ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Journal of Statistical Planning and Inference ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Keeping scores

David Collins, Tamás Lengyel*, Gregory P. Tollisen

Math. Dept., Occidental College, 1600 Campus Road, Los Angeles, CA90041, USA

Received 1 July 2006; received in revised form 2 July 2006; accepted 6 July 2006

Abstract

Consider a sports competition in which participants alternately perform a scored task (such as the distance a discus is thrown), and a list of the top m scores is updated throughout. We consider the average number and distribution of records among the top m throughout and at the end of the competition. We also answer questions concerning the number of times the list changes, when each change occurs, and the waiting times between changes. We touch on results concerning changes in l-records and the values within the list.

Keywords: Record positions and values; Distribution of records; Order statistics

1. Introduction

Playing video games one day, one of the authors, David Collins, was prompted to enter his name upon achieving a high score. Rather than his name, he entered "best" when the score was the highest yet attained up to that point and "other" if the score fell within the top $m=5$ but was not a record. He observed that the average number of "best" scores retained in the top m seemed to remain fixed over time. In fact, it was equal to the harmonic number $H_5=1+\frac{1}{2}+\cdots+\frac{1}{5}$. Of course, a devoted sports fan has an infinite appetite for records, however, overly refined, such as "second best" scores, etc., and needs no excuse for hairsplitting statistics, e.g., frequency of changes among the top m. Motivated by this observation, we investigate problems related to records and score keeping regimes.

We will use the following assumptions. The competition consists of a total of n turns. At all times, only the m highest scores are retained. Throughout the paper we assume that the scores are independent and identically distributed random variables with some continuous distribution. Thus, all the n scores are different and we do not consider the possibility of improvement in scoring with time. It is remarkable that only the record values are dependent on the score distribution and not the record times, inter-record times, or the number of records (cf. Arnold et al., 1998). A score qualifies to be "best" if her score, at the time she achieves it, is higher than any other score. In Theorem 1 we determine the distribution of the number of "best" scores when the top m scores are stored and kept updated. Complementary results and generalization are considered in Theorem 2 and in Section 5.

Records and their various generalizations have been extensively studied in the literature, e.g., in Glick (1978), Arnold et al. (1998), and Nevzorov (2001), often by applying advanced theory. Here we derive Theorem 5 about the first moment of the distribution of the so called mth record positions by a purely combinatorial argument.

* Corresponding author. Tel.: +1 323 259 2516; fax: +1 323 259 2704.
E-mail address: lengyel@oxy.edu (T. Lengyel).

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.07.006

2. The static approach: at the end

The average number of "best" scores within the top m scores is the harmonic number $H_m = \sum_{i=1}^{m} 1/i, m \geqslant 1$. In order to see this we denote the sequence of observations by $\{X_1, X_2, \ldots, X_n\}$ and define Y_n to be the number of "best" scores in the top m after n games. In fact, $Y_n = \sum_{i=1}^{m} V_i$ with indicator variable $V_i = 1$ exactly when the ith observation to enter the top m entries overall was observed before any greater value. More formally, let $t_1 < t_2 < \cdots < t_m$ be the positions of the top m entries $X_{t_1}, X_{t_2}, \ldots, X_{t_m}$ (i.e., $X_{t_i} > X_l$ for all $l \neq t_i$ $i = 1, 2, \ldots, m$). Clearly, $P(V_i = 1) = P(X_{t_i}$ is " best" at time of entry$) = 1/i$ and thus, $EY_n = \sum_{i=1}^{m} P(V_i = 1) = H_m$. This idea can be developed further to include the distribution of Y_n, as we will see in Theorem 1.

Let $s(m, k)$ denote the unsigned Stirling number of the first kind with parameters m and k, i.e., the number of permutations over an m-element set with exactly k cycles. We have the somewhat surprising:

Theorem 1. *Let Y_n denote the number of "best" scores among the top m after n games. The random variable $Y_n, n \geqslant m$, has a stationary distribution. In fact, $P(Y_n = j) = s(m, j)/m!$, $j = 1, 2, \ldots, m$.*

Proof. Regardless of $n, n \geqslant m$, we keep and update the top m scores. Therefore, at the end we have the m top scores overall. Now in fact, only the ranks and positions of *these* scores relative to each other matter, for any other score among the n, being smaller, has no impact on the record status (i.e., "best" vs. "other") of the m. Thus, it is enough to count the number of permutations on the set $\{1, 2, \ldots, m\}$ with j records. By a standard combinatorial argument based on the cycle-representation of permutations, the number of permutations with j records is equal to the number of permutations with j cycles, as observed by Rényi, cf. Blom et al. (1994), Comtet (1974), and Lovász (1979). The result follows. □

3. Successive observations

We are also interested in modeling the possible changes here by a Markov chain with nonstationary transition probabilities (often referred to as a nonhomogenous Markov chain). As above, let Y_n denote the number of "best" scores among the top m after n games. Clearly, if $m = 1$ then $Y_n = 1$ for all $n \geqslant 1$.

