Discrete Mathematics 22 (1978) 35–40.

ASCENDING SEQUENCES IN PERMUTATIONS

D.G. ROGERS
Department of Mathematics, The University of Western Australia, Nedlands, Australia

Received 13 July 1976

The problem of the number $p(n, r)$, $(1 \leq r \leq n)$, of permutations on the set $\{1, \ldots, n\}$ with longest ascending subsequence of given length r is considered. By placing further restrictions on the ascending subsequence, combinatorial identities are obtained which allow the explicit calculation of $p(n, r)$ in some cases.

1. Introduction, definitions and notation

Ulam [5] asks: what is the distribution of the length of the longest monotone subsequence of terms in a random permutation of the first $n^2 + 1$ natural numbers? Hammersley, in a recent discussion [1] of this problem, established the result: if π_n is a random permutation, uniformly distributed over the symmetric group S_n and if $l(\pi_n)$ is the length of the longest ascending sequence (q.v.) in π_n then, for some constant c, $n^{-1/2} l(\pi_n)$ converges to c in probability. Hammersley conjectured that $c = 2$. Support has been given to this by Logan and Shepp [3] who, by variational methods based on a relation between the probability distribution of $l(\pi_n)$ and Young tableaux [4], showed that $c \geq 2$. It is also known [2] that $c < 2.49$.

In this note we consider some of the combinatorial aspects of the problem. (In view of [4], this is equivalent to considering combinatorial properties of Young tableaux, in terms of which some of our results (e.g. (17)) are already known.) We begin with some definitions and notation.

Let S_n be the symmetric group of permutations of the set $\{1, \ldots, n\}$ where for π in S_n we write, as usual,

$$\pi = (\pi_1, \ldots, \pi_n). \tag{1}$$

The integers $i_1, \ldots, i_m$ give an ascending sequence in π if

$$1 \leq i_1 < i_2 < \cdots < i_m \leq n; \qquad \pi(i_1) < \pi(i_2) < \cdots < \pi(i_n) \tag{2}$$

and then $l(\pi)$, the length of the longest ascending sequence in π, is the largest integer for which (2) holds. We refer to the 'gap' between π_{i-1} and π_i $(i = 2, \ldots, n)$ in (1) as the ith gap, the first gap coming before π_1 and the $(n+1)$th or last gap coming after π_n: many of our arguments depend on our producing elements of S_{n+1} out of those of S_n by inserting $(n + 1)$ into a suitable gap.

Let $P(n, m)$, $(1 \leq m \leq n)$, be the set of permutations in S_n with longest ascending sequence of length m and let $p(n, m)$ be the number of permutations in $P(n, m)$. Knowledge of the $p(n, m)$ would enable us to determine the constant c of Hammersley's result. Following Hammersley, for π given by (1), suppose that $l(\pi) = l$ and for $i = 1, \ldots, l$ define $a_i = a_i(\pi)$ to be the greatest integer j such that $(\pi_1, \ldots, \pi_j)$ has a longest ascending sequence of length i, setting $a_0 = 0$. Let $C(\boldsymbol{a})$ be the set of permutations in S_n for which

$$\boldsymbol{a} = \boldsymbol{a}(\pi) = (a_0, a_1, \ldots, a_l)$$

and let $c(\boldsymbol{a})$ be the number of permutations in $C(\boldsymbol{a})$. Hammersley noted a number of relations which hold between the $c(\boldsymbol{a})$ of which the most important are

$$p(n, m) = \sum_{\boldsymbol{a} \in A_{n,m}} c(\boldsymbol{a}), \tag{3}$$

where

$$A_{n,m} = \{\boldsymbol{a} : 0 = a_0 < a_1 < \cdots < a_m = n\}$$

and

$$c(a_0, a_1, \ldots, a_m, n+1) + c(a_0, a_1, \ldots, a_m, n+1) \\ = (n+1) c(a_0, a_1, \ldots, a_m, n). \tag{4}$$

We now continue this work by determining, first of all, some of the $c(\boldsymbol{a})$.

