

# Probabilistic number theory and random permutations: Functional limit theory

Gutti Jogesh Babu

*Dedicated to Professor K. Ramachandra on his 70th birthday*

## Abstract

The ideas from Probabilistic Number Theory are useful in the study of measures on partitions of integers. Connection between the Ewens sampling formula in population genetics and the partitions of an integer generated by random permutations will be discussed. Functional limit theory for partial sum processes induced by Ewens sampling formula is reviewed. The results on limit processes with dependent increments are illustrated.

## 1 Introduction

In the last few decades, mathematical population geneticists have been exploring the mechanisms that maintain diversity in a population. In 1972, Ewens established a formula to describe the probability distribution of a sample of genes from a population that has evolved over many generations, by a family of measures on the set of permutations of the first  $n$  integers (equivalently on the set of partitions of  $n$ ). The Ewens formula can be used to test if the popular assumptions are consistent with data, and to estimate the parameters. The statistics that are useful in this connection will generally be expressed as functions of the sums of transforms of the ‘allelic partition’. Such statistics can be viewed as functions of a process on the permutation group of integers.

In a series of papers [3]-[7], Babu and Manstavičius, have developed necessary and sufficient conditions for the weak convergence of a partial sum process based on these measures to a process with independent increments. Under very general conditions, it has been shown that a partial sum process converges weakly in a function space if and only if a related process defined through sums of independent random variables converge. In this paper, the case where the limiting processes need not be processes with independent increments is considered. Thus, under Ewens sampling formula, the limiting process of the partial sums of dependent variables differs from that of the associated process defined through the partial sums of independent random variables. The basic ideas for proofs come from probabilistic number theory and analytic number theory. Integration over Hankel contour (see Corollary 2.1 in Section II.5.2 in [17]) plays an important role.

---

1991 *Mathematics Subject Classification*. 60F17, 60C05, 11K65.

*Key words and phrases*. Cycle, Ewens sampling formula, functional limit theorem, random partitions, slowly varying function, weak convergence.

Research of Babu was supported in part by NSF grants DMS-0101360, AST-0434234 and AST-0326524.

## 2 Probabilistic Number Theory

We shall start with a brief comparative analysis of the developments in probabilistic number theory and the theory of random permutations. The uniform probability measure

$$\nu_n(A) = \frac{1}{n} \#\{1 \leq m \leq n : m \in A\}$$

on integers, satisfies for  $k, l \geq 0$ ,

$$\nu_n(\alpha_p(m) = k) = \frac{1}{n} \left( \left[ \frac{n}{p^k} \right] - \left[ \frac{n}{p^{k+1}} \right] \right) \approx \frac{1}{p^k} \left( 1 - \frac{1}{p} \right), \quad k \geq 0,$$

and

$$\begin{aligned} \nu_n(\alpha_p(m) = k, \alpha_q(m) = l) &\approx \frac{1}{p^k} \left( 1 - \frac{1}{p} \right) \frac{1}{q^l} \left( 1 - \frac{1}{q} \right), \\ &\approx \nu_n(\alpha_p(m) = k) \nu_n(\alpha_q(m) = l), \end{aligned}$$

where  $m = \prod_p p^{\alpha_p(m)}$  is the unique representation of integer  $m$  as the product of prime powers. It follows that  $\alpha_p$  has asymptotically geometric distribution. In addition,  $\alpha_p$  and  $\alpha_q$  are asymptotically independent, where  $p$  and  $q$  are distinct primes.

