Journal of Combinatorial Theory, Series A 116 (2009) 1154–1166

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series A

www.elsevier.com/locate/jcta

Asymptotic behavior of permutation records

Igor Kortchemski

École Normale Supérieure, 75005 Paris, France

ARTICLE INFO

Article history:
Received 14 June 2008
Available online 8 April 2009

Keywords:
Records
Stirling numbers
Scaled asymptotic behavior
Independent random variables

ABSTRACT

We study the asymptotic behavior of two statistics defined on the symmetric group $\mathfrak{S}_n$ when n tends to infinity: the number of elements of $\mathfrak{S}_n$ having k records, and the number of elements of $\mathfrak{S}_n$ for which the sum of the positions of their records is k. We use a probabilistic argument to show that the scaled asymptotic behavior of these statistics can be described by remarkably simple functions.

1. Introduction

In this paper we will study records, also known in the literature as *left to right maxima*, *outstanding elements* or *strong records*. By definition, a *record* of a permutation $\sigma = a_1 \ldots a_n \in \mathfrak{S}_n$ is a number a_j such that $a_i < a_j$ for all $i < j$. The study of the records has been initiated by Rényi who proved that the number of elements of $\mathfrak{S}_n$ with exactly m cycles in their cycle decomposition is equal to the number of elements of $\mathfrak{S}_n$ having exactly m records, the latter being given by $c(n,m)$, the unsigned Stirling number of the first kind [1]. The asymptotic behavior of these numbers has first been studied by Jordan [2] who showed that for fixed m and large n, $c(n,m)/(n-1)! \sim (\ln n + \gamma)^{m-1}/(m-1)!$, with γ being the Euler constant. Moser and Wyman [3] considered three overlapping regions of the $(n,m)-$plane (for $n \geqslant m$) and obtained asymptotic formulae in each case. Wilf [4] gave an explicit asymptotic expansion of $c(n,m)$ with m fixed in terms of powers of n and $\ln n$. Finally, Temme [5] obtained an asymptotic formula for $c(n,m)$ in the limit $n \to \infty$, uniformly for $1 \leqslant m \leqslant n$ (including the case $m \sim n$). Hwang [6] established an explicit asymptotic expansion of $c(n,m)$ valid for $m = \mathcal{O}(\ln n)$. Later, Chelluri, Richmond and Temme [7] introduced the generalized Stirling numbers of the first kind and studied their asymptotic behavior. In particular, they re-derived Temme's previous result and showed that their results agree with those of Moser and Wyman.

Asymptotic properties of record statistics have been studied by Wilf who obtained the following results [8]:

E-mail address: igor.kortchemski@ens.fr.

0097-3165/$ – see front matter
doi:10.1016/j.jcta.2009.03.003

(i) for fixed r the average value of the rth records, over all permutations of $\mathfrak{S}_n$ that have that many, is asymptotic to $(\ln n)^{r-1}/(r-1)!$ when $n \to +\infty$,
(ii) the average value of a permutation σ at its rth record, among all permutations that have that many, is asymptotic to $(1-1/2^r)n$ when $n \to +\infty$,
(iii) Let $1 < j_1 < j_2 < \cdots < j_m$ be fixed integers, and suppose that we attach a symbol $s(j) =$ 'Y' or 'N' to each of these j's. Then the probability P that a permutation of $\mathfrak{S}_n$ does have a record at each of the j_v that is marked 'Y', and does not have a record at any of those that are marked 'N', is:

$$P = \prod_{s(j_v)=\text{'N'}} \left(1 - \frac{1}{j_v}\right) \prod_{s(j_v)=\text{'Y'}} \frac{1}{j_v}. \tag{1}$$

Myers and Wilf generalized the notion of records [13]. They studied strong and weak records defined on multiset permutations and words (a *weak record* of a word $w_1, \ldots, w_n$ is a term w_j such that $w_i \leqslant w_j$ for all $i < j$ [10]). For multiset permutations, they derived the generating function for the number of permutations of a fixed multiset M which contain exactly r strong (respectively weak) records. They also obtained the generating function of the probability that a randomly selected permutation of M has exactly r strong records. This gives the average number of strong (respectively weak) records among all permutations of M.

The notion of records has been extended to random variables (for a survey of some results see [14]). The asymptotic behavior of the two statistics, 'position' and 'value' of the rth record, have been studied in [9–11] for a sequence of i.i.d random variables which follow the geometric law of parameter p, which approach the model of random permutations in the limit $p \to 0$. For other interesting results concerning record statistics see [13].

In this article, we re-derive Wilf's result (1) by using of a *probabilistic argument*. This will allow us to perform a study of the asymptotic behavior of another statistic of records: the number of permutations of length n having k records in the limit $n \to +\infty$ with the ratio k/n fixed. In this limit the *re-scaled* number of records k/n takes values on the interval $[0,1]$. We also introduce the new statistic for a permutation called 'sum of the positions of its records'. We find the asymptotic behavior of the number of permutations of length n for which the sum of the positions of their records is k in the limit $n \to +\infty$ with the ratio $k/(\frac{n(n+1)}{2})$ fixed. More precisely, we show the following results:

(I) Let $c(n,k)$ be the number of permutations of length n having k records. Its generating function is given by $q(q+1)\cdots(q+n-1)$, so that $c(n,k)$ is the coefficient of q^k in the power expansion. For $n \geqslant 1$ and $x \in [0,1]$ define the function f_n by:

$$f_n(x) = \begin{cases} c(n,[nx]) & \text{if } x \geqslant \frac{1}{n}, \\ c(n,1) & \text{otherwise,} \end{cases} \tag{2}$$

where $[x]$ stands for the integer part of x. Then, when n tends to infinity, the sequence of functions $\{\frac{\ln(f_n)}{n\ln(n)}; n \in \mathbb{N}^+\}$ converges uniformly with respect to x on the interval $[0,1]$ to the function $x \mapsto 1-x$ with an accuracy $\mathcal{O}(\frac{1}{\ln n})$. In other words, there exists a constant C such that for all integer $n \geqslant 2$:

$$\sup_{x\in[0,1]} \left| \frac{\ln(f_n(x))}{n\ln n} - (1-x) \right| \leqslant \frac{C}{\ln n}.$$

