Professor Ross Pinsky

Selected Topics in Probability 106935

Winter Semester 2015-2016

The Theory of Random Permutations and some Connections to Probabilistic Number Theory

Let S_n denote the set of permutations of the set $\{1, 2, \ldots, n\}$. In this course we will study the asymptotic behavior of S_n, as $n \to \infty$, under a family of measures known as the Ewens sampling formula. Let $K^{(n)}(\sigma)$ denote the number of cycles in the permutation $\sigma \in S_n$. For each $\theta > 0$, the Ewens sampling formula with parameter θ is the probability measure $\mathbb{P}_\theta$ on S_n defined by $\mathbb{P}_\theta(\sigma) = \frac{\theta^{K^{(n)}(\sigma)}}{\theta^{(n)}}$, where $\theta^{(n)} = \sum_{\sigma \in S_n} \theta^{K^{(n)}(\sigma)}$ is the normalizing constant. One can show that $\theta^{(n)} = \theta(\theta+1)\cdots(\theta+n-1)$. When $\theta = 1$, one obtains the uniform distribution on S_n. When $\theta > 1$, there is a bias towards permutations with many cycles, and when $\theta < 1$ there is a bias towards permutation with few cycles.

Here are the main results we will prove:

I. Results Concerning Random Permutations

1. *Weak law of large numbers and central limit theorem for $K^{(n)}$, the number of cycles:* $\frac{K^{(n)}}{\theta \log n} \stackrel{\mathrm{w}}{\to} 1$, $\quad \frac{K^{(n)} - \theta \log n}{\sqrt{\theta \log n}} \stackrel{\mathrm{dist}}{\to} \mathrm{N}(0,1)$.

2. *Limiting distribution of the small cycles:*
For $\sigma \in S_n$, let $C_j^{(n)}(\sigma)$ denote the number of cycles of length j in σ. Then $(C_1^{(n)}, C_2^{(n)}, \ldots) \stackrel{\mathrm{dist}}{\to} (Z_1, Z_2, \ldots)$, where the $\{Z_j\}_{j=1}^\infty$ are independent random variables satisfying $Z_j \sim \mathrm{Pois}(\frac{\theta}{j})$.

3. *Limiting behavior of the distribution of the shortest cycle:* Let $Y_1^{(n)}(\sigma)$ denote the length of the shortest cycle in σ. Then for $u > 1$,

$$\mathbb{P}_\theta(Y_1^{(n)} > \frac{n}{u}) \sim \Gamma(\theta)(\frac{u}{n})^\theta \omega_\theta(u), \text{ as } n \to \infty.$$

where ω_θ is the generalized Buchstab function. The Buchstab function, $\omega \equiv \omega_1$, is continuous on $[1, \infty)$ and satisfies the delay differential equation $(u\omega(u))' = \omega(u-1)$, for $u > 2$; $u\omega(u) = 1$, for $1 \le u \le 2$; $\lim_{u\to\infty} \omega(u) =$

$e^{-\gamma}$, where γ is Euler's constant. In fact ω can be written explicitly as

$$\omega(u) = u^{-1}\Big(1 + \sum_{2\le k\le u} \frac{1}{k!}\int\cdots\int_{J_k(u)} \frac{dy_1\cdots dy_{k-1}}{y_1\cdots y_{k-1}(1-y_1-\cdots-y_{k-1})}\Big),$$

where $J_k(u) = \{u^{-1} < y_i < 1, 1 \le i \le k-1; u^{-1} < 1 - y_1 - \cdots - y_{k-1} < 1\}$.

