4-20-2016

Sequential Selection of a Monotone Subsequence from a Random Permutation

Peichao Peng

J. Michael Steele
University of Pennsylvania

Recommended Citation

Peng, P., & Steele, J. (2016). Sequential Selection of a Monotone Subsequence from a Random Permutation. *Proceedings of the American Mathematical Society, 144* 4973-4982. http://dx.doi.org/10.1090/proc/13104

Sequential Selection of a Monotone Subsequence from a Random Permutation

Abstract

We find a two term asymptotic expansion for the optimal expected value of a sequentially selected monotone subsequence from a random permutation of length *n*. A striking feature of this expansion is that it tells us that the expected value of optimal selection from a random permutation is quantifiably larger than optimal sequential selection from an independent sequence of uniformly distributed random variables; specifically, it is larger by at least (1/6) log $n + O(1)$.

Disciplines

Business | Mathematics | Statistics and Probability

SEQUENTIAL SELECTION OF A MONOTONE SUBSEQUENCE FROM A RANDOM PERMUTATION

PEICHAO PENG AND J. MICHAEL STEELE

Abstract. We find a two term asymptotic expansion for the optimal expected value of a sequentially selected monotone subsequence from a random permutation of length n. A striking feature of this expansion is that tells us that the expected value of optimal selection from a random permutation is quantifiably larger than optimal sequential selection from an independent sequences of uniformly distributed random variables; specifically, it is larger by at least $(1/6)\log n + O(1)$.

Key Words. Monotone subsequence problem, sequential selection, online selection, Markov decision problem, nonlinear recursion, asymptotics

Mathematics Subject Classification (2010). Primary: 60C05, 60G40, 90C40; Secondary: 60F99, 90C27, 90C39

1. Sequential Subsequence Problems

In the classical monotone subsequence problem, one chooses a random permutation $\pi : [1:n] \to [1:n]$, and one considers the length of its longest increasing subsequence,

$$L_n = \max\{k : \pi[i_1] < \pi[i_2] < \cdots < \pi[i_k] \quad \text{where } 1 \le i_1 < i_2 \cdots < i_k \le n\}.$$

On the other hand, in the *sequential* monotone subsequence problem one views the values $\pi[1]$, $\pi[2]$, ... as though they were presented over time to a decision maker who, when shown the value $\pi[i]$ at time i, must decide (once and for all) either to accept or reject $\pi[i]$ as element of the selected increasing subsequence.

The decision to accept or reject $\pi[i]$ at time i is based on just the knowledge of the time horizon n and the observed values $\pi[1], \pi[2], \ldots, \pi[i]$. Thus, in slightly more formal language, the sequential selection problems amounts to the consideration of random variables of the form

$$(1) \quad L_n^{\boldsymbol{\tau}} = \max\{k : \pi[\tau_1] < \pi[\tau_2] < \cdots < \pi[\tau_k] \quad \text{where } 1 \le \tau_1 < \tau_2 \cdots < \tau_k \le n\},$$

where the indices τ_i, $i = 1, 2, \ldots$ are stopping times with respect to the increasing sequence of σ-fields $\mathcal{F}_k = \sigma\{\pi[1], \pi[2], \ldots, \pi[k]\}$, $1 \le k \le n$. We call a sequence of such stopping times a *feasible selection strategy*, and, if we use $\boldsymbol{\tau}$ as a shorthand for such a strategy, then the quantity of central interest here can be written as

$$(2) \qquad s(n) = \sup_{\boldsymbol{\tau}} \mathbb{E}[L_n^{\boldsymbol{\tau}}],$$

where one takes the supremum over all feasible selection strategies.

Peichao Peng: Department of Statistics, The Wharton School, University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA, 19104. Email address: ppeichao@wharton.upenn.edu.
J. M. Steele: Department of Statistics, The Wharton School, University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA, 19104. Email address: steele@wharton.upenn.edu.

It was conjectured in Baer and Brock (1968) that

$$s(n) \sim \sqrt{2n} \quad \text{as } n \to \infty, \tag{3}$$

and a proof of this relation was first given in Samuels and Steele (1981). A much simpler proof of (3) was later given by Gnedin (2000) who made use of a recursion that had been used for numerical computations by Baer and Brock (1968). The main purpose of this note is to show that by a more sustained investigation of that recursion one can obtain a two term expansion.