2. An instance of Pascal's triangle

As a preliminary example of our method of proof we establish the identity (c.f. Pascal's triangle)

$$c(0, r, r+1, \ldots, n) = c(0, r, r+1, \ldots, n-1) + c(0, r-1, r, \ldots, n-1) \tag{5}$$

which, since

$$c(0, n) = 1 = c(0, 1, 2, \ldots, n), \tag{6}$$

yields (c.f. [1, 16.16] where a more direct proof is given).

$$c(0, r, r+1, \ldots, n) = \binom{n-1}{r-1}. \tag{7}$$

To prove (5), we observe that any element of $C(0, r, r+1, \ldots, n)$ may be obtained by inserting n into either the last gap of an element of $C(0, r, r+1, \ldots, n-1)$ or the first gap of an element of $C(0, r-1, r, \ldots, n-1)$.

3. Ascending sequences of length two

The $c(0, r, n)$ also satisfy a family of relations. For example, we may produce a member of $C(0, 1, n)$ by inserting n in the second gap of an element of $C(0, n-1)$ or of $C(0, r, n-1)$ for $r = 1, \ldots, n-2$; and since all the elements of $C(0, 1, n)$ may be produced in this way we have the set identity

$$C(0,1,n)=\left\{\bigcup_{r=1}^{n-2} C(0,r,n-1)\right\}\cup C(0,n-1), \qquad n\geq 2$$

and so

$$c(0,1,n)=c(0,n-1)+\sum_{r=1}^{n-2} c(0,r,n-1), \qquad n\geq 2 \tag{8}$$

(where we make the convention now and hereafter that $\bigcup_{r=a}^{b}=\phi$, $\sum_{r=a}^{b}=0$ whenever $b<a$). Similar insertion arguments show that

$$c(0,s,n)=c(0,n-1)+\sum_{r=s-1}^{n-2} c(0,r,n-1), \qquad n\geq s+1\geq 3, \tag{9}$$

and hence, in particular, we have from (8, 9),

$$c(0,1,n)=c(0,2,n), \qquad n\geq 3. \tag{10}$$

We may bring (8, 9) into a 'Pascal' form, similar to (5), but with different 'boundary' conditions, by writing

$$c(0,r,n)=f(n-1,n-r), \qquad n-1\geq r\geq 1 \tag{11}$$

and we then obtain, after some manipulation, the equation

$$f(n,r)=f(n,r-1)+f(n-1,r), \qquad n\geq r\geq 1 \tag{12}$$

gether with, from (10),

$$f(n,n)=f(n,n-1), \qquad n\geq 1 \tag{13}$$

where we also take (c.f. (6))

$$f(n,0)=1, \qquad n\geq 0. \tag{14}$$

The problem is thereby seen to be equivalent to another well known problem, that of determining the number of walks, on the non-negative quadrant of the integral square lattice in two dimensional Euclidean space with the restriction $y\leq x$, from the origin to the point (n,r), $0\leq r\leq n$. This follows since (12, 13, 14) are just the equations governing the number in question. Hence, we find (see e.g. [6, pps. 26–27, 169–184]) that

$$f(n,r)=\left(1-\frac{r}{n+1}\right)\binom{n+r}{r}, \qquad n\geq r\geq 0. \tag{15}$$

(The $f(n,r)$ have an interpretation also in terms of permutations as the number of π in $P(n+2,2)$ such that, with π as in (1),

$$\pi_1=r+1; \qquad \pi_2=n+2,$$

so

$$c(0,1,n)=\sum_{r=0}^{n-2} f(n-2,r)=f(n-1,n-1), \qquad n\geq 2.)$$

Now, from (11, 15):

$$c(0, r, n) = \frac{r}{n}\binom{2n-(r+1)}{n-r} \tag{16}$$

and so from (3, 10), noting that $c(0, n) = 1$, $(n \geqslant 1)$:

$$\begin{aligned} p(n, 2) &= \sum_{r=1}^{n-1} c(0, r, n) = c(0, 1, n+1) - c(0, n) \\ &= \frac{1}{(n+1)}\binom{2n}{n} - 1 \end{aligned} \tag{17}$$

(which confirms (15, 9) of [1]).