4. *Limiting distribution of the longest cycles:* Let $L_k^{(n)}(\sigma)$ denote the length of the k-th longest cycle in σ. Then $\frac{L_1^{(n)}}{n}$ converges in distribution to a random variable with density F_θ given by $F_\theta(x) = x^{\theta-1}\Gamma(\theta)\rho_\theta(\frac{1}{x})$, $x > 0$, where ρ_θ is the generalized Dickman function. The Dickman function, $\rho \equiv \rho_1$, solves the delay differential equation $u\rho'(u) + \rho(u-1) = 0$, $u > 0$; $\rho(u) = 0, u < 0; \rho(u) = 1, 0 \le u \le 1$. It can be written explicitly as

$$\rho(u) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!}\int\cdots\int_{I_k(u)} \frac{dy_1\cdots dy_k}{y_1\cdots y_k},$$

where $I_k(u) = \{uy_1 > 1, \ldots, uy_k > 1, y_1 + \cdots + y_k < 1\}$. More generally, we will prove that $(\frac{L_1^{(n)}}{n}, \frac{L_2^{(n)}}{n}, \cdots)$ converges in distribution to the Poisson-Dirichlet distribution with parameter θ.

5. *The ordered cycles:* There is a natural way, to construct random permutations, cycle by cycle. It is called the Feller coupling. This construction gives one an ordered sequence of cycle lengths, $(A_1^{(n)}, A_2^{(n)}, \cdots)$ for the random permutation, where $A_k^{(n)} = 0$ if there are fewer than k cycles. We will prove that $(\frac{A_1^{(n)}}{n}, \frac{A_2^{(n)}}{n}, \cdots)$ converges in distribution to the GEM distribution with parameter θ. The Poisson-Dirichlet distribution and the GEM distribution are related; the former is obtained as the decreasing order statistics of the latter.

The results concerning the limiting behavior of the shortest cycle and the limiting distribution of the longest cycle have interesting counterparts in number theory:

II. Results Concerning Number Theory

6. *Integers free of small prime factors:* Let $\Phi(x, y)$ denote the number of positive integers less than or equal to x with no prime factors less than y. Then $\Phi(x, x^{\frac{1}{u}}) \sim u\omega(u)\frac{x}{\log x}$. where ω is the Buchstab function. (Note that $u\omega(u) = 1$, for $u \in [1, 2]$, so this case reduces to the Prime Number Theorem.

We will use, but not prove, the Prime Number Theory in order to prove the result here.)

7. *Integers free of large prime factors:* Let $\Psi(x, y)$ denote the number of positive integers less than or equal to x which have no prime factor greater than y. Then for each $u > 1$, $\Psi(x, x^{\frac{1}{u}}) \sim x\rho(u)$, where ρ is the Dickman function. From this one can readily conclude that the asymptotic density of positive integers x with no prime factors greater than $x^{\frac{1}{u}}$ is $\rho(u)$. We can show that ρ decays very rapidly: $\rho(u) \leq \frac{1}{\Gamma(u+1)}$. Thus, in particular, the asymptotic density of positive integers x with no prime factor larger than $x^{\frac{1}{n}}$ is no more than $\frac{1}{n!}$. For the proof of the result here, we will use and prove (i) Chebyshev's theorem: $\pi(n)$, the number of primes less than or equal to n, satisfies $\pi(n) = O(\frac{n}{\log n})$; and (ii) Mertens' theorem: $\sum_{p \leq n} \frac{1}{p} = \log \log n + b + o(1)$, for some constant b.

Prerequisites: *No background in number theory is needed. A first course in probability, without measure theory, is almost a sufficient background in probability. The basic theory of weak convergence of measures is also needed, but I will review this at the beginning of the semester. Besides this, a strong background in basic mathematical analysis and some mathematical maturity are indispensable. Proofs of the results concerning random permutations will consist of a very integrated mixture of probability and analysis.*

Course Hours: Officially: Sunday: 12:30-14:30; Monday: 9:30-10:30, but I am flexible to making changes if necessary.

Books:

Arratia, R., Barbour, A.D. and Tavaré, S., *Logarithmic Combinatorial Structures: a Probabilistic Approach*, European Mathematical Society, 2003.

Montgomery, H. and Vaughn, R., *Multiplicative Number Theory. I. Classical Theory*, Cambridge University Press, 2007.

Pinsky R., *Problems from the Discrete to the Continuous*, Springer, 2014.