Theorem 1 (Sequential Selection from a Random Permutation)**.** *For $n \to \infty$ one has the asymptotic relation*

$$s(n) = \sqrt{2n} + \frac{1}{6}\log n + O(1). \tag{4}$$

Given what is known for some closely related problems, the explicit second order term $(\log n)/6$ presents something of a surprise. For comparison, suppose we consider sequential selection from a sequence of n independently uniformly distributed random variables X_1, X_2, ..., X_n. In this problem a feasible selection strategy $\boldsymbol{\tau}$ is again expressed by an increasing sequence of stopping times τ_j, $j = 1, 2, \ldots$, but now the stopping times are adapted to the increasing σ-fields $\widehat{\mathcal{F}}_j = \sigma\{X_1, X_2, \ldots, X_j\}$. The analog of (1) is then

$$\widehat{L}_n^{\boldsymbol{\tau}} = \max\{k : X_{\tau_1} < X_{\tau_2} < \cdots < X_{\tau_k} \quad \text{where } 1 \le \tau_1 < \tau_2 \cdots < \tau_k \le n\}, \tag{5}$$

and the analog of (2) is given by

$$\widehat{s}(n) = \sup_{\boldsymbol{\tau}} \mathbb{E}[\widehat{L}_n^{\boldsymbol{\tau}}].$$

It was proved by Bruss and Robertson (1991) that for $\widehat{s}(n)$ one has a uniform upper bound

$$\widehat{s}(n) \le \sqrt{2n} \quad \text{for all } n \ge 1, \tag{6}$$

so, by comparison with (4), we see there is a sense in which sequential selection of a monotone subsequence from a permutation is *easier* than sequential selection from an independent sequence. In part, this is intuitive; each successive observation from a permutation gives useful information about the subsequent values that can be observed. By (4) one quantifies how much this information helps.

Since (6) holds for all n and since (4) is only asymptotic, it also seems natural to ask if there is a relation between $\widehat{s}(n)$ and $s(n)$ that is valid for all n. There is such a relation if one gives up the logarithmic gap.

Theorem 2 (Selection for Random Permutations vs Random Sequences)**.** *One has for all $n = 1, 2, \ldots$ that*

$$\widehat{s}(n) \le s(n).$$

Here we should also note that much more is known about $\widehat{s}(n)$ than just (6); in particular, there are several further connections between $s(n)$ and $\widehat{s}(n)$. These are taken up in a later section, but first it will be useful to give the proofs of Theorems 1 and 2.

The proof of Theorem 1 takes most of our effort, and it is given over the next few sections. Section 2 develops the basic recurrence relations, and Section 3 develops stability relations for these recursions. In Section 4 we then do the calculations that support a candidate for the asymptotic approximation of $s(n)$, and we compete the proof of Theorem 1. Our arguments conclude in Section 5 with the brief — and almost computation free — proof of Theorem 2. Finally, in Section 6 we discuss further relations between $s(n)$, $\widehat{s}(n)$, and some other closely related quantities that motivate two open problems.

2. Recurrence Relations

One can get a recurrence relation for $s(n)$ by first step analysis. Specifically, we take a random permutation $\pi : [1 : n+1] \to [1 : n+1]$, and we consider its initial value $\pi[1] = k$. If we reject $\pi[1]$ as an element of our subsequence, we are faced with the problem of sequential selection from the reduced random permutation π' on an n-element set. Alternatively, if we choose $\pi[1] = k$ as an element of our subsequence, we are then faced with the problem of sequential selection for a reduced random permutation π'' of the set $\{k+1, k+2, \ldots, n+1\}$ that has $n+1-k$ elements. By taking the better of these two possibilities, we get from the uniform distribution of $\pi[1]$ that

$$s(n+1) = \frac{1}{n+1} \sum_{k=1}^{n+1} \max\{s(n), 1 + s(n+1-k)\}. \tag{7}$$

From the definition (2) of $s(n)$ one has $s(1) = 1$, so subsequent values can then be computed by (7). For illustration and for later discussion, we note that one has the approximate values:

n	1	2	3	4	5	6	7	8	9	10
$s(n)$	1	1.5	2	2.375	2.725	3.046	3.333	3.601	3.857	4.098
$\sqrt{2n}$	1.414	2	2.449	2.828	3.162	3.464	3.742	4	4.243	4.472.