4. Ascending sequences of length three and higher

The $c(\boldsymbol{a})$, for general $\boldsymbol{a}$, satisfy equations of the same kind as (8, 9) except that these equations, being now of greater complexity, are more difficult to write down than to prove. As an example we consider those for $c(0, t, s, n)$, $(s \geqslant t+1)$. It is not difficult to see, on the lines of (8), that, for $n-1 \geqslant s \geqslant 2$,

$$\begin{aligned} C(0, 1, s, n) = &\left\{ \bigcup_{s=1}^{s-2} C(0, r, s-1, n-1) \right\} \\ &\cup \left\{ \bigcup_{r=s}^{n-2} C(0, 1, r, n-1) \right\} \cup C(0, 1, n-1). \end{aligned}$$

For all the elements of $C(0, 1, s, n)$ are produced by inserting n into either the second gap of an element of $C(0, r, s-1, n-1)$, $r = 1, \ldots, s-2$ or the sth gap of an element of $C(0, 1, r, n-1)$, $r = s, \ldots, n-2$, or the s-th gap of an element of $C(0, 1, n-1)$. Hence

$$c(0, 1, s, n) = c(0, 1, n-1) + \sum_{r=1}^{s-2} c(0, r, s-1, n-1) + \sum_{r=s}^{n-2} c(0, 1, r, n-1),$$
$$n-1 \geqslant s \geqslant 2. \tag{18}$$

More generally, we may show that

$$c(0, t, s, n) = c(0, 1, n-1) + \sum_{r=t-1}^{s-2} c(0, r, s-1, n-1) + \sum_{r=s}^{n-2} c(0, t, r, n-1),$$
$$n-1 \geqslant s \geqslant t+1 \geqslant 3 \tag{19}$$

and equations (15), (16) are sufficient for us to prove inductively that

$$c(0, 1, r, n) = c(0, 2, r, n), \qquad n-1 \geqslant r \geqslant 3, \tag{20a}$$

$$c(0, r, r+1, n) = c(0, r, r+2, n), \qquad n-3 \geqslant r \geqslant 1. \tag{20b}$$

On the other hand (18, 19) have not, as yet, yielded closed expressions for $c(0,t,s,n)$ akin to (17). [Eqs. (20a, b) do, however, reveal an error in Table XIV of [1]: $c(0,1,7,9)$ should be 3751, the value correctly given for $c(0,2,7,9)$, and not 4168.]

The extension of such results to longer ascending sequences is straightforward and, for some $\boldsymbol{a}$, it is then easy to determine $c(\boldsymbol{a})$ explicitly. For example corresponding to (18, 19) we have, again by an insertion argument,

$$c(0,1,2,\ldots,n-2,n)=c(0,1,2,\ldots,n-2,n-1)+c(0,1,2,\ldots,n-3,n-1), \qquad n\geqslant 4$$

from which it follows, since

$$c(0,1,2,\ldots,n-1,n)=1, \qquad n\geqslant 2,$$

that

$$c(0,1,2,\ldots,n-2,n)=(n-1), \qquad n\geqslant 3.$$

Corresponding to (20) we find

$$c(\boldsymbol{a})=c(0,1,2,\ldots,n-2,n) \qquad \boldsymbol{a}\in A_{n,n-1},$$

and hence, combining these results with (3) we have (c.f. [1, 15.11])

$$p(n,n-1)=\sum_{\boldsymbol{a}\in A_{n,n-1}} c(\boldsymbol{a})=(n-1)(n-1), \qquad n\geqslant 2. \tag{21}$$

