J. Appl. Prob. **29**, 280–290 (1992)
Printed in Israel

A CONVERGENCE THEOREM FOR SYMMETRIC FUNCTIONALS OF RANDOM PARTITIONS

PAUL JOYCE,* *University of Idaho*
SIMON TAVARÉ,** *University of Southern California*

Abstract

This paper gives general conditions under which symmetric functionals of random partitions of the integer m converge in distribution as $m \to \infty$. The main result is used to settle a conjecture of Donnelly et al. (1991) to the effect that the mean of the sum of the square roots of the relative sizes of the components of a random mapping of m integers converges to $\pi/2$ as $m \to \infty$.

RANDOM MAPPINGS; RANDOM PERMUTATIONS; POISSON–DIRICHLET DISTRIBUTION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60C05
SECONDARY 60F05, 92D10

1. Introduction

This paper is motivated by a conjecture made by Donnelly et al. (1991) in their paper on functionals of random mappings.

Consider a random mapping of the set $\{1, 2, \cdots, m\}$ into itself where each of the m^m maps is equally likely. The mapping partitions the set $\{1, 2, \cdots, m\}$ into components, integers i and j being in the same component if some functional iterate of i equals some functional iterate of j. If $M_i^{(m)}$ is the size of the ith largest component of a random map and $Y_i^{(m)} = m^{-1}M_i^{(m)}$, then it was shown by Aldous (1985) that as $m \to \infty$

$$(Y_1^{(m)}, Y_2^{(m)}, \cdots) \Rightarrow (Y_1, Y_2, \cdots), \tag{1}$$

$\Rightarrow$ denoting weak convergence in the product topology on $\bar{\nabla} = \{(x_1, x_2, \cdots): x_1 \geqq x_2 \geqq \cdots \geqq 0, \Sigma x_i \leqq 1\}$. Further, $(Y_1, Y_2, \cdots)$ has the Poisson–Dirichlet distribution with parameter $\frac{1}{2}$. (For definitions and basic properties of the Poisson–Dirichlet distribution, see Kingman (1977).)

Received 10 September 1990; revision received 7 March 1991.
* Postal address: Department of Mathematics and Statistics, University of Idaho, Moscow, ID 83843, USA.
** Postal address: Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113, USA.
The authors were supported in part by NSF grant DMS90-05833.

Let $g:[0,1]\to[0,\infty)$ be a continuous function, and define $f_g:\bar{\nabla}\to[0,\infty]$ by $f_g(x)=\sum_{i=1}^{\infty} g(x_i)$. If $g(x)=o(x)$ as $x\to 0$, Kingman (1977) showed that f_g is continuous, and hence that $f_g(Y_1^{(m)}, Y_2^{(m)},\cdots)\Rightarrow f_g(Y_1, Y_2,\cdots)$ as $m\to\infty$, using the continuous mapping theorem. In the case $g(x)=\sqrt{x}$, the limiting behavior of $f_g(Y_1^{(m)}, Y_2^{(m)},\cdots)$ cannot be obtained immediately from this result, since $\sqrt{x}$ is not $o(x)$ as $x\to 0$. In addition, Lemma 4.2 of Donnelly et al. (1991) shows that if $g(x)$ is not $o(x)$ as $x\to 0$ then f_g is discontinuous at any $y\in\bar{\nabla}$ for which $f_g(y)$ is finite. Thus further attempts to use just the continuous mapping theorem must fail in such cases. However, computer simulations described in Donnelly et al. (1991) suggest that

$$\mathbb{E}\left(\sum_{i=1}^{\infty}\sqrt{Y_i^{(n)}}\right)\to\mathbb{E}\left(\sum_{i=1}^{\infty}\sqrt{Y_i}\right)=\int_0^1 \tfrac{1}{2}x^{-1/2}(1-x)^{-1/2}dx=\frac{\pi}{2}. \tag{2}$$

We show that (2) is indeed true. In fact, for a large collection of continuous functions g we show that $\mathbb{E}\sum_{i=1}^{\infty} g(Y_i^{(m)})\to\mathbb{E}\sum_{i=1}^{\infty} g(Y_i)$, and in addition $\sum_{i=1}^{\infty} g(Y_i^{(m)})\Rightarrow\sum_{i=1}^{\infty} g(Y_i)$. This result follows because the sizes of the components of a random mapping converge weakly to the Poisson–Dirichlet distribution in a particularly nice way. However, in Section 3, we show that it is possible to construct a sequence of distributions on partitions (other than those induced by random mappings) where the normalized sizes $(Y_1^{(m)}, Y_2^{(m)},\cdots)$ converge in distribution to a Poisson–Dirichlet distributed point $(Y_1, Y_2,\cdots)$ but $\mathbb{E}\sum_{i=1}^{\infty}\sqrt{Y_i^{(m)}}\not\to\mathbb{E}\sum_{i=1}^{\infty}\sqrt{Y_i}$.