Here we observe that for the 10 values in the table one has $s(n) \le \sqrt{2n}$, and, in fact, this relation persists for all $1 \le n \le 174$. Nevertheless, for $n = 175$ one has $\sqrt{2n} < s(n)$, just as (4) requires for all sufficiently large values of n.

We also know from (2) that the map $n \mapsto s(n)$ is strictly monotone increasing, and, as a consequence, the recursion (7) can be written a bit more simply as

$$\begin{aligned} s(n+1) &= \frac{1}{n+1} \max_{1 \le k \le n} \left\{ (n-k+1)s(n) + \sum_{i=n-k+1}^{n} \{s(i)+1\} \right\} \\ &= \frac{1}{n+1} \max_{1 \le k \le n} \left\{ (n-k+1)s(n) + k + \sum_{i=n-k+1}^{n} s(i) \right\}. \end{aligned} \tag{8}$$

In essence, this recursion goes back to Baer and Brock (1968, p. 408), and it is the basis of most of our analysis.

3. Comparison Principles

Given a map $g : \mathbb{N} \to \mathbb{R}$ and $1 \le k \le n$, it will be convenient to set

$$H(n,k,g) = k + (n-k+1)g(n) + \sum_{i=n-k+1}^{n} g(i), \tag{9}$$

so the optimality recursion (8) can be written more succinctly as

$$s(n+1) = \frac{1}{n+1} \max_{1 \le k \le n} H(n,k,s). \tag{10}$$

The next two lemmas make rigorous the idea that if g is almost a solution of (10) for all n, then $g(n)$ is close to $s(n)$ for all n.

Lemma 3 (Upper Comparison)**.** *If* $\delta : \mathbb{N} \to \mathbb{R}^+$, $1 \leq g(1) + \delta(1)$, *and*

$$\frac{1}{n+1} \max_{1 \leq k \leq n} H(n, k, g) \leq g(n+1) + \delta(n+1) \quad \textit{for all } n \geq 1, \tag{11}$$

then one has

$$s(n) \leq g(n) + \sum_{i=1}^{n} \delta(i) \quad \textit{for all } n \geq 1. \tag{12}$$

Proof. We set $\Delta(i) = \delta(1)+\delta(2)+\cdots+\delta(i)$, and we argue by induction. Specifically, using (12) for $1 \leq i \leq n$ we have

$$\begin{aligned} H(n, k, s) &= k + (n-k+1)s(n) + \sum_{i=n-k+1}^{n} s(i) \\ &\leq k + (n-k+1)(g(n) + \Delta(n)) + \sum_{i=n-k+1}^{n} \{g(i) + \Delta(i)\} \end{aligned}$$

so by monotonicity of $\Delta(\cdot)$ we have

$$\frac{1}{n+1} H(n, k, s) \leq \frac{1}{n+1} H(n, k, g) + \Delta(n).$$

Now, when we take the maximum over $k \in [1 : n]$, the recursion (8) and the induction condition (11), give us

$$\begin{aligned} s(n+1) &\leq \frac{1}{n+1} \max_{1 \leq k \leq n} H(n, k, g) + \Delta(n) \\ &\leq g(n+1) + \delta(n+1) + \Delta(n) = g(n+1) + \Delta(n+1), \end{aligned}$$

so induction establishes (12) for all $n \geq 1$. □

Naturally, there is a lower bound comparison principle that parallels Lemma 3. The statement has several moving parts, so we frame it as a separate lemma even though its proof can be safely omitted.

Lemma 4 (Lower Comparison)**.** *If* $\delta : \mathbb{N} \to \mathbb{R}^+$, $g(1) - \delta(1) \leq 1$, *and*

$$g(n+1) - \delta(n+1) \leq \frac{1}{n+1} \max_{1 \leq k \leq n} H(n, k, g) \quad \textit{for all } n \geq 0,$$

then one has

$$g(n) - \sum_{i=1}^{n} \delta(i) \leq s(n) \quad \textit{for all } n \geq 1.$$

4. An Approximation Solution

We now argue that the function $f : \mathbb{N} \to \mathbb{R}$ defined by

$$f(n) = \sqrt{2n} + \frac{1}{6} \log n, \tag{13}$$

gives one an approximate solution of the recurrence equation (8) for $n \mapsto s(n)$.

