



A Large Deviation Principle for the Order of a Random Permutation

Neil O'Connell
Basic Research Institute in the
Mathematical Sciences
HP Laboratories Bristol
HPL-BRIMS-96-18
July, 1996

We obtain a large deviation principle for the scaled logarithm of the order of a random permutation of a large number of objects, and give an explicit expression for the convex dual of the rate function.

1 Introduction

Let σ_n be a random permutation of n objects, and denote by $O(\sigma_n)$ the order of σ_n . In 1967, Erdős and Turán [7] proved that, as $n \rightarrow \infty$, the sequence

$$\frac{\log O(\sigma_n) - \frac{1}{2} \log^2 n}{\left(\frac{1}{3} \log^3 n\right)^{1/2}}$$

converges in distribution to a standard Gaussian random variable. Their original proof was direct; since then, as one might expect, several probabilistic proofs have appeared. In particular, Arratia and Tavaré [3] give a proof based on the *Feller coupling*, which we will describe later.

Somewhat surprisingly, a corresponding large deviation principle (LDP) does not seem to have appeared in the literature. In this paper we prove such an LDP and give an explicit expression for the convex dual of the rate function. To be more specific, we will show that for any Borel set B ,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \log P \left(\frac{\log O(\sigma_n)}{\log^2 n} \in B \right) = - \inf_{x \in B} \Lambda^*(x),$$

where

$$\Lambda^*(x) = \sup_{\theta} [\theta x - \Lambda(\theta)]$$

and

$$\Lambda(\theta) = \frac{e^{2\theta} - 1}{2\theta} - 1.$$

To prove this, we will also make use of the Feller coupling, which provides a means of comparing $\log O(\sigma_n)$ with the more accessible

$$\sum_{j=1}^n Z_j \log j,$$

where the Z_j 's are independent Poisson random variables with respective means $1/j$. It turns out that, when properly normalised, these sequences are exponentially equivalent. The result can thus be obtained by a routine application of the Gärtner-Ellis theorem to the latter sequence.

The outline of this paper is as follows. In section 2 we describe the Feller coupling, which forms the basis of our approach, and record the relevant estimates. We will state and prove the main result in section 3. In section 4 we discuss some more sophisticated LDP's, which incorporate cycle lengths and allow one to consider certain biased permutations.

2 The Feller coupling

This coupling was introduced by Feller [8] in 1945, to establish a central limit theorem for the number of cycles in a random permutation. The idea is that one can construct a random permutation of n objects using n independent Bernoulli random variables, with the property that the times between successive 1's in the sequence correspond to the cycle-lengths in the permutation.

Let X_1, X_2, \dots be a sequence of Bernoulli random variables with

$$P(X_j = 1) = 1 - P(X_j = 0) = 1/j.$$

For each n , we construct a permutation σ_n recursively using the sequence

$$(X_n, X_{n-1}, \dots, X_1).$$

Start with 1 in the first cycle. If $X_n = 1$, close the cycle and start a new one, with the next available integer (2, in the first instance). If $X_n = 0$, choose one of the $n - 1$ remaining integers at random and place to the right of 1 in the same cycle. Continuing in this way produces a permutation $\sigma_n \in S_n$. Remarkably, the distribution of σ_n is uniform on S_n . Moreover, if C_j^n denotes the number of cycles of length j in the permutation σ_n , then

$$C_j^n = Z_j - Z_{j,n-j} + X_{n-j+1}(1 - X_{n-j+2}) \cdots (1 - X_n), \quad (1)$$

where

$$Z_{jm} = \sum_{i=m+1}^{\infty} X_i(1 - X_{i+1}) \cdots (1 - X_{i+j-1})X_{i+j}$$

is the number of j -spacings in the sequence X after time m , and $Z_j = Z_{j0}$. Another remarkable fact is that the Z_j 's are independent Poisson random variables with respective means $1/j$.

Now for any sequence a , define

$$O_n(a) = \text{lcm}\{i : 1 \leq i \leq n, a_i > 0\} \leq P_n(a) = \prod_{i=1}^n i^{a_i}.$$

Then $O(\sigma_n) = O_n(C^n)$, and we have the following estimates. For any pair of sequences a and b ,

$$O_n(a) \leq O_n(b)n^{|a-b|}$$

where $|a - b| = \sum_{i=1}^n |a_i - b_i|$. It therefore follows from (1) that

$$\log O_n(Z) - (Y_n + 1) \log n \leq \log O(\sigma_n) \leq \log O_n(Z) + \log n, \quad (2)$$

where $Y_n = \sum_{j=1}^n Z_{jn}$. We also have [3]:

$$D_n := \log P_n(Z) - \log O_n(Z) = \sum_{p^* \leq n} \left(\sum_{j \leq n: p^* | j} Z_j - 1 \right)^+ \log p. \quad (3)$$

The first summation is over all prime powers less than, or equal to, n .