For the combined probabilities we get:

Theorem 2. *For $m \geqslant 2$, we have*

$$r_{j,j+a}^{(n)} = P(Y_{n+1} = j + a \cap Y_n = j) = \begin{cases} \dfrac{s(m-1, j)}{(m-2)!} \dfrac{1}{m(n+1)} & \text{if } a = 1 \text{ and } 1 \leqslant j \leqslant m-1, \\ \dfrac{s(m-1, j-1)}{(m-2)!} \dfrac{1}{m(n+1)} & \text{if } a = -1 \text{ and } 2 \leqslant j \leqslant m, \\ P(Y_n = j) - r_{j,j-1}^{(n)} - r_{j,j+1}^{(n)} & \text{if } a = 0 \text{ and } 2 \leqslant j \leqslant m-1, \\ \dfrac{n}{m(n+1)} & \text{if } a = 0 \text{ and } j = 1, \\ \dfrac{n-m+2}{m!(n+1)} & \text{if } a = 0 \text{ and } j = m. \end{cases}$$

Proof. To determine $P(Y_{n+1} = j + 1 \cap Y_n = j)$ we consider all permutations with j records over m (i.e., an m-element set) in which the smallest record is not the mth in rank (in fact, not of the lowest rank with respect to the overall 2nd, 3rd, $\ldots, m+1$th in the set of the $n+1$ elements): $s(m, j) - s(m-1, j-1) = (m-1)s(m-1, j)$, thus $P(Y_{n+1} = j + 1 \cap Y_n = j) = \frac{(m-1)s(m-1,j)}{m!} \frac{1}{n+1}$, for the $n+1$th element is of the highest rank among the $n+1$ elements.

To determine $P(Y_{n+1} = j - 1 \cap Y_n = j)$ we consider all permutations with j records over m with the smallest record being the rank m element which is exceeded by the $n+1$th element, therefore, the probability is $\frac{s(m-1,j-1)}{m!} \frac{m-1}{n+1}$.

All other cases can be derived in a similar direct fashion (or by Theorem 1). □

Please cite this article as: David Collins et al., Keeping scores, Journal of Statistical Planning and Inference (2006), doi: 10.1016/j.jspi.2006.07.006

We note some consequences of this theorem. Surprisingly, $P(Y_{n+1}=j+1\cap Y_n=j)=P(Y_{n+1}=j\cap Y_n=j+1)$ holds for all $j:1\leqslant j\leqslant m-1$ and $n\geqslant 1$. Since Y_n has the stationary distribution given in Theorem 1, we get $p_{j,j+1}=P(Y_{n+1}=j+1|Y_n=j)=\frac{(m-1)s(m-1,j)}{s(m,j)}\frac{1}{n+1}$, $p_{j+1,j}=P(Y_{n+1}=j\mid Y_n=j+1)=\frac{(m-1)s(m-1,j)}{s(m,j+1)}\frac{1}{n+1}$, and $p_{j,j}=P(Y_{n+1}=j\mid Y_n=j)=P(Y_n=j\mid Y_{n+1}=j)$ for the transition probabilities. With a little work, these time dependent probabilities also lead to the stationary distribution of Y_n found in Theorem 1. We also note that the above properties establish that the Markov chain is *reversible*.

One might be interested in calculating the expected number of new games until a change in the number Y_n of "best" scores occurs. By Theorem 2 we get that the expected waiting times are infinite. For example, given that $Y_n=1$ for some $n\geqslant m$, if it takes exactly $W=j$ new games to have two "best" scores in the top m then the corresponding probability of this happening is $\frac{n}{n+1}\frac{n+1}{n+2}\cdots\frac{n+j-2}{n+j-1}(1-\frac{n+j-1}{n+j})=\frac{n}{(n+j-1)(n+j)}$, and the conditional expected value of W is infinite. The proof is similar for $Y_n=j$ with $j:2\leqslant j\leqslant m$, yielding the infiniteness of the expected waiting time.

