High-water marks

Michi

Given a sequence of values $a_k{}_{k=1}^n$, the high-water marks are the values at which the running maximum increases. For example, given a sequence $[3, 5, 7, 8, 8, 5, 7, 9, 2, 5]$ with running maxima $[3, 5, 7, 8, 8, 8, 9, 9, 9]$, the high-water marks are $[3, 5, 7, 8, 9]$, which occur at $k = (1, 2, 3, 4, 8)$. For every sequence a_k there is a number of high-water marks N_{a_k}.

If now we consider a set σ of all permutations of n numbers $(1, \ldots, n)$ what is the analytical expression for the distribution of the number of high-water marks ($N_{a\in\sigma}$)? For example, if $n = 3$, the distribution vector is $\left(\frac{2}{3!}, \frac{3}{3!}, \frac{1}{3!}\right)$.

To solve this problem consider the cycle representation of a permutation (including the trivial one-element cycles). We have a bijection of the symmetric group with itself that maps between numbers of cycles and numbers of high-water marks. Thus the distribution of the number of high-water marks is the distribution of the number of cycles. This is given by the unsigned Stirling numbers of the first kind, which has a distribution

$$P_s\left(k; n\right) = \frac{|s\left(n; k\right)|}{n!}, \tag{1}$$

where

$$\begin{cases} s\left(n; 1\right) = (n-1)! \\ s\left(n; n-1\right) = \frac{n(n-1)}{2} \\ s\left(n; n\right) = 1 \end{cases} \tag{2}$$

Proof of the bijection

Let's proof the bijection of the symmetric group with itself that maps between numbers of cycles and numbers of high-water marks.

- Let $r\left(n, i\right)$ be the number of permutations of n elements with running maximum equal to i.
- Let $c\left(n, i\right)$ be the number of permutations of n elements whose cycle type has i parts.

We make two claims for all n and $i \leq n$:

$$\begin{cases} r\left(n, i\right) = r\left(n-1, i-1\right) + r\left(n-1, i\right)\left(n-1\right) \\ c\left(n, i\right) = c\left(n-1, i-1\right) + c\left(n-1, i\right)\left(n-1\right) \end{cases} \tag{3}$$

Since we know that $r\left(3, i\right) = c\left(3, i\right)$ for all $i \leq 3$ and $r\left(n, 1\right) = c\left(n, 1\right) = (n-1)!$, it follows from the two claims that $r\left(n, i\right) = c\left(n, i\right)$ holds for all n and all $i \leq n$.

We prove the first claim. Consider a permutation x of the elements $1, 2, \ldots, n-1$ and write x as a list. We can insert a new element 0 into x to obtain a permutation y of the elements $0, 1, 2, \ldots, n-1$. Going over all x and inserting 0 at each of the n possible positions gives all permutations of $0, 1, 2, \ldots, n-1$ exactly once. For example, let $n = 3$. Let $x = 12$. Then the possible y's we can obtain are $y = 012$, $y = 102$, $y = 120$. Let $x = 21$. Then the possible y's we can obtain are $y = 021$, $y = 201$, $y = 210$. Each y appears once and together they give all permutations of 3 elements. Suppose y has running maximum i. Then there are two options. Either x has running maximum $i-1$ and 0 was inserted at the first index. Or x has running maximum i, and 0 was inserted not at the first index. There are $r\left(n-1, i-1\right)$ possible x for the first option. There are $r(n-1,i)$ possible x and $(n-1)$ possible locations for 0 for the second option. Therefore, $r(n,i) = r(n-1,i-1) + r(n-1,i)(n-1)$.

We prove the second claim. Consider a permutation x of the elements $1, 2, \ldots, n-1$ and write x as a product of disjoint cycles. We can insert 0 in one of the existing cycles of x or as a new cycle to obtain a permutation y of $0, 1, 2, \ldots, n-1$. Going over all x and inserting 0 in every possible way gives all permutations of $0, 1, 2, \ldots, n-1$ exactly once. For example, let $n = 3$. Let $x = (1)(2)$. Then the possible y's we can obtain are $y = (0)(1)(2)$, $y = (01)(2)$, $y = (1)(02)$. Let $x = (12)$. Then the possible y's we can obtain are $y = (0)(12)$, $y = (012)$, $y = (102)$. Each y appears once and together they give all permutations of 3 elements. Suppose y has a cycle type with i parts. Then there are two options. Either x has a cycle type with $i-1$ parts and the new cycle (0) was inserted. Or x has a cycle type with i parts and 0 was inserted in one of the existing cycles. There are $c(n-1,i-1)$ possible x for the first option. There are $c(n-1,i)$ possible x and $(n-1)$ possible possible ways to insert 0 in the existing cycles for the second option. Therefore, $c(n,i) = c(n-1,i-1) + c(n-1,i)(n-1)$.