have the same number k_j of cycles of length j for all $1 \leqslant j \leqslant n$. Hence, this class can be identified by $\bar{k}$, and probability (1.1) can be assigned to it. We also get the same probability measure for the classes if we start with the definition $\nu_{n,\theta}(\{\sigma\}) := \theta^{w(\sigma)}/\theta_{(n)}$ for each $\sigma \in \mathbb{S}_n$. In what follows, we preserve the notation $\nu_{n,\theta}$ for this probability measure on $\mathbb{S}_n$.

It is well known (see, for example, [2]) that the asymptotic distribution with respect to $\nu_{n,\theta}$ of $k_j(.)$ for a fixed $j \geqslant 1$, as $n \to \infty$, is Poisson with parameter θ/j. Relation (1.2) makes $k_j(.)$, $1 \leqslant j \leqslant n$, a dependent sequence. The dependency is rather strong for $\varepsilon n \leqslant j \leqslant n$. Nevertheless, functional limit theorems similar to those for independent random variables can be established.

In the case $\theta = 1$ for the function $w(\sigma)$ in (1.3), the first functional limit theorem was established by DeLaurentis and Pittel [7]. The case of general θ for the function $w(\sigma)$ was examined by Hansen [11] and Donnelly *et al.* [8]. A short proof of Hansen's theorem is given in Section 2.C of [2]. The convergence of more general partial sum processes to the Brownian motion was investigated by the authors [4]. It was shown that an analogue of the Lindeberg condition is necessary and sufficient for the weak convergence of the processes, in contrast to the fact that this condition is not necessary for the one-dimensional central limit theorem. In this paper, we extend these results to a class of infinitely divisible limit laws that include stable laws.

This branch of statistical group theory is closely associated with distributions of arithmetic functions in probabilistic number theory. The idea of approximating truncated sums of dependent and the corresponding independent random variables in total-variation distance has been exploited and well documented in probabilistic number theory (see, for example, [3], [12], [13], [14]). Thus, for the proofs in this paper, it is rather natural to borrow some techniques from probabilistic number theory. In recent papers, R. Arratia and his colleagues (see [1], [2], and the references given in these papers) have exploited similar ideas in probabilistic combinatorics based on the approximation in the total-variation distance.

2. MAIN RESULTS

Let $h_j(k)$, $k \geqslant 0$, $j \geqslant 1$, be a real double sequence such that $h_j(0) = 0$ for each j. It naturally defines the additive function on $\mathbb{S}_n$

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

The number of cycles $w(\sigma)$ and the function defined by $h_j(k) = \mathbf{1}\{k \geqslant 1\} \log j$ and well approximating the logarithm of the group-theoretical order of $\sigma \in \mathbb{S}_n$ are the classical examples of additive functions. The function given by $h_j(k) = k\{\{\alpha j\}\}$, where $\{\{u\}\}$ denotes the fractional part of $u \in \mathbb{R}$, appears in the recent studies [17] and [10] on the distribution of eigenvalues of the permutation matrices. The weak convergence of the distribution functions $\nu_{n,1}(h(\sigma) - \alpha(n) < x\beta(n))$, $\alpha(n) \in \mathbb{R}$, $\beta(n) > 0$, was examined in [15]. It appears that the events $\{\sigma\colon\ k_j(\sigma) \geqslant 2\}$ are rather rear and, therefore, in the case $\beta(n) \to \infty$, the values $h_j(k)$, $k \geqslant 2$, $1 \leqslant j \leqslant n$, have no influence on the limit distribution. The same phenomenon has been observed in the invariance principle [4].

Set, for brevity, $a(j) = h_j(1)$, and $u^* = (1 \wedge |u|)\mathrm{sgn}u$, where $a \wedge b := \min\{a, b\}$. Here and in what follows, the limits are taken as $n \to \infty$. We sometimes use the notation $f(n) \ll g(n)$ instead of $f(n) = \mathrm{O}(g(n))$. Throughout this paper, we assume that the normalizing factors $\beta(n) > 0$ satisfy $\beta(n) \to \infty$. The sequence $\{\beta(n)\}$ need not be monotone. Define

$$B(u, n) = \sum_{j \leqslant u} \left(\frac{a(j)}{\beta(n)}\right)^{*^2} \frac{\theta}{j}, \qquad A(u, n) = \sum_{j \leqslant u} \left(\frac{a(j)}{\beta(n)}\right)^{*} \frac{\theta}{j},$$

and

$$y(t) := y_n(t) = \max\{l \leqslant n\colon\ B(l,n) \leqslant tB(n,n)\}, \quad t \in [0,1]. \tag{2.1}$$

We shall consider the weak convergence (denoted by $\Rightarrow$) of the process

$$H_n := H_n(\sigma, t) = \frac{1}{\beta(n)} \sum_{j \leqslant y(t)} h_j(k_j(\sigma)) - A(y(t), n), \quad t \in [0,1],$$

under the probability $\nu_{n,\theta}$ in the space $\mathbb{D}[0,1]$ endowed with the Skorokhod topology [6]. Having in mind the asymptotic distribution of $k_j(\sigma)$, we define the corresponding process X_n with independent increments

$$X_n := X_n(t) = \sum_{j \leqslant y(t)} X_{nj} - A(y(t), n), \quad t \in [0,1],$$

where $X_{nj} = a(j)\xi_j/\beta(n)$ and ξ_j, $1 \leqslant j \leqslant n$, are independent Poisson random variables with $\mathbf{E}(\xi_j) = \theta/j$. Since $B(u,n)$ is close to the sum of truncated second moments of the involved random variables, our choice of the time index function $y(t)$ seems to be natural. The quantity $A(u,n)$ then serves for the centralizing function.

