FPSAC 2022
Indian Institute of Science, Bangalore

Continuously Increasing Subsequences of Random Multiset Permutations

Alexander Clifton [1] Bishal Deb [2] Yifeng Huang [3] Sam Spiro [4] Semin Yoo [5]

[1]Dept. of Mathematics, Emory University [2]Dept. of Mathematics, University College London [3]Dept. of Mathematics, University of Michigan
[4]Dept. of Mathematics, University of California at San Diego [5]School of Computational Sciences, Korea Institute of Advanced Studies

Abstract

For a word π and integer i, we define $L^i(\pi)$ to be the length of the longest subsequence of the form $i(i+1)\cdots j$, and we let $L(\pi) := \max_i L^i(\pi)$. In this work we estimate the expected values of $L^1(\pi)$ and $L(\pi)$ when π is chosen uniformly at random from all words which use each of the first n positive integers exactly m times. We show that $\mathbb{E}[L^1(\pi)] \sim m$ if n is sufficiently larger in terms of m as m tends towards infinity, confirming a conjecture of Diaconis, Graham, He, and Spiro. We also show that $\mathbb{E}[L(\pi)]$ is asymptotic to the inverse gamma function $\Gamma^{-1}(n)$ if n is sufficiently large in terms of m as m tends towards infinity.

Introduction

- Let $\mathfrak{S}_{m,n}$ denote the set of words π which use each integer in $[n] := \{1, 2, \ldots, n\}$ exactly m times, and we will refer to $\pi \in \mathfrak{S}_{m,n}$ as a **multiset permutation**. For example,

$$\pi = 211323 \in \mathfrak{S}_{2,3}.$$

- Let $L^i_{m,n}(\pi)$ denote the length of the longest subsequence of π of the form $i(i+1)(i+2)\cdots j$, which we call an ***i*-continuously increasing subsequence**.
- A subsequence is a **continuously increasing subsequence** if it is an i-continuously increasing subsequence for some i. We define

$$L_{m,n}(\pi) = \max_i L^i_{m,n}(\pi)$$

to be the length of a longest continuously increasing subsesquence of π.
- For example, if $\pi = 2341524315$ then

$$L_{2,5}(\pi) = L^2_{2,5}(\pi) = 4 \qquad (\mathbf{2341}524315)$$

and

$$L^1_{2,5}(\pi) = 3 \qquad (23415\mathbf{24}31\mathbf{5}).$$

- We study $L^1_{m,n}(\pi)$ and $L_{m,n}(\pi)$ when π is chosen uniformly at random from $\mathfrak{S}_{m,n}$ and focus on the regime where n is much larger than m.

Brief Description of Main Results

- $\mathbb{E}[L^1_{m,n}(\pi)]$ — This quantity was briefly studied by Diaconis, Graham, He, and Spiro [2] due to its relationship with a certain card game. They proved $\mathbb{E}[L^1_{m,n}(\pi)] \le m + Cm^{3/4}\log m$ when n is sufficiently large in terms of m. They conjectured that this upper bound is asymptotically tight. We prove this conjecture in a strong form, by providing precise estimates.

$$\left| \lim_{n\to\infty} \mathbb{E}[L^1_{m,n}(\pi)] - \left(m + 1 - \frac{1}{m+2}\right)\right| \le O(e^{-\beta m}).$$

We use techniques from generating function theory to obtain exact expressions followed by some analysis of distribution of roots of the partial sums of the exponential function, which appear in our exact expression.
- $\mathbb{E}[L_{m,n}(\pi)]$ — We obtain precise bounds for this quantity:

$$\mathbb{E}[L_{m,n}(\pi)] = \Gamma^{-1}(n) + \Theta\left(1 + \frac{\log m}{\log(\Gamma^{-1}(n))}\Gamma^{-1}(n)\right).$$

We use first moment method to prove the upper bound and techniques from coding theory to prove the lower bound.

First Main Result

Theorem 1

a For any integer $m \ge 1$, let $\alpha_1, \ldots, \alpha_m$ be the zeroes of $E_m(x) := \sum_{k=0}^m \frac{x^k}{k!}$. If $\pi \in \mathfrak{S}_{m,n}$ is chosen uniformly at random, then the expected maximum length of a 1-continuously increasing subsequence of π satisfies

$$\lim_{n\to\infty} \mathbb{E}[L^1_{m,n}(\pi)] = -1 - \sum \alpha_i^{-1} e^{-\alpha_i}. \qquad (1)$$

b There exists an absolute constant $\beta > 0$ such that

$$\left| \lim_{n\to\infty} \mathbb{E}[L^1_{m,n}(\pi)] - \left(m + 1 - \frac{1}{m+2}\right)\right| \le O(e^{-\beta m}).$$