(II) Let $\mathcal{C}(n,k)$ be the number of permutations of length n for which the sum of the positions of their records is k. Its generating function is given by $q(q^2+1)(q^3+2)\cdots(q^n+n-1)$, so that $\mathcal{C}(n,k)$ is the coefficient of q^k in the power expansion. For $n \geqslant 1$ and $x \in [0,1]$ define the function ϕ_n by[1]:

$$\phi_n(x) = \begin{cases} \mathcal{C}(n,1) = (n-1)! & \text{if } x < \frac{6}{n(n+1)}, \\ \mathcal{C}(n, \frac{n(n+1)}{2}) = 1 & \text{if } x \geqslant 1 - \frac{2}{n(n+1)}, \\ \mathcal{C}(n, [\frac{n(n+1)}{2}x]) & \text{otherwise.} \end{cases} \tag{3}$$

[1] This definition is motivated by the fact that $\mathcal{C}(n,k) = 0$ if and only if $k = 2$ or $k = \frac{n(n+1)}{2} - 1$.

Then, when n tends to infinity, the sequence of functions $\{\frac{\ln(\phi_n)}{n\ln(n)}; n \in \mathbb{N}^+\}$ converges uniformly with respect to x on the interval $[0,1]$ to the function $x \mapsto \sqrt{1-x}$ with an accuracy $\mathcal{O}(\frac{1}{\ln n})$. In other words, there exists a constant $\widetilde{C}$ such that for all integer n:

$$\sup_{x\in[0,1]}\left|\frac{\ln(\phi_n(x))}{n\ln n} - \sqrt{1-x}\right| \leqslant \frac{\widetilde{C}}{\ln n}.$$

It is important to note that the functions f_n and ϕ_n are defined on the same interval $[0,1]$.

The paper is organized as follows. In Section 2 we derive some results concerning records which will be used in the rest of the paper. In Section 3 we prove assertion (I) and in Section 4 we show assertion (II). In Appendix A we prove that (I) is consistent with Temme's result [5] previously mentioned.

2. Notations and preliminary results

We endow the symmetric group $\mathfrak{S}_n$ on a set of n elements with the uniform law. We begin with some useful definitions.

Definition 2.1. Let $\sigma = a_1 \dots a_n \in \mathfrak{S}_n$. Recall that a *record* of σ is a number a_j such that $a_i < a_j$ for all $i < j$. We define $\mathrm{rec}(\sigma)$ as the number of records of σ. The generating function of this statistic is:

$$T_n(q) = \sum_{\sigma\in\mathfrak{S}_n} q^{\mathrm{rec}(\sigma)}.$$

Likewise we define $\mathrm{srec}(\sigma)$ as the sum of the positions of all records of σ. The generating function of this statistic:

$$P_n(q) = \sum_{\sigma\in\mathfrak{S}_n} q^{\mathrm{srec}(\sigma)}.$$

Let $X_k(\sigma)$ be the random variable which equals 1 if k is a position of a record of σ and 0 otherwise.

Example 2.2. The permutation of length 8, $\sigma = 4,7,5,1,6,8,2,3$ (i.e. σ sends 1 to 4, 2 to 7, etc.) has 3 records: 4, 7 and 8 so that $\mathrm{rec}(\sigma) = 3$ and $\mathrm{srec}(\sigma) = 1+2+6 = 9$.

Our work relies on the following proposition, first proved by Rényi [1].

Proposition 2.3. *The random variables $X_1, X_2, \dots, X_n$ are independent. Moreover, $\mathbb{P}(X_k = 1) = \frac{1}{k}$.*

Proof. This proposition is a consequence of Proposition 1.3.9 and Corollary 1.3.10 of [15,16]. It comes from the following remark. For a permutation $\sigma = a_1 \dots a_n$ and an integer n such that $1 \leqslant i \leqslant n$, define:

$$r_i(\sigma) = \mathrm{Card}\{j\colon\ j < i,\ a_j > a_i\}.$$

Then the mapping which sends a permutation $\sigma = a_1 \dots a_n$ on the n-tuple $(r_1(\sigma), \dots, r_n(\sigma))$ is as a bijection between $\mathfrak{S}_n$ and the n-tuples $(r_1, \dots, r_n)$ such that $0 \leqslant r_i \leqslant i-1$ for all i. Furthermore, $r_i(\sigma) = 0$ if, and only if, a_j is a record of σ. □

Example 2.4. By this bijection, the image of 4, 7, 5, 1, 6, 8, 2, 3 is $\{0,0,1,3,1,0,5,5\}$.

Definition 2.5. Define $\mathcal{C}(n,k)$ as the number of elements of $\mathfrak{S}_n$ for which the sum of the positions of their records is k.

Using Proposition 2.3 one can show the following results.

Proposition 2.6. *The generating functions of the statistics* 'rec' *and* 'srec' *are*:

$$T_n(q) = \sum_{\sigma \in \mathfrak{S}_n} q^{\mathrm{rec}(\sigma)} = q(q+1)\cdots(q+n-1) = \sum_{k=0}^{n} c(n,k)q^k, \tag{4}$$

$$P_n(q) = \sum_{\sigma \in \mathfrak{S}_n} q^{\mathrm{srec}(\sigma)} = q\big(q^2+1\big)\big(q^3+2\big)\cdots\big(q^n+n-1\big) = \sum_{k=1}^{\frac{n(n+1)}{2}} \mathcal{C}(n,k)q^k. \tag{5}$$

The generating function (4) is well known [16].

3. Asymptotic behavior of the coefficients $c(n,k)$

Let us first examine the asymptotic behavior of $c(n,k)$ when n tends to infinity and k/n is fixed. In this limit, it is not obvious that the coefficients $c(n,k)$ have a well defined asymptotic behavior. It is convenient to introduce the new scaling variable $x = k/n$ which takes values in the interval $[0,1]$. We introduce a new function f_n as:

$$f_n(x) = c(n, nx). \tag{6}$$

Note that when x is of the form k/n the two relations (2) and (6) coincide. For a fixed integer n, the variable x takes rational values. Moreover, in the limit $n \to \infty$ they run through a dense subset of $[0,1]$. One may wonder whether the function $f_n(x)$ is well defined in this limit. And if so, what is the limit function?

For this purpose we extend the function f_n to the whole interval $[0,1]$ as in Eq. (2). We now state the theorem which describes the asymptotic behavior of $c(n,k)$.

