Precise Large Deviations of the Ewens-Pitman Model

Mariia Khodiakova[a]

[a] University of Zurich.

This work is joint with Thomas Lehéricy, Pierre-Loic Méliot and Ashkan Nikeghbali.
The Ewens-Pitman model was introduced in [1] by Pitman as a generalization of the Ewens model in population genetics ([2]). For any $n \in \mathbb{N}$, we consider a random partition of $[n] = \{1, \dots, n\}$ into $K_n \in \{1, \dots, n\}$ partition sets, and $K_{r,n} \in \{0, 1, \dots, n\}$ is the number of partition subsets of size r, for any $r = 1, \dots, n$:

$$n = \sum_{r=1}^{n} r K_{r,n}, \quad K_n = \sum_{r=1}^{n} K_{r,n}.$$

For $\alpha \in [0, 1)$ and $\theta > -\alpha$, the Ewens-Pitman model assigns to $\mathbf{K}_n = (K_{1,n}, \dots, K_{n,n})$ the probability

$$\mathbb{P}_{\alpha,\theta}(\mathbf{K}_n = (k_1, \dots, k_n)) = n! \frac{\left(\frac{\theta}{\alpha}\right)^{\left(\sum_{i=1}^{n} k_i\right)}}{(\theta)^{(n)}} \prod_{i=1}^{n} \left(\frac{\alpha(1-\alpha)^{(i-1)}}{i!} \right)^{k_i} \frac{1}{k_i!},$$

where $(a)^{(m)} = a(a+1)\cdots(a+m-1)$ for all $a \in \mathbb{R}$ and $m \in \mathbb{N} \cup \{0\}$ ($(a)^{(0)} = 1$). Setting $\alpha = 0$ corresponds to the Ewens model.
There is a significant number of works about asymptotic behaviour of the number K_n of partition sets in the Ewens-Pitman model. Now, we are interested in large deviations. For $\alpha = 0$, K_n is a sum of independent Bernoulli random variables ([3]), from which it follows easily that (K_n/n) satisfies a large deviation principle with good rate function

$$I_\theta(x) = \begin{cases} x \ln\left(\frac{x}{\theta}\right) - x + \theta, & x > 0, \\ \theta, & x = 0, \\ +\infty, & x < 0. \end{cases}$$

For $\alpha \in (0, 1)$, this is not true. However, in [4], Feng and Hoppe proved large deviation principle with good rate function $I_\alpha(x) = \sup_{t \in \mathbb{R}}(xt - L_\alpha(t))$, where

$$L_\alpha(t) = \begin{cases} -\ln\left(1 - (1 - e^{-t})^{1/\alpha}\right) & t > 0, \\ 0, & t \le 0. \end{cases}$$

We obtain precise large deviation principle for (K_n/n) with $\alpha \in (0, 1)$, considering the moment-generating function of K_n with using some results of Bercu and Favaro ([5]). Our research is based on improving Chaganty-Sethuraman's result ([6]) about local limit theorems and precise large deviations for the lattice case that has been recently obtained by us with using the approach suggested in [7] by Delbaen, Kowalski, and Nikeghbali. Notice that, in this case, we do not have mod-φ convergence and we cannot apply some already known results of Large Deviation theory. Thus, one of the most important parts of our work is to weaken assumptions on the known general results and to extend their applications.
It is worth mentioning that, in [8], Peng and Zhou recently proved precise large deviation principle for the Ewens-Pitman model, but their approach is different from ours and does not allow us to clearly see the connection with Feng-Hoppe large deviation principle.

References:

[1] Pitman J. Exchangeable and partially exchangeable random partitions, *Probability theory and related fields*, 102.2 (1995), 145-158.

[2] Ewens W. J. The sampling theory of selectively neutral alleles, *Theoretical population biology*, 3.1 (1972), 87-112.

[3] Korwar R. M., Hollander M. Contributions to the theory of Dirichlet processes, *The Annals of Probability*, 1.4 (1973), 705-711.

[4] Feng S., Hoppe F. M. Large deviation principles for some random combinatorial structures in population genetics and Brownian motion, *The Annals of Applied Probability*, 8.4 (1998), 975-994.

[5] Bercu B., Favaro S. A new look on large deviations and concentration inequalities for the Ewens-Pitman model, *Preprint*, 2025.

[6] Chaganty N. R., Sethuraman J. Strong large deviation and local limit theorems, *The Annals of Probability*, (1993), 1671-1690.

[7] Delbaen F., Kowalski E., Nikeghbali A. Mod-φ convergence, *International Mathematics Research Notices*, 2015.11 (2015), 3445-3485.

[8] Peng Z., Zhou Y. Precise Deviations for the Ewens-Pitman Model, *Preprint*, 2025.