The main results are established under the following assumption on a random element $X = \{X(t)\colon 0 \leqslant t \leqslant 1\}$ in $\mathbb{D}[0,1]$.

Condition A. For some $0 < \eta < 1$ and for all but countably many $t \in (\eta, 1)$, the distribution of $X(1) - X(t)$ is absolutely continuous with respect to the Lebesgue measure on the real line.

THEOREM 1. *Suppose that X is a random element in $\mathbb{D}[0,1]$ for which Condition A holds and that the distribution of $X(1)$ is nondegenerate. Then $H_n \Rightarrow X$ if and only if $X_n \Rightarrow X$.*

Remark 1. The limiting process X in Theorem 1 is necessarily a process with independent increments and satisfies $P(X(0) = 0) = 1$. It is interesting to note that the convergence of the process defined through the partial sums of dependent random variables is equivalent to the convergence of the process defined through the partial sums of the corresponding independent random variables. This holds in spite of the strong dependent structure on $\{k_j(\sigma)\colon \frac{1}{2}n \leqslant j \leqslant n\}$. Such a result is not possible for the one-dimensional convergence of $\nu_{n,\theta} \cdot X(1)^{-1}$. Here and in what follows, $P \cdot Y^{-1}$ denotes the distribution of a random element Y with respect to the probability measure P.

Remark 2. Suppose that X is a process with independent increments. Since $X(t)$ and $X(1) - X(t)$ are independent and $X(1) - X(t)$ has an absolutely continuous distribution for some $0 < t < 1$, it follows that $X(1)$ has an absolutely continuous distribution. Hence, the distribution of $X(1)$ is nondegenerate. Thus, if $X_n \Rightarrow X$, then the assumption of nondegeneracy of the distribution of $X(1)$ is redundant and can be dropped from the 'if' part of Theorem 1. Instead of Condition A, Theorem 1 can be established under the assumptions that X is a process with independent increments and that the characteristic function ϕ_{t1} of $X(1) - X(t)$ satisfies $\inf_\lambda |\phi_{t1}(\lambda)| = 0$ for all $t \in (\eta, 1)$ for some $0 < \eta < 1$.

In the following theorem, we confine ourselves to the stable limit processes.

THEOREM 2. *Let X be a process with independent increments satisfying $P(X(0) = 0) = 1$ and let*

$$\phi(\lambda; t) = \mathbf{E}\left(e^{i\lambda X(t)}\right) = \exp\left\{ta_1 \int_{-\infty}^{0} (e^{i\lambda u} - 1 - i\lambda u^*)\,d(|u|^{-\alpha}) - ta_2 \int_{0}^{\infty} (e^{i\lambda u} - 1 - i\lambda u^*)\,d(u^{-\alpha})\right\},$$

$$a_1, a_2 \geqslant 0, \quad a_1 + a_2 > 0, \quad 0 < \alpha < 2, \quad 0 \leqslant t \leqslant 1, \quad \lambda \in \mathbb{R}.$$

In order that $H_n \Rightarrow X$, it is necessary and sufficient that, for every $u > 0$,

$$\sum_{\substack{j \leqslant n \\ a(j) < -u\beta(n)}} j^{-1} \to a_1\theta^{-1}u^{-\alpha}, \qquad \sum_{\substack{j \leqslant n \\ a(j) > u\beta(n)}} j^{-1} \to a_2\theta^{-1}u^{-\alpha}, \tag{2.2}$$

and

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \beta^{-2}(n) \sum_{\substack{j \leqslant n \\ |a(j)| < \varepsilon\beta(n)}} a(j)^2 j^{-1} = 0. \tag{2.3}$$

Remark 3. Since, in Theorem 2, $X(s)$ and $X(t) - X(s)$ are independent and $X(t) = X(s) + (X(t) - X(s))$ for $0 \leqslant s \leqslant t \leqslant 1$, the characteristic function of $X(t) - X(s)$ equals $\phi(.;t)/\phi(.;s) = \phi(.;t-s)$. Hence, the marginal distributions of $\{X(t); 0 \leqslant t \leqslant 1\}$ completely determine the measure induced by X on $\mathbb{D}[0,1]$.

Remark 4. The choice of the 'time' index function $\{y(t)\colon\ 0 \leqslant t \leqslant 1\}$ makes possible to derive the functional limit result from the one-dimensional weak convergence. However, $H_n(1) \Rightarrow X(1)$ does not imply $X_n(1) \Rightarrow X(1)$. The counterexample given in [4] illustrates this in the case X is the Brownian motion. The following example illustrates this fact when $\theta = 1$ and $X(1)$ has a stable distribution. The example shows that $H_n(1)$ weakly converges to a stable distribution, while the corresponding independent version $X_n(1)$ converges to an entirely different distribution. The details of the example are available in [5].