Proposition 5. *There is a constant* $0 < B < \infty$ *such that for all* $n \geq 1$, *one has*

$$-Bn^{-3/2} \leq \frac{1}{n+1}\left\{\max_{1\leq k\leq n} H(n,k,f)\right\} - f(n+1) \leq Bn^{-3/2}. \tag{14}$$

First Step: Localization of the Maximum

To deal with the maximum in (14), we first estimate

$$k^*(n) = \mathrm{locmax}_k\ H(n,k,f).$$

From the definition (9) of $H(n,k,f)$ we find

$$H(n,k+1,f) - H(n,k,f) = 1 - f(n) + f(n-k),$$

and, from the definition (13) of f, we see this difference is monotone decreasing function of k; accordingly, we also have the representation

$$k^*(n) = 1 + \max\{k : 0 \leq 1 - f(n) + f(n-k)\}. \tag{15}$$

Now, for each $n = 1, 2, \ldots$ we then consider the function $D_n : [0,n] \to \mathbb{R}$ defined by setting

$$D_n(x) = 1 - f(n) + f(n-x) = 1 - \{\sqrt{2n} - \sqrt{2(n-x)}\} - \frac{1}{6}\{\log n - \log(n-x)\}.$$

This function is strictly decreasing with $D_n(0) = 1$ and $D_n(n) = -\infty$, so there is a unique solution of the equation $D_n(x) = 0$. For $x \in [0,n)$ we also have the easy bound

$$D_n(x) = 1 - \frac{1}{2}\int_{2(n-x)}^{2n} \frac{1}{\sqrt{u}}\,du - \frac{1}{6}\log(n/(n-x)) \leq 1 - \frac{x}{\sqrt{2n}}.$$

This gives us $D_n(\sqrt{2n}) \leq 0$, so by monotonicity we have $x_n \leq \sqrt{2n}$.

To refine this bound to an asymptotic estimate, we start with he equation $D_n(x_n) = 0$ and we apply Taylor expansions to get

$$\begin{aligned} 1 &= \sqrt{2n}\left\{1 - (1 - x_n/n)^{1/2}\right\} - \frac{1}{6}\log(1 - x_n/n) \\ &= \sqrt{2n}\left\{\frac{x_n}{2n} + O(x_n^2/n^2)\right\} + O(x_n/n). \end{aligned}$$

By simplification, we then get

$$\sqrt{2n} = x_n + O(x_n^2/n) + O(x_n/n^{1/2}) = x_n + O(1), \tag{16}$$

where in the last step we used our first bound $x_n \leq \sqrt{2n}$.

Finally, by (16) and the characterization (15), we immediately find the estimate that we need for $k^*(n)$.

Lemma 6. *There is a constant* $A > 0$ *such that for all* $n \geq 1$, *we have*

$$\sqrt{2n} - A \leq k^*(n) \leq \sqrt{2n} + A. \tag{17}$$

Remark 7. The relations (16) and (17) can be sharpened. Specifically, if we use a two-term Taylor series with integral remainders, then one can show $\sqrt{2n} - 2 \leq x_n$. Since we already know that $x_n \leq \sqrt{2n}$, we then see from the characterization (15) and integrality of $k^*(n)$ that we can take $A = 2$ in Lemma 6. This refinement does not lead to a meaningful improvement in Theorem 1, so we omit the details of the expansions with remainders.

Completion of Proof of Proposition 5

To prove Proposition 5, we first note that the definition (9) of $H(n,k,f)$ one has for all $1 \le k \le n$ that

$$\frac{1}{n+1}H(n,k,f) = f(n) + \frac{1}{n+1}\left\{k - \sum_{i=1}^{k-1}(f(n)-f(n-i))\right\} \tag{18}$$

The task is to estimate the right-hand side of (18) when $k = k^*(n)$ and $k^*(n)$ is given by (15).

