Longest increasing subsequences in pattern-restricted permutations

Emeric Deutsch
Polytechnic University
Brooklyn, NY 11201
deutsch@duke.poly.edu

A. J. Hildebrand
University of Illinois
Urbana, IL 61801
hildebr@math.uiuc.edu

Herbert S. Wilf
University of Pennsylvania
Philadelphia, PA 19104-6395
wilf@math.upenn.edu

Submitted: Apr 8, 2003; Accepted: Apr 21, 2003; Published: May 3, 2003
MR Subject Classifications: 05A16, 05A05

Abstract

Inspired by the results of Baik, Deift and Johansson on the limiting distribution of the lengths of the longest increasing subsequences in random permutations, we find those limiting distributions for pattern-restricted permutations in which the pattern is any one of the six patterns of length 3. We show that the (132)-avoiding case is identical to the distribution of heights of ordered trees, and that the (321)-avoiding case has interesting connections with a well known theorem of Erdős-Szekeres.

1 Introduction

A great deal of spectacular work has been done on determining the distribution of the length of the longest increasing subsequence, $\mathrm{LIS}(p)$, in a random permutation p. This work culminated in the paper of Baik, Deift and Johansson [2] which found the limiting distribution function.

Historically, the *average* length, $\mu(n)$, of the LIS had been found to be $\geq \sqrt{n}/2$ [6] and then $\geq 2\sqrt{n}$ [12]. The correct order of magnitude was established in [18], where it was shown to be $\sim 2\sqrt{n}$. Next the standard deviation $\sigma(n)$ had been estimated [13] to be approximately $n^{1/6}$ but was not proved to be so. Finally, in [2], the complete limiting distribution function of the normalized random variable $(\mathrm{LIS}-\mu(n))/n^{1/6}$ was determined.

We can ask the same questions in some given subset of the set of all permutations. For example, what is the distribution of the length of the longest increasing subsequence in a randomly chosen permutation from the set of those that avoid the pattern[1] (231),

[1] A permutation p avoids (132) (resp. (231), (321)) if there do not exist $1 \leq i < j < k \leq n$ s.t. $p(i) < p(k) < p(j)$ (resp. $p(k) < p(i) < p(j)$, $p(i) > p(j) > p(k)$).

or the pattern (321)? The tools are available for the study of such problems. A paper of Reifegerste [14] finds a simple expression for the frequencies of different LIS lengths in the (231)-avoiding case. For (132)-avoiding permutations it turns out that the distributions are identical with known distribution functions of ordered trees by height, so we can simply quote the known results. Finally, for (321)-avoiding permutations the problem is related to a celebrated theorem of Erdős-Szekeres. In this case we exhibit the relevant statistics and find explicitly the limiting distribution function.

We remark that for all six of the patterns of length 3 it is true that the number of permutations of n letters that avoid the pattern is the Catalan number $C_n = \binom{2n}{n}/(n+1)$. So for each of the six patterns that we will study here, we will be asking what fraction of the roughly $c4^n/n^{3/2}$ permutations that avoid that pattern have a LIS of some given length.

Our main results are the following.

Theorem 1. *In the class of* (231)*-avoiding permutations of* n *letters, and in the class of* (312)*-avoiding permutations of* n *letters, the length of the longest increasing subsequence has mean* $(n+1)/2$*, and standard deviation* $\sim \sqrt{n}/2$*. Moreover, the random variable* $(LIS(p) - (n+1)/2)/(\sqrt{n}/2)$ *has asymptotically the standard normal distribution.*

Theorem 2. *In the class of* (132)*-avoiding permutations of* n *letters, and in the class of* (213)*-avoiding permutations of* n *letters, the length of the longest increasing subsequence has mean* $= \sqrt{\pi n} + O(n^{\frac{1}{4}})$ *and standard deviation*

$$= c_1\sqrt{n} + O(n^{\frac{1}{4}}) \qquad \left(c_1 = \sqrt{\pi\left(\frac{\pi}{3} - 1\right)} = 0.38506..\right). \tag{1.1}$$

Moreover, the normalized random variable

$$X_n(p) \stackrel{def}{=} \frac{\mathrm{LIS}(p) - \sqrt{\pi n}}{\sqrt{n}},$$

defined for n*-permutations* p *that avoid the pattern (132), satisfies*

$$\lim_{n\to\infty} \mathrm{Prob}(X_n(p) \le \theta) = \sum_{t=-\infty}^{\infty} (1 - 2t^2(\theta + \sqrt{\pi})^2)e^{-(\theta+\sqrt{\pi})^2 t^2} \qquad (\theta > -\sqrt{\pi}). \tag{1.2}$$