Counterexample. Let $0 < \alpha < 2$. Let F denote the distribution function of the stable law with characteristic function ϕ_α given by $\phi_\alpha(s) = \mathrm{e}^{-|s|^\alpha}$. Define

$$h_j(1) = \begin{cases} j^{1/\alpha}F^{-1}(\{j\sqrt{2}\}) & \text{if} \quad |F^{-1}(\{j\sqrt{2}\})| \leqslant j^{1/\alpha}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\{x\}$ denotes the fractional part of x. If $\beta(n) = n^{1/\alpha}$, then, for the completely additive function $h(\sigma) = \sum_{j=1}^n h_j(1)k_j(\sigma)$, we have

$$\nu_{n,1}\left(\frac{1}{\beta(n)}h(\sigma) - A(n,n) < x\right) \longrightarrow F(x)$$

for all x. However, the distribution of

$$X_n(1) = \frac{1}{\beta(n)} \sum_{j \leqslant n} h_j(1)\xi_j - A(n,n)$$

converges to the distribution with characteristic function ϕ given by

$$\phi(\lambda) = \exp\left\{ \int_0^1 \frac{1}{y}\Big(\phi_\alpha(\lambda y^{1/\alpha}) - 1\Big)\mathrm{d}y \right\}.$$

As an extension of Theorem 1 of [4], we have:

THEOREM 3. *For the weak convergence of H_n to a standard Brownian motion, it is necessary and sufficient that*

$$\sum_{\substack{j \leqslant n \\ a(j) \leqslant u\beta(n)}} \left(\frac{a(j)}{\beta(n)}\right)^{*^2} \frac{\theta}{j} \longrightarrow \begin{cases} 1 & \textit{for} \quad u > 0, \\ 0 & \textit{for} \quad u < 0. \end{cases}$$

3. AUXILIARY RESULTS

We shall show later that our problem easily reduces to considering processes with $h_j(k) = kh_j(1) = ka(j)$ for all $j \geqslant 1$ and $k \geqslant 0$. Let

$$\widehat{H}_n := \widehat{H}_n(\sigma, t) := \frac{1}{\beta(n)} \sum_{j \leqslant y(t)} a(j) k_j(\sigma) - A\big(y(t), n\big). \tag{3.1}$$

For $0 \leqslant r \leqslant n$, let $\widehat{H}_n^r := \widehat{H}_n^r(\sigma, t)$, $H_n^r := H_n^r(\sigma, t)$, and $X_n^r := X_n^r(t)$ be the processes obtained from $\widehat{H}_n$, H_n, and X_n, respectively, by substituting $y(t) \wedge r$ for $y(t)$. We now present the basic techniques in the following lemmas.

LEMMA 1. *Suppose that, for some* $0 < \delta < 1 < C$, *we have* $\beta(\delta n) \leqslant C\beta(n)$ *for all* $n \geqslant 1$ *and* $B(n, n) - B(\delta n, n) = \mathrm{o}(1)$. *Then*

$$B(n, n) - B(\varepsilon n, n) = \mathrm{o}(1) \tag{3.2}$$

for each $0 < \varepsilon < 1$.

Proof. Clearly, for any b and d satisfying $0 \leqslant b \leqslant dC$, we have $b^* \leqslant d^* C$. Hence, for any $j \geqslant 1$,

$$|(a_j/\beta(n))^*| \leqslant |(a_j/\beta(\delta n))^*| C.$$

This implies that

$$\begin{aligned} B(n, n) - B(\delta^2 n, n) &= B(\delta n, n) - B(\delta^2 n, n) + B(n, n) - B(\delta n, n) = B(\delta n, n) - B(\delta^2 n, n) + \mathrm{o}(1) \\ &\leqslant \big(B(\delta n, \delta n) - B(\delta^2 n, \delta n)\big) C^2 + \mathrm{o}(1) = \mathrm{o}(1). \end{aligned}$$

The lemma follows by a repeated use of this argument.

LEMMA 2. *For any* $\gamma > 0$ *and* $0 \leqslant r \leqslant n$, *we have*

$$P_{n,r}(\gamma) := P\left(\sup_t |X_n(t) - X_n^r(t)| \geqslant \gamma \right) \ll B(n, n) - B(r, n) + \mathrm{o}(1). \tag{3.3}$$

Proof. Note that, as $\beta(n) \to \infty$,

$$\left| \sum_{j \leqslant u} \mathbf{E} X_{nj}^* - A(u, n) \right| \leqslant \theta \sum_{j=1}^{\infty} \left| \frac{a(j)}{\beta(n)} \right|^* \frac{|\mathrm{e}^{-\theta/j} - 1|}{j} + \sum_{j=1}^{\infty} \sum_{k=2}^{\infty} \left| \frac{a(j)k}{\beta(n)} \right|^* \left(\frac{\theta}{j} \right)^k \frac{\mathrm{e}^{-\theta/j}}{k!} = \mathrm{o}(1) \tag{3.4}$$

holds uniformly in $u \geqslant 1$, which yields

$$\max_{0 \leqslant r \leqslant n} \max_{r \leqslant k \leqslant n} \left| \sum_{r < j \leqslant k} \mathbf{E} X_{nj}^* - (A(k, n) - A(r, n)) \right| = \mathrm{o}(1).$$