For the moment, we assume that one has $k \le D\sqrt{n}$ where $D > 0$ is constant. With this assumption, we find that after making Taylor expansions we get from explicit summations that

$$\begin{aligned}\sum_{i=1}^{k-1}(f(n)-f(n-i)) &= \sum_{i=1}^{k-1}\left(\sqrt{2n}-\sqrt{2(n-i)}\right) + \sum_{i=1}^{k-1}\left(\frac{\log n}{6} - \frac{\log(n-i)}{6}\right)\\ &= \sum_{i=1}^{k-1}\left(\frac{i}{\sqrt{2n}} + \frac{i^2}{4n\sqrt{2n}} + O\left(\frac{i^3}{n^{5/2}}\right)\right) + \sum_{i=1}^{k-1}\left(\frac{i}{6n} + O\left(\frac{i^2}{n^2}\right)\right)\\ &= \frac{(k-1)k}{2\sqrt{2n}} + \frac{(k-1)k(2k-1)}{24n\sqrt{2n}} + \frac{(k-1)k}{12n} + O(n^{-1/2}), \end{aligned} \tag{19}$$

where the implied constant of the remainder term depends only on D.

We now define $r(n)$ by the relation $k^*(n) = \sqrt{2n} + r(n)$, and we note by (17) that $|r(n)| \le A$. Direct algebraic expansions then give us the elementary estimates

$$\frac{(k^*(n)-1)k^*(n)}{12n} = \frac{1}{6} + O(n^{-1/2})$$

and

$$\frac{(k^*(n)-1)k^*(n)(2k^*(n)-1)}{24n\sqrt{2n}} = \frac{1}{6} + O(n^{-1/2}),$$

where in each case the implied constant depends only on A.

Estimation of the first summand of (19) is slightly more delicate than this since we need to account for the dependence of this term on $r(n)$; specifically we have

$$\begin{aligned}\frac{(k^*(n)-1)k^*(n)}{2\sqrt{2n}} &= \frac{\left(\sqrt{2n}+r(n)-1\right)\left(\sqrt{2n}+r(n)\right)}{2\sqrt{2n}}\\ &= \sqrt{n/2} + r(n) - \frac{1}{2} + O(n^{-1/2}).\end{aligned}$$

Now, for a pleasing surprise, we note from the last estimate and from the definition of $k^*(n)$ and $r(n)$ that we have cancelation of $r(n)$ when we then compute the critical sum; thus, one has simply

$$k^*(n) - \sum_{i=1}^{k^*(n)-1}(f(n)-f(n-i)) = \sqrt{n/2} + \frac{1}{6} + O(n^{-1/2}). \tag{20}$$

Finally, from the formula (13) for $f(\cdot)$, we have the Taylor expansion

$$f(n+1) - f(n) = \frac{1}{\sqrt{2n}} + \frac{1}{6n} + O(n^{-3/2}), \tag{21}$$

so, when we return to the identity (18), we see that the estimates (20) and (21) give us the estimate

$$\frac{1}{n+1}\left\{\max_{1\le k\le n} H(n,k,f)\right\} - f(n+1)$$
$$= \frac{1}{n+1}\left(\sqrt{n/2} + \frac{1}{6} + O(n^{-1/2})\right) + f(n) - f(n+1) = O(n^{-3/2}).$$

Here the implied constant is absolute, and the proof of Proposition 5 is complete.

Completion of Proof of Theorem 1

Lemmas 3 and 4 combine with Proposition 5 to tell us that by summing the sequence $n^{-3/2}$, $n = 1, 2, \ldots$ and by writing $\zeta(z) = 1 + 2^{-z} + 3^{-z} + \cdots$ one has

$$|s(n) - f(n)| \le \zeta(3/2)B \le (2.62)B \quad \text{for all } n \ge 1.$$

This is slightly more than one needs to complete the proof of Theorem 1.

5. Proof of Theorem 2

The sequential monotone selection problem is a finite horizon Markov decision problem with bounded rewards and finite action space, and for such problems it is known one cannot improve upon an optimal deterministic strategy by the use of strategies that incorporate randomization, (cf. Bertsekas and Shreve, 1978, Corollary 8.5.1). The proof Theorem 2 exploits this observation by constructing a randomized algorithm for the sequential selection of a monotone subsequence from a random permutation.