The trade-off is clear. We can assume weaker conditions on g if we can establish stronger conditions on how the relative sizes converge. Establishing such conditions, which are easily verified in cases of interest, is the central idea of this paper. In the next section we prove a general result which has the conditions necessary to establish convergence of f_g under very mild assumptions on g. In the third section we verify these conditions for two important cases: the components of random mappings, and random partitions distributed according to the Ewens sampling formula.

2. A convergence theorem

For each positive integer m, let $(A_1^{(m)}, A_2^{(m)},\cdots, A_m^{(m)})$ be a random partition of m, so that $A_i^{(m)}$ is the number of parts of size i, and $\sum_{i=1}^{m} iA_i^{(m)}=m$. Let $N_1^{(m)}, N_2^{(m)},\cdots$ be a listing of the parts in decreasing order, and define $X_i^{(m)}=m^{-1}N_i^{(m)}$. For example, if $m=10$ and $A_1^{(10)}=2$, $A_2^{(10)}=1$ and $A_6^{(10)}=1$, then $N_1^{(10)}=6$, $N_2^{(10)}=2$, $N_3^{(10)}=1$ and $N_4^{(10)}=1$. To make $(X_1^{(m)}, X_2^{(m)},\cdots)$ a point in $\bar{\nabla}$, we use the convention that $X_i^{(m)}=0$ if it has not already been defined. Another way to label the parts is to label at random according to their sizes; that is, to size-bias the sizes $N_1^{(m)}, N_2^{(m)},\cdots$. We let $C_1^{(m)}, C_2^{(m)},\cdots$ denote the parts in size-biased order, and define $Z_i^{(m)}=m^{-1}C_i^{(m)}$.

Theorem 1. (i) *Suppose that as* $m\to\infty$, $(X_1^{(m)}, X_2^{(m)},\cdots)\Rightarrow(X_1, X_2,\cdots)$, *a random point in* $\bar{\nabla}$.

Denote the size-biased permutation of $(X_1, X_2,\cdots)$ *by* $(Z_1, Z_2,\cdots)$.

(ii) *Assume that the distribution of* Z_1 *has a probability density function* $h(x)$, $x\in[0,1]$ *that is continuous for* $x\in(0,1)$.

Let $h_m : (0, 1] \to [0, \infty)$ be defined in such a way that

$$h_m\left(\frac{j}{m}\right) = j\mathbb{E}(A_j^{(m)}), \qquad j = 1, \cdots, m. \tag{3}$$

(iii) *Assume that there exists a $t_0 \in (0, 1)$ such that*

$$A := \sup_m \sup_{x \in (0, t_0]} h_m(x) < \infty, \tag{4}$$

and such that h_m converges uniformly to h on $[s, t_0]$ for $0 < s < t_0$.

(iv) *Suppose $g \in C[0, 1]$, $g \geqq 0$ and that for some $\delta > 0$ there exists a non-increasing function g^* satisfying*

$$g(x) \leqq x g^*(x), \qquad 0 < x < \delta, \tag{5}$$

$$\int_0^\delta g^*(x)dx < \infty. \tag{6}$$

Then as $m \to \infty$,

$$\mathbb{E}\left(\sum_{i=1}^\infty g(X_i^{(m)})\right) \to \mathbb{E}\left(\sum_{i=1}^\infty g(X_i)\right) = \int_0^1 \frac{g(x)}{x} h(x)dx, \tag{7}$$

and

$$\sum_{i=1}^\infty g(X_i^{(m)}) \Rightarrow \sum_{i=1}^\infty g(X_i). \tag{8}$$

Remark 1. It follows from (4) and (6) that

$$\int_0^\delta g^*(x)h(x)dx < \infty. \tag{9}$$

Remark 2. To motivate the conditions in the statement of the theorem and the proof we consider two formal arguments that establish (7).