For all $\gamma > 0$, by the standard arguments we obtain that

$$\begin{aligned} P_{n,r}(\gamma) &\leqslant P\left(\bigcup_{r < j \leqslant n} (X_{nj} \neq X_{nj}^*) \right) + P\left(\max_{r < k \leqslant n} \left| \sum_{r < j \leqslant k} (X_{nj}^* - \mathbf{E} X_{nj}^*) \right| \geqslant \gamma/2 \right) \\ &\quad + \mathrm{o}(1) \ll B(n, n) - B(r, n) + \mathrm{o}(1). \end{aligned}$$

This completes the proof.

LEMMA 3. *Suppose that Condition A holds for a random element X in $\mathbb{D}[0,1]$. If $X_n \Rightarrow X$, then* (3.2) *holds and*

$$X_n^{\varepsilon n} \Rightarrow X \tag{3.5}$$

for each $0 < \varepsilon < 1$.

Proof. Let $0 < \varepsilon < 1$ and $\varepsilon n > 1$. If $\varepsilon n \leqslant y(t) = y_n(t) \leqslant n$, then, for all real λ,

$$\begin{aligned} \left|\mathbf{E}e^{i\lambda(X_n(1)-X_n(t))}\right| &= \left|\exp\Big\{\theta \sum_{y(t)<j\leqslant n} \frac{1}{j}\big(e^{ia(j)\lambda/\beta(n)}-1\big)\Big\}\right| = \exp\Big\{\theta \sum_{y(t)<j\leqslant n} \frac{1}{j}\big(\cos(a(j)\lambda/\beta(n))-1\big)\Big\} \\ &\geqslant \exp\Big\{-2\theta \sum_{y(t)<j\leqslant n}\frac{1}{j}\Big\} \geqslant \exp\Big\{-2\theta\int_{y(t)}^{n}\frac{dx}{x}\Big\} \geqslant \exp\Big\{-2\theta\int_{n\varepsilon}^{n}\frac{dx}{x}\Big\} \geqslant \varepsilon^{2\theta}. \end{aligned} \tag{3.6}$$

Note that $X_n \Rightarrow X$ implies that X is a process with independent increments and that $X_n(1)-X_n(t) \Rightarrow X(1)-X(t)$ for all but countable $t \in [0,1]$ ([6], p. 124). Define $t_n = \max\{s \leqslant 1\colon\ y(s) \leqslant \varepsilon n\}$. If $t_n \to t' < 1$ for some subsequence $n := n' \to \infty$, then $y_n(t) > \varepsilon n$ for some $t \in (\max\{\eta, t'\}, 1)$, for which $X_n(1) - X_n(t) \Rightarrow X(1) - X(t)$. But, in this case, the distribution of $X(1) - X(t)$ is absolutely continuous. Consequently, by the Riemann–Lebesgue lemma, its characteristic function ϕ_{t1} satisfies $\phi_{t1}(\lambda) \to 0$ as $\lambda \to \infty$. This contradicts (3.6) and yield, $t_n \to 1$. Thus, we have

$$1 + o(1) = t_n \leqslant \frac{B(y(t_n)+1, n)}{B(n,n)} \leqslant \frac{B(\varepsilon n+1, n)}{B(n,n)} \leqslant 1. \tag{3.7}$$

Since $X_n(1) \Rightarrow X(1)$, by Lemma 9 of Chapter 4 of [16], it follows that $B(n,n)$ is bounded. Since $0 \leqslant B(\varepsilon n+1, n) - B(\varepsilon n, n) = o(1)$, (3.2) now follows from (3.7). Finally (3.2) and Lemma 2 yield (3.5). This completes the proof of the lemma.

LEMMA 4 ([1], Theorem 3). *For $0 < \varepsilon < 1$ and $2 \leqslant r \leqslant \varepsilon n$, we have*

$$\left\|\nu_{n,\theta}\cdot \widehat{H}_n^{r^{-1}} - P\cdot X_n^{r^{-1}}\right\| \leqslant \varepsilon\theta\big(\theta + (1-\varepsilon)^{-1}\big),$$

where $\|\cdot\|$ denotes the total-variation norm.

Let, for brevity, $\mathcal{L}(I)$ be the linear space of real functions g on $I \subset \mathbb{R}$ with the finite supremum norm.

LEMMA 5 ([4]). *Let*

$$h(\sigma,t) = h_1(k_1(\sigma),t) + \cdots + h_n(k_n(\sigma),t),$$

where $h_j(k,t)$, $t \in I \subset \mathbb{R}$, be a set of real arrays of functions on I such that $h_j(0,t) = 0$ and $h_j(k,\cdot) \in \mathcal{L}(I)$ for $k \geqslant 0$, $j \leqslant n$, and $t \in I$. Denote $\Xi_n(t) = h_1(\xi_1,t) + \cdots + h_n(\xi_n,t)$. Then, for any $g \in \mathcal{L}(I)$ and $x \geqslant 0$,

$$\nu_{n,\theta}\Big(\sup_{t\in I}|h(\sigma,t)-g(t)| \geqslant x\Big) \leqslant C(\theta)\Big(P^{\theta\wedge 1}\Big(\sup_{t\in I}|\Xi_n(t)-g(t)| \geqslant x/3\Big) + n^{-\theta}\Big).$$

Here $C(\theta)$ is a positive constant depending only on θ.