Theorem 3.1. *The sequence of functions* $\{\frac{\ln(f_n)}{n\ln(n)}; n \in \mathbb{N}^+\}$ *converges uniformly when n tends to infinity with respect to x on the interval* $[0,1]$ *to the function* $x \mapsto 1-x$ *with an accuracy* $\mathcal{O}(\frac{1}{\ln n})$. *In other words, there exists a constant C such that for all integer n:*

$$\sup_{x\in[0,1]} \left| \frac{\ln(f_n(x))}{n\ln(n)} - (1-x) \right| \leqslant \frac{C}{\ln n}. \tag{7}$$

The proof relies on a combinatorial and probabilistic interpretation of the coefficients $c(n,k)$. In the rest of this section we consider n to be an integer.

Lemma 3.2. *Let* $\mathbb{P}(\mathrm{rec}=k) = \frac{c(n,k)}{n!}$. *Then:*

$$\mathbb{P}(rec = k) = \sum_{\substack{v_1<v_2<\cdots<v_k\leqslant n \\ v_1=1}} \frac{1}{v_1 v_2 \cdots v_k}\left(1-\frac{1}{v_{k+1}}\right)\cdots\left(1-\frac{1}{v_n}\right) \tag{8}$$

under the additional conditions $v_{k+1} < \cdots < v_n \leqslant n$ *and* $v_i \neq v_j$ *for* $i \neq j$.

Proof. Choosing a permutation with k records is equivalent to choosing the positions of its k records as $1 = v_1 < v_2 < \cdots < v_k \leqslant n$. By Proposition 2.3, a position v_i is chosen as a record position with probability $\frac{1}{v_i}$ and is not chosen as a record position with probability $1-\frac{1}{v_i}$. Furthermore this choice is independent for each position. This yields the result of the lemma. □

This gives a probabilistic proof of Wilf's result (1) mentioned in the Introduction.

Lemma 3.3. *Let $x \in [\frac{1}{n}, 1]$ and $k = [nx]$. The following double inequality holds:*

$$\frac{(n-[nx])!}{n(n!)} \leqslant \mathbb{P}(rec = k) \leqslant 2^n \frac{1}{[nx]!}.$$

Proof. This is a consequence of Lemma 3.2 after taking into account the properties:

(i) the number of k-tuples $(v_1, \ldots, v_k)$ such that $v_1 < v_2 < \cdots < v_k \leqslant n$, $v_1 = 1$, $v_i \neq v_j$ for $i \neq j$ and $v_{k+1} < \cdots < v_n \leqslant n$ is less than 2^n,

(ii) we have:

$$\frac{1}{n} = \left(1 - \frac{1}{2}\right) \cdots \left(1 - \frac{1}{n}\right) \leqslant \left(1 - \frac{1}{v_{k+1}}\right) \cdots \left(1 - \frac{1}{v_n}\right) \leqslant 1,$$

(iii) and for integers $1 = v_1 < v_2 < \cdots < v_k \leqslant n$:

$$k! \leqslant v_1 v_2 \cdots v_k \leqslant \frac{n!}{(n-k)!}. \qquad \square$$

All the essential ingredients have now been gathered, and we turn to the proof of Theorem 3.1.

Proof of Theorem 3.1. Using the same variables as in Lemma 3.3 and the function f_n defined in (2), one obtains:

$$\frac{\ln((n-[nx])!)}{n \ln n} - \frac{1}{n} \leqslant \frac{\ln(f_n(x))}{n \ln(n)} \leqslant \frac{\ln 2}{\ln n} + \frac{\ln n!}{n \ln n} - \frac{\ln([nx]!)}{n \ln n}. \tag{9}$$

Define $M_n(x) = \ln n |\frac{\ln(f_n(x))}{n \ln(n)} - (1-x)|$. Then:

$$\sup_{x \in [0,1]} M_n(x) \leqslant \sup_{x \in [0, \frac{1}{n}]} M_n(x) + \sup_{x \in [\frac{1}{n}, 1]} M_n(x).$$

Using $c(n,1) = (n-1)!$ (see Eq. (4)) and Lemma 3.3 one obtains:

$$\sup_{x \in [0,1]} M_n(x) \leqslant \sup_{x \in [0, \frac{1}{n}]} \ln n \left| \frac{\ln(n-1)!}{n \ln n} - (1-x) \right| + \sup_{x \in [\frac{1}{n}, 1]} \ln n \left| \frac{\ln(f_n(x))}{n \ln(n)} - (1-x) \right|.$$

Stirling's formula ($\ln n! = n \ln n + \mathcal{O}(n)$) shows that the first term is bounded when n tends to infinity. Now denote B_n the second term of the right-hand side of the above inequality. By Eq. (9) there exist constants C_1 and C_2 such that:

$$B_n \leqslant C_1 + \sup_{x \in [\frac{1}{n}, 1]} \left| \frac{\ln n!}{n} - \frac{\ln([nx]!)}{n} - (1-x) \ln n \right| + \sup_{x \in [\frac{1}{n}, 1]} \left| \frac{\ln((n-[nx])!)}{n} - (1-x) \ln n \right| \leqslant C_2,$$

where in the second inequality we used Stirling's formula. This concludes the proof. $\square$

4. Asymptotics of the coefficients $\mathcal{C}(n,k)$

We now investigate the asymptotic behavior of the coefficients $\mathcal{C}(n,k)$ (see Definition 2.5) for large n. When n is a fixed integer, the coefficients $\mathcal{C}(n,k)$ (with $1 \leqslant k \leqslant \frac{n(n+1)}{2}$) are positive integers. As before, we are interested in their asymptotic behavior in the limit $n \to +\infty$ with the ratio $k/(\frac{n(n+1)}{2})$ fixed. To this end we introduce the scaling variable $x = 2k/(n(n+1))$ which takes values in the interval $[0,1]$ as well as the function ϕ_n:

$$\phi_n(x) = \mathcal{C}\left(n, \frac{n(n+1)}{2} x\right). \tag{10}$$

Then, the problem reduces to finding the asymptotic behavior of $\phi_n(x)$ in the limit $n \to \infty$. For this purpose we extend the function ϕ_n to the whole interval $[0,1]$ as in Eq. (3). We now state the main theorem which describes the asymptotic behavior of the coefficients $\mathcal{C}(n,k)$.