Formal argument A. Note that if the parts of the partition of m are given in decreasing order $X_1^{(m)}, X_2^{(m)}, \cdots$ then the ith largest part is chosen to be labeled 1 under the size-biased labeling with probability $X_i^{(m)}$. Therefore if $Z_1^{(m)}$ is the part chosen first under the size-biased labeling,

$$\mathbb{E}\left(\frac{g(Z_1^{(m)})}{Z_1^{(m)}} \,\middle|\, X_1^{(m)}, X_2^{(m)}, \cdots\right) = \sum_{i=1}^\infty g(X_i^{(m)}), \tag{10}$$

implying that

$$\lim_{m\to\infty} \mathbb{E}\left(\sum_{i=1}^\infty g(X_i^{(m)})\right) = \lim_{m\to\infty} \mathbb{E}\left(\frac{g(Z_1^{(m)})}{Z_1^{(m)}}\right).$$

It was shown by Donnelly and Joyce (1989) that if the relative sizes in decreasing order converge in distribution, then the relative sizes under size-biased labeling also converge in distribution. If we formally interchange limit with expectation the limit becomes $\mathbb{E}(g(Z_1)/Z_1)$. If we replace Conditions (iii) and (iv) with the condition that $g(Z_1^{(m)})/Z_1^{(m)}$ be uniformly integrable, then interchanging expectation and limit would be justified. However, as Donnelly et al. (1991) point out, this is difficult to verify directly in cases of interest, since $x^{-1}g(x)$ is typically unbounded near 0. However, once Conditions (iii) and (iv) are verified, uniform integrability will follow as a consequence of this theorem.

Formal argument B. Note that

$$\sum_{j=1}^{m} \mathbb{E}g(X_j^{(m)}) = \sum_{j=1}^{m} g(j/m)\mathbb{E}(A_j^{(m)}) = \sum_{j=1}^{m} \frac{g(j/m)}{j/m} h_m(j/m)\frac{1}{m}.$$

Formally, if we replace h_m with h in the last equality, then the limiting mean would be $\int_0^1 x^{-1}g(x)h(x)dx$. It is difficult to make this argument precise, because $g(x)h(x)/x$ may be unbounded near $x=0$ and $x=1$. Our proof actually combines the two formal arguments. Note that in order for arguments A and B to be consistent, $h(x)$ must be the density of Z_1.

Proof. Note first that by (5) of Assumption (iv) we must have $g(0)=0$. Let t_0 be as given in Assumption (iii). Since the distribution function of each X_i is continuous except at a countable number of points, the set

$$U \equiv \bigcup_{i=1}^{\infty} \{s>0 : \mathbb{P}(X_i = s) > 0\} \tag{11}$$

is countable. For $t \in (0, t_0) \setminus U$ we have

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{\infty} g(X_i^{(m)})\right) &= \mathbb{E}\left(\sum_{i=1}^{m} A_i^{(m)} g(i/m)\right) \\
&= \mathbb{E}\left(\sum_{i=1}^{m} A_i^{(m)} g(i/m) I\{i \leqq mt\}\right) + \mathbb{E}\left(\sum_{i=1}^{m} A_i^{(m)} g(i/m) I\{i > mt\}\right) \\
&= \sum_{i=1}^{[mt]} g(i/m)\mathbb{E}(A_i^{(m)}) + \sum_{i=1}^{m} \mathbb{E}(g(X_i^{(m)}) I\{X_i^{(m)} > t\}) \\
&= \sum_{i=1}^{[mt]} \frac{g(i/m)}{i/m} h_m(i/m)\frac{1}{m} + \mathbb{E}\left(\frac{g(Z_1^{(m)})}{Z_1^{(m)}} I\{Z_1^{(m)} > t\}\right),
\end{aligned}$$

the last equality following from (10). Next note that $x^{-1}g(x)I\{x>t\}$ is bounded by $t^{-1}\sup_{x\in[0,1]} g(x)$. The dominated convergence theorem and Conditions (i) and (ii) then guarantee that

$$\lim_{m\to\infty} \mathbb{E}\left(\frac{g(Z_1^{(m)})}{Z_1^{(m)}} I\{Z_1^{(m)} > t\}\right) = \mathbb{E}\left(\frac{g(Z_1)}{Z_1} I\{Z_1 > t\}\right).$$

