Running minimum in the best-choice problem

Alexander Gnedin 1 · Patryk Kozieł 2 · Małgorzata Sulkowska 2,3

Received: 7 April 2022 / Revised: 10 November 2022 / Accepted: 11 November 2022 / Published online: 29 November 2022 © The Author(s) 2022

Abstract

The full-information best choice problem asks one to find a strategy maximising the probability of stopping at the minimum (or maximum) of a sequence \( X_1, \ldots, X_n \) of i.i.d. random variables with continuous distribution. In this paper we look at more general models, where independent \( X_j \)'s may have different distributions, discrete or continuous. A central role in our study is played by the running minimum process, which we first employ to re-visit the classic problem and its limit Poisson counterpart. The approach is further applied to two explicitly solvable models: in the first the distribution of the \( j \)th variable is uniform on \( \{ j, \ldots, n \} \), and in the second it is uniform on \( \{ 1, \ldots, n \} \).

Keywords
Best-choice problem · Running minimum process · Poisson point process · Full-information problem

AMS 2000 Subject Classifications: Primary 60G40 · Secondary 60G55

1 Introduction

This article deals with the stochastic optimisation task of selecting the minimum value \( M_n = \min(X_1, \ldots, X_n) \) from a sequence of independent random variables \( X_1, \ldots, X_n \) with known distributions. The course of action is as follows: the sample
values of $X_j$’s are shown sequentially one at a time, and the proceedings may be stopped by selecting the number being shown. Neither foresight nor return to past observations is allowed. Let $v_n$ be the maximal achievable probability of selecting a sample value equal to $M_n$. The problem is to find $v_n$ and an optimal strategy that yields this probability.

The case of i.i.d. $X_j$’s with continuous distribution was first studied in the seminal paper by Gilbert and Mosteller (1966) (where selection of the maximum sample value was considered, which is equivalent). Bojdecki (1978) and Sakaguchi (1978) formulated the problem in terms of the embedded Markov chain of records and showed in this framework that the optimal strategy is myopic, that is prescribes to select a record observation as soon as the probability of success by stopping immediately is not less than by waiting and stopping at the next record (if available). This basic version of the problem was dubbed the ‘full-information’ best choice (aka ‘secretary’) problem, to contrast with the ‘no-information’ problem where the distribution is unknown and the decisions depend only on the relative ranks of observations (Ferguson 1989; Samuels 1991b). Many generalisations have appeared in the literature, with various assumptions on the underlying random sequence, information available to the observer and the class of admissible selection strategies (Berezovsky and Gnedin 1984; Kuchta 2017; Pinsky 2022; Porosiński 1987; Porosiński and Szajowski 2000). We mention in passing ‘partial information’ models with the data sampled from partly known distribution, which may be itself random or known up to a few indefinite parameters (Campbell 1982; Ferguson 1989; Goldenshluger and Zeevi 2021; Petruccelli 1980; Stewart 1978).

In the continuous i.i.d. setting the optimal strategy is determined by $n$ thresholds $b_1 \leq \cdots \leq b_n$. The optimal success probability, further denoted $v_n$, does not depend on the distribution (which in Gilbert and Mosteller (1966) was assumed uniform), decreases in $n$ and converges to a value $v \approx 0.580164$ for which there is an explicit formula due to Samuels (1991b). The limit value $v$ has an insightful interpretation as the success probability in an optimal selection problem associated with a planar Poisson process (Gnedin 1996, 2004; Gnedin and Miretskiy 2007), very much in line with the theory of extremal processes (Resnick 2008).

Getting away from the classic i.i.d. instance by allowing for different and not necessarily continuous distributions of $X_j$’s, the problem becomes mainly different in two aspects. Firstly, the optimal strategy and $v_n$ depend in a complex way on the distributions. Secondly, if the distributions are not continuous there could be ‘ties’, in particular multiple sample elements may coincide with the minimum $M_n$. Some past work is relevant. Hill and Kennedy (1992) showed that using single-threshold strategies one can achieve the success probability at least $(1 - 1/n)^{n-1}$ and that this lower bound on $v_n$ is sharp. Faller and Rüschendorf (2012) studied large-$n$ approximations of the optimal strategy and $v_n$ in terms of a stopping problem for the planar Poisson process. Scherwentke (1989) studied the structure of the optimal strategy for i.i.d. observations with discrete distribution (see comments on this paper in Samuels (1991a)).

In the present paper we first observe that for arbitrary independent $X_j$’s the optimal strategy is given by a nondecreasing sequence of thresholds, as in the i.i.d. continuous case. For the general i.i.d. case we bound the success probability as $v_n \leq v_n \leq v_n + \delta_n$, where $\delta_n$ is the probability of a tie for the minimum. The structure
of the optimal strategy allows us to represent $v_n$ in terms of a first passage problem for the running minimum process $M_j = \min(X_1, \ldots, X_j)$, $j \leq n$, which is our main technical novelty. The representation involves two components which distinguish whether the stopping domain is entered due to the time factor only or due to the occurrence of a new record. Re-visiting the continuous i.i.d. problem, we apply the approach to derive a new formula for the success probability $v_n$. Then we resort to two discrete models which possess explicitly solvable limit Poisson counterparts. In the first the distribution of $X_j$ is uniform on $\{j, \ldots, n\}$ (the triangular model). Interestingly, the limit success probability about 0.703128 has appeared in a problem of selection from partially ordered data (Gnedin 2007). In the second model the distribution is uniform on $\{1, \ldots, n\}$ (the rectangular model), and the limit success probability is found to be about 0.761260. These instances are chosen to contrast two situations where the ties vanish or persist in the large-$n$ limit. Both models admit tractable parametric generalisations.

All the numerical values presented in the article are truncated to 6 decimal places.

2 Strategies and bounds

We introduce first some terminology. In the event $X_j = M_j$ we say that $X_j$ is a (lower) record. Equivalently, a record at index $j$ occurs if $X_j \leq M_{j-1}$ (where $M_0 := \infty$). The records we consider here are sometimes called weak records, as opposed to strong records distinguished by the strict inequality $X_j < M_{j-1}$.

An admissible selection strategy $\tau$ is a stopping time which takes values in the set $\{1, \ldots, n, \infty\}$ and is adapted to the natural filtration of the sequence $X_1, \ldots, X_n$. That is to say, each event $\tau = j$ is determined by the values of $X_1, \ldots, X_j$ only, for $j \leq n$. Stopping with $\tau$ is regarded as a success if $X_\tau = M_\tau$. In the event $X_\tau = \infty$ we set $\tau = \infty$, and regard this as ‘no selection’.

The general discrete-time optimal stopping theory ensures existence of a stopping time achieving the maximum success probability

$$v_n := \sup_\tau \mathbb{P}(X_\tau = M_n),$$

see Chow et al. (1991), Theorem 3.2. Since $M_n$ is the minimal record value, the optimal strategy is among the strategies that only select records, i.e. satisfy $X_\tau = M_\tau$ (provided $\tau < \infty$). However, we will also make use of stopping times that violate this condition.