In the necessity part of Theorem 1, we will use an estimate of the mean values of multiplicative functions defined on permutations having only long cycles.

LEMMA 6 ([4]). *For $b(j) \in \mathbb{C}$, $1 \leqslant j \leqslant n$, and $\sigma \in \mathbb{S}_n$, let*

$$f(\sigma) = \prod_{j=1}^{n} b(j)^{k_j(\sigma)}, \quad 0^0 := 1.$$

If $b(j) = 1$ for all but $j \in J \subset (n/2, n]$, then

$$M_n(f) := \frac{1}{\theta_{(n)}} \sum_{\sigma \in \mathbb{S}_n} \theta^{w(\sigma)} f(\sigma) = 1 + \theta \sum_{j \in J} \frac{b(j) - 1}{j} \frac{n!}{\theta_{(n)}} \frac{\theta_{(n-j)}}{(n-j)!}.$$

Moreover, if $|b(j)| \leqslant 1$ and $J \subset \big((1-\delta)n, n\big]$ for $0 < \delta < \delta(\theta) < 1/2$, then

$$|M_n(f)| > c(\theta) > 0,$$

provided that $\delta(\theta)$ is sufficiently small and n is sufficiently large, $n > n(\theta)$.

The next lemma shows that the values $h_j(k)$ for $k \geqslant 2$ and $1 \leqslant j \leqslant n$ can be neglected.

LEMMA 7. *The measures $\nu_{n,\theta} \cdot H_n^{-1}$ and $\nu_{n,\theta} \cdot \widehat{H}_n^{-1}$ can converge only simultaneously and to the same limit.*

Proof. Since $\beta(n) \to \infty$, we have by Lemma 5 that, for any $\gamma > 0$ and $K > 2$,

$$\begin{aligned}
\nu(\gamma; n, \theta) &:= \nu_{n,\theta}\Big(\sup_t \big| H_n(\sigma, t) - \widehat{H}_n(\sigma, t) \big| > \gamma \Big) \\
&\leqslant C(\theta) \bigg(P^{\theta \wedge 1} \bigg(\sup_t \bigg| \sum_{j \leq y(t)} \big(h_j(\xi_j) - a(j)\xi_j \big) \bigg| \geqslant \gamma\beta(n)/3 \bigg) + n^{-\theta} \bigg) \\
&\ll P^{\theta \wedge 1}\big(\exists j \leqslant K\colon\ \xi_j \geqslant K \big) + P^{\theta \wedge 1}\big(\exists j \geqslant K\colon\ \xi_j \geqslant 2 \big) \\
&\quad + P^{\theta \wedge 1}\bigg(\sum_{j \leqslant K} \big(|h_j(\xi_j)| + |a(j)|\xi_j \big) \geqslant \gamma\beta(n)/3,\ \xi_i \leqslant K\ \forall\, i \leqslant K \bigg) + \mathrm{o}(1) \\
&\ll \bigg(\sum_{j \leqslant K} \sum_{k \geqslant K} \mathrm{e}^{-\theta/j} \frac{\theta^k}{j^k\, k!} \bigg)^{\theta \wedge 1} + \bigg(\sum_{j \geqslant K} \sum_{k \geqslant 2} \mathrm{e}^{-\theta/j} \frac{\theta^k}{j^k\, k!} \bigg)^{\theta \wedge 1} + \mathrm{o}_K(1) \\
&\ll \bigg(\sum_{j \leqslant K} \frac{\theta^K}{j^K\, K!} \bigg)^{\theta \wedge 1} + \bigg(\sum_{j \geqslant K} \frac{\theta^2}{j^2} \bigg)^{\theta \wedge 1} + \mathrm{o}_K(1) \ll K^{-\theta \wedge 1} + \mathrm{o}_K(1).
\end{aligned}$$

Here $\mathrm{o}_K(1)$ denotes the estimate $\mathrm{o}(1)$ depending on K. Hence $\nu(\gamma; n, \theta) = \mathrm{o}(1)$ for arbitrary $\gamma > 0$. Consequently, the measures $\nu_{n,\theta} \cdot H_n^{-1}$ and $\nu_{n,\theta} \cdot \widehat{H}_n^{-1}$ can converge only simultaneously and to the same limit.

4. PROOFS OF THE RESULTS

Without loss of generality, by Lemma 7 we can further assume that

$$h_j(k_j(\sigma)) = a(j)k_j(\sigma).$$

Proof of Theorem 1. Suppose that $X_n \Rightarrow X$, $0 < \varepsilon < 1$, and $r = \varepsilon n$. Then by Lemmas 2, 3, and 5, we get that, for any $\gamma > 0$,

$$\nu_{n,\theta}\Big(\sup_t |H_n(\sigma, t) - H_n^r(\sigma, t)| \geqslant \gamma \Big) \leqslant C(\theta) P_{n,r}^{\theta \wedge 1}(\gamma/3) + \mathrm{o}(1) = \mathrm{o}(1). \tag{4.1}$$

The weak convergence $H_n \Rightarrow X$ now follows from (3.5) of Lemma 3 and Lemma 4.