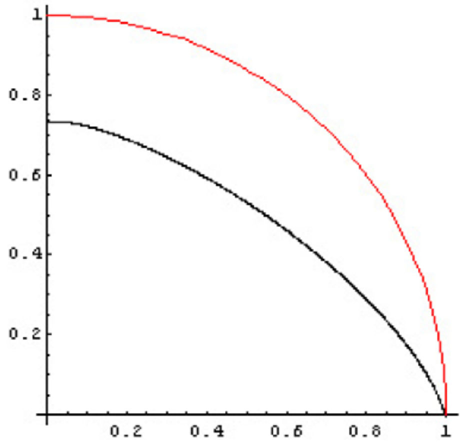

Fig. 1. Graph of ψ_{50} and $\sqrt{1-x}$.

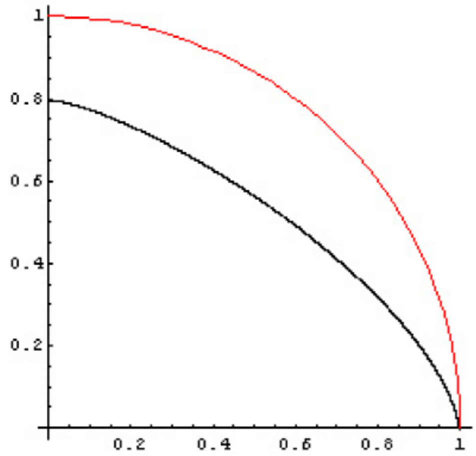

Fig. 2. Graph of ψ_{150} and $\sqrt{1-x}$.

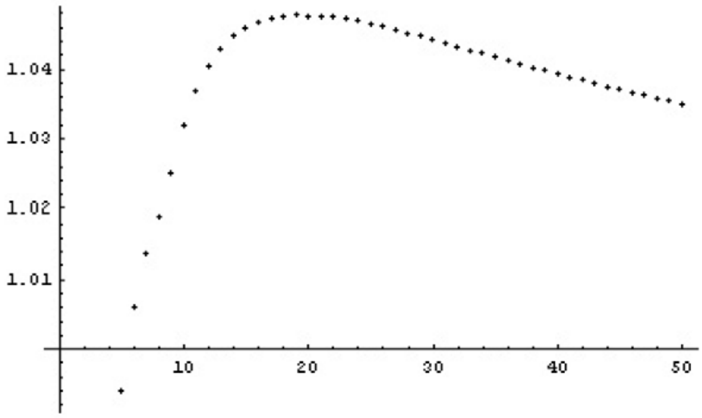

Fig. 3. Graph of the function τ for $n = 2, \ldots, 50$.

Theorem 4.1. *The sequence of functions* $\{\frac{\ln(\phi_n)}{n\ln(n)}; n \in \mathbb{N}^+\}$ *converges uniformly when n tends to infinity with respect to x on the interval* $[0,1]$ *to the function* $x \mapsto \sqrt{1-x}$ *with an accuracy* $\mathcal{O}(\frac{1}{\ln n})$*. In other words, there exists a constant* $\widetilde{C}$ *such that for all integer n*:

$$\sup_{x\in[0,1]}\left|\frac{\ln(\phi_n(x))}{n\ln(n)} - \sqrt{1-x}\right| \leqslant \frac{\widetilde{C}}{\ln n}.$$

To illustrate this theorem we plot the two functions $x \mapsto \sqrt{1-x}$ and $\psi_n = \frac{\ln(\phi_n)}{n\ln(n)}$ for different values of n (for $n = 50$ in Fig. 1 and for $n = 150$ in Fig. 2) obtained with `Mathematica`. In agreement with Theorem 4.1, ψ_n converges to the function $\sqrt{1-x}$. We also plot the function:

$$\tau:\quad n \longmapsto \ln n \cdot \sup_{x\in[0,1]}\left|\frac{\ln(\phi_n(x))}{n\ln(n)} - \sqrt{1-x}\right|$$

in Fig. 3 for $n = 2, 3, \ldots, 50$. In agreement with Theorem 4.1, this function is bounded by a constant $\widetilde{C}$.

The proof of this theorem is based on a combinatorial and probabilistic interpretation of the coefficients $\mathcal{C}(n,k)$ which gives us Lemma 4.2.

Lemma 4.2. *Let* $\mathbb{P}(\mathrm{srec}=k)=\frac{\mathcal{C}(n,k)}{n!}$. *Then*:

$$\mathbb{P}(\mathrm{srec}=k)=\sum_{\substack{v_1+v_2+\cdots+v_r=k\\ v_1=1,\,r\leqslant n}}\frac{1}{v_1v_2\cdots v_r}\left(1-\frac{1}{v_{r+1}}\right)\cdots\left(1-\frac{1}{v_n}\right) \tag{11}$$

under the additional conditions $v_1<v_2<\cdots<v_r\leqslant n$, $v_{r+1}<\cdots<v_n\leqslant n$ *and* $v_i\neq v_j$ *for* $i\neq j$.

Proof. The proof is similar to the one of Lemma 3.2. Choosing an element of $\mathfrak{S}_n$ for which the sum of the positions of its records is k, is choosing the position of its records $1=v_1<v_2<\cdots<v_r\leqslant n$ such that $v_1+v_2+\cdots+v_r=k$. By Proposition 2.3, a position v_i is chosen as a record position with probability $\frac{1}{v_i}$ and is not chosen as a record position with probability $1-\frac{1}{v_i}$. Furthermore this choice is independent for each position. This yields the result of the lemma. □

Definition 4.3. We say that the r-tuple $(v_1,\ldots,v_r)$ satisfies the conditions $(C_{k,n})$ if:

$$(C_{k,n}):\quad v_1=1,\ r\leqslant n,\ v_1<v_2<\cdots<v_r\leqslant n \text{ and } v_1+v_2+\cdots+v_r=k. \tag{12}$$

Note that there exists an r-tuple $(v_1,\ldots,v_r)$ satisfying the conditions $(C_{k,n})$ if and only if $\mathcal{C}(n,k)>0$, that is $k\neq 2$ and $k\neq\frac{n(n+1)}{2}-1$. We now isolate the greatest term in the sum of formula (11). This motivates the following definition.