2.1 Structure of the stopping domain

For independent $X_1, \ldots, X_n$, both the running minimum $(M_j, j \leq n)$ and the bivariate sequence $((X_j, M_j), j \leq n)$ are Markov processes; these facts will be tacitly used in what follows along with the natural monotonicity $M_1 \geq \cdots \geq M_n$. Let $F_j$ denote the right-continuous c.d.f. of $X_j$, and let
\[
s(j, x) := \prod_{k=j+1}^{n} (1 - F_k(x)),
\]
\[
v(j, x) := \sup_{\{\tau: \tau \ge j, X_{\tau} \le x\}} \mathbb{P}(X_{\tau} = \min\{X_{j+1}, \ldots, X_n\}).
\]

In particular, \(s(n, x) = 1\) and \(v(n, x) = 0\). If no choice has been made from the first \(j - 1\) observations, then conditionally on record \(X_j = x\) stopping at \(X_j\) is successful with probability \(s(j, x)\), regardless of the previously observed values \(X_1, \ldots, X_{j-1}\). Likewise, the maximum success probability achievable by skipping the record is \(v(j, x)\), as follows from the Markov property (Chow et al. 1991, Theorem 5.1). Given a record observation, the functions \(s(j, x)\) and \(v(j, x)\) measure the reward from stopping and the continuation value, respectively.

The function \(s(j, x)\) is nonincreasing and left-continuous in \(x\), and nondecreasing in \(j\). Similarly, \(v(j, x)\) is nondecreasing and right-continuous in \(x\). Discontinuity may only occur if some of the distributions \(F_{j+1}, \ldots, F_n\) have an atom at \(x\), in which case the jump of \(s(j, x)\) is equal to the probability of \(\min\{X_{j+1}, \ldots, X_n\} = x\), and the jump of \(v(j, x)\) is not bigger than that.

Let \(B := \{(j, x): s(j, x) \ge v(j, x)\}\), seen as a subset of \([1, \ldots, n] \times \mathbb{R}\). The optimality principle for problems with finitely many decision steps (Chow et al. 1991, Theorem 3.2) entails that the stopping time

\[
\tau_n := \min\{j \le n: X_j = M_j, (j, X_j) \in B\}
\]

\((\min \emptyset = \infty)\) is optimal.

Define a threshold \(b_j = \sup\{x: s(j, x) \ge v(j, x)\}\) for \(j < n\), and set \(b_n = \infty\). For \(j < n\) the threshold is finite, because for \(x \to \infty\) it holds that \(s(j, x) \to 0\) and \(\liminf v(j, x) \ge (n - j)^{-1}\) (this bound is the limit probability of success by selecting from \(X_{j+1}, \ldots, X_n\) at random). By the monotonicity, for \(j < n\) we have \((j, x) \in B\) if \(x < b_j\), and \((j, x) \notin B\) if \(x > b_j\). Care is needed regarding the edge point \((j, b_j)\), which may or may not belong to \(B\), as the next example illustrates.

**Example** For \(n = 2\) let \(X_1, X_2\) be i.i.d. with c.d.f.

\[
F(x) = \begin{cases} 
  x, & \text{for } 0 \le x < x_0, \\
  x + \delta, & \text{for } x_0 \le x \le 1 - \delta
\end{cases}
\]

for some \(0 < x_0 < x_0 + \delta < 1\). This is a mixture of the uniform distribution on \([0, 1 - \delta]\) and the Dirac mass at \(x_0\). Clearly, \(s(1, x) = 1 - F(x-0)\) and \(v(1, x) = F(x)\). Under the optimal strategy, the condition for stopping at \(X_1 = x\) is \(F(x) \le 1/2\) if \(x \neq x_0\), and \(F(x_0-0) = x_0 \le (1 - \delta)/2\) if \(x = x_0\). Now, for \(x_0 = 1/8, \delta = 1/2\) the stopping condition is \(x \le 1/8\), and for \(x_0 = 1/2, \delta = 1/4\) it is \(x < 1/2\). In both cases the threshold is \(b_1 = x_0\).

The following example showing an extreme possibility is the ‘Bernoulli pyramid’ of Hill and Kennedy (1992).

**Example** Suppose \(X_1 = 1\), and that for \(j = 2, \ldots, n\) the independent observations are two-valued with \(\mathbb{P}(X_j = 1/j) = 1 - \mathbb{P}(X_j = j) = p\). There are no ties and the events \(X_j = M_j\) are independent. From this one computes readily that \(s(j, x) = (1 - p)^{n-j}\) and
\(v(j, x) = (n-j)p(1-p)^{n-j-1}\), hence it is optimal to stop if \(1-p \geq (n-j)p\). That is to say, \(b_j = -\infty\) for \(j < n - (1-p)/p\) and \(b_j = \infty\) for \(j \geq n - (1-p)/p\). The worst-case scenario for the observer is the parameter value \(p = 1/n\), when the optimal best-choice probability is \(v_n = (1-1/n)^{n-1}\).

The lower bound \((1-1/n)^{n-1}\) for the success probability also holds if the data are sampled from a continuous distribution but the observation order is controlled by adversary (Gnedin and Krengel 1995). Recently, Nuti (2020) showed the lower bound \(v_n\) for the model where the observed sequence is a random permutation of independent \(X_j\)’s drawn from continuous \(F_j\)’s.

**Proposition 1** The thresholds defining the optimal stopping time \(\tau_n\) satisfy

\[
b_1 \leq \cdots \leq b_{n-1} < b_n = \infty.
\]

For the extended running minimum process \(((j, M_j), j \leq n)\) the set \(B\) is closed (no-exit).

**Proof** Arguing by backward induction suppose \(b_{j+1} \leq \cdots \leq b_{n-1}\) and that \(B\) is no-exit for the paths entering on step \(j+1\) or later. If \(1 - F_{j+1}(b_{j+1}) = 0\) then obviously \(s(j, x) = 0\) for \(x > b_{j+1}\), whence \(b_j \leq b_{j+1}\). Suppose \(1 - F_{j+1}(b_{j+1}) > 0\). For \(x > b_{j+1}\) a continuity point of \(F_{j+1}\) we have

\[
v(j, x) \geq [1 - F_{j+1}(x)]v(j+1, x) + \int_{-\infty}^{x} s(j+1, y)dF_{j+1}(y) \geq s(j+1, x) \geq s(j, x),
\]

and \(v(j+1, x) > s(j+1, x)\) implies that the inequality is strict whenever \(F_{j+1}(x) < 1\), hence \(b_j \leq x\). Letting \(x \downarrow b_{j+1}\) gives \(b_j \leq b_{j+1}\), which by the virtue of natural monotonicity of the running minimum yields the induction step if either \(b_j < b_{j+1}\), or \(b_j = b_{j+1}\) and \(s(j, b_j) < v(j, b_j)\).