The Fundamental Theorem of probabilistic number theory [15] states that

$$\nu_n(\alpha_p(m) = k_p, p \leq r) = P(\xi_p = k_p, p \leq r) + o(1),$$

where  $r \leq n^\varepsilon$  for each  $\varepsilon > 0$ , and  $\{\xi_p\}$  are independent geometric random variables

$$P(\xi_p = k) = \frac{1}{p^k} \left( 1 - \frac{1}{p} \right), \quad k \geq 0.$$

If  $h$  is an additive arithmetic function,  $h(mn) = h(m) + h(n)$ ,  $(m, n) = 1$ , then  $h$  can be represented as  $h(m) = \sum_p h(p^{\alpha_p(m)})$ . As  $\alpha_p(m) = 1$  for a prime  $\sqrt{n} < p \leq n$  implies  $\alpha_q(m) = 0$  for all primes  $q \neq p$ ,  $\sqrt{n} < q \leq n$ , it follows that they are not independent even asymptotically. To establish the limiting distribution of  $h$  under  $\nu_n$ , one uses the decomposition  $h(m) = h_r(m) + h^r(m)$ , where

$$h_r(m) = \sum_{p \leq r} h(p^{\alpha_p(m)}), \quad h^r(m) = \sum_{p > r} h(p^{\alpha_p(m)}).$$

Then the fundamental theorem of probabilistic number theory is used to approximate  $\nu_n(h_r(m) \leq x)$  by  $P(\sum_{p \leq r} f_p(\xi_p) \leq x)$  and showing that the contribution of  $h^r$  is negligible, where  $f_p(k) = h(p^k)$ .

Similar ideas are used in obtaining functional limit theorems by Babu [2] for the partial sum process

$$X_n(t) = (1/\sqrt{B(n, n)}) \sum' h(q), \quad B(n, k) = \sum_{p \leq k} \frac{1}{p} h^2(p), \quad t \in [0, 1],$$

where the sum  $\sum'$  is taken over all primes  $q \leq n$  satisfying  $B(n, q) \leq tB(n, n)$ .

### 3 Statistical Group Theory

Similar approach can be used in the study of statistical group theory and in particular random permutations. Let  $\mathbf{S}_n$  denote the group of permutations on  $\{1, \dots, n\}$ . Each permutation  $\sigma$  can be decomposed as  $\sigma = \kappa_1 \dots \kappa_\omega$  where  $\omega(\sigma)$  denotes the number of cycles of  $\sigma$ , and  $\kappa_i$  denote the independent cycles. For example, the permutation  $\tau$  that maps  $\{1, 2, 3, 4, 5, 6, 7, 8\}$  to  $\{5, 3, 6, 1, 8, 2, 7, 4\}$  has three cycles  $(1\ 5\ 8\ 4)$ ,  $(2\ 3\ 6)$ ,  $(7)$  and can be represented as  $\tau = (1\ 5\ 8\ 4)(2\ 3\ 6)(7)$ . Thus  $Ord(\tau) = 12$ , where the order  $Ord(\sigma)$  of permutation  $\sigma$  is defined to be the smallest  $k$  such that  $\sigma^k = \text{identity permutation}$ . If  $k_j(\sigma)$  denotes the number of cycles of length  $j$  of  $\sigma$ , then  $\omega(\sigma) = k_1(\sigma) + \dots + k_n(\sigma)$  and  $Ord(\sigma) = \text{l.c.m.}\{j \leq n : k_j(\sigma) > 0\}$ .

Goncharov, Erdos-Turan, and others contributed to the theory. In 1942, V. L. Goncharov [12] has shown that

$$\frac{1}{n!} \#\{\sigma \in \mathbf{S}_n : \omega(\sigma) - \log n < x \sqrt{\log n}\} \rightarrow \Phi(x),$$

where  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}u^2} du$ . In 1965, Erdős and Turán [10] have established that

$$\frac{1}{n!} \#\left\{\sigma \in \mathbf{S}_n : \log Ord(\sigma) - \frac{1}{2} \log^2 n \leq \frac{x}{\sqrt{3}} \log^{3/2} n\right\} \rightarrow \Phi(x).$$

However, in these and other early works, there is no trace of the use of ideas from probabilistic number theory, though the functions are similar.

The equivalent relation,  $\sigma \sim \tau$  if  $k_j(\sigma) = k_j(\tau)$  for all  $j$ , partitions  $\mathbf{S}_n$  into equivalence classes, known as conjugate classes. Hence we can identify  $\sigma$  with the vector  $\bar{k} = (k_1(\sigma), \dots, k_n(\sigma))$ , where  $1k_1(\sigma) + \dots + nk_n(\sigma) = n$ . This leads to random partitions of integer  $n$ .