It remains to prove that

(12) $$\lim_{m\to\infty}\left|\sum_{i=1}^{[mt]}\frac{g(i/m)}{i/m}h_m(i/m)\frac{1}{m}-\int_0^t\frac{g(x)}{x}h(x)dx\right|=0.$$

To establish (12), choose $s\in(0,\delta)$ and use the triangle inequality:

$$\left|\sum_{i=1}^{[mt]}\frac{g(i/m)}{i/m}h_m(i/m)\frac{1}{m}-\int_0^t\frac{g(x)}{x}h(x)dx\right|$$
$$\leqq A\left(\sum_{i=1}^{[ms]}\frac{g(i/m)}{i/m}\frac{1}{m}\right)+\left(\int_0^s\frac{g(x)}{x}h(x)dx\right)$$
$$+\sup_{x\in[s,t]}|h_m(x)-h(x)|\left(\sum_{i=[ms]+1}^{[mt]}\frac{g(i/m)}{i/m}\frac{1}{m}\right)$$
$$+\left|\sum_{i=[ms]+1}^{[mt]}\frac{g(i/m)}{i/m}h(i/m)\frac{1}{m}-\int_s^t\frac{g(x)}{x}h(x)dx\right|$$

where $A\equiv\sup_m\sup_{x\in(0,t_0]}h_m(x)$ is finite by Assumption (iii). Note that $x^{-1}g(x)h(x)$ and $x^{-1}g(x)$ are bounded and continuous on $[s,t]$ and therefore Riemann integrable there. Also note that $h_m\to h$ uniformly on $[s,t]$ by Assumption (iii). Thus the third and fourth terms of the above inequality tend to 0 as $m\to\infty$. Next observe that by monotonicity of g^*,

$$0\leqq\sum_{i=1}^{[ms]}\frac{g(i/m)}{i/m}\frac{1}{m}\leqq\int_0^s g^*(x)dx,$$

which may be made arbitrarily small by letting $s\to 0$. Finally, the term $\int_0^s x^{-1}g(x)h(x)dx\to 0$ as $s\to 0$, by (9). This completes the proof of (7).

To establish (8) we first employ Skorohod's theorem (cf. Ethier and Kurtz (1986), p. 102) to conclude that $(X_1^{(m)},X_2^{(m)},\cdots)\to(X_1,X_2,\cdots)$ as $m\to\infty$ almost surely on some probability space.

Let $\varepsilon>0$ be given. Using (11) choose $s\equiv s(\varepsilon)\in(0,t_0)\setminus U$ so small that

$$\int_0^s\frac{g(x)}{x}h(x)dx<\frac{\varepsilon^2}{2},$$

and choose $R\equiv R(s,\varepsilon)$ so large that $\mathbb{P}(X_R>s)<\varepsilon/2$ and

(13) $$\mathbb{P}\left(\sum_{i=R+1}^{\infty}g(X_i)>\varepsilon\right)<\varepsilon.$$

This last choice is possible because $\mathbb{E}\sum_{i=1}^{\infty}g(X_i)\equiv\mathbb{E}g(Z_1)/Z_1<\infty$ by (9). Since $\mathbb{P}(X_R=s)=0$, $\mathbb{P}(X_R^{(m)}>s)\to\mathbb{P}(X_R>s)$, so that there exists N_0 such that for all $m>N_0$

(14) $$\mathbb{P}(X_R^{(m)}>s)<\frac{\varepsilon}{2}.$$

By (12),

$$\lim_{m\to\infty} \sum_{i=1}^{[ms]} \mathbb{E}(A_i^{(m)}) g(i/m) = \int_0^s \frac{g(x)}{x} h(x) dx,$$

so that there exists N_1 such that for all $m > N_1$

$$\mathbb{P}\left(\sum_{i=1}^{[ms]} A_i^{(m)} g(i/m) > \varepsilon\right) \leqq \sum_{i=1}^{[ms]} \frac{\mathbb{E}(A_i^{(m)}) g(i/m)}{\varepsilon} < \frac{\varepsilon}{2}. \tag{15}$$

