LMD mokslo darbai, 49–54
1998, Vilnius

A functional limit theorem for random mappings

E. Manstavičius* (VU)

1. Introduction

Let $\mathbf{T}_N$ be the set of all mappings φ from the set $\{1, \dots, N\}$ into itself and $\nu_N(\dots)$ be the uniform probability measure on $\mathbf{T}_N$. We are interested in structural properties of a random φ which can be described in terms of its functional graph G_φ, e.g., a labelled directed graph on N vertices. We recall that an edge from i to j exists in the graph G_φ if and only if $\varphi(i) = j$. Suppose that G_φ has the component structure $\bar{k} = (k_1, \dots, k_N)$, where $k_j = k_j(\varphi)$ denotes the number of connected components of size j, $1k_1 + \dots + Nk_N = N$. Denote $w(\varphi) = k_1 + \cdots k_N$ the number of connected components in a mapping φ defined as that for the graph G_φ. Let further the limits are taken as $N \to \infty$.

In 1969 V. E. Stepanov [8] proved the central limit theorem for $w(\varphi)$. V. F. Kolchin [6] determined, for fixed m, the limiting distribution of the size of the m-th largest connected component. D. Aldous [1] improved this result by proving a global limit theorem for the component structure of a random mapping. He showed that the ordered sequence of sizes of components can be described by the Poisson–Dirichlet distribution with the parameter 1/2 on the set $\{(x_1, x_2, \dots)\colon x_1, x_2, \dots \geq 0, x_1 + x_2 + \dots = 1\}$. J.C.Hansen [5] considered the number $V_N(\varphi, t)$ of connected components in G_φ of size less than or equal to N^t, where $0 \leq t \leq 1$. To present her result, we set

$$W_N := W_N(\varphi, t) = (V_N(\varphi, t) - (t/2)/\log N)/\sqrt{(1/2)\log N}.$$

For a fixed $\varphi \in \mathbf{T}_N$, the function $W_N(\varphi, .)$ is an element of $\mathbf{D}[0, 1]$, the space of right-continuous functions with left limits on $[0, 1]$. Let $\mathcal{D}$ be the Borel σ-field of subsets of $\mathbf{D}[0, 1]$ with respect to the uniform topology, and $\nu_N \cdot W_N^{-1}$ be the distribution of the process W_N. Denote by W the Wiener measure.

THEOREM A [5]. *The measures $\nu_N \cdot W_N^{-1}$ weakly converge to W.*

We will generalize this theorem by establishing an invariance principle for *additive functions* (decomposable statistics) defined on the set $\mathbf{T}_N$. By definition such a function $h\colon \mathbf{T}_N \to \mathbf{R}$ has the decomposition

$$h(\varphi) = \sum_{j=1}^{N} h_j(k_j(\varphi)) \tag{1}$$

*Partially supported by Grant from Lithuanian Foundation of Studies and Science.

for each $\varphi \in \mathbf{T}_N$, where $h_j(k),\ j \geq 1,\ k \geq 1$, is some double sequence in $\mathbf{R}$ such that $h_j(0) = 0,\ j \geq 1$. If $h_j(k) = kh_j(1)$ for all $1 \leq j \leq N$ and $k \geq 0$, then h is called *a completely additive function* (linear statistics).

It follows from [6] that, for a fixed j, $k_j(\varphi)$ asymptotically behaves like the Poisson random variable (r.v.) ξ_j with parameter

$$\lambda_j := \frac{e^{-j}}{j}\sum_{s=0}^{j-1}\frac{j^s}{s!}$$

as $j \to \infty$. Since

$$|\lambda_j - 1/(2j)| \leq 8j^{-3/2}, \quad j \geq 1 \tag{2}$$

(see [5]), it is natural to use the following normalizing sequences

$$A(N) := \frac{1}{2}\sum_{j=1}^{N}\frac{a(j)}{j}, \qquad B^2(N) := \frac{1}{2}\sum_{j=1}^{N}\frac{a(j)^2}{j},$$

where $a(j) := h_j(1)$. Let

$$H_N := H_N(\varphi, t) = \frac{1}{B(N)}\Bigg(\sum_{j \leq y(t)} h_j(k_j(\varphi)) - A(y(t))\Bigg),$$

where

$$y(t) := y_N(t) = \max\{u\colon\ B^2(u) \leq tB^2(N)\} \qquad t \in [0, 1].$$

In the present remark we prove the following theorem.