3 The large deviation principle

We say a sequence of random variables X_n satisfies the LDP with rate a_n and rate function I if, for any Borel set B ,

$$-\inf_{x \in B^o} I(x) \leq \liminf_n a_n^{-1} \log P(X_n \in B) \leq \limsup_n a_n^{-1} \log P(X_n \in B) \leq -\inf_{x \in \bar{B}} I(x).$$

Theorem 1 *The sequence $\log O(\sigma_n) / \log^2 n$ satisfies the LDP with rate $\log n$ and rate function given by the convex dual of*

$$\Lambda(\theta) = \frac{e^{2\theta} - 1}{2\theta} - 1.$$

Recalling the estimates (2) and (3), the statement of Theorem 1 follows from the next three lemmas.

Lemma 1 *For any $c > 0$,*

$$\limsup_n \frac{1}{\log n} \log P(Y_n > c \log n) = -\infty.$$

Proof. Set $N_{mn} = \sum_{j=m+1}^n X_j$ and observe that

$$Y_n \leq A_n + B_n + C_n,$$

where

$$A_n = \sum_{k=1}^{\infty} N_{kn, (k+1)n} I(N_{kn, (k+1)n} > 1),$$

$$B_n = \sum_{k > 0 \text{ even}} I(N_{kn, (k+1)n} > 0, N_{(k+1)n, (k+2)n} > 0),$$

and

$$C_n = \sum_{k>0 \text{ odd}} I(N_{kn, (k+1)n} > 0, N_{(k+1)n, (k+2)n} > 0).$$

Note that A_n , B_n and C_n are each sums of independent Bernoulli random variables, so we can use the decomposition

$$P(Y_n > c \log n) \leq P(A_n > \frac{c}{3} \log n) + P(B_n > \frac{c}{3} \log n) + P(C_n > \frac{c}{3} \log n),$$

apply Chebyshev's inequality (with the convex function $x \mapsto e^{\theta x}$) to each term, then consider the normalised logarithmic limit and let $\theta \rightarrow \infty$ for the result. \square

Lemma 2 For any $c > 0$,

$$\limsup_n \frac{1}{\log n} \log P(D_n > c \log^2 n) = -\infty.$$

Proof. First we write

$$D_n = S_n - f(n) + R_n$$

where $S_n = \sum_{j \leq n} a(j)Z_j$, $a(j) = \sum_{p^* | j} \log p$, $f(n) = \sum_{p^* \leq n} \log p$ and

$$R_n = \sum_{p^* \leq n} I \left(\sum_{j \leq n: p^* | j} Z_j = 0 \right) \log p.$$

($\sum_{p^* | j}$ denotes the sum over all prime powers that divide j .) Set $g(n) = \sum_{j \leq n} \log j$. For any $\theta > 0$ we have

$$\log E e^{\theta(S_n - f(n)) / \log n} \leq -g(n).$$

It follows that, for $\lambda > 1/\theta$,

$$\begin{aligned} \frac{1}{\log n} \log P(S_n - f(n) + \lambda g(n) \log n > \frac{c}{2} \log^2 n) &\leq -\theta c/2 + (\theta \lambda - 1)g(n)/\log n \\ &\rightarrow -\infty, \end{aligned}$$

as $n \rightarrow \infty$. Also,

$$\begin{aligned} &\frac{1}{\log n} \log P(R_n - \lambda g(n) \log n > \frac{c}{2} \log^2 n) \\ &\leq \frac{1}{\log n} \log P(R_n > \lambda g(n) \log n) \end{aligned}$$