It remains to exclude the possibility \(b_j = b_{j+1}\) when stopping at record \(X_k = b_j\) is optimal for \(k = j\) and not optimal for \(k = j+1\). But in the case \(b_j = b_{j+1}\) and \(s(j, b_j) \geq v(j, b_j)\), for \(z = b_j\) we obtain

\[
s(j+1, z) \geq s(j, z) \geq v(j, z) = [1 - F_{j+1}(z)]v(j+1, z) + [F_{j+1}(z) - F_{j+1}(z-)] \max\{s(j+1, z), v(j+1, z)\} + \int_{-\infty}^{z-} s(j+1, y)dF_{j+1}(y) \geq v(j+1, z),
\]

therefore if stopping at record \(X_j = b_j\) is optimal this also holds for \(X_{j+1} = b_j\). This completes the induction step.

Implicit to the proof is the formula for the continuation value in \(B\)
\[ v(j, x) = \sum_{k=j+1}^{n} \left( \prod_{i=j+1}^{k-1} (1 - F_i(x)) \right) \int_{-\infty}^{x} s(k, y) \, dF_k(y), \quad (j, x) \in B. \] (3)

For arbitrary \( x \), the expression in the right-hand side of (3) gives the probability of success, conditional on \( M_j = x \), when stopping occurs at the first subsequent record (if available) from \( X_{j+1}, \ldots, X_n \). Comparing this with (2) gives a constructive way to determine the thresholds, however the resulting inequality does not simplify (as noticed in the review by Samuels (1991a) on the discrete i.i.d. setting of Scherwentke (1989)).

If the \( X_j \)'s are continuously distributed or, more generally, if the probability of a tie \( X_i = X_j \) for \( i \neq j \) is zero, the selection problem can be formulated as stopping at the last strict record, which is the last change of value in the running minimum process \( M_j, \ j \leq n \). But if ties are possible, there is a nuisance that the coincidence \( M_{j-1} = M_j \) does not distinguish between the no-record event \( X_j > M_j \) and record \( X_j = M_j \). This can be circumvented by considering the (extended) running minimum process as a Markov chain with two types of states, where a state \((j, x)\) encodes the event \( X_j > M_j = x \) (the running minimum is \( x \), no record at index \( j \)), and a state \((j, x)^{\circ} \) stays for the record \( X_j = M_j = x \). The transition probabilities are straightforward in terms of the \( F_j \)'s, for instance a transition from either \((j, x)\) or \((j, x)^{\circ}\) to \((j+1, x)^{\circ}\) occurs with probability \( F_{j+1}(x) - F_{j+1}(x^-) \).

Accordingly, we shall say that a first passage into \( B \) occurs by \emph{jump} if the running minimum enters the set at some record index (in a state \((j, x)^{\circ}\)), and that it occurs by \emph{drift} otherwise (in a state \((j, x)\)). Given such event, the optimal success probability is \( s(j, x) \) if the first passage occurs in \((j, x)^{\circ}\) and it is \( v(j, x) \) given by (3) for \((j, x)\).

### 2.2 Bounds in the i.i.d. case

Suppose \( X_1, \ldots, X_n \) are i.i.d with arbitrary distribution \( F \). By breaking ties with the aid of auxiliary randomisation we shall connect the problem to the i.i.d. continuous case.

It is a standard fact that \( F(X_j) \) has uniform distribution if \( F \) is continuous. In general \( F(X_j) \) has a ‘uniformised’ distribution on \([0, 1]\) which does not charge the open gaps in the range of \( F \), see Gnedin (1997). To spread a mass over the gaps, take \( U_1, \ldots, U_n \) i.i.d. uniform-[0, 1] and independent of \( X_j \)'s to define the random variables

\[ Y_j := F(X_j) - [F(X_j) - F(X_{j^-})]U_j, \] (4)

which are also i.i.d. uniform-[0, 1]. Moreover, they satisfy \( F^{-1}(Y_j) \) \( X_j \) where \( F^{-1} \) is the generalised inverse. By monotonicity, \( Y_j = \min(Y_1, \ldots, Y_n) \) implies \( X_j = M_n = \min(X_1, \ldots, X_n) \), therefore for any stopping time \( \tau \) adapted to the \( Y_j \)'s

\[ \P(Y_{\tau} = \min(Y_1, \ldots, Y_n)) \leq \P(X_{\tau} = M_n) \leq v_n. \]

The second inequality follows since such \( \tau \) is a randomised stopping time, hence cannot improve upon the optimal strategy (Chow et al. 1991). Taking supremum in the left-hand side gives the lower bound \( v_n \leq v_{n'} \).

The probability of a tie for the minimum in sequence \( X_1, \ldots, X_n \) is given by
Running minimum in the best-choice problem

If there is no tie for the minimum, that is \( X_j = M_n \) holds for exactly one \( j \leq n \), then \( X_j = M_n \) implies \( Y_j = \min\{Y_1, \ldots, Y_n\} \). Thus we obtain

\[
P(X_\tau = M_n) - \delta_n \leq P(X_\tau = M_n, \text{ no tie}) \leq P(Y_\tau = \min\{Y_1, \ldots, Y_n\}) \leq v_n
\]

for each \( \tau \) adapted to the \( X_j \)'s (hence also adapted to the \( Y_j \)'s). Choosing the optimal \( \tau = \tau_n \) in the left-hand side gives \( v_n \leq v_n + \delta_n \). To summarise, with the account of monotonicity of \( v_n \)'s (Samuels 1991b) we have the following result.

**Proposition 2** In the i.i.d. case \( v_n \leq v_n \leq v_n + \delta_n \), hence \( v_n > v \) implies the universal sharp bound \( v_n > v \).

The probability of a tie for the first place (maximum) is a thoroughly studied subject (Baryshnikov et al. 1995). With the obvious adjustment these results are applicable to infer about possible asymptotic behaviours of (5).

### 2.3 Poisson limit

Let \( N \) be a Poisson point process in \([0, T) \times \mathbb{R}\) with \( 0 < T \leq \infty \) and some nonatomic intensity measure \( \mu \), which satisfies

\[
\mu([t] \times \mathbb{R}) = 0, \quad 0 \leq t < T,
\mu([t, T) \times [0, \infty)) = \infty, \quad 0 < t < T,
\mu([0, t) \times [0, \infty)) = \infty, \quad 0 < t < T,
\mu([0, T) \times (-\infty, x)) < \infty, \quad x \in \mathbb{R}.
\]

The generic point \((t, x)\) of \( N \) is thought of as mark \( x \) observed at time \( t \). This is regarded as a strong record if \( \mathcal{N}([0, t) \times [0, x]) = 1 \) (point \((t, x)\) itself), and a weak record if \( \mathcal{N}([0, t) \times [0, x]) = 0 \). The first assumption excludes multiple points, but several points of \( N \) can share the same mark \( x \) if the measure is singular with \( \mu([0, T) \times \{x\}) > 0 \). Furthermore, there is a well defined running minimum process \( (Z_\tau, \tau \in [0, T]) \), where \( Z_\tau \) is the minimal mark observed within time interval \([0, \tau]\). The information of observer accumulated by time \( t \) is the configuration of \( N \)-atoms on \([0, t] \times \mathbb{R}\). The task is to stop with the highest possible probability at observation with the minimal mark \( Z_T := \lim_{t \to T} Z_t \).