4. Relaxed version

In many athletic events, in track and field for instance, the top $m=3$ scores are displayed, rather than just the current record. This has the advantage of making the event more interesting to spectators, for there will be considerably more changes on the leader board than if only the top score ($m=1$) is updated. In fact, we will see that the expected "waiting times" for changes among the top m scores are finite if $m\geqslant 2$ in contrast with the case $m=1$. (A comprehensive study of the case $m=1$ can be found, e.g., in Arnold et al., 1998, pp. 25–28.)

Accordingly, we turn to a *relaxed* version of the original problem and study Z_n, the cumulative number of changes with respect to the top m entries with no regard to the actual qualification "best" or "other," i.e., we define Z_n to be the number of times the membership of the top m items, i.e., the number of times the mth order statistic changes. If $m\geqslant 2$ then the membership can change without changing the actual "best" vs. "other" make up. We note that the event that the mth order statistic changes is often described by saying that an mth record (cf. Nevzorov, 2001) or a Type 2 m-record (cf. Arnold et al., 1998) is established. In other words, these records are the mth largest yet seen.

We also define N_i as the index (position) of the ith change in the sequence $Z_n, n\geqslant 1$. Clearly, $Z_i=N_i=i$, $i=1,2,\ldots,m$, and $P(N_{m+1}>k)=1/\binom{k}{m}$ if $k\geqslant m$ since $N_{m+1}>k$ (i.e., the $m+1$th change has not yet occurred when k elements have been ordered successively) occurs exactly if the highest m elements among the k elements are observed before the other $k-m$ elements which happens with probability $1/\binom{k}{m}$. This yields

$$EN_{m+1}=\sum_{k\geqslant 0}P(N_{m+1}>k)=m+\sum_{k\geqslant m}\frac{1}{\binom{k}{m}}=\frac{m^2}{m-1},\tag{1}$$

where the last equation follows by a standard identity for the reciprocal sum of binomial coefficients. It follows that EN_{m+1} is finite which is markedly different from the case $m=1$ in which $EN_{1+i}=\infty$ for all $i\geqslant 1$, cf. Glick (1978).

Finding the distribution of the index of the $m+i$th change, $N_{m+i}, i>1$, which will be addressed momentarily, requires the distribution of Z_n. We derive this distribution by introducing indicator variables. Let $Z_n=\sum_{i=1}^n U_i$ with $U_i=1$ if the ith observation caused a change in the current set of top m. Note that $U_i=1$, if $i\leqslant m$, and

$$P(U_i=1)=\frac{m}{i}\quad \text{if } i\geqslant m,\tag{2}$$

for this happens if the ith observation is among the top m. In fact, the U_is are *independent* random variables with $P(U_i=1)=\frac{m(i-1)!}{i!}=\frac{m}{i}$, cf. Blom et al. (1994), and

$$E(Z_n)=m\left(1+\frac{1}{m+1}+\cdots+\frac{1}{n}\right)\sim m\ln n$$

as $n\to\infty$, cf. Glick (1978).

Now we establish the probability generating function of Z_n for $n\geqslant m+1$: $E(s^{Z_n})=s^m\prod_{i=m+1}^n E(s^{U_i})=s^m\prod_{i=m+1}^n(1-\frac{m}{i}+\frac{m}{i}s)=s^{m-1}\frac{ms}{m}\frac{(ms+1)(ms+2)\cdots(ms-m+n)}{(m+1)(m+2)\cdots n}=s^{m-1}\binom{ms-m+n}{n-m+1}/\binom{n}{n-m+1}$ by identity (2). Note that the generating function (a polynomial) for the unsigned Stirling numbers of the first kind is $(s+n-1)_n=\binom{s+n-1}{n}n!=\sum_{i=1}^n s(n,i)s^i$

giving $E(s^{Z_n}) = s^{m-1}\sum_{i=1}^{n-m+1}\frac{s(n-m+1,i)}{(n)_{n-m+1}}(ms)^i$, with $(a)_b$ denoting the falling factorial $a(a-1)\cdots(a-b+1)$, after substituting ms and $n-m+1$ for s and n, respectively. This leads to:

Theorem 3. *We have $Z_n = n$ for $1 \leqslant n \leqslant m$ while for $i \geqslant 1$ and $n \geqslant m+i-1$ we have that*

$$P(Z_n = m+i-1) = \frac{m^i s(n-m+1,i)}{(n)_{n-m+1}}. \tag{3}$$

In particular, $P(Z_n = i) = \frac{s(n,i)}{n!}$, if $m = 1$ which is of course the same as Theorem 1 but with n rather than m. We can also derive for $k \geqslant m+1$ that

$$\begin{aligned} P(N_{m+i} = k) &= P(Z_{k-1} = m+i-1)P(U_k = 1) \\ &= P(Z_{k-1} = m+i-1)(m/k), \end{aligned} \tag{4}$$

and thus, $EN_{m+i} = \sum_{k=m+i}^{\infty} kP(N_{m+i} = k) = m\sum_{k=m+i}^{\infty} P(Z_{k-1} = m+i-1)$ results in

$$EN_{m+i} = m^{i+1}\sum_{k=m+i}^{\infty}\frac{s(k-m,i)}{(k-1)_{k-m}} \tag{5}$$

which holds for every $i \geqslant 0$ by (3) and setting $s(0,0) = 1$.

We can now return to the task of determining the distribution of N_{m+i}, $i \geqslant 1$. The distribution of N_{m+i} is described by identities (3) and (4):

Theorem 4. *For $i \geqslant 0$, $k \geqslant m+i$ and with $s(0,0) = 1$, we have*

$$P(N_{m+i} = k) = \frac{m^{i+1}s(k-m,i)}{(k)_{k-m+1}}.$$

Remark. For $m = 1$ we get $N_1 = 1$, and furthermore $\sum_{k=i+1}^{n} kP(N_{1+i} = k) = \sum_{k=i+1}^{n}\frac{k\,s(k-1,i)}{k!} = \frac{s(n,i+1)}{(n-1)!}$. Taken asymptotically as $n \to \infty$, we get

$$EN_{1+i} \sim \frac{(\ln n)^i}{i!}, \quad i = 1, 2, \ldots, \tag{6}$$

in agreement with Wilf (1995).

In contrast, we will be able to determine the (finite) expected value of N_{m+i} for $m \geqslant 2$ exactly in Theorem 5.

Example. If $m = 2$ then we get $EN_1 = 1$, $EN_2 = 2$, and $EN_3 = 2\sum_{n=3}^{\infty} P(Z_{n-1} = 2) = 2\sum_{n=3}^{\infty} 2s(n-2,1)/(n-1)! = 4$ which agrees with $2 + \sum_{k=2}^{\infty} 1/\binom{k}{2} = 4$ in (1). Note that $P(N_{2+i} = n) = 2^{i+1}s(n-2,i)/n!$. We also get (Comtet, 1974, p. 217) that $EN_4 = 2\sum_{n=4}^{\infty} P(Z_{n-1} = 3) = 2\sum_{n=4}^{\infty} 4s(n-2,2)/(n-1)! = 8\sum_{n=4}^{\infty}\frac{H_{n-3}}{(n-1)(n-2)} = 8$ and $EN_5 = 2\sum_{n=5}^{\infty} P(Z_{n-1} = 4) = 16\sum_{n=5}^{\infty} s(n-2,3)/(n-1)! = 16\sum_{n=5}^{\infty}\frac{1}{2(n-1)(n-2)}(H_{n-3}^2 - \sum_{i=1}^{n-3}\frac{1}{i^2}) = 16$ by (3)–(5).

The pattern seen above is formalized in the general:

Theorem 5. *For $m \geqslant 2$ and $i \geqslant 0$, we have that*

$$EN_{m+i} = \frac{m^{i+1}}{(m-1)^i}.$$

Proof. To give the main idea of the proof, we start with the case of $m = 2$ stating that $EN_{2+i} = 2^{i+1}$. (Clearly, $EN_1 = 1$ and $EN_2 = 2$.) If $i \geqslant 1$ then Eqs. (3) and (4) with $m = 1$ imply the identity $\sum_{n \geqslant i} P(N_i = n) = 1 = \sum_{n \geqslant i}\frac{s(n-1,i-1)}{n!}$. Now $EN_{2+i} = 2^{i+1}\sum_{n \geqslant i+2}\frac{s(n-2,i)}{(n-1)!} = 2^{i+1}$ follows by (5).