### 4 Ewens Sampling Formula

The family of probability measures on the symmetric group  $\mathbf{S}_n$  of permutations on  $\{1, \dots, n\}$ , induced by the Ewens sampling formula (see [11]) are given by

$$\nu_{n,\theta}(\bar{k}) := \frac{n!}{\theta_{(n)}} \prod_{j=1}^n \binom{\theta}{j}^{k_j} \frac{1}{k_j!}, \quad \bar{k} := (k_1, \dots, k_n) \in \mathbf{Z}^{+n},$$

for the partition  $n = 1k_1 + \dots + nk_n$ ,  $n \in \mathbf{N}$ , and 0 otherwise, where  $\theta > 0$ , and  $\theta_{(n)} = \theta(\theta+1) \dots (\theta+n-1)$ . The quantity  $\nu_{n,\theta}(\bar{k})$  can also be viewed as the probability measure on the class of conjugate elements  $\sigma \in \mathbf{S}_n$ , all having  $k_j(\sigma) = k_j$  cycles of length  $j$ ,  $1 \leq j \leq n$ . The probability measure  $\nu_{n,\theta}$  is induced by the measure  $\nu'_{n,\theta}$  on  $\mathbf{S}_n$ , that assigns a mass proportional to  $\theta^{w(\sigma)}$  for  $\sigma \in \mathbf{S}_n$ , where  $w(\sigma) = k_1(\sigma) + \dots + k_n(\sigma)$  denotes the total number of cycles of  $\sigma$ . This can be seen from

$$\nu'_\theta(\sigma) = \theta^{w(\sigma)} \left( \sum_{\tau \in \mathbf{S}_n} \theta^{w(\tau)} \right)^{-1} = \frac{\theta^{w(\sigma)}}{\theta_{(n)}}.$$

Thus, we use this probability measure on  $\mathbf{S}_n$  and leave the same notation  $\nu_{n,\theta}$  for it.

The case  $\theta = 1$  corresponds to the measure induced by the uniform probability  $(1/n!) \# \{\sigma \in \mathbf{S}_n : \dots\}$  on  $\mathbf{S}_n$ . If  $k_j(\sigma) = 1$  for some  $\frac{n}{2} < j \leq n$ , then  $k_i(\sigma) = 0$  for all  $\frac{n}{2} < i \leq n, i \neq j$ .

As mentioned in the introduction §1, the Ewens formula describes the probability law of a sample of  $n$  genes from a population that has evolved over many generations. The domain of  $\nu_{n,\theta}$ ,  $\{(k_1, \dots, k_n) : 1k_1 + \dots + nk_n = n\}$  is same as that of the Allelic Partition  $\bar{k} = (k_1, \dots, k_n)$ , where  $k_j$  denotes the number of alleles appearing  $j$  times. The distribution of *Allelic Partition* has all the information available in the sample of  $n$  genes. Hence, the Ewens formula can be used to test if the popular assumptions are consistent with data, and to estimate the parameters.

This motivates consideration of additive functions on  $\mathbf{S}_n$ . A function  $h : \mathbf{S}_n \rightarrow \mathbf{R}$  is called additive if  $h_j(0) = 0$  and  $h(\sigma) = \sum_{j=1}^n h_j(k_j(\sigma))$ ,  $\sigma \in \bar{k}$ . Kolchin and Chistyakov [14] showed that  $\nu_{n,1}(\sum_{j \leq r} a_{jn} k_j(\sigma) - A_r < x)$  converges for some sequence  $\{A_r\}$  if and only if  $P(\sum_{j \leq r} a_{jn} Y_j - A_r < x)$  converges, where  $Y_j$  are independent Poisson random variable with mean  $1/j$ ,  $r = r(n) \rightarrow \infty$ , and  $r \log r = o(n)$ .