We refer to Faller and Rüschendorf (2012) for more formal and complete treatment of the optimal stopping problem under slightly different assumptions on \( \mu \). In particular, they only consider the case \( T = 1 \), but this is not a substantial constraint, because increasing transformations of scales do not really change the problem. Part (a) of their Theorem 2.1 implies that there exists a nondecreasing function \( b : [0, T) \to (-\infty, \infty) \) such that it is optimal to stop at the first record falling in the domain \( B = \{(t, x) : x \leq b(t)\} \). Equation (2.7) of Faller and
Rüschendorf (2012) gives a multiple integral expression for the probability of success under the optimal stopping time.

A connection with the discrete time stopping problem is the following. Let $X_1, \ldots, X_n$ be independent, possibly with different distributions that may depend on $n$. Consider the scatter of $n$ points $\{(1, X_1), \ldots, (n, X_n)\}$, subject to a suitable monotone coordinate-wise scaling, as a finite point process $N_0$ on the plane, and suppose that $N_n$ converges weakly to $N$. Then, by part (b) of the cited theorem from Faller and Rüschendorf (2012), $v_n$ converges to the optimal stopping value for $N$. Part (c) asserts that stopping on the first record of $N_n$ falling in $B$ is asymptotically optimal for the discrete time problem.

The choice of scaling is dictated by the vague convergence of the intensity measure $N_n(\cdot)$ to a Radon measure on the plane. In the i.i.d. case this is typically a linear scaling from the extreme-value theory (Resnick 2008). See Falk et al. (2011) for examples of scaling for non-i.i.d. data.

3 The classic problem re-visited

3.1 The discrete time problem

Let $X_1, \ldots, X_n$ be i.i.d. uniform-[0, 1]. The ties have probability zero and all records are strong. Consider stopping time of the form

$$\tau = \min\{j \leq n : X_j \leq b_j, X_j = M_j\},$$

where $0 \leq b_1 \leq \cdots \leq b_n \leq 1$ are arbitrary thresholds, not necessarily optimal. We aim to decompose the success probability $\mathbb{P}(X_\tau = M_n)$ according to the distribution of the first passage time for the running minimum

$$\sigma := \min\{j \leq n : M_j \leq b_{j+1}\}.$$

Obviously $\sigma \leq \tau$: namely $\sigma = \tau$ if $\sigma$ is a record time, and otherwise $\sigma < \tau$ and $\tau$ is the first record time (if any) after $\sigma$.

Suppose $\sigma = j, M_j = x$. If $x \leq b_j$ then $M_{j-1} \leq b_j$ is impossible (otherwise we would have $\sigma < j$), hence $M_{j-1} > b_j$ implying that $\tau = j$ is a record time; we qualify this event as a passage by jump. Alternatively, in the case $b_j < x \leq b_{j+1}$ the passage is by drift and $\tau > j$ coincides with the first subsequent record time (if any). Accordingly, we have for the joint distribution of $\sigma$ and $M_\sigma$

$$\mathbb{P}(\sigma = j, M_j \in [x, x + dx]) = \begin{cases} (1 - b_j)^{j-1} \, dx, & 0 \leq x \leq b_j, \\ j(1 - x)^{j-1} \, dx, & b_j < x \leq b_{j+1}. \end{cases}$$

In the first case the (conditional) success probability is $(1 - x)^{n-j}$, and in the second case

$$\sum_{k=j+1}^{n} (1 - x)^{k-j-1} \int_{0}^{x} (1 - y)^{n-k} \, dy = \sum_{k=1}^{n-j} \frac{(1 - x)^{n-j-k} - (1 - x)^{n-j}}{k}.$$

\(\circled{\text{Springer}}\)
Integrating yields
\[
\mathbb{P}(\sigma = j, X_\tau = M_n) = \frac{(1 - b_j)^{n-j-1} - (1 - b_j)^n}{n-j+1} + \sum_{k=j}^{n-1} \sum_{i=1}^{j} \left[ \frac{(1 - b_j)^k - (1 - b_{j+1})^k}{k(n-k)} - \frac{(1 - b_j)^n - (1 - b_{j+1})^n}{n(n-k)} \right].
\]
This can be compared with Equation (3c-1) for \(\mathbb{P}(\tau = j, X_j = M_n)\) in Gilbert and Mosteller (1966) (where \(d_j\) is our \(1 - b_j\)). Summation gives a new formula for the success probability
\[
\mathbb{P}(X_\tau = M_n) = \frac{1}{n} \left[ 1 - \sum_{j=1}^{n} (1 - b_j)^n \right] + \sum_{j=1}^{n-1} \sum_{i=1}^{j} \left[ \frac{(1 - b_j)^i}{j(n-j)} - \frac{(1 - b_j)^n}{n(n-j)} \right],
\]
with two parts corresponding to the passage by jump and drift.

The optimal threshold \(b_j\) is a solution to
\[
\sum_{i=1}^{n-j} \frac{(1 - x)^{i-1} - 1}{i} = 1, \quad x \in [0, 1].
\]
Using this Sakaguchi (1973) obtained a nice formula
\[
v_n = \frac{1}{n} \left( 1 + \sum_{j=1}^{n-1} \sum_{k=j}^{n} \frac{(1 - b_j)^k}{j} \right),
\]
which also gives the mean \(\mathbb{E}[\tau_n/n]\) for the optimal \(\tau_n\), as was shown by Tamaki (2015).

### 3.2 Poissonisation

Let \(N\) be a planar Poisson process in \([0, 1] \times [0, \infty)\) with the Lebesgue measure as intensity. It is well known and easy to verify that for \(X_1, \ldots, X_n\) i.i.d. uniform-[0, 1], the planar point process with atoms \((j/n, nX_j), j \leq n\), converges weakly to \(N\).

There exists a still closer connection through embedding of the finite sample in \(N\) (Gnedin 1996). Consider instead of the uniform distribution the mean-exponential. A sample from this can be implemented by splitting \([0, 1]\) evenly in \(n\) subintervals, and identifying the \(j\)th point of \(N_n\) with the point of \(N\) having the minimal mark among the arrivals within the time interval \([(j-1)/n, j/n)\). By this coupling the inequality \(v_n > v\) becomes immediate by noting that the stopping problem for \(N_n\) is equivalent to a choice from \(N\) where the observer at time \(j/n\) has the power to revoke any arrival within \([(j-1)/n, j/n)\).