We first recall that if e_i, $i = 1, 2, \ldots, n+1$ are independent exponentially distributed random variables with mean 1 and if one sets

$$Y_i = \frac{e_1 + e_2 + \cdots + e_i}{e_1 + e_2 + \cdots + e_{n+1}},$$

then the vector $(Y_1, Y_2, \ldots, Y_n)$ has the same distribution as the vector of order statistics $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ of an i.i.d. sample of size n from the uniform distribution. Next we let $\mathcal{A}$ denote an optimal algorithm for sequential selection from an independent sample $X_1, X_2, \ldots, X_n$ from the uniform distribution, and we let $\boldsymbol{\tau}(\mathcal{A})$ denote the associated sequence of stopping times. If $\widehat{L}_n^{\boldsymbol{\tau}(\mathcal{A})}$ denotes the length of the subsequence that is chosen from from $X_1, X_2, \ldots, X_n$ when one follows the strategy $\boldsymbol{\tau}(\mathcal{A})$ determined by $\mathcal{A}$, then by optimality of $\mathcal{A}$ for selection from $X_1, X_2, \ldots, X_n$ we have

$$\widehat{s}(n) = \sup_{\boldsymbol{\tau}} \mathbb{E}[\widehat{L}_n^{\boldsymbol{\tau}}] = \mathbb{E}[\widehat{L}_n^{\boldsymbol{\tau}(\mathcal{A})}].$$

We use the algorithm $\mathcal{A}$ to construct a new randomized algorithm $\mathcal{A}'$ for sequential selection of an increasing from a random permutation $\pi : [n] \mapsto [n]$. First, the decision maker generates independent exponential random variables e_i, $i = 1, 2, \ldots, n+1$ as above. This is done off-line, and this step can be viewed as an internal randomization.

Now, for $i = 1, 2, \ldots, n$, when we are presented with $\pi[i]$ at time i, we compute $X_i = Y_{\pi[i]}$. Finally, if at time i the value X_i would be accepted by the algorithm $\mathcal{A}$, then the algorithm $\mathcal{A}'$ accepts $\pi[i]$. Otherwise the newly observed value $\pi[i]$ is rejected. By our construction we have

$$\mathbb{E}[L_n^{\boldsymbol{\tau}(\mathcal{A}')}] = \mathbb{E}[\widehat{L}_n^{\boldsymbol{\tau}(\mathcal{A})}] = \widehat{s}(n). \tag{22}$$

Moreover, $\mathcal{A}'$ is a randomized algorithm for construction an increasing subsequence of a random permutation π. By definition, $s(n)$ is the expected length of a monotone subsequence selected from a random permutation by an optimal deterministic algorithm, and by our earlier observation, the randomized algorithm $\mathcal{A}'$ cannot do better. Thus, from (22) one has $\widehat{s}(n) \le s(n)$, and the proof of Theorem 1 is complete.

6. Further Connections and Considerations

As we noted before, Bruss and Robertson (1991) discovered the uniform bound

$$\widehat{s}(n) \le \sqrt{2n} \quad \text{for all } n \ge 1, \tag{23}$$

and their proof depended on a general bound for the expected value of the partial sums of the smallest order statistics of a uniformly distributed random sample. Gnedin (1999) later gave a much different proof of (23) and generalized the bound in a way that accommodate random samples with random sizes. More recently, Arlotto, Mossel and Steele (2015) obtained yet another proof (23) as a corollary to bounds on the *quickest selection problem*, which is an informal dual to the traditional selection problem.

Since the bound (23) is now well understood from several points of view, it is reasonable to ask about the possibility of some corresponding uniform bound on $s(n)$. The numerical values that we noted after the recursion (23) and the relation

$$s(n) = \sqrt{2n} + \frac{1}{6}\log n + O(1) \tag{24}$$

from Theorem 1 both tell us that one cannot expect a uniform bound for $s(n)$ that is as simple as that for $\widehat{s}(n)$ given by (23). Nevertheless, numerical evidence suggest that the $O(1)$ term in (24) is always negative. The tools used here cannot confirm this conjecture, but the multiple perspectives available for (23) give one hope.

A closely related issue arises for $\widehat{s}(n)$ when one considers lower bounds. Here the first steps were taken by Bruss and Delbaen (2001) who considered i.i.d. samples of size N_ν where N_ν is an independent random variable with the Poisson distribution with mean ν. In the natural (but slightly overburdened) notation, they proved that there is a constant $c > 0$ such that

$$\sqrt{2\nu} - c\log\nu \le \widehat{s}(\nu);$$

moreover, Bruss and Delbaen (2004) subsequently proved that for the optimal feasible strategy $\boldsymbol{\tau}_* = (\tau_1, \tau_2, \ldots)$ the random variable