Choose $N = \max\{N_0, N_1\}$. Note that $\Sigma_{i=1}^{[ms]} A_i^{(m)} g(i/m) \leqq \varepsilon$ and $X_R^{(m)} \leqq s$ imply that

$$\sum_{i=R+1}^{\infty} g(X_i^{(m)}) \leqq \sum_{\{i: X_i^{(m)} \leqq s\}} g(X_i^{(m)}) = \sum_{i=1}^{[ms]} A_i^{(m)} g(i/m) \leqq \varepsilon.$$

It follows from (14) and (15) that if $m > N$ then

$$\mathbb{P}\left(\sum_{i=R+1}^{\infty} g(X_i^{(m)}) > \varepsilon\right) < \varepsilon. \tag{16}$$

The proof may be completed by using the inequality

$$\begin{aligned}
&\mathbb{P}\left(\left|\sum_{i=1}^{\infty} g(X_i^{(m)}) - \sum_{i=1}^{\infty} g(X_i)\right| > 3\varepsilon\right) \\
&\leqq \mathbb{P}\left(\sum_{i=R+1}^{\infty} g(X_i^{(m)}) > \varepsilon\right) + \mathbb{P}\left(\sum_{i=R+1}^{\infty} g(X_i) > \varepsilon\right) \\
&\quad + \mathbb{P}\left(\left|\sum_{i=1}^{R} g(X_i^{(m)}) - \sum_{i=1}^{R} g(X_i)\right| > \varepsilon\right).
\end{aligned}$$

The first term is controlled by (16) and the second by (13), and the third may be made small for large enough m by Condition (i) of Theorem 1.

Remark 3. For general g, if we assume that both g^+ and g^- satisfy the conditions of the theorem, then its conclusion will hold for $g = g^+ - g^-$.

Remark 4. If $g : [0, 1]^k \to [0, \infty)$ is continuous, and $f_g : \bar{\nabla} \to [0, \infty)$ is given by

$$f_g(x_1, x_2, \cdots) = \sum_{i_1, \cdots, i_k \text{ distinct}} g(x_{i_1}, \cdots, x_{i_k}),$$

then similar arguments may be used to study weak convergence and convergence of means in this setting.

3. Some examples

Example 1. Consider a random partition $(A_1^{(m)}, A_2^{(m)}, \cdots, A_m^{(m)})$ of the integer m with probability distribution given by the Ewens sampling formula (Ewens (1972)):

$$\mathbb{P}(A_1^{(m)} = a_1, \cdots, A_m^{(m)} = a_m) = \frac{m!}{\theta_{(m)}} \prod_{j=1}^{m} \left(\frac{\theta}{j}\right)^{a_j} \frac{1}{a_j!}, \tag{17}$$

where $\theta_{(m)} = \theta(\theta+1)\cdots(\theta+m-1)$, $\theta > 0$, and $\Sigma_{j=1}^{m} ja_j = m$. Let $N_1^{(m)}, N_2^{(m)}, \cdots$ be a listing of the parts in decreasing order and $X_i^{(m)} = m^{-1}N_i^{(m)}$. Kingman (1977) showed that as $m \to \infty$

$$(X_1^{(m)}, X_2^{(m)}, \cdots) \Rightarrow (X_1, X_2, \cdots), \tag{18}$$

where $(X_1, X_2, \cdots)$ has the Poisson–Dirichlet distribution with parameter θ. We will verify Conditions (iii) of Theorem 1 by showing that $h_m(x) \to h(x)$ uniformly on the interval $[0, t]$ for all $0 < t < 1$. In this case $h(x) = \theta(1-x)^{\theta-1}$, which is the density function of Z_1. Watterson (1974) showed that

$$\mathbb{E}(A_i^{(m)}) = \frac{\theta}{i} \frac{m(m-1)\cdots(m-i+1)}{(\theta+m-1)\cdots(\theta+m-i)},$$

so from (3) we may take

$$h_m(x) = \frac{\theta\Gamma(m+1)\Gamma(m(1-x)+\theta)}{\Gamma(m+\theta)\Gamma(m(1-x)+1)}.$$

Corollary 1. *If g satisfies Assumption* (iv) *of Theorem* 1, *and* $(X_1^{(m)}, X_2^{(m)}, \cdots)$ *are the normalized parts from a random partition of m with distribution given by* (17), *then*

$$\lim_{m\to\infty} \mathbb{E}\left(\sum_{i=1}^{\infty} g(X_i^{(m)})\right) = \int_0^1 \frac{g(x)}{x}\, \theta(1-x)^{\theta-1} dx,$$

and

$$\sum_{i=1}^{\infty} g(X_i^{(m)}) \Rightarrow \sum_{i=1}^{\infty} g(X_i) \quad \textit{as } m \to \infty.$$