The running minimum \((Z_t, t \in [0, 1])\) is defined to be the lowest mark of arrival on \([0, t]\). With every point \((t, x) \in [0, 1] \times [0, \infty)\) we associate a rectangular box \([[t, 1] \times [0, x])\) to the south-east with the corner point \((t, x)\) excluded. If there is an atom of \(N\) at location \((t, x)\) we regard it as record if there are no other atoms.
south-west of \((t, x)\), and it is the last record (with mark \(Z_t\)) if the box contains no atoms. (See Fig. 1.)

An admissible strategy in this setting is a stopping time which takes values in the random set of record times and may also assume value 1 (the event of no selection). With every nondecreasing continuous function \(b : [0, 1] \rightarrow [0, \infty), b(1-) = \infty\), we associate a stopping time

\[
\tau = \inf\{t : (t, Z_t) \text{ is record }, Z_t \leq b(t)\},
\]

where \(\inf \emptyset = 1\). The associated first passage time into \(B\) is defined as

\[
\sigma = \inf\{t : Z_t \leq b(t)\}.
\]

Clearly, \(\sigma\) is adapted to the natural filtration of \(N\), \(\sigma < 1\) a.s. and \(\sigma \leq \tau\). However, \(\sigma\) is not admissible, because in the event \(Z_\sigma = b(\sigma)\) of the boundary crossing by drift, there is arrival at time \(\sigma\) with probability zero.

The modes of the first passage, jump and drift, can be distinguished in geometric terms. Move the rectangular frame spanned on \((0, 0)\) and \((t, b(t))\) until one of the sides meets a point of \(N\). If the point falls on the eastern side of the frame, the running minimum crosses the boundary by a jump and \(\sigma = \tau\), if the point appears on the northern side, the boundary is hit by a drift and \(\sigma < \tau\). See Figs. 2 and 3.

We aim to express the success probability of \(\tau\) in terms of \((\sigma, Z_\sigma)\). A key observation which leads to explicit formulas is the self-similarity property: two boxes with the same area \(z\) can be mapped to one another by a bijection which preserves both

---

**Fig. 1** Records in a planar Poisson process in \([0, 1] \times [0, \infty)\):

- - record (no Poisson atoms south-west of it).
- - the last record (its box contains no Poisson atoms)
Running minimum in the best-choice problem

measure and coordinate-wise order. Consider a box with apex at \((t, x)\) hence size \(z = (1 - t)x\). Then

(i) if stopping occurs on a record at \((t, x)\) it is successful with probability \(e^{-z}\),

(ii) if stopping occurs on a record at \((t, Ux)\), for \(U\) distributed uniformly on \([0, 1]\), it is successful with probability

\[
J(z) := \int_{0}^{1} e^{-\alpha u} \, du = \frac{1 - e^{-z}}{z},
\]

Fig. 2 Event \(\sigma = \tau\), the boundary \(b(t)\) is crossed by a jump, hence the record is caught by the eastern side of a frame spanned on \((0, 0)\) and \((t, b(t))\)

Fig. 3 Event \(\sigma < \tau\) - the boundary \(b(t)\) hit by a drift. Record caught by the northern side of a frame spanned on \((0, 0)\) and \((t, b(t))\)
(iii) if stopping occurs at the earliest arrival inside the box (if any) it is successful with probability

\[ D(z) := \int_0^z e^{-s} J(z - s) \, ds = e^{-z} \int_0^z \frac{e^s - 1}{s} \, ds. \]

These formulas are most easily derived for standard box \([0, z] \times [0, 1]\).

Now, the running minimum crosses the boundary \(b\) at time \([t, t + dt]\) by jump (hence \(\sigma = \tau\)) if there are no Poisson atoms south-west of the point \((t, b(t))\), and there is an arrival below \(b(t)\). Given such arrival, the distribution of the record value \(Z_t\) is uniform on \([0, b(t)]\), therefore this crossing event contributes to the success probability

\[ e^{-tb(t)} b(t) J((1 - t)b(t)) \, dt. \]

Alternatively, \((Z_t)\) drifts into the stopping domain at time \([t, t + dt]\) (hence \(\sigma < \tau\)) and \(\tau\) wins with the next available record with probability

\[ e^{-tb(t)} t(b(t + dt) - b(t)) D((1 - t)b(t)). \]

We write the success probability with \(\tau\) as a functional of the boundary

\[ \mathcal{P}(b) := \int_0^1 e^{-tb(t)} b(t) J((1 - t)b(t)) \, dt + \int_0^\infty e^{-tb(t)} t D((1 - t)b(t)) \, db(t). \quad (7) \]

Note that the distribution of \(\sigma\) is given by

\[ \mathbb{P}(\sigma \leq t) = \int_0^t e^{-sb(s)} b(s) \, ds + \int_0^{b(t)} e^{-sb(s)} s \, db(s), \]

where the terms correspond to two types of boundary crossing.

We may view maximising the functional (7) as a problem from the calculus of variations. Recalling that the box area at record arrival is the only statistic which matters, suggests to try the hyperbolic shapes

\[ b(t) = \frac{\beta}{1 - t}. \quad (8) \]

Indeed, equating (i) and (iii), \(e^{-z} = D(z)\), we see that the balance between immediate stopping and stopping at the next record is achieved for \(\beta^* = 0.804352\) solving the equation

\[ \int_0^z \frac{e^s - 1}{s} \, ds = 1. \]

The optimal stopping time is defined by domain \(B\) with the hyperbolic boundary and \(\beta^*\) (see Fig. 4), because \(B\) is no-exit domain for the running minimum process, hence the monotone case of optimal stopping (Chow et al. 1991) applies.

For (8), the functional simplifies enormously, and the success probability becomes
Running minimum in the best-choice problem

Finally, for the optimal $\beta^*$ we get with the account of $D(\beta^*) = \exp(-\beta^*)$

$$\sup_b \mathcal{P}(b) = e^{-\beta^*} + (e^{\beta^*} - 1 - \beta^*) \int_{\beta^*}^{\infty} e^{-s} \frac{ds}{s} = 0.580164$$

which is the formula due to Samuels (1991b).

Historically, the first study of the full-information best-choice problem with arrivals by Poisson process was Sakaguchi (1976). In that paper the marks are uniform-$[0, 1]$ and the process runs with finite horizon $T$. To obtain a sensible $T \to \infty$ limit one needs to resort to the equivalent model of Poisson process in $[0, 1] \times [0, T]$. The finite-$T$ problem can then be interpreted as $N$ conditioned on the initial record at point $(0, T)$, then for $T \geq \beta^*$ the optimal success probability is given by the above formula but with the upper limit $T$ in the exponential integral.

4 A uniform triangular model

In the models of this section the background process lives in the domain $x \geq t$. These have some appeal for applications in scheduling, where interval $[t, x]$ represents the time span needed to process a job by a server, and exactly one job is to be
chosen by a stopping strategy. The optimisation task is to maximise the probability of choosing the job with the earliest completion time.

