

YOUNG TABLEAUX AND LONGEST MONOTONE SUBSEQUENCES: AN INEQUALITY AND A CONJECTURE

by A. D. GORDON

(Received 9th March 1984, revised 11th June 1984)

Much work has been done on the following problem, which is sometimes referred to as Ulam’s problem: what is the distribution of $\alpha_N[\bar{\alpha}_N]$, the length (i.e. number of terms) of a longest monotone increasing [decreasing] subsequence of (not necessarily consecutive) terms in a random permutation of the first N integers? For example, it has been shown that $N^{-1/2}\alpha_N$ converges almost surely to 2 [6, 7, 9]. In some cases, it is important to know the value of $\beta_N(j)$, the number of permutations for which $\alpha_N=j$. Thus, Gordon [4] describes a test statistic for the comparison of two rankings of the same set of objects, whose null distribution is that of α_N over all $N!$ permutations; large values of α_N lead to rejection of the null hypothesis of no concordance between the rankings, hence it is relevant to obtain values, or inequalities, for

$$P_N(k) \equiv \sum_{j=0}^k \beta_N(N-j)$$

the number of permutations for which $\alpha_N \geq N-k$.

Relevant information about $\{\beta_N(j)\}$ can be obtained by noting a link between Ulam’s problem and Young tableaux. Let $\mathbf{i} \equiv (i_1, i_2, \dots, i_m)$, where $i_1 \geq i_2 \geq \dots \geq i_m$, be a partition of N . Consider a two dimensional array with m rows, having i_r cells in the r th row ($r=1, \dots, m$). A Young tableau of order N and shape \mathbf{i} is an arrangement of the first N integers in this array such that all rows and columns form increasing sequences. For each cell (r, s) in the array, the hook length h_{rs} is defined to be the number of cells (a, b) for which either (i) $a=r$ and $b \geq s$, or (ii) $a \geq r$ and $b=s$. Frame, Robinson and Thrall [3] proved that $\Phi_N(\mathbf{i})$, the number of Young tableaux of order N and shape \mathbf{i} is given by

$$\Phi_N(\mathbf{i}) = N! / \prod h_{rs} \tag{1}$$

where the product is taken over all cells in the tableau. As Greene, Nijenhuis and Wilf [5] remark, it is difficult to explain why the hook lengths h_{rs} appear; they do not seem to arise naturally in any combinatorial sense. An alternative expression for $\Phi_N(\mathbf{i})$ is conjectured in this paper.

Schensted [8] related Ulam’s problem to the problem of determining the number of Young tableaux of a given shape. A permutation of the first N integers can be mapped into a Young tableau. In general, the mapping is not 1–1: $\Phi_N(\mathbf{i})$ different permutations are mapped into a Young tableau of shape \mathbf{i} with the same set of cell entries; hence, $[\Phi_N(\mathbf{i})]^2$ permutations map into a Young tableau of shape \mathbf{i} . Further, $\alpha_N = i_1$, the number

of columns in the tableau, and $\bar{\alpha}_N = m$, the number of rows in the tableau. For small values of N , the shape \mathbf{i} will be uniquely determined by specifying i_1 and m , but for larger values of N several different shapes are compatible with specified values of i_1 and m . Thus, $\beta_N(N - j)$ can be expressed as

$$\beta_N(N - j) = \sum [\Phi_N(N - j, i_2, \dots, i_m)]^2 \tag{2}$$

where the summation is over all shapes of tableau for which

$$\sum_{r=2}^m i_r = j.$$

This result allows one to obtain explicit expressions, and Baer and Brock [1] evaluate $\{\beta_N(j)\}$ for $N = 4(1)36$, in this manner.

An upper bound for $\beta_N(N - j)$ can be obtained as follows: since the integer “1” must appear in cell (1, 1) of the tableau, there are at most $\binom{N-j-1}{j}$ ways of selecting the j integers which occupy cells in the sub-tableau comprising the second and higher rows of the tableau. Some of these choices will give cell entries which do not satisfy the condition that columns form increasing sequences, so the following inequality will hold:

$$\Phi_N(N - j, i_2, \dots, i_m) \leq \binom{N-j-1}{j} \Phi_j(i_2, \dots, i_m). \tag{3}$$

Hence, from (2) and (3),

$$\begin{aligned} \beta_N(N - j) &\leq \binom{N-j-1}{j}^2 \sum [\Phi_j(i_2, \dots, i_m)]^2 \\ &= \binom{N-j-1}{j}^2 j! \end{aligned} \tag{4}$$

since the summation is over all possible tableaux of order j .