Similarly we may show that

$$c(0,3,4,\ldots,n)=\binom{n-1}{2}, \qquad n\geqslant 3,$$

$$c(0,1,2,\ldots,n-3,n)=2\binom{n-1}{2}-1, \qquad n\geqslant 4,$$

$$c(\boldsymbol{a})=c(0,1,2,\ldots,n-3,n), \quad \boldsymbol{a}\in A_{n,m-2}, \quad a_1<3, \qquad n\geqslant 4$$

and so (c.f. [1, 15.14])

$$\begin{aligned} p(n,n-2)&=\sum_{\boldsymbol{a}\in A_{n,n-2}} c(\boldsymbol{a})=c(0,3,4,\ldots,n)+\sum_{\substack{|\boldsymbol{a}|=n-2\\ a_1<3}} c(\boldsymbol{a})\\ &=\binom{n-1}{2}+\left[\binom{n-1}{2}-1\right]\left[2\binom{n-1}{2}-1\right], \qquad n\geqslant 3. \end{aligned} \tag{22}$$

Considering $\boldsymbol{a}$ in $A_{n,n-3}$ we obtain, after the same manner,

$$c(\boldsymbol{a})=c(0,1,2,\ldots,n-4,n), \quad \boldsymbol{a}\in A_{n,n-3}, \quad a_1=1, a_2<5, \qquad n\geqslant 5$$

$$c(\boldsymbol{a})=c(0,2,3,\ldots,n-3,n), \quad \boldsymbol{a}\in A_{n,n-3}, \quad a_1=2, a_2<5, \qquad n\geqslant 5$$

$$c(0,1,5,6,\ldots,n)=c(0,2,5,6,\ldots,n), \qquad n\geqslant 5$$

$$c(\boldsymbol{a})=c(0,3,4,\ldots,n-2,n), \quad \boldsymbol{a}\in A_{n,n-3}, \quad a_1=3, \qquad n\geqslant 5.$$

and, again, we may determine the $c(a)$ explicitly, if cumbersomely, for example

$$c(0,1,2,\ldots,n-4,n)=c(0,2,3,\ldots,n-3,n)=3!\binom{n-1}{3}-(3n-5),\qquad n\geq 5,$$

$$c(0,4,5,\ldots,(n-1),n)=\binom{n-1}{3},\qquad n\geq 4.$$

It follows, collecting the terms together as before using (3), that $p(n,n-3)$ is a polynomial of degree six in n (disproving (15.13) of [1]), the formula given there being one of degree seven in n).

5. Open problems

It remains a challenge to find more elegant formulae for the $c(a)$ and in particular to give an explicit solution for (18, 19).

Eq. (17), involving as it does the Catalan numbers C_n,

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

reveals an association with other well known combinatorial problems, as, for example, the walk problem already mentioned, in which these numbers also occur. The association is already known in the case of Young tableaux but a direct proof would be welcome as it might suggest other ways of calculating $c(a)$.

The permutations in S_n may be further restricted by taking account of longer descending sequences as well as longest ascending sequences: what then are the analogues of the present results?

References

[1] J.M. Hammersley, A Few Seedlings of Research. Proc. Sixth Berkeley Symp, Math. Statist. Prob. 1, (Univ. of California Press, 1972), pp. 345–394.

[2] J.F.C. Kingman, Subadditive Ergodic theory, Ann. Probab. 1 (1973) 883–909.

[3] B.F. Logan and L.A. Shepp, A variational problem for random Young tableaux (to appear).

[4] C Schensted, Longest increasing and decreasing subsequences, Can. J. Math. 13 (1961) 179–191.

[5] S.M. Ulam, Monte Carlo calculations in problems of mathematical physics, in: E.F. Beckenbach, ed., Modern Mathematics for the Engineer, Second Series (McGraw-Hill, New York, 1961).

[6] A.M. Yaglom and I.M. Yaglom, Challenging Problems in Mathematics with Elementary Solutions, Vol. 1 (Holden Day, San Francisco, 1964).