THEOREM. *Let $B(N) \to \infty$. The measures $\nu_N \cdot H_N^{-1}$ weakly converge to W if and only if*

$$\Lambda_N(\varepsilon) := \frac{1}{B^2(N)}\sum_{\substack{j=1\\ |a(j)|\geq\varepsilon B(N)}}^{N}\frac{a(j)^2}{j} = o(1) \tag{3}$$

for each $\varepsilon > 0$.

This result is analoguous to the functional limit theorem for additive functions on permutations established in our paper [4] written jointly with Gutti J.Babu. In this investigation, for a probability measure on the symmetric group, we have used the Ewens sampling formula which, if the parameter equals 1/2, is close to the distribution of component vector of a random mapping from $\mathbf{T}_N$ (see [3] for the details). Thus some similarity with the paper [4] is unavoidable, and, by this reason, our proof is fairly sketchy.

2. Proof of Theorem

Sufficiency. As in [4], the problem can be reduced to that for completely additive functions. Let ξ_j, be the independent Poisson r.vs with parameters λ_j, $1 \le j \le N$, given on some probability space $\{\Omega, \mathcal{F}, P\}$. Set $a \wedge b = \min\{a, b\}$,

$$X_N(t) = \frac{1}{B(N)}\Big(\sum_{j \le y(t)} a(j)\xi_j - A(y(t))\Big),$$

$$X_N^r(t) = \frac{1}{B(N)}\Big(\sum_{j \le y(t)\wedge r} a(j)\xi_j - A(y(t) \wedge r)\Big),$$

and

$$H_N^r := H_N^r(\varphi, t) = \frac{1}{B(N)}\Big(\sum_{j \le y(t)\wedge r} a(j)k_j(\varphi) - A(y(t) \wedge r)\Big), \qquad 1 \le r \le N.$$

Let $||\cdot||$ denote the total variation distance on the set of probability measures on $\mathcal{D}$.

LEMMA 1. *We have*

$$||\nu_N \cdot (H_N^r)^{-1} - P \cdot (X_N^r)^{-1}|| = o(1)$$

for an arbitrary sequence $r = r(N) \to \infty$, $r = o(N)$. *Moreover, if*

$$B(N) - B(r) = o(B(N)) \tag{4}$$

for some sequence $r = r(N) \to \infty$, *then*

$$P(\varepsilon) := P\Big(\sup_{0 \le t \le 1} |X_N(t) - X_N^r(t)| \ge \varepsilon\Big) = o(1)$$

and

$$\nu_N(\varepsilon) := \nu_N\Big(\sup_{0 \le t \le 1} |H_N(\varphi, t) - H_N^r(\varphi, t)| \ge \varepsilon\Big) = o(1)$$

for each $\varepsilon > 0$.

Proof. The first assertion is a corollary of Theorem 10 in [2] or Theorem 1.3 in [7]. The estimate for the processes defined in terms of independent r.vs follows from Levy's inequality. Further we can use the inequality

$$\nu_N(\varepsilon) \ll_c (P(\varepsilon/3) + N^{-1})^c,$$

with arbitray $0 < c < 1/2$, following from Lemma A of the paper [4]. Lemma 1 is proved.

We now proceed with the following remark. Traditionally, in the partial sum processes the time parameter t is involved through the variances of the summands. So, in the definition of $X_N(t)$, we should have used

$$\bar{y}(t) := \max\left\{u: \sum_{j\le u}\lambda_j a(j)^2 \le u\sum_{j\le N}\lambda_j a(j)^2\right\}$$

instead of $y(t)$. By (2) this change corresponds to the shift of t by the factor $1+o(1)$ with the uniform in t error estimate. Since the processes $X_N(t)$ and $X_N(t(1+o(1)))$ can converge only simultaneously, we may use $1/2j$ instead of λ_j. Similarly, one can observe that the Lindeberg condition for the r. vs $a(j)\xi_j$ is equivalent to (3). It implies (4) and also gives weak convergence of X_N to the standard Brownian motion. Further an application of Lemma 1 completes the proof of sufficiency.