Definition 4.4. Let k,n be integers such that $1\leqslant k\leqslant\frac{n(n+1)}{2}$, $k\neq 2$ and $k\neq\frac{n(n+1)}{2}-1$. Define:

$$m(n,k)=\min_{\substack{r\leqslant n,\,(v_1,\ldots,v_r)\\ \text{satisfies } (C_{k,n})}} v_1v_2\cdots v_r. \tag{13}$$

This minimum gives us an inequality satisfied by the coefficients:

Proposition 4.5. *We have*:

$$\frac{1}{nm(n,k)}\leqslant\mathbb{P}(\mathrm{srec}=k)\leqslant\frac{2^n}{m(n,k)}. \tag{14}$$

Proof. Note that the total number of r-tuples $(1\leqslant r\leqslant n)$ satisfying the conditions $(C_{k,n})$ is less then 2^n, so that:

$$\frac{1}{nm(n,k)}=\frac{1}{m(n,k)}\left(1-\frac{1}{2}\right)\cdots\left(1-\frac{1}{n}\right)\leqslant\mathbb{P}(\mathrm{srec}=k)\leqslant\frac{2^n}{m(n,k)}.$$

This concludes the proof. □

We will have to study three cases: $3\leqslant k\leqslant n$, $n\leqslant k<\frac{n(n-1)}{2}$ and $\frac{n(n-1)}{2}\leqslant k\leqslant\frac{n(n+1)}{2}$. For each case, we will find either the expression for the minimum $m(n,k)$ defined in (13) or a lower and upper bound for $m(n,k)$. These expressions will be useful in finding the asymptotic behavior. In the following, n will be considered as an integer greater than 3 and k as an integer.

4.1. Case $3\leqslant k\leqslant n$

This case is the easiest one as shows the following lemma.

Proposition 4.6. *Let* $3\leqslant k\leqslant n$. *Then*:

$$\frac{n!}{(k-1)n}\leqslant\mathcal{C}(n,k)\leqslant\frac{2^n n!}{k-1}. \tag{15}$$

Proof. One can verify that the minimum $m(n,k)$, defined in (13), is realized by the r-tuple $(v_1,\dots,v_r)$ for $r=2$, when $v_1=1$ and $v_2=k-1$. In this case, $m(n,k)=k-1$, and Proposition 4.5 yields the result. □

It is important to note that when $n+1 \leqslant k$ this argument cannot be applied anymore. We would like to take $v_2=k-1$, this is impossible since the conditions (12) require $v_2 \leqslant n$.

4.2. *Case* $n+1 \leqslant k < \frac{n(n-1)}{2}$

Let k be an integer such that $n+1 \leqslant k < \frac{n(n-1)}{2}$. In this case, one expects the minimum $m(n,k)$ to be realized when most of the records are at the last positions:

Lemma 4.7. *Let $(v_1,\dots,v_r)$ be an r-tuple which realizes the minimum $m(n,k)$. Let $i_0=i_0(n,k)$ be the greatest integer such that*:

$$k-1 \geqslant n+(n-1)+\cdots+(n-i_0). \tag{16}$$

Then for $0 \leqslant i \leqslant i_0(n,k)$, $n-i$ is equal to one of the v_j.

Proof. For the sake of contradiction assume that i is the smallest integer such that $n-i$ does not appear in $v_1,\dots,v_r$. Then either $1<v_2<n-i$, or

$$v_1+v_2+\cdots+v_r = 1+(n-i+1)+\cdots+(n-1)+n \leqslant 1+(k-1)-(n-i) < k,$$

which contradicts the conditions (12). So let $j \geqslant 2$ be the greatest integer such that $v_j<n-i$. The desired contradiction will arise if we find an r-tuple such that the product of its elements will be less than $m(n,k)$, therefore contradicting the minimality of $m(n,k)$. A few cases have to be studied. It is important to remember that the v_i have to be all different.

Case 1: $v_j \neq n-i-1$. If there exists $l>1$ such that $1<v_l \leqslant (n-i)-v_j$, we delete v_l and v_j from our r-tuple and replace them by v_l+v_j. This contradicts the minimality of $m(n,k)$ since $v_lv_j>v_l+v_j$. Moreover, if $j=2$ then:

$$1+v_j+(n-i+1)+\cdots+n < 1+(n-i)+(n-i+1)+\cdots+n \leqslant k,$$

which contradicts (12). Hence $1<v_2<v_j$ and $v_2>(n-i)-v_j \geqslant 2$, implying $v_2-1 \geqslant 2$. Replace v_2 by v_2-1 and v_j by v_j+1. This contradicts the minimality of $m(n,k)$ since $v_2v_j>(v_2-1)(v_j+1)$.

Case 2: $v_j=n-i-1$. If $j=2$, as before:

$$\begin{aligned}1+(n-i-1)+(n-i+1)+\cdots+n &\leqslant (n-i-1)+k-(n-i)-(n-i-1)-\cdots-(n-i_0)\\ &< k,\end{aligned}$$

which contradicts (12). Thus $1<v_2<v_j$.

Subcase 2.1: $v_2>2$. Remplace v_2 by v_2-1 and v_j by v_j+1 to get a contradiction since $v_2v_j>(v_2-1)(v_j+1)$.

Subcase 2.2: $v_2=2$. Since $k<\frac{n(n-1)}{2}$ (this inequality is crucial here), there exists an integer $1<u<v_j$ such that u is not one of the v_i. Consider the greatest integer l such that $v_l<u$. The minimality of $m(n,k)$ is finally contradicted by replacing in the r-tuple $\{2,v_l,v_j\}$ by $\{v_l+1,v_j+1\}$ since $2v_lv_j>(v_l+1)(v_j+1)$. □

Lemma 4.8. *Following the notations of Lemma* 4.7, *we have*:

$$\frac{\Gamma(n+1)}{\Gamma(n-i_0)} \leqslant m(n,k) \leqslant \frac{\Gamma(n+1)}{\Gamma(n-i_0)}e^n, \tag{17}$$

where $\Gamma(n)$ is Euler's Gamma function.