Please cite this article as: David Collins et al., Keeping scores, Journal of Statistical Planning and Inference (2006), doi: 10.1016/j.jspi.2006.07.006

We can proceed similarly if $m>2$. In identity (5) we decrease both parameters n and m, i.e., we set $n=n'+1$ and $m=m'+1$ for n and m, respectively. A closer look at the summation part in (5), $\sum_{n\geqslant m+i}\frac{s(n-m,i)}{(n-1)_{n-m}}=\sum_{n'\geqslant m'+i}\frac{s(n'-m',i)}{(n')_{n'-m'}}=m'\sum_{n'\geqslant m'+i}\frac{s(n'-m',i)}{(n')_{n'-m'+1}}=\frac{m'}{(m')^{i+1}}\sum_{n'\geqslant m'+i}P(N_{m'+i}=n')$ reveals that it is equal to $m'/(m')^{i+1}=1/(m')^i=1/(m-1)^i$ byTheorem 4. The proof is complete. □

Note that Theorem 5 generalizes well, and for the factorial moment of order $r=1,2,\ldots$, $E((N_{m+i})_r)=E(N_{m+i}(N_{m+i}-1)\cdots(N_{m+i}-r+1))$, the identity

$$E((N_{m+i})_r)=\frac{m^{i+1}(m-1)!}{(m-r)^{i+1}(m-r-1)!}$$

holds according to Nevzorov (2001, p. 89), by Martingale theory.

Theorem 5 means that on average, changes in the membership list of the top m scores, i.e., in Z_n, are observed at a slower and slower pace: more precisely, exponentially slowly since $EN_{m+i}=m(1+\frac{1}{m-1})^i$ as i grows. The waiting times, $\Delta_{i+1}=N_{m+i+1}-N_{m+i}$, $i\geqslant 0$, also have finite expected values if $m\geqslant 2$. In fact, we get that

$$E\Delta_{i+1}=\left(1+\frac{1}{m-1}\right)^{i+1},\quad i\geqslant 0. \tag{7}$$

The waiting time, i.e., the time that the mth order statistic spends at a value, is called the *sojourn* in that value, cf. Bunge and Goldie (2001). Note that $N_{m+i}=k$ if the ith sojourn ends at time k.

5. Other records, record values and positions

In this section we try to generalize Theorem 2 where we considered the "best" scores only. We have only a partial result which makes use of the notion of l-records that will be discussed presently. When processing sequentially, relative ranks (other than the top) and their related quantities, e.g., the positions and the values of the second best (also called near-records or 2-records), third best, etc. observations also can be studied. For a general $l\geqslant 1$, we refer to these records as Type 1 l-records (cf. Arnold et al., 1998 or simply l-records), i.e., the Nth observation X_N is an l-record if exactly $l-1$ previous observations are larger than X_N. (In this nomenclature, the 1-records are simply the records, while l-records without reference to l, $l\geqslant 2$, are often called partial records.)

From now on let $l\geqslant 1$, $r\geqslant 1$, and $N_{l,r}$ denote the position of the rth l-record.

We note that by a beautiful result of Ignatov (1981), the value sets $\mathscr{R}_l$ (with elements $X_{N_{l,1}}<X_{N_{l,2}}<\cdots$) of the l-records, $l=1,2,\ldots$, are independent and identically distributed random sets (i.e., $X_{N_{l_1,i}}$ and $X_{N_{l_2,i}}$ are *i.i.d.* random variables for $i=1,2,\ldots,l_1\neq l_2$). Thus, the smallest elements of the value sets follow the original distribution. For other values of these sets a transformation comes to the rescue, for record values submit to another invariance principle through a clever transformation. Assume that the observations follow the continuous distribution $F(x)$ which is strictly increasing over its support set. One can analyze the record values through the monotone transformation $G(X)=-\ln(1-F(X))$. The transformed record value $G(X_{N_{1,r}})$ has gamma distribution with r degrees of freedom, cf. Glick (1978). For example, if the X_is have exponential distribution with parameter 1 then $X_{N_{1,r}}$ has gamma distribution with parameters r and 1. Note that for the random permutation $(X_1,X_2,\ldots,X_n)$ of $\{1,2,\ldots n\}$, we have $EX_{N_{1,r}}\sim n(1-2^{-r})$ by Wilf (1995), while for the continuous uniform distribution over the interval $[0,n]$, the invariance and some calculation lead to the equality $EX_{N_{1,r}}=n(1-2^{-r})$.

We can combine the first m record value sets $\mathscr{R}_1,\mathscr{R}_2,\ldots,\mathscr{R}_m$, and reorder their entries to get the sequence $Q_1^{(m)}<Q_2^{(m)}<\cdots$. By Theorem 4.1 in Bunge and Goldie (2001) we have that

$$E(\Delta_{i+1}\mid Q_{i+1}^{(m)})=\frac{1}{1-F(Q_{i+1}^{(m)})}=\mathrm{e}^{G(Q_{i+1}^{(m)})}$$

and this fact combined with the previous transformation can be used to derive an alternative proof of identity (7). A similar derivation follows by extending identity (2.6.4) (Arnold et al., 1998, p. 29) to mth records.