$$\widehat{L}^{\boldsymbol{\tau}_*}_{N_\nu} = \max\{k : X_{\tau_1} < X_{\tau_2} < \cdots < X_{\tau_k} \quad \text{where } 1 \le \tau_1 < \tau_2 \cdots < \tau_k \le N_\nu\},$$

also satisfies a central limit theorem. Arlotto, Nguyen and Steele (2015) considered the de-Poissonization of these results, and it was found that one has the corresponding CLT for $\widehat{L}^{\boldsymbol{\tau}_*}_n$ where the sample size n is deterministic. In particular, one has the bounds

$$\sqrt{2n} - c\log n \le \widehat{s}(n) \le \sqrt{2n}.$$

Now, by analogy with (24), one strongly expects that there is a constant $c > 0$ such that

$$\widehat{s}(n) = \sqrt{2n} - c\log n + O(1). \tag{25}$$

Still, a proof this conjecture is reasonably remote, since, for the moment, there is not even a compelling candidate for the value of c.

For a second point of comparison, one can recall the *non-sequential* selection problem where one studies

$$\ell(n) = \mathbb{E}[\max\{k : X_{i_1} < X_{i_2} < \ldots < X_{i_k},\ 1 \le i_1 < i_2 < \cdots < i_k \le n\}].$$

Through a long sequence of investigations culminating with Baik, Deift and Johansson (1999), it is now known that one has

$$\ell(n) = 2\sqrt{n} - \alpha n^{1/6} + o(n^{1/6}), \tag{26}$$

where the constant $\alpha = 1.77108...$ is determined numerically in terms of solutions of a Painlevé equation of type II. Romik (2014) gives an elegant account of the extensive technology behind (26), and there are interesting analogies between $\ell(n)$ and $s(n)$. Nevertheless, a proof of the conjecture (25) seems much more likely to come from direct methods like those used here to prove (24).

Finally, one should note that the asymptotic formulas for $n \mapsto \ell(n)$, $n \mapsto s(n)$, and $n \mapsto \widehat{s}(n)$ all suggest that these maps are concave, but so far only $n \mapsto \widehat{s}(n)$ has been proved to be concave (cf. Arlotto, Nguyen and Steele (2015, p. 3604)).

References

Arlotto, A., Mossel, E. and Steele, J. M. (2015), 'Quickest online selection of an increasing subsequence of specified size', to appear in *Random Structures and Algorithms (eprint arXiv:1412.7985)* .

Arlotto, A., Nguyen, V. V. and Steele, J. M. (2015), 'Optimal online selection of a monotone subsequence: a central limit theorem', *Stochastic Process. Appl.* **125**(9), 3596–3622.

Baer, R. M. and Brock, P. (1968), 'Natural sorting over permutation spaces', *Math. Comp.* **22**, 385–410.

Baik, J., Deift, P. and Johansson, K. (1999), 'On the distribution of the length of the longest increasing subsequence of random permutations', *J. Amer. Math. Soc.* **12**(4), 1119–1178.

Bertsekas, D. P. and Shreve, S. E. (1978), *Stochastic optimal control: the discrete time case*, Vol. 139 of *Mathematics in Science and Engineering*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, NY.

Bruss, F. T. and Delbaen, F. (2001), 'Optimal rules for the sequential selection of monotone subsequences of maximum expected length', *Stochastic Process. Appl.* **96**(2), 313–342.

Bruss, F. T. and Delbaen, F. (2004), 'A central limit theorem for the optimal selection process for monotone subsequences of maximum expected length', *Stochastic Process. Appl.* **114**(2), 287–311.

Bruss, F. T. and Robertson, J. B. (1991), '"Wald's lemma" for sums of order statistics of i.i.d. random variables', *Adv. in Appl. Probab.* **23**(3), 612–623.

Gnedin, A. V. (1999), 'Sequential selection of an increasing subsequence from a sample of random size', *J. Appl. Probab.* **36**(4), 1074–1085.

Gnedin, A. V. (2000), 'A note on sequential selection from permutations', *Combin. Probab. Comput.* **9**(1), 13–17.

Romik, D. (2014), *The Surprising Mathematics of Longest Increasing Subsequences*, Cambridge University Press, Cambridge.

Samuels, S. M. and Steele, J. M. (1981), 'Optimal sequential selection of a monotone sequence from a random sample', *Ann. Probab.* **9**(6), 937–947.