4.1 The discrete time problem

Let \( X_1, \ldots, X_n \) be independent, with \( X_j \) having discrete uniform distribution on \( \{j, \ldots, n\} \). Obviously, we may focus on the states of the running minimum within the lattice domain \( \{(j, x) : j \leq x \leq n\} \).

By Proposition 1 the optimal stopping time is given by a set of nondecreasing thresholds \( b_j \). Stopping at record \( (j, x) \) is successful with probability

\[
s(j, x) = \prod_{i=0}^{x-j-1} \frac{n-x+1}{n-j-i}.
\]

Given the running minimum \( M_j = x \) with \( x \leq b_j \), the continuation value is a specialisation of (3), assuming the form

\[
v(j, x) = \sum_{i=1}^{x-j} \left( \prod_{k=0}^{i-2} \frac{n-x}{n-j-k} \right) \cdot \frac{1}{n-j+1} \cdot \sum_{y=j+i}^{x} s(j+i, y).
\]

The success probability splits in two components, \( v_n = J_n + D_n \), where \( J_n \) results from the running minimum breaking into \( B \) by jump, while \( D_n \) relates to drifting into \( B \). Explicitly,

\[
J_n = \sum_{j=1}^{n} \left( \prod_{k=0}^{j-2} \frac{n-b_j}{n-k} \right) \cdot \frac{1}{n-j+1} \cdot \sum_{x=j}^{b_j} s(j, x)
\]

and

\[
D_n = \sum_{j=2}^{n} \left( \prod_{i=0}^{j-2} \frac{n-b_{j-1}}{n-i} - \prod_{i=0}^{j-2} \frac{n-b_j}{n-i} \right) \cdot \sum_{y=b_{j-1}+1}^{b_j} v(j, y) \cdot \frac{b_j - b_{j-1}}{b_j - b_{j-1}}
\]

\( b_j \) is the biggest \( x \) with \( s(j, x) \geq v(j, x) \) and the latter are given by (9) and (10).

The computed values plotted in Fig. 5 suggest that \( v_n \) monotonically decreases to a limit 0.703128 (check the next subsection for its exact derivation).

4.2 Poissonisation

The right scaling is guessed from the Rayleigh distribution limit

\[
P(M_n > x\sqrt{n}) = \prod_{j\leq x\sqrt{n}} \left( 1 - \frac{j}{n - \lfloor x\sqrt{n} \rfloor + j} \right) \to e^{-x^2/2}, \ x > 0.
\]

 Springer
We truncated the product since $X_j > x\sqrt{n}$ for $j > x\sqrt{n}$. Thus we define $N_n$ to be the point process with atoms $(j/\sqrt{n}, X_j/\sqrt{n})$, $j \leq n$.

Now, we assert that the process $N_n$ converges in distribution to a Poisson process $N$ with unit rate in the sector $\{(t, x) : 0 \leq t \leq x < \infty\}$. A pathway to proving this is the following. For $x > 0$, convergence of the reduced point process of scaled times $\{j/\sqrt{n} : X_j \leq x\sqrt{n}\}$ is established in line with Chapter 9 of Falk et al. (2011): this includes convergence of the mean measure and the avoidance probabilities akin to (11) with $j$ in the bounds $t_1 < j/\sqrt{n} < t_2$. Then the convergence of the planar process $N_n$ restricted to $[0, x] \times [0, x]$ follows by application of the theorem about marked Poisson processes. Sending $x \to \infty$ completes the argument (see Fig. 6).

The best-choice problem for $N$ is very similar to the one in the previous section. Under a box with apex $(t, x)$ we shall understand now the isosceles triangle with one side lying on the diagonal and two other sides being parallel to the coordinate axis. The box area is equal to $z := (x - t)^2/2$. Equal-sized boxes can be mapped to one another by sliding along the diagonal (see Fig. 7).

In these terms, the basic functions are defined as follows:

(i) if stopping occurs on a record at $(t, x)$ it is successful with probability $e^{-z}$,

(ii) if stopping occurs on a record at random location $(t, t + U(x - t))$, for $U$ distributed uniformly on $[0, 1]$, it is successful with probability

$$J(z) := \int_0^1 e^{-zu^2} du = \frac{\sqrt{\pi} \text{erf}(\sqrt{z})}{2\sqrt{z}}$$
Fig. 6  Records in a planar Poisson process above the diagonal $t = x$ of the positive quadrant: $\blacktriangle$ - record (no Poisson atoms south-west of it), $\bigcirc$ - the last record (its box contains no Poisson atoms)

Fig. 7  The linear boundary $b(t) = t + \sqrt{2}\beta^*$
Running minimum in the best-choice problem

(recall that \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt \)),

(iii) if stopping occurs on the earliest arrival inside the box (if any) it is successful with probability

\[
D(z) := \int_0^{\sqrt{2z}} \exp\left( -\sqrt{2z} s + s^2 / 2 \right) \left( \sqrt{2z} - s \right) \, ds = e^{-z} \int_0^{\sqrt{2z}} \int_0^u e^{(u^2-v^2)/2} \, dv \, du.
\]

The boundaries that come in question in this problem are non-decreasing functions \( b : [0, \infty) \rightarrow [0, \infty) \) that satisfy \( b(t) \geq t \). The analogue of (7) becomes

\[
P(b) := \int_0^{\infty} e^{-tb(t)+t^2/2}(b(t) - t) J[(b(t) - t)^2/2] \, dt + \int_0^{\infty} e^{-tb(t)+t^2/2} t D[(b(t) - t)^2/2] \, db(t).
\]

In the view of self-similarity the maximiser should be a linear function

\( b(t) = t + \sqrt{2\beta} \),

and then the success probability simplifies as

\[
P(b) = D(\beta) + (J(\beta) - D(\beta)) P(\sigma = \tau)
\]

\[
= D(\beta) + (J(\beta) - D(\beta)) \sqrt{2\beta} \int_0^{\infty} \exp\left( -\frac{t^2}{2} - \sqrt{2\beta} t \right) \, dt
\]

\[
= D(\beta) + (J(\beta) - D(\beta)) \sqrt{\pi \beta} \, e^\beta \, \text{erfc}(\sqrt{\beta})
\]

(recall that \( \text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} \, dt \)). Equation \( e^{-z} = D(z) \) becomes

\[
\int_0^{\sqrt{2z}} \int_0^u e^{(u^2-v^2)/2} \, dv \, du = 1,
\]

which by monotonicity has a unique solution

\( \beta^* = 0.760660 \).

The stopping strategy with boundary \( b(t) = t + \sqrt{2\beta^*} \) is overall optimal, and yields the success probability

\[
\sup_b P(b) = e^{-\beta^*} + \left( \frac{e^{\beta^*} \sqrt{\pi}}{2 \sqrt{\beta^*}} \text{erf}(\sqrt{\beta^*}) - 1 \right) \sqrt{\pi \beta^*} \, \text{erfc}(\sqrt{\beta^*})
\]

\[
= 0.703128,
\]

confirming the limit we obtained numerically.
4.3 The box-area jump chain and extensions

The limit (12) has appeared previously in a context of generalised records from partially ordered data (Gnedin 2007). The source of coincidence lies in the structure of the one-dimensional box area process associated with the running minimum. This is an interesting connection deserving some comments.