Proof. By Lemma 4.7, it is clear that $\frac{\Gamma(n+1)}{\Gamma(n-i_0)} \leqslant m(n,k)$. For the second inequality, let j be the greatest integer such that $v_j < n - i_0$. By definition of i_0:

$$k - 1 < n + (n-1) + \cdots + (n - i_0) + (n - i_0 - 1).$$

Since $(v_1, \ldots, v_r)$ satisfies (12):

$$1 + \sum_{i=2}^{j} v_i + (n - i_0) + (n - i_0 + 1) + \cdots + (n-1) + n = k.$$

Hence:

$$\sum_{i=2}^{j} v_i < n - i_0 - 1 \leqslant n.$$

The arithmetic–geometric mean inequality and a brief study of the function $x \mapsto (\frac{n}{x})^x$ yield:

$$\prod_{i=2}^{j} v_i \leqslant \left(\frac{n}{j-1}\right)^{j-1} \leqslant e^n, \tag{18}$$

implying the result of the lemma. □

Proposition 4.9. *For $n+1 \leqslant k < \frac{n(n-1)}{2}$ we have*:

$$\frac{1}{ne^n}\Gamma(n - i_0) \leqslant \mathcal{C}(n,k) \leqslant 2^n \Gamma(n - i_0). \tag{19}$$

Proof. Combine Proposition 4.5 and Lemma 4.8, and use the fact that $\mathcal{C}(n,k) = \mathbb{P}(\text{srec} = k)/n!$. □

4.3. Case $\frac{n(n-1)}{2} \leqslant k \leqslant \frac{n(n+1)}{2}$

Let k be an integer such that $\frac{n(n-1)}{2} \leqslant k \leqslant \frac{n(n+1)}{2}$ and $k \neq \frac{n(n+1)}{2} - 1$. Let us prove that the result of Proposition 4.9 holds in this case too (note that we will not use Lemma 4.7).

Lemma 4.10. *Let $(v_1, \ldots, v_r)$ be an r-tuple which realizes the minimum $m(n,k)$ defined in* (13)*. Let $i_0 = i_0(n,k)$ be the greatest integer such that*:

$$k - 1 \geqslant n + (n-1) + \cdots + (n - i_0).$$

Then:

$$\frac{\Gamma(n+1)}{\Gamma(n-i_0)} \leqslant m(n,k) \leqslant \frac{\Gamma(n+1)}{\Gamma(n-i_0)} e^n, \tag{20}$$

where Γ is Euler's Gamma function.

Proof. It goes along the same lines as the proof of Lemma 4.7. If, for $0 \leqslant i \leqslant i_0$, $n - i$ is equal to one of the v_j, then we continue as in the previous subsection. If not, let i be the smallest integer such that $n - i$ does not appear in $v_1, \ldots, v_r$. We can also assume that we are in Subcase 2.2 of the proof of Lemma 4.7 (if not, we obtain a contradiction as in the proof of Lemma 4.7). Consequently, $v_2 = 2$, there exists an integer j such that $v_j < n - i$ and if $1 < u < v_j$ then u is one of the v_i. In other words, all positions but $n - i$ are records. For convenience, we introduce $u = \frac{n(n+1)}{2} - k$, where $k = v_1 + v_2 + \cdots + v_r$, so that $3 \leqslant u \leqslant n$. Thus:

$$u = \frac{n(n+1)}{2} - k = n - i \quad \text{and} \quad m(n,k) = \frac{n!}{n-i} = \frac{n!}{u}.$$

By definition of i_0, Eq. (16), one gets:

$$i_0(n,k) = \left[\frac{2n-1-\sqrt{4n^2+4n-8k+9}}{2}\right] = \left[n - \frac{1}{2} - \sqrt{2u+\frac{9}{4}}\right]. \tag{21}$$

Let us prove the first inequality in (20). First note that it is equivalent to $u \leqslant \Gamma(n-i_0)$. For $x>0$ define $E(x)$ to be x if $x \in \mathbb{N}$ and $[x]+1$ otherwise, so that $n-[n-x]=E(x)$. In virtue of (21), the inequality $u \leqslant \Gamma(n-i_0)$ is equivalent to:

$$u \leqslant \Gamma\left(E\left(\frac{1}{2}+\sqrt{2u+\frac{9}{4}}\right)\right).$$

This inequality is verified for $u=3$ and for all integers $u \geqslant 4$ one has:

$$u \leqslant \Gamma\left(\frac{1}{2}+\sqrt{2u+\frac{9}{4}}\right),$$

so that the first inequality is proved.

Let us now prove the second inequality in (20). For all integer u such that $3 \leqslant u \leqslant n$, we have:

$$\Gamma(n-i_0) \leqslant \Gamma\left(\left[\frac{1}{2}+\sqrt{2u+\frac{9}{4}}\right]+1\right) \leqslant ue^u \leqslant ue^n,$$

where the first inequality follows from the properties of the Γ function and the second one from the fact that $u \leqslant n$. □

Using Proposition 4.5 and Lemma 4.10 we finally extend Proposition 4.9 to:

Proposition 4.11. *For an integer k such that* $n+1 \leqslant k \leqslant \frac{n(n+1)}{2}$ *and* $k \neq \frac{n(n+1)}{2}-1$ *we have*:

$$\frac{1}{ne^n}\Gamma(n-i_0) \leqslant \mathcal{C}(n,k) \leqslant 2^n\Gamma(n-i_0). \tag{22}$$

4.4. Proof of the main theorem

Let n be an integer such that $n \geqslant 4$ and $x \in [0,1]$. Define $k=k(n,x)=[x\frac{n(n+1)}{2}]$. Note that $3 \leqslant k \leqslant n$ if and only if $\frac{6}{n(n+1)} \leqslant x < \frac{2}{n}$. When $x \geqslant \frac{2}{n}$, define $i_0(n,x) := i_0(n,k(n,x))$ as in Eq. (16).

Lemma 4.12. *The following inequality holds*:

$$\forall n \in \mathbb{N},\ n \geqslant 4,\ \forall x \in \left[\frac{2}{n}, 1-\frac{2}{n(n+1)}\right], \quad \left|n-i_0(n,x)-n\sqrt{1-x}\right| \leqslant 3.$$

Proof. It can be deduced from Eq. (21) by using the fact that $k=[x\frac{n(n+1)}{2}]$. □

Proof of Theorem 4.1. Define $K_n(x) = \ln n\left|\frac{\ln(\phi_n(x))}{n\ln(n)} - \sqrt{1-x}\right|$ (see Eq. (3) for the definition of ϕ_n). Then:

$$\sup_{x\in[0,1]} K_n(x) \leqslant \sup_{x\in[0,\frac{6}{n(n+1)}]} K_n(x) + \sup_{x\in[\frac{6}{n(n+1)},\frac{2}{n})} K_n(x) + \sup_{x\in[\frac{2}{n},1]} K_n(x).$$

Denote the first term on the right-hand side of this inequality as A_n, the second one as B_n and the third one as C_n. We prove that they are all bounded by a constant independent of n.