Theorem 3. *For permutations* p *in the class of* (321)*-avoiding permutations of* n *letters, define the random variable*

$$X_n(p) \stackrel{def}{=} \frac{\mathrm{LIS}(p) - \frac{n}{2}}{\sqrt{n}}.$$

Then we have

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Prob}(X_n(p) \le \theta) &= \frac{2}{\sqrt{\pi}} \int_0^{4\theta^2} u^{1/2} e^{-u} du \\ &= \frac{\Gamma(3/2, 4\theta^2)}{\Gamma(3/2)},\end{aligned}$$

where $\Gamma(z,w)$ *is the incomplete Gamma function*

$$\Gamma(z,w) = \int_0^w u^{z-1}e^{-u}du.$$

2 Avoiding the pattern (231)

One situation where it is simple to deduce the limiting distribution is that of (231)-avoiding permutations. This is because Reifegerste [14] has shown that the number of (231)-avoiding permutations whose longest increasing subsequence has length exactly k is

$$e(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}. \tag{2.1}$$

The number of all (231)-avoiding permutations of n letters is C_n, the Catalan number. From this it is easy to check that the mean length of the LIS in this family is $(n+1)/2$ and the standard deviation is $\sqrt{n}/2 + O(1)$. Finally it is a straightforward exercise in Stirling's formula to find that

$$\lim_{n\to\infty} \sqrt{n}\frac{e\left(n, \frac{n+1}{2} + \theta\frac{\sqrt{n}}{2}\right)}{C_n} = \frac{2}{\sqrt{\pi}}e^{-\theta^2}.$$

This says that for (231)-avoiding permutations p, the random variable

$$X(p) = \frac{LIS(p) - \frac{n+1}{2}}{\frac{1}{2}\sqrt{n}}$$

has asymptotically the standard normal distribution.

It seems noteworthy that in this case the mean is quite large, i.e., typically such a permutation contains an increasing subsequence whose length is about half that of the permutation itself.

It also is noteworthy that (2.1) shows that the number of (231)-avoiding permutations of n letters whose longest increasing subsequence has length k is the same as the number whose longest increasing subsequence has length $n-k+1$. It is arresting that the number of (231)-avoiding permutations of, say, 1000 letters, whose LIS has length 980 is equal to the number whose LIS has length 21.

3 Avoiding the pattern (132)

The case of (132)-avoiding permutations is handled by the following sequence of observations.

- The number $f(n,k)$ of n-permutations that avoid (132) and whose longest increasing subsequence has length $< k$ is equal to the number of ordered trees of n edges whose height is $< k$. This follows from a bijection of Jani and Rieper [10].

- It is known from the theory of ordered trees [5] that this number is

$$f(n,k) = 2\sum_{t=-\infty}^{\infty}\left(\binom{2n}{n+t(k+1)} - \frac{1}{4}\binom{2n+2}{n+1+t(k+1)}\right) \qquad (k=1,2,3,\ldots). \tag{3.1}$$

- To see that the normalized random variable

$$X(p_n) = \frac{\mathrm{LIS}(p_n) - \sqrt{\pi n}}{\sqrt{n}},$$

defined on (132)-avoiding permutations p_n of n letters, has a nontrivial limiting distribution, as $n \to \infty$, we consider

$$\begin{aligned}\mathrm{Prob}(X(p_n) < \theta) &= \mathrm{Prob}(\mathrm{LIS}(p_n) \leq (\theta + \sqrt{\pi}\,)\sqrt{n}) \\ &= \frac{f(n, \sqrt{n}(\theta + \sqrt{\pi}\,))}{C_n},\end{aligned}$$

where

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

is the nth Catalan number. But by (3.1) we see that

$$f(n, \sqrt{n}(\theta+\sqrt{\pi}\,)) = \frac{1}{C_n}\sum_{t=-\infty}^{\infty}\phi(n,t),$$

in which the general t-th term of the sum is

$$\phi(n,t) = \frac{n+1}{\binom{2n}{n}}\binom{2n}{n+tc\sqrt{n}}\left(2 - \frac{(n+1)(2n+1)}{(n+1)^2 - t^2c^2n}\right). \qquad (c = \theta + \sqrt{\pi})$$

It is an easy exercise in Stirling's formula to see that

$$\lim_{n\to\infty}\phi(n,t) = (1-2c^2t^2)e^{-c^2t^2}$$

and the proof is complete.