Necessity. We need a result on the mean value $M_N(f)$ of a completely multiplicative function $f\colon \mathbf{T}_N \to \mathbf{C}$. By definition, similarly to (1), such a function has the decomposition

$$f(\varphi) = \prod_{j=1}^{N} b(j)^{k_j(\varphi)}$$

for each $\varphi \in \mathbf{T}_N$, where $b(j)$, $j \ge 1$, is some sequence in $\mathbf{C}$.

LEMMA 2. *Let $f\colon \mathbf{T}_N \to \mathbf{C}$ be a completely multiplicative function defined by $b(j) = 1$ for all but $j \in J \subset (N/2, N]$. Then*

$$M_N(f) = 1 + \frac{N!e^N}{N^N}\sum_{j\in J}(b(j)-1)\lambda_j\frac{e^{-(N-j)}(N-j)^{N-j}}{(N-j)!}.$$

Moreover, if $|b(j)| \le 1$ and $J \subset \big((1-\delta)N, N\big]$ with sufficiently small $\delta > 0$, then

$$|M_N(f)| > c(\delta) > 0 \tag{5}$$

provided N is sufficiently large, $N > N(\delta)$.

Proof. Grouping the mappings of $\mathbf{T}_N$ into the classes with *a fortiori* prescribed component structure $\bar{k} = (k_1, \dots, k_N)$, $1k_1 + \cdots + Nk_N = N$, we obtain

$$M_N(f) = \frac{1}{N^N}\sum_{\varphi\in\mathbf{T}_N} f(\varphi) = \frac{N!e^N}{N^N}\sum_{\bar{k}}\prod_{j=1}^{N}\frac{(b(j)\lambda_j)^{k_j}}{k_j!}.$$

Note that, if $k_j \geq 1$ for some $j \in J$, then $k_j = 1$ and $k_l = 0$ for the remaining $l \neq j$ and $l \in J$. Hence

$$\begin{aligned} M_N(f) &= \frac{N!e^N}{N^N}\Bigg(\sum_{\substack{\bar{k} \\ k_l=0\,\forall l \in J}} \prod_{l=1}^{N} \frac{\lambda_l^{k_l}}{k_l!} + \sum_{j \in J} b(j) \sum_{\substack{\bar{k} \\ k_j=1}} \prod_{l=1}^{N} \frac{\lambda_l^{k_l}}{k_l!}\Bigg) \\ &= 1 + \frac{N!e^N}{N^N} \sum_{j \in J} \big(b(j) - 1\big) \sum_{\substack{\bar{k} \\ k_j=1}} \prod_{l=1}^{N} \frac{\lambda_l^{k_l}}{k_l!} \\ &= 1 + \sum_{j \in J} \big(b(j) - 1\big)\Bigg(1 - \frac{N!e^N}{N^N} \sum_{\substack{\bar{k} \\ k_j=0}} \prod_{l=1}^{N} \frac{\lambda_l^{k_l}}{k_l!}\Bigg) \\ &= 1 + \sum_{j \in J} \big(b(j) - 1\big)\Bigg(1 - \frac{N!e^N}{N^N} d_j(N)\Bigg), \end{aligned} \tag{6}$$

where

$$d_j(N) = \sum_{\substack{\bar{k} \\ k_j=0}} \prod_{l=1}^{N} \frac{\lambda_l^{k_l}}{k_l!}.$$

From the identities

$$\sum_{N \geq 0} d_j(N) z^N = \prod_{l \geq 1,\, l \neq j} e^{\lambda_l z^l} = e^{-\lambda_j z^j}\Bigg(1 + \sum_{N \geq 1} \frac{N^N e^{-N}}{N!} z^N\Bigg)$$

we have

$$d_j(N) = \sum_{\substack{k,n \geq 0 \\ jk+n=N}} (-1)^k \frac{\lambda_j^k}{k!} \frac{n^n e^{-n}}{n!} = \frac{N^N e^{-N}}{N!} - \lambda_j \frac{(N-j)^{N-j} e^{-(N-j)}}{(N-j)!}$$

provided $j \in J$. Inserting this into (6), we obtain the first assertion of Lemma 2.

Using (2) and the inequalities

$$\sqrt{2\pi} n^{n+1/2} e^{-n} < n! < 2\sqrt{2\pi} n^{n+1/2} e^{-n}, \quad n \geq 1$$

from the expression of $M_N(f)$ we get its lower estimate. Lemma 2 is proved.