Proof. Let $r_m(y) = \Gamma(my+\theta)/\Gamma(my+1)$ and define

$$\lambda(x) = \frac{\Gamma(x+1)}{\sqrt{2\pi}\, x^{x+1/2} e^{-x}}. \tag{19}$$

We may write

$$r_m(y) = \exp(-(\theta-1))\left(1 + \frac{\theta-1}{my}\right)^{my+\theta-1/2} (my)^{\theta-1} \frac{\lambda(my+\theta-1)}{\lambda(my)}.$$

It is known (cf. Abramowitz and Stegun (1965)) that

$$\lambda(x) = \exp\left(\frac{\alpha}{12x}\right) \tag{20}$$

for some $0 < \alpha < 1$. It follows from (20) that $\lambda(m(1-x)+\theta-1)$ and $\lambda(m(1-x))$ converge uniformly to 1 for all $x \in [0, t]$. Since $(1+\alpha/z)^z$ increases to e^α as $z \to \infty$, we see that $(1+(\theta-1)/(m(1-x)))^{m(1-x)+\theta-1/2}$ converges uniformly for $x \in [0, t]$ to $e^{\theta-1}$. Therefore

$$h_m(x) = \frac{\theta r_m(1-x)}{r_m(1)} \to \theta(1-x)^{\theta-1}$$

as $m \to \infty$, uniformly on $[0, t]$ for all $0 < t < 1$.

There are numerous functions g that satisfy the conditions of Corollary 1 that are not $o(x)$ as $x \to 0$. For example, if $g(x) = \sqrt{x}$ we may take $g^*(x) = x^{-1/2}$ and conclude that if $\boldsymbol{X} = (X_1, X_2, \cdots)$, then

$$\mathbb{E} f_g(X) = \frac{\sqrt{\pi}\Gamma(\theta+1)}{\Gamma(\theta+\frac{1}{2})}.$$

As $\theta \to \infty$, $\mathbb{E} f_g(X) \sim \sqrt{\pi\theta}$.

A more interesting example has $g(x) = -x \log(x)$, in which case

$$f_g(X) = -\sum_{i=1}^{\infty} X_i \log(X_i),$$

the entropy of the random measure $\boldsymbol{X} = (X_1, X_2, \cdots)$. It is straightforward to calculate

$$\mathbb{E} f_g(X) = \sum_{n=1}^{\infty} \frac{\theta}{n(n+\theta)} = \psi(1+\theta) + \gamma, \tag{21}$$

where ψ is the digamma function, and γ is Euler's constant (Abramowitz and Stegun (1965), p. 258). For large values of θ, it follows that

$$\mathbb{E} f_g(X) = \log(\theta) + \gamma + \frac{1}{2\theta} + O(\theta^{-2}), \qquad \theta \to \infty.$$

Example 2. We now return to our random mapping problem. Recall the definitions given at the beginning of the paper: $M_i^{(m)}$ is the size of the ith largest component of a random mapping on m integers, and $Y_i^{(m)} = m^{-1} M_i^{(m)}$. Also recall that $\boldsymbol{Y} = (Y_1, Y_2, \cdots)$ has the Poisson–Dirichlet distribution with parameter $\frac{1}{2}$.

Corollary 2. *If g satisfies condition* (iv) *of Theorem* 1, *then*

$$\lim_{m\to\infty} \mathbb{E}\left(\sum_{i=1}^{\infty} g(Y_i^{(m)})\right) = \int_0^1 \frac{g(x)}{x} \tfrac{1}{2}(1-x)^{-1/2} dx$$

and

$$\sum_{i=1}^{\infty} g(Y_i^{(m)}) \Rightarrow \sum_{i=1}^{\infty} g(Y_i) \qquad \textit{as } m \to \infty.$$