To facilitate the study of the limiting distributions of additive functions on  $\mathbf{S}_n$ , Arratia and Tavaré [1] developed a result similar to fundamental result of Kubilius in probabilistic number theory, which states that

$$\nu_{n,1}(k_j(\sigma) = k_j, j \leq r) = P(Y_j = k_j, j \leq r) + O_\delta \left( \exp \left\{ -(1-\delta) \frac{n}{r} \log \frac{n}{r} \right\} \right).$$

The measure  $\nu_{n,\theta}$  can be represented using independent Poisson random variables  $\xi_j$  with  $\mathbf{E}(\xi_j) = \frac{\theta}{j}$ , as

$$\nu_{n,\theta}(\bar{k}) = P(\xi_1 = k_1, \dots, \xi_n = k_n \mid \zeta = n), \quad \zeta = 1\xi_1 + \dots + n\xi_n.$$

## 5 Functional Limit Theorems: Processes with independents

To state the functional limit theorems, let  $h(\sigma) = \sum_{i=1}^n h_j(k_j(\sigma))$  denote an additive function on  $\mathbf{S}_n$ . Let

$$A(u) = \theta \sum_{j \leq u} \frac{1}{j} h_j(1), \quad B^2(u) = \theta \sum_{j \leq u} \frac{1}{j} h_j(1)^2,$$

and

$$y_n(t) = \max\{u : B^2(u) \leq tB^2(n)\}.$$

The first functional limit theorem for  $\theta = 1$  and  $\omega(\sigma)$  was obtained by DeLaurentis and Pittel [8]. This was extended to general  $\theta$  by Hansen [13] and Donnelly *et al.* [9]. The following theorem for general additive functions is from [3].

**Theorem 1 (Babu and Manstavičius [3]).** *Suppose  $B(n) \rightarrow \infty$ , and*

$$H_n(\sigma, t) = \frac{1}{B(n)} \left( \sum_{j \leq y_n(t)} h_j(k_j(\sigma)) - A(y_n(t)) \right).$$

*Then  $\nu_{n,\theta} \cdot H_n^{-1} \Rightarrow W$  if and only if for each  $\varepsilon > 0$ ,*

$$\Lambda_n(\varepsilon) = \frac{1}{B^2(n)} \sum_{|h_j(1)| \geq \varepsilon B(n)} \frac{1}{j} h_j(1)^2 \rightarrow 0,$$

*where  $W$  denotes the Brownian Motion on  $[0, 1]$ .*

This result leads, via invariance principle of the probability theory, to the limiting distributions of functions of partial sums of  $h_j$ . These results throw new light on the partitions of integers. For example the functional limit theorem implies,

$$\nu_{n,\theta}(h(\sigma) - A(n) \leq xB(n)) \rightarrow \Phi(x),$$

$$\begin{aligned} \nu_{n,\theta} \left( \sup_{k \leq n} \left( \sum_{j \leq k} h_j(k_j(\sigma)) - A(k) \right) \leq xB(n) \right) &\rightarrow P \left( \sup_{0 \leq t \leq 1} W(t) \leq x \right) \\ &= 2\Phi(x), \quad x \geq 0 \\ \nu_{n,\theta} \left( \sup_{k \leq n} \left| \sum_{j \leq k} h_j(k_j(\sigma)) - A(k) \right| \leq xB(n) \right) &\rightarrow P \left( \sup_{0 \leq t \leq 1} |W(t)| \leq x \right) \\ &= \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-\pi^2(2k+1)^2/8x^2}. \end{aligned}$$