(i) We have:

$$A_n = \sup_{x\in[0,\frac{6}{n(n+1)}]} \ln n\left|\frac{\ln(n-1)!}{n\ln n} - \sqrt{1-x}\right| \leqslant \widetilde{C}_1$$

by Stirling's formula.

(ii) By Proposition 4.6 we have:

$$\begin{aligned}
B_n &\leqslant \sup_{x\in[\frac{6}{n(n+1)},\frac{2}{n})}\left|\frac{\ln(n!)-\ln([x\frac{n(n+1)}{2}]-1)-\ln n}{n}-\sqrt{1-x}\ln n\right| \\
&\quad+\sup_{x\in[\frac{6}{n(n+1)},\frac{2}{n})}\left|\frac{n\ln(2)+\ln(n!)-\ln([x\frac{n(n+1)}{2}]-1)}{n}-\sqrt{1-x}\ln(n)\right| \\
&\leqslant \widetilde{C}_2+2\sup_{x\in[\frac{6}{n(n+1)},\frac{2}{n})}\left|\frac{\ln(n!)}{n}-\sqrt{1-x}\ln(n)\right| \\
&\leqslant \widetilde{C}_3.
\end{aligned}$$

(iii) By Lemma 4.12 for $\frac{2}{n}\leqslant x\leqslant 1-(\frac{5}{n})^2$, one has $n\sqrt{1-x}\geqslant 5$ so that $n-i_0(n,x)\geqslant n\sqrt{1-x}-3$ and $n\sqrt{1-x}-3\geqslant 2$. For $1-(\frac{5}{n})^2\leqslant x<1-\frac{2}{n(n+1)}$ one has $n-i_0(n,x)\geqslant 2$. Note that for x such that $x\geqslant 1-\frac{2}{n(n+1)}$ one has $\phi_n(x)=1$. Hence by Proposition 4.11:

$$\begin{aligned}
C_n &\leqslant \sup_{x\in[\frac{2}{n},1]}\left|\ln 2+\frac{\ln(\Gamma(n\sqrt{1-x}+3))}{n}-\sqrt{1-x}\ln n\right| \\
&\quad+\sup_{x\in[\frac{2}{n},1-(\frac{5}{n})^2]}\left|-\frac{\ln n}{n}-1+\frac{\ln(\Gamma(n\sqrt{1-x}-3))}{n}-\sqrt{1-x}\ln n\right| \\
&\quad+\sup_{x\in[1-(\frac{5}{n})^2,1-\frac{2}{n(n+1)})}\left|-\frac{\ln n}{n}-1-\sqrt{1-x}\ln n\right|+\sup_{x\in[1-\frac{2}{n(n+1)},1]}|\sqrt{1-x}\ln n| \\
&\leqslant \widetilde{C}_4,
\end{aligned}$$

and this concludes the proof. □

5. Conclusion

We have studied the asymptotic behavior of the integers $c(n,k)$ (respectively $\mathcal{C}(n,k)$) equal to the number of elements of $\mathfrak{S}_n$ having k records (respectively for which the sum of the positions of their records are k) by using a probabilistic argument. One can note that these integers can be defined outside of any combinatorial background since $c(n,k)$ appears as the coefficient of q^k in the polynomial $q(q+1)\cdots(q+n-1)$ and $\mathcal{C}(n,k)$ appears as the coefficient of q^k in the polynomial $q(q^2+1)(q^3+2)\cdots(q^n+n-1)$. Thus studying the asymptotic behavior of these numbers seems delicate, but the probabilistic interpretation gave us a convenient formula defining these integers. Surprisingly, the scaled asymptotic behavior of these rather complicated numbers can be described by a remarkably simple function.

Note added

Recently, based on the result obtained in the present paper, the statistic 'sum of the position of records' has also been considered in the case of the geometric law in [12].

Acknowledgments

I am deeply indebted to Richard Stanley for his helpful advice as well as to Philippe Biane for useful discussions and for sharing the idea which lead to the proof of the main theorem. I thank Herbert S. Wilf for useful comments and for pointing out additional references.

Appendix A

In this appendix we show that our Theorem 3.1 is consistent with Temme's result [5] (see also [7]). Let m, n be positive integers such that $m \leqslant n$. Define:

$$\phi(u) = \ln\big((u+1)(u+2)\cdots(u+n)\big) - m\ln u. \tag{23}$$

Let u_1 be the unique positive solution of the equation $\phi'(u) = 0$ (see [5] for the proof that u_1 is unique). Let $t_1 = m/(n-m)$ and $B = \phi(u_1) - n\ln(1+t_1) + m\ln t_1$. Finally let $g(t_1) = u_1^{-1}[m(n-m)/(n\phi''(u_1))]^{1/2}$.

Theorem A.1 *(Temme). The relation*

$$c(n,m) \sim e^B g(t_1)\binom{n}{m} \tag{24}$$

holds uniformly for $1 \leqslant m \leqslant n$ *in the limit* $n \to \infty$.

We show that our Theorem 3.1 is a consequence of Theorem A.1. In other words, we deduce from Temme's formula the fact that the coefficients $c(n,m)$ have a scaled asymptotic behavior in the limit $n \to \infty$ with the ratio m/n fixed, which is not clear a priori.

Proposition A.2. *Let x be a real number such that* $0 < x < 1$. *Using the previous definitions with* $m = [nx]$, *the following relation holds*:

$$\lim_{n\to\infty} \frac{\ln(e^B g(t_1)\binom{n}{[nx]})}{n\ln n} = 1 - x. \tag{25}$$

To prove this, the following lemma will be useful. It shows that u_1/n has a nice behavior for large n.