Proof. Because of (1) we need only verify Conditions (ii) and (iii) of Theorem 1. It is known (cf. Lemma 2.1 of Donnelly et al. (1991)) that

$$\mathbb{E}(A_j^{(m)}) = \binom{m}{j}\left(\frac{j}{m}\right)^j\left(\frac{m-j}{m}\right)^{m-j} \sum_{i=1}^{j} \binom{j-1}{i-1} \frac{(i-1)!}{j^i}.$$

Writing this in the form

$$\mathbb{E}(A_j^{(m)}) = \frac{m!}{j(m-j)!}\left(\frac{1}{m}\right)^j\left(\frac{m-j}{m}\right)^{m-j}\sum_{i=0}^{j-1}\frac{j^i}{i!}$$

and recalling (3), we may define

$$h_m(x) = \frac{m!}{\Gamma(m(1-x)+1)}\frac{(1-x)^{m(1-x)}}{m^{mx}}\sum_{k=0}^{\infty}\frac{(mx)^k}{k!}I\{k \leqq mx-1\}.$$

It is straightforward to show that $h_m(x) \to \frac{1}{2}(1-x)^{-1/2}$ pointwise. To establish uniform convergence away from the endpoints of (0, 1) a little extra care is needed. A calculation using (19) shows that

$$h_m(x) = \frac{\lambda(m)}{\lambda(m(1-x))}(1-x)^{-1/2}\sum_{k=0}^{\infty}\exp(-mx)\frac{(mx)^k}{k!}I\{k \leqq mx-1\}.$$

If $\{N(y), y \geqq 0\}$ is a Poisson process with rate 1, then

$$h_m(x) = \frac{\lambda(m)}{\lambda(m(1-x))}(1-x)^{-1/2}\mathbb{P}(N(mx) \leqq mx-1).$$

Choose $0 < s < t < 1$. Since $\lambda(m)/\lambda(m(1-x))$ converges uniformly to 1 on $[0, t]$, we need only show that as $m \to \infty$

$$\mathbb{P}(N(mx) \leqq mx-1) = \mathbb{P}\left(\frac{N(mx)-mx}{\sqrt{mx}} \leqq \frac{-1}{\sqrt{mx}}\right) \to \frac{1}{2} \tag{22}$$

uniformly on $[s, 1]$ to verify that $h_m(x)$ converges to $h(x) = \frac{1}{2}(1-x)^{-1/2}$ uniformly on $[s, t]$. The following argument, due to the referee, simplifies our original proof.

Let $\varepsilon > 0$ be given, and assume that $x \geqq s > 0$. Choose M sufficiently large that $1/\sqrt{Ms} < \varepsilon$ and for $y \geqq Ms$

$$\left|\mathbb{P}\left(\frac{N(y)-y}{\sqrt{y}} \leqq \pm\varepsilon\right) - \frac{1}{2}\right| < \varepsilon.$$

This last is possible by the central limit theorem, and the fact that the standard normal distribution function $\Phi(z)$ satisfies $|\Phi(\pm\varepsilon) - \frac{1}{2}| < \varepsilon/\sqrt{2\pi}$ if $\varepsilon > 0$. If $m \geqq M$ and $x \geqq s$, then $-\varepsilon < -1/\sqrt{mx} < \varepsilon$, so that the probability p on the right side of (22) satisfies $|p - \frac{1}{2}| < \varepsilon$ if $m \geqq M$ and $x \geqq s$. Hence $h_m(x) \to h(x)$ uniformly on $[s, t]$, as required. Since it is clear that $\sup_m \sup_{x\in(0,t]} h_m(x) < \infty$, the result follows by applying Theorem 1.

The particular case $g(x) = \sqrt{x}$ given in Donnelly et al. (1991) satisfies the conditions of Corollary 2, from which we verify that the limiting mean $\mathbb{E} f_g(Y)$ has value $\pi/2$. For the information function $g(x) = -x\log(x)$, we deduce from (21) that $\mathbb{E} f_g(Y) = 2 - \log(4) \approx 0.61371$.

Example 3. We conclude with an example where the relative sizes converge to the Poisson–Dirichlet distribution but Condition (iii) of Theorem 1 fails. We then show that the sum of the square roots of the relative sizes does not converge appropriately. The idea is to construct a random partition that has too many small components; individually they do not affect the (Poisson–Dirichlet) limit but the sum of the square roots of the relative sizes *is* affected by them. The example involves random permutations, the special case of random mappings in which the cycles of the permutation make up the components of the mapping.