Consider a piecewise deterministic, decreasing Markov process \( P \) on \( \mathbb{R}_+ \), which drifts to zero at unit speed and jumps at unit rate. When the jump occurs from location \( z \), the new state is \( zY \), where \( Y \) is random variable with given distribution on \((0, 1)\). The state 0 is terminal. Thus if \( P \) starts from \( z > 0 \), in one drift-and-jump cycle the process moves to \((z - E)_+ Y\), where \( E \) is independent exponential random variable. The associated optimisation problem amounts to stopping at the last state before absorption.

A process of this kind describes a time-changed box area associated with the running minimum. For the Poisson process of Sect. 3.2, the variable \( Y \) is uniform-[0, 1], and in the triangular model it is beta(1/2, 1). Two different modes of the first passage by the running minimum occur when \( P \) enters \([0, \beta^*]\) by drift or by jump, where \( \beta^* \) is the optimal parameter of the boundary.

More generally, for \( Y \) following beta \((\theta, 1)\) distribution, Equation (9) from Gnedin (2007) gives the success probability as

\[
\mathcal{P}(\beta^*) = \Gamma(-\theta + 1, \beta^*, \infty) \left(-\beta^* \theta + e^{\beta^* \theta} \Gamma(\theta, 0, \beta^*)\right) + e^{-\beta^*},
\]

(13)

where

\[
\Gamma(a, b, c) = \int_b^c e^{-t} t^{a-1} dt.
\]

One can verify analytically that for \( \theta = 1/2 \) the formula agrees with our (12). Indeed, (13) specialises as

\[
\mathcal{P}(\beta^*) = \Gamma(1/2, \beta^*, \infty) \left(-\sqrt{\beta^*} + e^{\beta^*} \theta \Gamma(\theta, 0, \beta^*)\right) + e^{-\beta^*},
\]

(14)

so observing

\[
\Gamma(1/2, \beta^*, \infty) = \int_{\beta^*}^{\infty} e^{-t} t^{1/2} dt = 2 \int_{\beta^*}^{\infty} e^{-x^2} dx = \sqrt{\pi} \text{erfc} \left(\sqrt{\beta^*}\right)
\]

(15)

and similarly

\[
\Gamma(1/2, 0, \beta^*) = \sqrt{\pi} \text{erf} \left(\sqrt{\beta^*}\right)
\]

(16)

we obtain (12) by substituting (15) and (16) into (14)

We also considered other processes of independent observations with linear trend that give the same limit best-choice probability (12):
(i) \( X_1, \ldots, X_n \) independent, with \( X_j \) distributed uniformly on \( \{j, \ldots, j+n-1\} \). Here again the limit distribution is Rayleigh, \( \mathbb{P}[M_n > x\sqrt{n}] \to e^{-x^2/2} \), and the point process with atoms \((j/\sqrt{n}, X_j/\sqrt{n})\) converges in distribution to a Poisson process \( N \) with unit rate in the sector \( \{(t,x) : 0 \leq t \leq x < \infty\} \).

(ii) \( X_j = j + \rho n U_j \), where \( \rho > 0 \) is a parameter and \( U_1, U_2, \ldots \) are i.i.d. uniform-[0, 1]. This time \( \mathbb{P}[M_n > x\sqrt{\rho n}] \to e^{-x^2/2} \) and the point process with atoms \((j/\sqrt{\rho n}, X_j/\sqrt{\rho n})\) converges weakly to the same Poisson \( N \).

On the other hand,

(iii) \( X_j = j + n U_j^1/\theta \), where \( \theta > 0 \) is a parameter and \( U_1, U_2, \ldots \) are i.i.d. uniform-[0, 1], leads to (13). Here \( \mathbb{P}[M_n > x\sqrt{n/\theta}] \to e^{-x^2/(\theta + 1)} \) which is a Weibull distribution with shape parameter \((\theta + 1)\) and scale parameter \((\theta + 1)^{\frac{1}{\theta + 1}}\). The point process with atoms \((j/n^{\frac{1}{\theta + 1}}, X_j/n^{\frac{1}{\theta + 1}})\) converges weakly to the Poisson process which is not homogeneous, rather has intensity measure \( \theta(x-t)^{\theta-1}dtdx, \ 0 \leq t \leq x \).

5 A uniform rectangular model

According to Proposition 2, the limit best choice probability for i.i.d. observations is \( v \), provided the probability of a tie for the sample minimum approaches 0 as \( n \to \infty \). For fixed, not depending on \( n \), discrete distribution this may or may not be the case. Moreover, when \( (-X_j)'s \) are geometric, the probability of a tie does not converge, but undergoes tiny fluctuations (Kirschenhofer and Prodinger 1996); in this setting one can expect that the best choice probability has no limit as well. In this section we consider a discrete uniform distribution, and achieve a positive limit probability of a tie for the sample minimum by letting the support of the distribution to depend on \( n \).

5.1 The discrete time problem

Let \( X_1, \ldots, X_n \) be independent, all distributed uniformly on \( \{1, \ldots, n\} \). The generic state of the running minimum is a pair \((j, x)\), where \( j, x \in \{1, \ldots, n\} \). In this setting the probability of a tie for a particular value does not go to 0 with \( n \to \infty \). In particular, the number of 1’s in the sequence of \( n \) observations is Binomial\((n, 1/n)\), hence approaching the Poisson\((1)\) distribution, so the strategy which just waits for the first 1 to appear succeeds with probability \( 1 - (1 - 1/n)^n \to 1 - 1/e = 0.632120 \), which already exceeds noticeably the universal sharp bound \( v = 0.580164 \) of Proposition 2.

Again, by Proposition 1 the optimal stopping time is determined by a set of non-decreasing thresholds \( b_1 \leq \cdots \leq b_n = \infty \). Stopping at record \((j, x)\) is successful with probability

\[
s(j, x) = \left( \frac{n-x+1}{n} \right)^{n-x}.
\]
Conditionally on the running minimum \( M_j = x \) with \( x < b_j \), the continuation value given by (3) reads as

\[
v(j, x) = \sum_{i=1}^{n-t} \left( \frac{n-x}{n} \right)^{j-1} \frac{1}{n} \sum_{y=1}^{x} s(t + i, y).
\]

The success probability may be again decomposed into terms \( v_n = J_n + D_n \), referring to the running minimum entering \( B \) by jump or by drift, respectively. We get

\[
J_n = \sum_{j=1}^{n} \left( \frac{n-b_j}{n} \right)^{j-1} \frac{1}{n} \sum_{x=1}^{b_j} s(j, x),
\]

and

\[
D_n = \sum_{j=2}^{n} \left[ \left( \frac{n-b_{j-1}}{n} \right)^{j-1} - \left( \frac{n-b_j}{n} \right)^{j-1} \right] \sum_{y=b_{j-1}+1}^{b_j} \frac{v(j, y)}{b_j - b_{j-1}},
\]

where \( b_j \) is defined as the biggest \( x \) with \( s(j, x) \geq v(j, x) \), and these are given by (9) and (10).