Lemma A.3. *For n sufficiently large we have*:

$$\frac{x^2}{6(4/3-x)} \leqslant \frac{u_1}{n} \leqslant \frac{x}{1-x} + \frac{1}{n}. \tag{26}$$

Proof. Let $f(u) = 1/(u+1) + 1/(u+2) + \cdots + 1/(u+n)$. Recall that u_1 satisfies the relation

$$\phi'(u_1) = f(u_1) - \frac{[nx]}{u_1} = 0. \tag{27}$$

For the second inequality in (26), note that:

$$\frac{n}{u_1+n} \leqslant \frac{1}{u_1+1} + \frac{1}{u_1+2} + \cdots + \frac{1}{u_1+n} = \frac{[nx]}{u_1} \leqslant \frac{nx+1}{u_1},$$

so that $u_1 \leqslant (n(x+1/n))/(1-(x+1/n)) \leqslant nx/(1-x) + 1$, where the last inequality holds for n sufficiently large.

For the first inequality in (26), let $\alpha = x^2/(6(4/3-x))$ so that $x/2 + (1-x/2)\alpha/(\alpha + x/2) = 3x/4$ and $p = [xn/2]$. Since the function $u \mapsto uf(u)$ is increasing for positive u, it is sufficient to show that $f(\alpha n) \leqslant (nx-1)/(\alpha n)$. Then:

$$\begin{aligned} f(\alpha n) &= \frac{1}{\alpha n+1} + \frac{1}{\alpha n+2} + \cdots + \frac{1}{\alpha n+n} \\ &\leqslant \frac{p}{\alpha n} + \frac{n-p}{\alpha n+p} \\ &\leqslant \frac{xn/2}{\alpha n} + \frac{n-(xn/2-1)}{\alpha n + xn/2 - 1} \end{aligned}$$

$$\leqslant \frac{x}{2\alpha} + \frac{1 - x/2 - 1/n}{\alpha + x/2 - 1/n} = \frac{3x}{4\alpha} - \frac{1 - x/2}{\alpha + x/2} + \frac{1 - x/2 - 1/n}{\alpha + x/2 - 1/n}$$
$$\leqslant \frac{x}{\alpha} - \frac{1}{n\alpha}, \tag{28}$$

where the last inequality holds for n sufficiently large. □

Proof of Proposition 5.2. Let us show that only the term $\phi(u_1)$ entering the expression for B provides the dominant contribution on the right-hand side of Eq. (25). More precisely:

- Using Stirling's formula, it is easy to see that $\lim_{n\to\infty} \frac{\ln\binom{n}{[nx]}}{n\ln n} = 0$.
- To show that $\lim_{n\to\infty} \frac{\ln g(t_1)}{n\ln n} = 0$, it is sufficient to show that $\lim_{n\to\infty} \frac{\ln \phi''(u_1)}{n\ln n} = 0$. It is convenient to write $\phi''(u_1)$ as $\phi''(u_1) = \psi'(u_1+n+1) - \psi'(u_1+1) + [nx]/u_1^2$, where ψ is the psi function. Lemma A.3 and the fact that $\psi'(x) \sim 1/x$ for large x (see e.g. [17]) give the result.
- It is clear that $\lim_{n\to\infty} \frac{\ln|-n\ln(1+t_1)+[nx]\ln t_1|}{n\ln n} = 0$.
- Finally, observe that $\phi(u_1) = \ln\Gamma(u_1+n+1) - \ln\Gamma(u_1+1) - [nx]\ln u_1$. Stirling's formula and the fact that $\ln(u_1+n+1) = \ln(u_1+1) + o(\ln n)$ (which is a consequence of Lemma A.3) show that $\lim_{n\to\infty} \frac{\phi(u_1)}{n\ln n} = 1 - x$. This concludes the proof. □

Thus we have reproduced the scaled asymptotic behavior of $c(n,[nx])$ using Temme's result. However, it seems difficult to also reproduce by this means the error estimate stated in Theorem 3.1. To this end, it would be necessary to give a more precise asymptotic behavior of u_1.

References

[1] A. Rényi, Théorie des éléments saillants d'une suite d'observations, Ann. Fac. Sci. Univ. Clermont-Ferrand 8 (1962) 7–13.
[2] C. Jordan, The Calculus of Finite Differences, second ed., Chelsea, 1947.
[3] L. Moser, M. Wyman, Asymptotic development of the Stirling numbers of the first kind, J. London Math. Soc. 33 (1958) 133–146.
[4] H.S. Wilf, The asymptotic behavior of the Stirling numbers of the first kind, J. Combin. Theory Ser. A 64 (1993) 344–349.
[5] N.M. Temme, Asymptotic estimates of Stirling numbers, Stud. Appl. Math. 89 (1993) 233–243.
[6] H. Hwang, Asymptotic expansions for the Stirling numbers of the first kind, J. Combin. Theory Ser. A 71 (2) (1995) 343–351.
[7] R. Chelluri, L.B. Richmond, N.M. Temme, Asymptotic estimates for generalized Stirling numbers, Analysis 20 (2000) 1–13.
[8] H.S. Wilf, On the outstanding elements of permutations, http://www.math.upenn.edu/~wilf/website/outstelmts.pdf, 1995.
[9] H. Prodinger, Combinatorics of geometrically distributed random variables: Left-to-right maxima, Discrete Math. 153 (1996) 253–270.
[10] A. Knopfmacher, H. Prodinger, Combinatorics of geometrically distributed random variables: Value and position of the rth left-to-right maximum, Discrete Math. 226 (2001) 255–267.
[11] H. Prodinger, Combinatorics of geometrically distributed random variables: Value and position of large left-to-right maxima, Discrete Math. 254 (2002) 459–471.
[12] H. Prodinger, Records in geometrically distributed words: Sum of positions, Appl. Anal. Discrete Math. 2 (2008) 234–240.
[13] A.N. Myers, H.S. Wilf, Left-to-right maxima in words and multiset permutations, Israel J. Math. 166 (2008) 167–183.
[14] N. Glick, Breaking records and breaking boards, Amer. Math. Monthly 85 (1978) 2–26.
[15] R.P. Stanley, Private communication, unpublished.
[16] R.P. Stanley, Enumerative Combinatorics, vol. 1, Cambridge University Press, Cambridge, England, 1999.
[17] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs and Mathematics Tables, National Bureau of Standards Appl. Ser., vol. 55, US Government Printing Office, Washington, DC, 1964.