The limiting cumulative distribution function $F(\theta)$, in (1.2), is related to the Jacobi theta function

$$\Theta(x) = \sum_{t=-\infty}^{\infty} e^{-t^2\pi x}.$$

In fact, $F(\theta) = \Theta(u) + 2u\Theta'(u)$, where $u = (1+\theta/\sqrt{\pi})^2$. It is well known [8] that this is indeed the limiting distribution, from the theory of ordered trees, but the derivation is simple enough that we have included it above. Indeed [8] shows that this same distribution function, originally considered by Rényi and Szekeres [15], is the limiting distribution function for many classes of trees.

We remark that although the Jani-Rieper bijection [10] is the most explicit mapping between 132-avoiding permutations of n letters whose LIS has length k and ordered trees of n edges whose height is k, there are other, somewhat less explicit, bijections in the literature. Indeed, in Krattenthaler [11], we find an explicit bijection between such permutations and Dyck paths. This mapping has the property that the length of the LIS corresponds to the height of the image path. In view of well known bijections between Dyck paths by height and ordered trees by height, we have, by composition, the mapping that we need here. Likewise in Chow, West [4] there are remarks in section 5 that suggest, less explicitly, that such bijections exist. Other, equivalent bijections with Dyck paths have been given by Fulmek [9], Reifegerste [14], and Bandlow and Killpatrick [3].

4 The (321) case and the Erdős-Szekeres theorem

Finally we ask for the limiting distribution function for the length of the longest increasing subsequence in the class $S_n(321)$ of (321)-avoiding permutations of n letters. Let $f(n,k)$ denote the number of such permutations whose longest increasing subsequence has length $\leq k$. A famous theorem of Erdős-Szekeres states that every permutation of $(r-1)(s-1)+1$ or more letters contains either a decreasing subsequence of length r or an increasing subsequence of length s. If we take $r=3$ we see that every permutation in the class $S_{2k+1}(321)$ has an increasing subsequence of length $k+1$, i.e. $f(n,k)=0$ for all $n \geq 2k+1$.

Even though, for each fixed k, $f(n,k)$ vanishes for all large enough n, the question of the limiting distribution function of the LIS remains.

Theorem 4. *The number $a(n,k)$ of (321)-avoiding permutations having longest increasing subsequence of length $=k$ is given by*

$$a(n,k) = \left(\frac{2k-n+1}{n+1}\binom{n+1}{k+1}\right)^2 \tag{4.1}$$

for $\lfloor (n+1)/2 \rfloor \leq k \leq n$, and vanishes otherwise.

These permutations correspond 1-1, under the Schensted insertion algorithm (see, e.g., [16]), to pairs of Young tableaux with n cells, at most two rows, and k columns, so the number of them is B^2, where B is the number of such tableaux. A brief computation with the hook formula shows that

$$B = \frac{2k-n+1}{n+1}\binom{n+1}{n-k},$$

completing the proof. □

It follows that the number $f(n,k)$ of permutations in $S_n(321)$ whose LIS has length $\leq k$ is

$$f(n,k) = \sum_{\lfloor (n+1)/2 \rfloor \leq j \leq k} \left(\frac{2j-n+1}{n+1}\binom{n+1}{j+1}\right)^2. \tag{4.2}$$

Now fix $t > 0$. The probability that a permutation of n letters has LIS of length at most $n/2 + t\sqrt{n}$ is

$$\frac{f(n, n/2 + t\sqrt{n})}{\frac{1}{n+1}\binom{2n}{n}}.$$

Hence the limiting probability distribution function that we seek is

$$\begin{aligned} F(t) &= \lim_{n\to\infty} \frac{1}{\frac{1}{n+1}\binom{2n}{n}} \sum_{\lfloor (n+1)/2 \rfloor \le j \le n/2 + t\sqrt{n}} \left(\frac{2j - n + 1}{n+1} \binom{n+1}{j+1} \right)^2 \\ &= \lim_{n\to\infty} \frac{4}{(n+1)\binom{2n}{n}} \sum_{0 \le k \le t\sqrt{n} + 1/2} k^2 \binom{n+1}{(n+1)/2 + k}^2. \end{aligned}$$

By an easy application of Stirling's formula the relation

$$\frac{\binom{2m}{m-k}}{\binom{2m}{m}} \sim e^{-k^2/m}$$

holds uniformly for all $|k| < t\sqrt{m}$, and we have

$$F(t) = \frac{32}{\sqrt{\pi}} \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{0 \le k \le t\sqrt{n}} \frac{k^2}{n} e^{-4k^2/n}.$$