Proof Sketch

- Let $h_m(n)$ denote the number of words $\pi \in \mathfrak{S}_{m,n}$ for which $L^1_{m,n}(\pi) = n$. Horton and Kurn [3] gave an exact formula for $h_m(n)$.
- We plug this in the identity

$$\lim_{k\to\infty} \mathbb{E}[L^1_{m,k}(\pi)] = \lim_{k\to\infty} \sum_{n=1}^{k} \Pr[L^1_{m,k}(\pi) \ge n] = \lim_{k\to\infty} \sum_{n=1}^{k} \frac{h_m(n)}{|\mathfrak{S}_{m,n}|} = \sum_{n=1}^{\infty} \frac{h_m(n)}{|\mathfrak{S}_{m,n}|}$$

followed by some careful computation to obtain

$$\lim_{n\to\infty} \mathbb{E}[L^1_{m,n}(\pi)] = \Phi\left(\frac{1}{m! x^m E_m(1/x)}\right)\Bigg|_{x=-1} \qquad (2)$$

where Φ is the Formal Laplace Transform for formal power series in x, defined on monomials by $\Phi(x^n) = x^n/n!$. Expanding Equation 2 into partial fractions we obtain Theorem 1(a).
- For Theorem 1(b), we show that

$$\sum_{i=1}^{m} \alpha_i^{-1} e^{-\alpha_i} = -m - 2 + \frac{1}{m+2} + O(e^{-\beta m})$$

for some positive constant β. We follow the approach used by Conrey and Ghosh [1], where a similar exponential sum was estimated.

Second Main Result

Theorem 2

If n is sufficiently large in terms of m, then

$$\mathbb{E}[L_{m,n}(\pi)] = \Gamma^{-1}(n) + \Theta\left(1 + \frac{\log m}{\log(\Gamma^{-1}(n))}\Gamma^{-1}(n)\right),$$

where $\Gamma^{-1}(n)$ is the inverse of the gamma function when restricted to $x \ge 1$.

The Upper Bound

- We obtain the inequality

$$\mathbb{E}[L_{m,n}(\pi)] = \sum_{K=1}^{\infty} \Pr[L_{m,n}(\pi) \ge K] \le k + \sum_{K>k} \Pr[L_{m,n}(\pi) \ge K] \le k + \sum_{K>k} \frac{nm^K}{K!}$$

where the last inequality is obtained by using Markov's inequality on a slightly different random variable.
- We then plug in $k \approx \Gamma^{-1}(n) + \frac{2\log m}{\log \Gamma^{-1}(n)}\Gamma^{-1}(n)$ to obtain an upper bound of the desired form.

The Lower Bound

- For a fixed integer k and for $0 \le j < \lfloor n/k \rfloor$, let $A_j(\pi)$ be the event that π contains the subsequence $(jk+1)(jk+2)\cdots((j+1)k)$ and let $p_k = \Pr[A_j(\pi)]$ where $\sigma \in \mathfrak{S}_{m,k}$ is chosen uniformly at random. By using the independence of the $A_j(\pi)$, we obtain the inequality

$$\Pr[L_{m,n}(\pi)] \ge 1 - (1 - p_k)^{\lfloor n/k \rfloor}.$$

- Using ideas from coding theory, we obtain the inequality for $m \ge 2$

$$p_k \ge (m/1.03)^k/(2n \cdot k!)$$

- We take partial sum in the identity $\mathbb{E}[L_{m,n}(\pi)] = \sum_{K=1}^{\infty} \Pr[L_{m,n}(\pi) \ge K]$ and the above inequalities to obtain a lower bound of the desired form.

Concluding Remarks and Open Problems

Based off of computational evidence, we conjecture the following for higher moments:

Conjecture

For all $r \ge 1$, if n is sufficiently large in terms of m, then

$$\mathbb{E}[(L^1_{m,n}(\pi) - \mu)^r] = c_r m^{\lfloor r/2 \rfloor} + O(m^{\lfloor r/2 \rfloor - 1}),$$

where $\mu = \mathbb{E}[L^1_{m,n}(\pi)]$ and

$$c_r = \begin{cases} \frac{r!}{2^{r/2}(r/2)!} & r \text{ even}, \\ \frac{r!}{3 \cdot 2^{(r-1)/2}((r-3)/2)!} & r \text{ odd}. \end{cases}$$

This fact would imply that $(L^1_{m,n}(\pi) - \mu)/\sigma$ converges in distribution to a standardized normal distribution where σ is the standard deviation of $L^1_{m,n}(\pi)$.

Our computational data also suggests the following conjecture for the non-centralized moments:

Conjecture

For all $r \ge 1$, if n is sufficiently large in terms of m, then

$$\mathbb{E}[L^1_{m,n}(\pi)^r] = m^r + \binom{r+1}{2} m^{r-1} + O(m^{r-2}).$$

It is natural to consider all subsequences in multiset permutations and ask the question:

Question

For $\pi \in \mathfrak{S}_{m,n}$, let $\widetilde{L}_{m,n}(\pi)$ denote the length of a longest increasing subsequence in π. What is $\mathbb{E}[\widetilde{L}_{m,n}(\pi)]$ asymptotic to when m is fixed?

When $m = 1$ it is well known that $\mathbb{E}[\widetilde{L}_{1,n}(\pi)] \sim 2\sqrt{n}$ (see [4]).

References

[1] Brian Conrey and Amit Ghosh. On the zeros of the Taylor polynomials associated with the exponential function. *American Mathematical Monthly*, 95(6):528–533, 1988.
[2] Persi Diaconis, Ron Graham, Xiaoyu He, and Sam Spiro. Card guessing with partial feedback. *Combinatorics, Probability and Computing*, pages 1–20, 2021.
[3] J.D. Horton and Andrew Kurn. Counting sequences with complete increasing subsequences. *Congressus Numerantium*, pages 75–80, 1981.
[4] Dan Romik. *The surprising mathematics of longest increasing subsequences*. Cambridge University Press, 2015.