By using a slowly varying function to scale  $h$ , Babu and Manstavičius obtained convergence to a stable processes and to general processes with independent increments. Let  $\beta(n) \rightarrow \infty$  and  $x^* = \min(|x|, 1) \text{ sign}(x)$ . For an additive function  $h$ , let

$$c(u, n) = \sum_{j \leq u} \frac{1}{j} \left( \frac{h_j(1)}{\beta(n)} \right)^{*2}, \quad A(u, n, \beta) = \theta \sum_{j \leq u} \frac{1}{j} \left( \frac{h_j(1)}{\beta(n)} \right)^*,$$

$$s_n(t) = \max\{l \leq n : c(l, n) \leq t c(n, n)\},$$

$$R_{n,h}(t) = \frac{1}{\beta(n)} \sum_{j \leq s_n(t)} h_j(k_j(\sigma)) - A(s_n(t), n, \beta)$$

and

$$X_{n,h}(t) = \frac{1}{\beta(n)} \sum_{j \leq s_n(t)} h_j(1) \xi_j - A(s_n(t), n, \beta).$$

The following result is from Babu and Manstavičius [7].

**Theorem 2.** *In order that  $R_{n,h} \Rightarrow X$ , where  $X$  is a process with independent increments and the distribution of  $X(1)$  is non-degenerate, it is necessary and sufficient that  $X_{n,h} \Rightarrow X$  and  $\beta(n)$  is slowly varying.*

If  $X_{n,h} \Rightarrow X$ , then the limiting process  $X$  is necessarily a process with independent increments and it 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. The result holds in spite of the strong dependent structure on  $\{k_j(\sigma) : \frac{1}{2}n \leq j \leq n\}$ .

**Counter Example.** In the one-dimensional case,  $X_{n,h}(1)$  and  $R_{n,h}(1)$  need not have the same limit. To see this, let  $\theta = 1$ ,  $0 < \alpha < 2$  and let  $F$  be the stable law with characteristic function  $\phi_\alpha(s) = e^{-|s|^\alpha}$ . Let  $\gamma_j$  denote the fractional part of  $j\sqrt{2}$ ,

$$a(j) = \begin{cases} j^{1/\alpha} F^{-1}(\gamma_j) & \text{if } |F^{-1}(\gamma_j)| \leq j^{1/\alpha} \\ 0 & \text{otherwise,} \end{cases}$$

$h(\sigma) = \sum_{j=1}^n k_j(\sigma)a(j)$ , and  $\beta(n) = n^{1/\alpha}$ . Then  $(h(\cdot)/\beta(n)) - A(n, n, \beta) \Rightarrow F$ . But  $X_{n,h}(1) \Rightarrow G$ , where the characteristic function of  $G$  is

$$\phi(s) = \exp \left\{ \int_0^1 \frac{1}{y} (e^{-y|s|^\alpha} - 1) dy \right\}.$$

## 6 Limit processes with dependent increments

The example above illustrates that if  $\beta$  is not slowly varying function, the limit process may have dependent increments. To study such limits, a generalization of the Main Lemma in [16] that involves integration over Hankel contour of the type given in Figure 1 is needed.

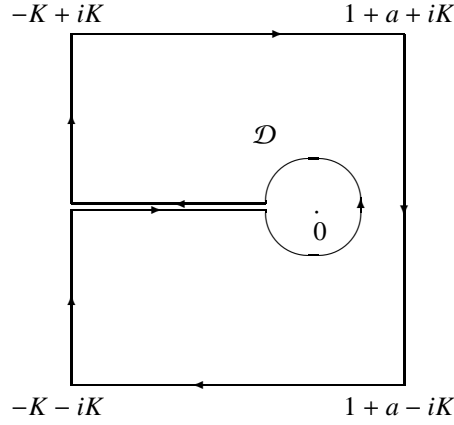


Figure 1: Contour with  $a \geq 0$  and  $K > 0$

However, to facilitate the discussion we consider an example with  $\beta(n) = n^\rho$ ,  $\rho > 0$ . Let the additive function  $h$  on  $\mathbf{S}_n$  be given by  $h(\sigma) = \sum_{j=1}^n k_j(\sigma)j^\rho$ . Let the processes  $H_n$  based on  $h$  be given by,

$$H_n(t) := H_n(t, \sigma) = (1/\beta(n)) \sum_{j \leq nt} k_j(\sigma)j^\rho = \sum_{j \leq nt} k_j(\sigma)(j/n)^\rho, \quad 0 \leq t \leq 1.$$