Conversely, suppose thst $H_n \Rightarrow X$. Then $H_n(., 1) - H_n(., t) \Rightarrow X(1) - X(t)$ for all but countable $t \in (0, 1)$. We shall first prove that $B(n, n)$ is bounded. Let $s'_n = (1 + B(n, n))^{-1/2}$ and $r = r(n) = y(s'_n)$. Suppose that $r > \varepsilon n$ for some $0 < \varepsilon < 1$. Then we clearly have

$$B(\varepsilon n, n) \leqslant B(y(s'_n), n) \leqslant s'_n B(n, n) \leqslant B(n, n)^{1/2}$$

and

$$B(n, n) - B(\varepsilon n, n) \leqslant 1 + \log(1/\varepsilon) \ll 1.$$

These inequalities together imply that

$$B(n, n) - B(n, n)^{1/2} \ll 1$$

and, hence, $B(n, n) \ll 1$. Thus, if $B(n, n) \to \infty$ for some subsequence $n := n_k \to \infty$, then $r(n) = \mathrm{o}(n)$ and $s'_n \to 0$. Consequently, Lemma 4 and the tightness of the family of measures $\nu_{n,\theta} \cdot H_n^{-1}$ (see [6]) imply

$$\begin{aligned} P\Bigg(\Bigg|\sum_{j \leqslant r} X_{nj} - A(r, n)\Bigg| \geqslant \gamma\Bigg) &= \nu_{n,\theta}\Bigg(\Bigg|\sum_{j \leqslant r} a(j)k_j(\sigma)/\beta(n) - A(r, n)\Bigg| \geqslant \gamma\Bigg) + \mathrm{o}(1) \\ &= \nu_{n,\theta}\big(\big|H_n(\sigma, s'_n)\big| \geqslant \gamma\big) + \mathrm{o}(1) = \mathrm{o}(1) \end{aligned}$$

for any $\gamma > 0$. Hence (see Theorem 1 on p. 258 of [14]), $B(r, n) = \mathrm{o}(1)$. But since $\beta(n) \to \infty$, we have

$$B(r, n) = B(r + 1, n) + \mathrm{o}(1) \geqslant B(n, n)^{1/2} + \mathrm{o}(1),$$

which contradicts $\limsup_{n \to \infty} B(n, n) = \infty$. This establishes $B(n, n) \ll 1$.

We shall now establish (3.2). Let the constant $\delta(\theta) > 0$ in Lemma 6 be sufficiently small. Let $0 < \delta < \delta(\theta)$ and $\tau_n = \sup\{u \leqslant 1\colon\ y(u) \leqslant (1 - \delta)n\}$. Suppose that $\tau_n \to t_0 < 1$ for some subsequence $n := n' \to \infty$. Then $y(t) > (1 - \delta)n$ for n sufficiently large and for some $t \in (\max\{\eta, t_0\}, 1)$ such that $H_n(., 1) - H_n(., t) \Rightarrow X(1) - X(t)$. Using Lemma 6 with

$$b(j) = \begin{cases} \exp\{i\lambda a(j)/\beta(n)\} & \text{if } \quad y(t) < j \leqslant n, \\ 1, & \text{otherwise,} \end{cases}$$

and $\lambda \in \mathbb{R}$, we note that the characteristic function ϕ_{t1} of $X(1) - X(t)$ satisfies

$$|\phi_{t1}(\lambda)| > c(\theta) > 0$$

uniformly in $\lambda \in \mathbb{R}$. By the Riemann–Lebesgue lemma, this contradicts the assumption that the distribution of $X(1) - X(t)$ is absolutely continuous. Thus, from the definitions of $y(t)$ and the sequence τ_n it follows that

$$1 + \mathrm{o}(1) \leqslant \tau_n \leqslant \frac{B\big(y(\tau_n) + 1, n\big)}{B(n, n)} \leqslant \frac{B\big((1 - \delta)n + 1, n\big)}{B(n, n)} \leqslant \frac{B\big(un, n\big)}{B(n, n)} \leqslant 1$$

for any $1 - \delta/2 \leqslant u \leqslant 1$ and $n > 2/\delta$. These inequalities and $B(n, n) \ll 1$ together imply

$$B(n, n) - B(un, n) = \mathrm{o}(1) \tag{4.2}$$

for any $1 - \delta/2 \leqslant u \leqslant 1$. Since $X(1)$ has a nondegenerate distribution, it follows that there exists c_0 such that

$$0 < c_0 < B(n, n) \tag{4.3}$$

for all large n. We shall now establish

$$\beta(un) \ll \beta(n) \tag{4.4}$$

for $u \in [1-\delta/2, 1]$. Note that, in this case, $2u > 1$. Substituting nu for n several times, we arrive at

$$\sum_{u^K n \leqslant j \leqslant n} \left(\frac{a(j)}{\beta(n,K)} \right)^{*2} \frac{1}{j} = \mathrm{o}(1) \tag{4.5}$$

for any fixed $K \geqslant 1$ with $\beta(n,K) := \max_{0 \leqslant l \leqslant K} \beta(u^l n)$. From (4.5) we can get a strictly increasing sequence of integers $n_m > 2^m$ such that, for all $n \geqslant n_m$,

$$\sum_{u^m n \leqslant j \leqslant n} \left(\frac{a(j)}{\beta(n,m)} \right)^{*2} \frac{1}{j} < \frac{1}{m}. \tag{4.6}$$