We now return to the processes. If $\nu_N \cdot H_N^{-1} \Rightarrow W$, then for each $0 \leq t < 1$, the distribution of the difference $H_N(\varphi, 1) - H_N(\varphi, t)$ converges weakly to the normal law with zero mean and variance $1 - t$. Let $\phi_N(u, t)$, $u \in \mathbf{R}$, denote the characteristic function of $H_N(\varphi, 1) - H_N(\varphi, t)$. Define $b(j) = \exp\{iua(j)/B(N)\}$ if $y(t) < j \leq N$ and $b(j) = 1$ elsewhere. For the completely multiplicative function f defined via $f_j(1) = b(j)$, we have

$$|\phi_N(u, t)| = |M_N(f)| \leq e^{-u^2/2(1-t)} + o(1) \tag{7}$$

for $u \in \mathbf{R}$ and $0 < t < 1$. For t close to 1, we will apply Lemma 2. Let δ be sufficiently small and $\tau_N = \sup\{t\colon\ y(t) \le (1-\delta)N\}$. Observe that $\tau_N \to 1$. Indeed, if $\tau_N \to t_0 < t_1 < 1$ for some subsequence $N := N' \to \infty$, then $y(t_1) \ge (1-\delta)N$ for N sufficiently large. Estimate (5) now yields $|\phi_N(u, t_1)| > c(\delta) > 0$ uniformly in $u \in \mathbf{R}$, contradicting to (7). Thus from the definitions of $y(t)$ and the sequence τ_N, it follows that

$$1 + o(1) \le \tau_N \le \frac{B^2\big(y(\tau_N)+1\big)}{B^2(N)} \le \frac{B^2\big((1-\delta))N+1\big)}{B^2(N)} \le 1.$$

Hence $B(uN) \sim B(N)$ for each $u \in [(1-(\delta/2))N, N]$ and some $\delta > 0$. Substituting $(1-(\delta/2))N$ for N repeatedly, we deduce the existence of $r = r(N) \to \infty$ such that $r = o(N)$ and $B(r) \sim B(N)$. Now repeating the arguments of the proof of the sufficiency part we obtain that $\nu_N\big(H_N^r(\sigma, 1) < x\big)$ converge to the standard normal law. This together with Lemma 1 leads to convergence of $P\big(X_N(1) < x\big)$ to the same law. Since $\xi_j/B(N)$, $j \le N$, form an infinitesimal array of random variables, and since $B(N) \to \infty$, the necessity of (3) follows from the Lindeberg–Feller theorem. This completes the proof of Theorem 1.

REFERENCES

[1] A. Aldous, *Exchangeability and Related Topics*, Lecture Notes in Math., 1117, 1985, Springer, Berlin.

[2] R. Arratia and S. Tavaré, Limit theorems for combinatorial structures via discrete process approximations, *Random Structures and Algorithms*, **3**(3) (1992), 321–345.

[3] R. Arratia, D. Stark, and S. Tavaré, Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies, *Ann. of Probab.*, **23**(3) (1995), 1347–1388.

[4] G. J. Babu and E. Manstavičius, Brownian motion and random permutations, *Ann. of Applied Probab.*, (submitted) 1998, 15 p.).

[5] J. C. Hansen, A functional central limit theorem for random mappings, *Ann. of Probab.*, **17**(1) (1989), 317–332.

[6] V. F. Kolchin, A problem of the allocation of particles in cells and random mappings, *Theory of Probab. Appl.*, **21** (1976), 48–63.

[7] D. Stark, Explicit limits of total variation distance in approximations of random logarithmic assemblies by related Poisson process, *Combinatorics, Probab., and Computing*, **6** (1997), 87–105.

[8] V. E. Stepanov, Limit distributions for certain characteristics of random mappings, *Theory of Probab. Appl.*, **14** (1969), 612–626.

Atsitiktinių atvaizdžių funkcinė ribinė teorema

E. Manstavičius (*VU*)

Naudojant adityviąsias funkcijas, apibrėžtas baigtinių aibių atvaizdžių aibėje, modeliuojamas Brown'o judesys. Rastos būtinosios ir pakankamosios sąlygos, kada atitinkama tikimybinių matų seka silpnai konverguoja į Wiener'io matą.