Note that  $H_n(1, \sigma) = 1$  for all  $\sigma \in \mathbf{S}_n$  if  $\rho = 1$ . We now present preliminary notation and results needed in illustrating the limiting process. We restrict to the case  $\theta = 1$ . First, we shall consider the mean values of multiplicative functions  $g : \mathbf{S}_n \rightarrow \mathbf{C}$  defined via  $g(\sigma) = \prod_{j=1}^n f(j)^{k_j(\sigma)}$ , where  $f(j)$ ,  $j \geq 1$  are complex numbers that may depend on  $n$  or other parameters. Its mean value with respect to the measure  $\nu_{n,1}$  equals

$$M_n(g) := \frac{1}{n!} \sum_{\sigma \in \mathbf{S}_n} g(\sigma) = \sum_{\substack{k_1, \dots, k_n \geq 0 \\ 1k_1 + \dots + nk_n = n}} \prod_{j=1}^n \left( \frac{1}{j} f(j) \right)^{k_j} \frac{1}{k_j!}.$$

The behavior of  $M_n(g)$  is examined in the next Lemma (see [3]) for large cycles.

**Lemma 1.** Let  $g : \mathbf{S}_n \rightarrow \mathbf{C}$  be a multiplicative function defined via  $f$  such that  $f(j) = 1$  for all but  $j \in J \subset (n/2, n]$ . Then

$$M_n(g) = 1 + \sum_{j \in J} \frac{1}{j} (f(j) - 1).$$

**Proof:** Observe that, if  $k_j \geq 1$  for some  $j \in J$ , then  $1k_1 + \cdots + nk_n = n$  implies  $k_j = 1$  and  $k_l = 0$  for the remaining  $l \neq j$  and  $l \in J$ . Let  $\Sigma_{(0)}$  denotes the sum over all  $(k_1, \dots, k_n)$  satisfying  $1k_1 + \cdots + nk_n = n$  and  $k_l = 0$  for all  $l \in J$ , and  $\Sigma_{(j)}$  denotes the sum over all  $(k_1, \dots, k_n)$  satisfying  $1k_1 + \cdots + nk_n = n$  and  $k_j = 1$ . Hence

$$\begin{aligned} M_n(g) &= \sum_{(0)} \prod_{l=1}^n \left(\frac{1}{l}\right)^{k_l} \frac{1}{k_l!} + \sum_{j \in J} f(j) \sum_{(j)} \prod_{l=1}^n \left(\frac{1}{l}\right)^{k_l} \frac{1}{k_l!} \\ &= 1 + \sum_{j \in J} (f(j) - 1) \sum_{(j)} \prod_{l=1}^n \left(\frac{1}{l}\right)^{k_l} \frac{1}{k_l!}. \end{aligned}$$

The Lemma follows now as the last sum is  $(1/j)$ .

The characteristic function  $\phi_{n,s,t}$  of  $H_n(t) - H_n(s)$  for  $\frac{1}{2} < s < t \leq 1$  is given by

$$\phi_{n,s,t}(\eta) = M_n \left( e^{i\eta(H_n(t) - H_n(s))} \right), \quad \eta \in \mathbf{R}.$$

We apply Lemma 1 with  $f(j) = \exp(i\eta(j/n)^\rho)$ ,  $\eta \in \mathbf{R}$ , to get

$$\begin{aligned} \phi_{n,s,t}(\eta) &= 1 + \sum_{s < (j/n) \leq t} \frac{1}{j} (e^{i\eta(j/n)^\rho} - 1) \\ &\rightarrow 1 + \int_s^t \frac{1}{v} (e^{i\eta v^\rho} - 1) dv \\ &= 1 + \frac{1}{\rho} \int_{s^\rho}^{t^\rho} \frac{1}{u} (e^{i\eta u} - 1) du =: \phi_{s,t}(\eta). \end{aligned}$$