Consider the set of all permutations of the set $\{1, 2, \cdots, m^2 - m\}$. Choose a permutation uniformly and at random from this set and call it Π_m. Define the permutation Π_m^* on the set $\{1, 2, \cdots, m^2\}$ by

$$\Pi_m^*(i) = \begin{cases} \Pi_m(i), & i \leqq m^2 - m \\ i, & m^2 - m < i \leqq m^2. \end{cases}$$

We consider the partitions of $m^2 - m$ and m^2 induced by these permutations. Let $Y_{i,m}$ (respectively, $Y_{i,m}^*$) be the size of the ith largest cycle of Π_m (respectively, Π_m^*), and let $X_{i,m} = Y_{i,m}/(m^2 - m)$, $X_{i,m}^* = Y_{i,m}^*/m^2$.

The distribution of the partition induced by Π_m is the Ewens sampling formula with $\theta = 1$ (cf. Joyce and Tavaré (1987)). Relation (18) shows that as $m \to \infty$

$$(23) \qquad (X_{1,m}, X_{2,m}, \cdots) \Rightarrow (X_1, X_2, \cdots)$$

where $(X_1, X_2, \cdots)$ has the Poisson–Dirichlet distribution with $\theta = 1$.

If F_m is the total number of cycles of Π_m then $F_m \to \infty$ in probability. Note that for fixed r

$$(24) \qquad \mathbb{P}(d((X_{1,m}, \cdots, X_{r,m}), (X_{1,m}^*, \cdots, X_{r,m}^*)) > 0) < \mathbb{P}(F_m \leqq r)$$

where d is the usual Euclidean metric on $\mathbb{R}^r$. It follows by (23) and (24) that $(X_{1,m}^*, X_{2,m}^*, \cdots)$ also converges weakly to the Poisson–Dirichlet distribution with parameter $\theta = 1$. But

$$\sum_{i=1}^{\infty} \mathbb{E}(\sqrt{X_{i,m}^*}) = \sqrt{\frac{m^2 - m}{m^2}} \sum_{i=1}^{\infty} \mathbb{E}(\sqrt{X_{i,m}}) + m(\sqrt{1/m^2}),$$

and so

$$\lim_{m \to \infty} \sum_{i=1}^{\infty} \mathbb{E}(\sqrt{X_{i,m}^*}) = \sum_{i=1}^{\infty} \mathbb{E}(\sqrt{X_i}) + 1.$$

Notice that the hypothesis (iii) is not satisfied for the partition induced by Π_m^*, since

$$h_{m^2}(1/m^2) = \mathbb{E}(A_1^{(m^2)}) \to \infty \qquad \text{as } m \to \infty.$$

Acknowledgements

We should like to thank Peter Donnelly, Warren Ewens and S. Padmadisastra for permission to quote their paper prior to publication. We also thank the referee for uncovering an error in an earlier draft of this paper.

References

ABRAMOWITZ, M. AND STEGUN, I. A. (1965) *Handbook of Mathematical Functions*. Dover, New York.

ALDOUS, D. J. (1985) Exchangeability and related topics. In *École d'été de Probabilités de Saint-Flour XIII*, 1983, ed. P. L. Hennequin, pp. 1–198. Lecture Notes in Mathematics **1117**, Springer-Verlag, Berlin.

DONNELLY, P. AND JOYCE, P. (1989) Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stoch. Proc. Appl.* **31**, 89–103.

DONNELLY, P., EWENS, W. J. AND PADMADISASTRA, S. (1991) Random functions: exact and asymptotic results. *Adv. Appl. Prob.* **23**, 437–455.

ETHIER, S. N. AND KURTZ, T. G. (1986) *Markov Processes: Characterization and Convergence*. Wiley, New York.

EWENS, W. J. (1972) The sampling theory of selectively neutral alleles. *Theoret Popn. Biol.* **3**, 87–112.

JOYCE, P. AND TAVARÉ, S. (1987) Cycles, permutations and the structure of the Yule process with immigration. *Stoch. Proc. Appl.* **25**, 309–314.

KINGMAN, J. F. C. (1977) The population structure associated with the Ewens sampling formula. *Theoret Popn. Biol.* **11**, 274–283.

WATTERSON, G. A. (1974) Models for logarithmic species abundance distributions. *Theoret. Popn. Biol.* **6**, 217–250.