This yields an upper bound for $P_N(k)/N!$ which is markedly sharper than the one presented in [4], but which still becomes blunter as N increases.

An examination of the joint distribution of α_N and $\bar{\alpha}_N$ suggested an alternative expression to (1) for $\Phi_N(N - j, i_2, \dots, i_m)$; this is given explicitly in Table 1 for values of $j \leq 5$, in which γ_r denotes $\binom{N-r-1}{r}$. It can be seen that the leading term in the expression is the one given on the right hand side of inequality (3). For $\alpha_N + \bar{\alpha}_N < N + 1$, the “correction factor”, to be subtracted from the leading term, has the following noteworthy features:

- (i) It is always of the form $\sum_{r=2}^j a_{j-r} \gamma_{j-r}$.
- (ii) The non-zero elements of the set of coefficients alternate in sign, with $a_{j-2} > 0$.
- (iii) $\sum_{r=2}^j |a_{j-r}|$ is equal to the coefficient of γ_r .

It has been verified that the expressions given in Table 1 are correct; it is conjectured

Table 1. Values of $\Phi_N(N-j, i_2, \dots, i_m)$ for $j=0, 1, \dots, 5$; γ_r denotes $\binom{N-r}{r}$.

	$\alpha_N = N-5$	$\alpha_N = N-4$	$\alpha_N = N-3$	$\alpha_N = N-2$	$\alpha_N = N-1$	$\alpha_N = N$
$\bar{\alpha}_N = 1$						$\Phi_N(N)$ $\gamma_0 - \gamma_{-2}$
$\bar{\alpha}_N = 2$	$\Phi_N(N-5, 5)$ $\gamma_5 - \gamma_3$	$\Phi_N(N-4, 4)$ $\gamma_4 - \gamma_2$	$\Phi_N(N-3, 3)$ $\gamma_3 - \gamma_1$	$\Phi_N(N-2, 2)$ $\gamma_2 - \gamma_0$	$\Phi_N(N-1, 1)$ $\gamma_1 - \gamma_{-1}$	
$\bar{\alpha}_N = 3$	$\Phi_N(N-5, 4, 1)$ $4\gamma_5 - 3\gamma_3 + \gamma_2$	$\Phi_N(N-4, 3, 1)$ $3\gamma_4 - 2\gamma_2 + \gamma_1$	-----			
	$\Phi_N(N-5, 3, 2)$ $5\gamma_5 - 3\gamma_3 + 2\gamma_2$	$\Phi_N(N-4, 2, 2)$ $2\gamma_4 - \gamma_2 + \gamma_1$	$\Phi_N(N-3, 2, 1)$ $2\gamma_3 - \gamma_1 + \gamma_0$	$\Phi_N(N-2, 1, 1)$ $\gamma_2 + \gamma_{-1}$		
$\bar{\alpha}_N = 4$	$\Phi_N(N-5, 3, 1, 1)$ $6\gamma_5 - 3\gamma_3 + 2\gamma_2 - \gamma_1$	-----				
	$\Phi_N(N-5, 2, 2, 1)$ $5\gamma_5 - 2\gamma_3 + 2\gamma_2 - \gamma_1$	$\Phi_N(N-4, 2, 1, 1)$ $3\gamma_4 - \gamma_2 + \gamma_1 - \gamma_0$	$\Phi_N(N-3, 1, 1, 1)$ $\gamma_3 - \gamma_{-1}$			
$\bar{\alpha}_N = 5$	$\Phi_N(N-5, 2, 1, 1, 1)$ $4\gamma_5 - \gamma_3 + \gamma_2 - \gamma_1 + \gamma_0$	$\Phi_N(N-4, 1, 1, 1, 1)$ $\gamma_4 + \gamma_{-1}$				
$\bar{\alpha}_N = 6$	$\Phi_N(N-5, 1, 1, 1, 1, 1)$ $\gamma_5 - \gamma_{-1}$					

that properties (i)–(iii) are true generally for all $\Phi_N(i)$ for which $\alpha_N + \bar{\alpha}_N < N + 1$. This general result has not been proved, but the following partial proof lends support to the conjecture.