If the limiting process of  $H_n$  has independent increments, then for all  $\frac{1}{2} < s < t < 1$ , and  $\eta \in \mathbf{R}$ ,

$$\phi_{s,1}(\eta) = \phi_{s,t}(\eta) \phi_{t,1}(\eta). \quad (1)$$

Hence for (1) to hold we must have,

$$\left( \int_{s^\rho}^{t^\rho} \frac{1}{u} (e^{i\eta u} - 1) du \right) \left( \int_{t^\rho}^1 \frac{1}{u} (e^{i\eta u} - 1) du \right) = 0$$

for all  $\frac{1}{2} < s < t < 1$  and  $\eta \in \mathbf{R}$ , which is impossible. This shows that the limiting process of  $H_n$  is not a process with independent increments.

The weak convergence of processes for general  $h$  and non-slowly varying  $\beta$  will be addressed elsewhere.  $\square$

## References

- [1] R. Arratia and S. Tavaré, The cycle structure of random permutations, *Annals of Probability*, **20** (1992) 1567–1591.
- [2] G.J. Babu, A note on the invariance principle for additive functions, *Sankhyā, A* **35** (1973), 3, 307–310.
- [3] G.J. Babu and E. Manstavičius, Brownian motion and random permutations, *Sankhyā, A* **61** (1999), 3, 312–327.
- [4] G.J. Babu and E. Manstavičius, Random permutations and the Ewens sampling formula in genetics, In: *Probab. Theory and Math. Stat.*, B.Grigelionis *et al.* (Eds), VSP/TEV, Vilnius/Utrecht, 1999, 33–42.
- [5] G.J. Babu and E. Manstavičius, Limit theorems for random permutations, In: *Paul Erdős and His Mathematics*, János Bolyai Mathematical Society, 1999, 19–22.
- [6] G.J. Babu and E. Manstavičius, Infinitely divisible limit processes for the Ewens sampling formula, *Lithuanian Math. J.* **42** (2002), 3, 232–242.
- [7] G.J. Babu and E. Manstavičius, Limit Processes with independent increments for the Ewens sampling formula, *Ann. Inst. Stat. Math.* **54** (2002), 3, 607–620.
- [8] J.M. DeLaurentis and B.G. Pittel, Random permutations and the Brownian motion, *Pacific J. Math.* **119** (1985), 287–301.
- [9] P. Donnelly, T.G. Kurtz and S. Tavaré, On the functional central limit theorem for the Ewens Sampling Formula, *Ann. Appl. Probab.* **1** (1991), 539–545.
- [10] P. Erdős and P. Turán. On some problems of a statistical group theory I. *Z. Wahrsch. Verw. Gebiete* **4** (1965) 175–186.
- [11] W.J. Ewens, The sampling theory of selectively neutral alleles, *Theor. Pop. Biol.* **3** (1972), 87–112.
- [12] V.L. Goncharov, On the distribution of cycles in permutations. *Dokl. Acad. Nauk SSSR* **35** (1942) 299–301 (Russian).
- [13] J.C. Hansen, A functional central limit theorem for the Ewens Sampling Formula, *J. Appl. Probab.* **27** (1990), 28–43.
- [14] V.F. Kolchin and V.P. Chistyakov, On the cycle structure of random permutations. *Matem. Zametki*, **18** (1975) 929–938 (Russian).
- [15] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*. Translations of Mathematical Monographs, **11**, AMS, Providence, Rhode Island, 1964.
- [16] E. Manstavičius, Additive and multiplicative functions on random permutations, *Lith. Math. J.* **36** (1996), 400–408.

- [17] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge studies in advanced mathematics **46**, Cambridge University Press, Cambridge, 1995.

**Gutti Jogesh Babu**

Department of Statistics,  
# 319 Thomas Building,  
The Pennsylvania State University,  
University Park, PA 16802, USA

email: [babu@stat.psu.edu](mailto:babu@stat.psu.edu)

<http://www.stat.psu.edu/babu>

<http://astrostatistics.psu.edu>