Let $K = K(n) := m$ for $n_m \leqslant n < n_{m+1}$ and $\tilde{\beta}(n) = \beta(n, K)$. Since $u^K n \geqslant u^m n_m \geqslant (2u)^m$, it is clear that $u^K n \to \infty$ and (4.5) holds with $K = K(n) \to \infty$. Note that $\tilde{\beta}(n) =: \beta(u^{l_0} n)$ for some $1 \leqslant l_0 = l(n,K) \leqslant K$ and $r = r(n) := u^K n = \mathrm{o}(n)$. If $\tilde{\beta}(n)/\beta(n) \to \infty$ for a subsequence $n := n' \to \infty$, then $\frac{\beta(n)}{\tilde{\beta}(n)} H_n(\cdot, 1) \Rightarrow 0$ and, by (4.5),

$$\frac{\beta(n)}{\tilde{\beta}(n)} H_n^r(\cdot, 1) - A_n \Rightarrow 0 \tag{4.7}$$

for some sequence of real numbers A_n. By (4.5), (4.7), and Lemma 4, it follows that $\tilde{\beta}(n)^{-1} \sum_{j=1}^n a(j)\xi_j - c_n \Rightarrow 0$ for some sequence of real numbers c_n. Thus, we have, by Theorem 4 of Chapter 9 of [16], that

$$\mathrm{o}(1) = \sum_{j \leqslant n} \left(\frac{a(j)}{\tilde{\beta}(n)} \right)^{*2} \frac{1}{j} \geqslant B(u^{l_0} n, u^{l_0} n),$$

which contradicts (4.3). This proves the boundedness of $\tilde{\beta}(n)/\beta(n)$, which, in turn, implies (4.4). Hence (3.2) holds by Lemma 1. Thus, (3.3) and another application of Lemma 5 yield (4.1) for $r = \varepsilon n$, for all $0 < \varepsilon < 1$ and $\gamma > 0$. So, by the assumption that $H_n \Rightarrow X$, we obtain that $H_n^{\varepsilon n} \Rightarrow X$ for all $0 < \varepsilon < 1$. Now Lemma 4 and another application of (3.3) complete the proof of Theorem 1.

Proof of Theorem 2. We first note that $P(X(0) = 0) = 1$ and that $X(1)$ has a nondegenerate distribution. Hence, by Theorem 1, it suffices to establish that $X_n \Rightarrow X$ if and only if (2.2) and (2.3) hold. Since $\beta(n) \to \infty$, we have, by the Markov inequality and without any additional assumptions, that

$$\max_{j \leqslant n} P(|X_{nj}| \geqslant \varepsilon) \leqslant \max_{j \leqslant n} \left\{ \frac{\theta}{\varepsilon j} \left| \frac{a_j}{\beta(n)} \right|^* \right\} \longrightarrow 0 \tag{4.8}$$

for $0 < \varepsilon < 1$. So $\{X_{nj}, 1 \leqslant j \leqslant n\}$ is an infinitesimal array. We now use Theorem 1 of [13] to establish $X_n \Rightarrow X$ under (2.2) and (2.3). By (2.3) we clearly have

$$\lim_{\tau \to 0} \limsup_{n \to \infty} \sum_{j=1}^{n} \left(\int_{|x| < \tau} x \mathrm{d}P(X_{nj} < x) \right)^2 = 0.$$

To verify the remaining assumptions of Theorem 1 of [13], let

$$Q_{nl}(u) = \theta \sum_{\substack{j \leqslant l \\ a(j) \leqslant u\beta(n)}} \left(\frac{a(j)}{\beta(n)} \right)^{*2} \frac{1}{j}, \quad u \in \overline{\mathbb{R}},\ 1 \leqslant l \leqslant n.$$

We have

$$\Psi_{nl}(u) := \sum_{j\leqslant l}\int_{-\infty}^{u} x^{*^2}\,\mathrm{d}P(X_{nj} < x) = \theta \sum_{\substack{j\leqslant l\\ a(j)\leqslant u\beta(n)}} \left(\frac{a(j)}{\beta(n)}\right)^{*^2}\frac{1}{j} + \mathrm{O}\Biggl(\sum_{j=1}^{\infty}\left(\frac{a(j)}{\beta(n)}\right)^{*^2}\frac{|\mathrm{e}^{-\theta/j}-1|}{j}\Biggr)$$

$$+\,\mathrm{O}\Biggl(\sum_{j=1}^{\infty}\sum_{k=2}^{\infty}\left(\frac{ka(j)}{\beta(n)}\right)^{*^2}\left(\frac{\theta}{j}\right)^{k}\frac{1}{k!}\Biggr) = \theta \sum_{\substack{j\leqslant l\\ a(j)\leqslant u\beta(n)}} \left(\frac{a(j)}{\beta(n)}\right)^{*^2}\frac{1}{j} + \mathrm{o}(1) = Q_{nl}(u) + \mathrm{o}(1),$$