Defining e_r to be a vector whose elements are all zero apart from the r th which is one, we have [2, Chapter 2]:

$$\Phi_N(i) = \sum_{r=1}^m \Phi_{N-1}(i - e_r).$$

Hence, with $i' \equiv (i_2, i_3, \dots, i_m)$,

$$\Phi_N(N-j, i') - \Phi_{N-1}((N-1)-j, i') = \sum_{r=2}^m \Phi_{N-1}((N-1)-(j-1), i' - e_r). \tag{5}$$

The two terms on the left hand side of equation (5) are of the same form, with j elements in the second and subsequent rows of the tableau, whereas all terms on the right hand side have $(j-1)$ elements in the second and subsequent rows. This suggests that, having verified the results for the rightmost few columns of Table 1, one can use induction to obtain an expression for $\Phi_N(N-j, i')$ from expressions for $\Phi_N(N-j+1, i' - e_r)$. The combinatorial identity

$$\binom{N-1}{j} - \binom{N-2}{j} = \binom{N-2}{j-1} \tag{6}$$

is very helpful in this respect. Obtaining expressions for $\Phi_N(i)$ recursively in this manner preserves the properties (i)–(iii) noted above.

As an illustration, consider deriving an expression for $\Phi_N(N-4, 3, 1)$ when all terms to the right of it in Table 1 are known. With $M \equiv N-1$,

$$\begin{aligned}\Phi_N(N-4, 3, 1) - \Phi_M(M-4, 3, 1) &= \Phi_M(M-3, 2, 1) + \Phi_M(M-3, 3) \\ &= \{2\binom{M-1}{3} - \binom{M-1}{1} + \binom{M-1}{0}\} + \{\binom{M-1}{3} - \binom{M-1}{1}\} \\ &= 3\binom{M-1}{3} - 2\binom{M-1}{1} + \binom{M-1}{0}\end{aligned}$$

Using identity (6), one obtains the solution

$$\Phi_N(N-4, 3, 1) = 3\binom{N-1}{4} - 2\binom{N-1}{2} + \binom{N-1}{1} + \text{constant}.$$

Provided that one defines the zero correction factors for the diagonal entries (for which $\alpha_N + \bar{\alpha}_N = N+1$) as shown in Table 1, and includes them in the algebra, the constant factor is zero for all expressions contained in the table. It has not been proved that this result is generally true, but it is conjectured that this is the case. If the result is true, it would suggest that there is further structure yet to be discovered in Young tableaux; it might also be possible to use symbolic manipulation programs to obtain terms outside the range of Table 1, which could then be used in the numerical evaluation of $\Phi_N(i)$.

Acknowledgements. I am grateful to B. M. Brown and R. M. Cormack for helpful discussions, and to a referee for constructive criticism of an earlier version of the paper.

REFERENCES

1. R. M. BAER and P. BROCK, Natural sorting over permutation spaces, *Math. Comp.* **22** (1968), 385–410.
2. C. BERGE, *Principles of Combinatorics* (Academic Press, New York and London, 1971).
3. J. S. FRAME, G. DE B. ROBINSON and R. M. THRALL, The hook graphs of the symmetric group, *Can. J. Math.* **6** (1954), 316–324.
4. A. D. GORDON, A measure of the agreement between rankings, *Biometrika* **66** (1979), 7–15.
5. C. GREENE, A. NIJENHUIS and H. S. WILF, A probabilistic proof of a formula for the number of Young tableaux of a given shape, *Adv. in Math.* **31** (1979), 104–109.
6. J. M. HAMMERSLEY, A few seedlings of research, *Proc. 6th Berkeley Symp.* **1** (1972), 345–394.
7. H. KESTEN, Contribution to discussion of a paper by J. F. C. Kingman, *Ann. Prob.* **1** (1973), 903.
8. C. SCHENSTED, Longest increasing and decreasing subsequences, *Can. J. Math.* **13** (1961), 179–191.
9. A. M. VERSHIK and S. V. KEROV, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables, *Sov. Math. Dokl.* **18** (1977), 527–531 (*Dokl. Akad. Nauk SSSR* **233** (1977), 1024–1027).

DEPARTMENT OF STATISTICS
UNIVERSITY OF ST. ANDREWS