As far as the positions of the l-records are concerned, there is another nice invariance property for the transition probabilities for the consecutive positions of the l-records. Independently of l and r, $P(N_{l,r+1}=j|N_{l,r}=k)=\frac{k}{(j-1)j}$

for all $j > k \geqslant l$, though with different initial distributions: $P(N_{1,1}=1)=1$ and $P(N_{l,1}=j)=\frac{l-1}{j(j-1)}$ with $j \geqslant l \geqslant 2$, cf. Blom et al. (1990). For the unconditional distribution $P(N_{l,r}=j)=\frac{l-1}{j}\sum_{l \leqslant i_1 < \cdots < i_r = j} \frac{1}{\prod_{t=1}^{r}(i_t - 1)}$ for $j-r+1 \geqslant l \geqslant 2$ by Blom et al. (1990).

If we generalize Y_n of Sections 2 and 3 by defining $Y_{n,l}$ (with $Y_n = Y_{n,1}$) to be the number of l-records that made the final top m entries after n games, then clearly we have that $Y_{n,l} \leqslant m-l+1,\ m \geqslant l \geqslant 1$. (Note that $Y_{n,l}$ can drop to zero if $l \geqslant 2$.) Let $X_{i:n}$, $i=1,2,\ldots,n$, denote the ith order statistic of the original scores $X_1, X_2, \ldots, X_n$, i.e., $X_{1:n} > X_{2:n} > \ldots > X_{n:n}$. The relation between l-records and the mth order statistic is that $Y_{n,l}=k$ exactly if $k = |\{r \mid X_{N_{l,r}} \geqslant X_{m:n}, N_{l,r} \leqslant n\}|$. As a potential follow up problem, the interested reader is invited to extend Theorems 1 and 2 to find the distribution of $Y_{n,l}$ for $l \geqslant 2$. It is fairly easy to show that $\{P(Y_{n,l}=k)\}_{k=0}^{m-l+1} = \{\frac{m-1}{m}, \frac{1}{m}\}$ for $l=m \geqslant 2$, and $\{v, \frac{v}{m-1}, \frac{v}{2(m-1)}\}$ for $l=m-1 \geqslant 2$ with $v=(1+(m-1)^{-1}+(2(m-1))^{-1})^{-1}$, by using transition probabilities. Note that in general, tridiagonal matrices of transition probabilities correspond to reversible Markov chains, making the determination of the stationary distribution easy.

References

Arnold, B.C., Balakrishnan, N., Nagaraja, N.H., 1998. Records. Wiley, New York.

Blom, G., Holst, L., Sandell, D., 1994. Problems and Snapshots from the World of Probability. Springer, New York, Berlin, Heidelberg.

Blom, G., Thorburn, D., Vessey, T.A., 1990. The distribution of the record position and its applications. Amer. Statist. 44, 151–153.

Bunge, J., Goldie, C.M., 2001. Record sequences and their applications. In: Shanbhag, D.N., Rao, C.R. (Eds.), Handbook of Statistics, Stochastic Processes: Theory and Methods, vol. 19. North-Holland, Amsterdam, New York, pp. 277–308.

Comtet, L., 1974. Advanced Combinatorics. D. Reidel, Dordrecht.

Glick, N., 1978. Breaking records and breaking boards. Amer. Math. Monthly 85, 2–26.

Ignatov, Z., 1981. Point processes generated by order statistics and their applications. In: Colloq. Math. Soc. János Bolyai, vol. 24, Point Processes and Queuing Problems (Colloq., Keszthely, 1978), North-Holland, Amsterdam, New York, pp. 109–116.

Lovász, L., 1979. Combinatorial Problems and Exercises. North-Holland, Akadémiai Kiadó, Amsterdam, New York, Budapest.

Nevzorov, V., 2001. Records: Mathematical Theory, Translations of Mathematical Monographs, vol. 194. American Mathematical Society, Providence, RI.

Wilf, H.S., 1995. On the outstanding elements of permutations. ⟨http://www.math.upenn.edu/~wilf/website/outstelmts.pdf⟩.

Please cite this article as: David Collins et al., Keeping scores, Journal of Statistical Planning and Inference (2006), doi: 10.1016/j.jspi.2006.07.006