$$\begin{aligned}
&\leq -\lambda\theta + \frac{1}{\log n} \log E e^{\theta R_n/g(n)} \\
&\leq -\lambda\theta + \frac{1}{\log n} \log E \exp \left(\theta \frac{\log n}{g(n)} \sum_{p^s \leq n} I(Z_{p^s} = 0) \right) \\
&= -\lambda\theta + \frac{1}{\log n} \sum_{p^s \leq n} \log \left(\exp[-1/p^s + \theta \log n/g(n)] + 1 - e^{-1/p^s} \right) \\
&\leq -\lambda\theta + \frac{1}{\log n} \sum_{p^s \leq n} \left(\exp[-1/p^s + \theta \log n/g(n)] - e^{-1/p^s} \right) \\
&= -\lambda\theta + \frac{1}{\log n} \left(\exp[-1/p^s + \theta \log n/g(n)] - 1 \right) \sum_{p^s \leq n} e^{-1/p^s} \\
&\leq -\lambda\theta + \frac{\theta n/g(n)}{1 + \theta \log n/g(n)} \\
&\rightarrow -\lambda\theta
\end{aligned}$$

as $n \rightarrow \infty$. Now let $\theta \rightarrow \infty$ for the result. \square

Lemma 3 *The sequence $\log P_n(Z)/\log^2 n$ satisfies the LDP with rate $\log n$ and rate function given by the convex dual of*

$$\Lambda(\theta) = \frac{e^{2\theta} - 1}{2\theta} - 1.$$

Proof. This is a routine application of the Gärtner-Ellis Theorem (see, for example, [6]), by which it suffices to check that

$$\lim_n \frac{1}{\log n} \log E \exp(\theta \log P_n(Z)/\log n) = \Lambda(\theta).$$

This limit can be evaluated by noting that

$$\log P_n(Z) = \sum_{j \leq n} Z_j \log j.$$

\square

4 Spin-off's

If $K(\sigma_n)$ denotes the number of cycles in the random permutation, that is,

$$K(\sigma_n) = \sum_{j=1}^n C_j^n,$$

then it follows from (2.1) and Lemmas 3.2 and 3.3 that the pair

$$(K(\sigma_n)/\log n, \log O(\sigma_n)/\log^2 n)$$

is exponentially equivalent (at rate $\log n$) to the pair

$$\left(\sum_{j \leq n} Z_j / \log n, \sum_{j \leq n} Z_j \log j / \log^2 n \right).$$

The Gärtner-Ellis Theorem can thus be applied to the latter sequence to yield an LDP for the former, with rate $\log n$, and rate function given by the convex dual of

$$\Lambda_2(s, t) = e^s \frac{e^{2t} - 1}{2t} - 1.$$

We can also deduce an LDP the logarithm of the order of a *biased* random permutation, the law of which is given by the ‘tilted’ measure

$$P_s(\sigma) \propto e^{sK(\sigma)}.$$

An LDP for the sequence $K(\sigma_n)/\log n$, under the law P_s , can either be deduced from the above or proved directly using well-known generating functions for Stirling numbers of the first kind; this LDP also has rate $\log n$ and the rate function given by

$$x \mapsto s - x + x \log(x/s).$$

Note that this is a Poisson rate function, and so this LDP is consistent with (but not implied by) the Poisson approximation results in the papers of Arratia, Barbour, Stark and Tavaré [1, 2, 4, 5]. For related work on the number of components in a random combinatorial structure see Hwang [9] and Maxwell [10], and references therein.

References

- [1] R.A. Arratia, A.D. Barbour and S. Tavaré. “Poisson process approximations for the Ewens sampling formula,” *Ann. Prob.* 2(3), pp. 519–535, 1992.
- [2] R.A. Arratia, Dudley Stark and S. Tavaré. “Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies,” *Ann. Prob.* 23(3), pp. 1347–1388, 1995.

- [3] R.A. Arratia and S. Tavaré. "Limit theorems for combinatorial structures via discrete process approximations," *Rand. Struct. Alg.* 3, pp. 321–345, 1992.
- [4] R.A. Arratia and S. Tavaré. "The cycle structure of random permutations," *Ann. Prob.* 20(3), pp. 1567–1591, 1992.
- [5] R.A. Arratia and S. Tavaré. "Independent process approximations for random combinatorial structures," *Adv. Math.* 104(1), pp. 90–154, 1994.
- [6] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Jones and Bartlett, 1993.
- [7] P. Erdős and P. Turán. "On some problems of a statistical group theory, III," *Acta Math. Acad. Sci. Hungar.* 18, pp. 309–320, 1967.
- [8] W. Feller. "The fundamental limit theorems in probability," *Bull. Amer. Math. Soc.* 51, pp. 800–832, 1945.
- [9] Hsien-Kuei Hwang. "Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques," PhD Thesis, Ecole Polytechnique, 1994.
- [10] Mark M. Maxwell. "A note on rational structures and their asymptotics," preprint.