uniformly in $u \in \overline{\mathbb{R}}$ and $1 \leqslant l \leqslant n$. As a result, even though $y(t)$ in (2.1) is defined through $B(l, n)$ instead of $\Psi_{nl}(+\infty)$, the proof of Theorem 1 of [14] goes through in view of its equation (9). The choice of centralizing constants $A(u, n)$ is guided by (3.4). According to Theorem 1 of [14], to establish the sufficiency of conditions (2.2) and (2.3) for the weak convergence, it suffices to show that

$$\Psi_{nn} \rightrightarrows \Psi, \tag{4.9}$$

where $\rightrightarrows$ means the weak convergence of nondecreasing functions including the convergence at the points $\pm\infty$, and

$$\Psi(u) = \begin{cases} a_1|u|^{-\alpha} & \text{for} \quad u \leqslant -1,\\ a_1 + a_1\alpha(2-\alpha)^{-1}(1-|u|^{2-\alpha}) & \text{for} \quad -1 < u \leqslant 0,\\ 2a_1(2-\alpha)^{-1} + \alpha a_2(2-\alpha)^{-1}u^{2-\alpha} & \text{for} \quad 0 < u \leqslant 1,\\ 2(a_1+a_2)(2-\alpha)^{-1} - a_2u^{-\alpha} & \text{for} \quad u > 1. \end{cases}$$

To verify (4.9) for $u < 0$, we have by the dominated convergence theorem that

$$\Psi_{nn}(u) = Q_{nn}(u) + \mathrm{o}(1) = \theta\int_{-\infty}^{u} x^{*^2}\,\mathrm{d}\sum_{\substack{j\leqslant n\\ a(j)<x\beta(n)}}\frac{1}{j} + \mathrm{o}(1) = \Psi(u) + \mathrm{o}(1).$$

The remaining cases are treated similarly. Thus, $X_n \Rightarrow X$.

If $X_n \Rightarrow X$, then clearly $X_n(1) \Rightarrow X(1)$ and X is a process with independent increments. As noted in (4.8), $\{X_{nj};\, 1 \leqslant j \leqslant n\}$ is an infinitesimal array. Hence, the necessity of (4.9) or, equivalently, that of conditions (2.2) and (2.3) for $X_n(1) \Rightarrow X(1)$ is well known (see Chapter 4 of [16]). This completes the proof of Theorem 2.

Theorem 3 is proved by similar arguments.

REFERENCES

1. R. Arratia, A. D. Barbour, and S. Tavaré, Poisson process approximations for the Ewens sampling formula, *Ann. Applied Probab.*, **2**, 519–535 (1992).
2. R. Arratia and S. Tavaré, Limit theorems for combinatorial structures via discrete process approximations, *Random Structures and Algorithms*, **3**, 321–345 (1992).
3. G. J. Babu, A note on the invariance principle for additive functions, *Sankhyā A*, **35**, 307–310 (1973).
4. G. J. Babu and E. Manstavičius, Brownian motion for random permutations, *Sankhyā A*, **61**, 312–327 (1999).
5. G. J. Babu and E. Manstavičius, Random permutations and the Ewens sampling formula in genetics, in: *Probab. Theory and Math. Statist.*, *Proc. 7th Vilnius Conf.* (1998), B. Grigelionis *et al.* (Eds.), VSP, Utrecht/TEV, Vilnius (1999), pp. 33–42.
6. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York (1968).
7. J. M. DeLaurentis and B. G. Pittel, Random permutations and the Brownian motion, *Pacific J. Math.*, **119**, 287–301 (1985).
8. P. Donnelly, T. G. Kurtz, and S. Tavaré, On the functional central limit theorem for the Ewens sampling formula, *Ann. Appl. Probab.*, **1**, 539–545 (1991).

9. W. J. Ewens, The sampling theory of selectively neutral alleles, *Theoret. Pop. Biol.*, **3**, 87–112 (1972).
10. B. M. Hambly, P. Keevash, N. O'Connell, and D. Stark, The characteristic polynomial of a random permutation matrix, *Stoch. Proc. Appl.*, **90**, 335–346 (2000).
11. J. C. Hansen, A functional central limit theorem for the Ewens sampling formula, *J. Appl. Probab.*, **27**, 28–43 (1990).
12. J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Amer. Math. Soc., Providence, RI (1964).
13. E. Manstavičius, Arithmetic simulation of stochastic processes, *Lith. Math. J.*, **24**(3), 276–285 (1984).
14. E. Manstavičius, Additive functions and stochastic processes, *Lith. Math. J.*, **24**, 52–61 (1985).
15. E. Manstavičius, Additive and multiplicative functions on random permutations, *Lith. Math. J.*, **36**, 276–285 (1996).
16. V. V. Petrov, *Sums of Independent Random Variables*, Springer, New York (1975).
17. K. Wieand, Eigenvalue Distributions of Random Matrices in the Permutation Group and Compact Lie Groups, PhD thesis, Harvard University (1998).

(*Translated by G. J. Babu and E. Manstavičius*)