The quantity whose limit is being taken is a Riemann sum that approximates the integral

$$\int_0^t x^2 e^{-4x^2}$$

with grid size $1/\sqrt{n}$, and so the desired limit is

$$\begin{aligned} F(t) &= \frac{32}{\sqrt{\pi}} \int_0^t x^2 e^{-4x^2} dx \\ &= \frac{2}{\sqrt{\pi}} \int_0^{4t^2} u^{1/2} e^{-u} du \\ &= \frac{\Gamma(3/2, 4t^2)}{\Gamma(3/2)}. \end{aligned}$$

The limiting distribution $F(t)$ is well-known in statistics: Except for a linear change of variables, it is a χ^2 distribution with three degrees of freedom, which arises as the distribution of the length of a 3-vector whose entries are independent standard normal random variables (see, for example, [1]). It would be interesting to find a heuristic explaining this connection.

5 Other patterns of three letters

We have so far dealt with the statistics of the length of the longest increasing subsequence in permutations that avoid any one of the patterns (132), (231), or (321). There remain three other patterns of three letters, viz. (123), (312), and (213).

The (123) case is trivial. If an n-permutation avoids (123) then its LIS has length exactly 2 except for the reversal of the identity permutation, whose LIS has length 1.

The (312) case is identical with the (231). Indeed, the length of the LIS of a (312)-avoiding permutation is equal to the length of the longest decreasing subsequence of its reversal, which is (213)-avoiding, and that is equal to the length of the LIS of the complement of its reversal, which is (231)-avoiding, as claimed.

Similarly, by reversal and complementation, the (213) case is identical with the (132).

We thank Christian Krattenthaler, Brendan McKay and Julian West for informative discussions related to this paper.

References

[1] M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, Dover Publ., New York, 1965.

[2] J. Baik, P. Deift and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, J. Amer. Math. Soc. **12** (1999) 1119-1178.

[3] Jason Bandlow and Kendra Killpatrick, An Area-to-Inv Bijection Between Dyck Paths and 312-avoiding Permutations, Elec. J. Combinatorics **8** (2001), #R40.

[4] T. Chow and J. West, Forbidden subsequences and Chebyshev polynomials, Discr. Math. **204** (1999), 119–128.

[5] De Bruijn, N.G., Knuth, D.E., and Rice, S.O., The average height of planted plane trees, In *Graph Theory and Computing* (1972), R.C. Read, Ed., Academic Press, 15–22.

[6] P. Erdős and G. Szekeres, A combinatorial theorem in geometry, Compositio Math., **2** 463-470, (1935).

[7] Philippe Flajolet, Xavier Gourdon and Philippe Dumas, Mellin transforms and asymptotics: harmonic sums, 55 pp., see, e.g.,the website of Flajolet.

[8] Philippe Flajolet and Andrew Odlyzko, The average height of binary trees and other simple trees, *J. Computer and System Sciences* **25** (1982), 171–213.

[9] Markus Fulmek, Enumeration of permutations containing a prescribed number of occurrences of a pattern of length three, arXiv: math.CO/0112092, Apr. 2002.

[10] Mahendra Jani and Robert Rieper, Continued fractions and Catalan problems, *Electronic J. Combinatorics* **7** (2000), #R45.

[11] C. Krattenthaler, Permutations with restricted patterns and Dyck paths, Adv. Appl. Math. **27** (2001), 510-530.

[12] B.F.Logan and L.A.Shepp, A variational problem for random Young tableaux, Advances in Math., **26**, 206-222, (1977).

[13] A.M.Odlyzko and E.M.Rains, On longest increasing subsequences in random permutations, Analysis, geometry, number theory: the mathematics of Leon Ehrenpreis (Philadelphia, PA, 1998), 439–451, Contemp. Math., **251**, Amer. Math. Soc., Providence, RI, 2000.

[14] Astrid Reifegerste, On the diagram of 132-avoiding permutations, arXiv: math. CO/ 0208006, Aug. 2002.

[15] A. Rényi and G. Szekeres, On the height of trees, *J. Austral. Math. Soc.* **7** (1967), 497–507.

[16] B. E. Sagan, The Symmetric Group, 2nd edition, Springer, New York, 2001.

[17] C. Schensted, Longest increasing and decreasing subsequences, Canad. J. Math., 13, 1961, 179-191.

[18] A.M.Vershik and S.V.Kerov, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables, Soviet Math. Dokl., **18**, 527-531, (1977).