The computed values, as presented in Fig. 8, suggest that \( v_n \) decreases monotonically to a limit 0.761260. Using the Poisson approximation we shall obtain an explicit expression in terms of the roots of certain equations.

### 5.2 Poissonisation

The point process \((j/n, X_n)\) converges to a Poisson process on \([0, 1] \times \mathbb{Z}_{\geq 0}\) with the intensity measure being the product of Lebesgue measure and the counting measure on integers. Hence to find the limit success probability we may work directly with the setting of this Poisson process. We prefer, however, to stay within the continuous framework of previous sections, and to work with the planar Poisson point process \(N\) in \([0, 1] \times [0, \infty)\) with the Lebesgue measure as intensity.

To that end, we just modify the ranking order. Let \(X_1, \ldots, X_n\) be i.i.d. uniform-\([0, n]\). Two values with \([X_j] = [X_i]\) will be treated as order-indistinguishable. In particular, we call \(X_j\) a (weak) record if \([X_j] \leq [X_i]\) for all \(i < j\). For \(n\) large, the distribution of \([M_j]\) is close to Geometric\((1 - 1/e)\).

Now, the planar point process with atoms \((j/n, X_j), j \leq n\), converges in distribution to \(N\). The running minimum \((Z_t, t \in [0, 1])\) is the lowest mark of arrival on \([0, t]\). Marks \(x, y\) with the same integer ceiling \([x] = [y]\) will be considered as order-indistinguishable. Accordingly, arrival \((t, x)\) is said to be a (weak) record if \([0, t] \times [0, [x]]\) contains no Poisson atoms. The role of a box is now played by the rectangle \([t, 1] \times [0, [x]]\).

The basic functions are defined as follows:

(i) if stopping occurs on a record at \((t, x)\) with \([x] = k\) it is successful with probability
(ii) if stopping occurs on the earliest arrival (if any) inside the box $[t, 1] \times [0, \lceil x \rceil)$ with $\lceil x \rceil = k$ it is successful with probability
\[
\int_t^1 e^{-(k-1)(1-t)} \sum_{j=1}^k e^{-(j-1)(1-t)} ds = e^{-k(1-t)} \sum_{j=1}^k \frac{e^{(1-t)} - 1}{j}.
\]

For $k = 1$ stopping is optimal for all $t$; we set $t_1 = 0, z_1 = e$ and for $k \geq 2$ the equality is achieved for $t_k$ defined to be the root of equation
\[
e^{-(k-1)(1-t)} = e^{-k(1-t)} \sum_{j=1}^k \frac{e^{(1-t)} - 1}{j}.
\]
Letting $z_k := e^{1-h_k}, z_k$ is a solution to
\[
\sum_{j=2}^k \frac{e^j}{j} = h_k, \quad h_k := \sum_{j=1}^k \frac{1}{j}.
\]

By monotonicity there exists a unique positive solution, and the roots are decreasing, so that $z_1 = e$ and $z_k \downarrow 1$.

It follows that the optimal stopping time is
\[\tau = \inf \{ t : (t, Z_t) \text{ is a record, such that } t \geq 1 - \log (z_k) \text{ for } k = \lceil Z_t \rceil \}.\]
That is, the stopping boundary is

\[ b(t) = \sum_{k=1}^{\infty} k 1(t_k < t \leq t_{k+1}) = \sum_{k=1}^{\infty} k 1(z_{k+1} \leq e^{1-t} < z_k). \]

The cutoffs are readily computable from (17), for instance

\[ z_1 = e = 2.718281, \]
\[ z_2 = \sqrt{3} = 1.732050, \]
\[ z_3 = 1.381554, \]
\[ z_4 = 1.258476, \]
\[ z_5 = 1.195517, \]
\[ z_{10} = 1.088218, \]
\[ z_{15} = 1.056969, \]
\[ z_{20} = 1.042069. \]

The associated hitting time for the running minimum is

\[ \sigma = \inf\{t : [Z_t] \leq b(t)\}. \]

The success probability again decomposes in terms corresponding to the events \( \sigma < \tau \) and \( \sigma = \tau \). The first term related to jump through the boundary becomes

\[ J := \sum_{k=1}^{\infty} \int_{t_k}^{t_{k+1}} e^{-kt} \sum_{j=1}^{k} e^{-(j-1)(1-t)} dt \]
\[ = \sum_{k=1}^{\infty} \sum_{j=1}^{k} \frac{e^{-k}}{j} (e^{(1-t_k)} - e^{(1-t_{k+1})}) \]
\[ = \sum_{k=1}^{\infty} \sum_{j=1}^{k} \frac{e^{-k}}{j} (z_k^j - z_{k+1}^j) \]
\[ = e^{-1}(z_1 - z_2) + \sum_{k=2}^{\infty} e^{-k} \left( z_k - z_{k+1} + \frac{z_{k+1}^{k+1} - 1}{k+1} \right), \]

where for the last equality we used (17) in the form

\[ \sum_{j=1}^{k} \frac{z_k^j}{j} = h_k + z_k, \]
\[ \sum_{j=1}^{k} \frac{z_{k+1}^j}{j} = h_{k+1} + z_{k+1} - \frac{z_{k+1}^{k+1}}{k+1}. \]
Note that if the ceiling of the running minimum $[Z]$ drifts into the boundary point $(t_k, k)$, the optimal success probability from this time on is the same as from stopping as if a record occurred at $(t_k, k)$. Hence the contribution of the event $\sigma < \tau$ becomes
\[
D := \sum_{k=2}^{\infty} (e^{-(k-1)t_k} - e^{-kt_k}) e^{-(k-1)(1-t_k)} = \sum_{k=2}^{\infty} e^{-k} (e - z_k).
\]

Putting the parts together, the optimal success probability after some cancellation and series work becomes
\[
\mathbb{P}(Z_{\tau} = \min_{t \in (0, 1)} Z_t) = J + D
\]
\[
= 2 - e + \frac{1}{e-1} + \frac{1 - 2\sqrt{3}}{2e} + e \log(e-1) - \sum_{k=2}^{\infty} e^{-k} \left( z_{k+1} - \frac{z_{k+1}^{k+1}}{k+1} \right)
\]
\[
= 0.761260.
\]

The general boundary For the general boundary defined by nondecreasing cut-offs $t_k$, the jump term
\[
J := \sum_{k=1}^{\infty} \sum_{j=1}^{k} \frac{e^{-k}}{j} (z_k^j - z_{k+1}^j)
\]
should be computed with $z_k = e^{1-t_k}$, and the drift term written as
\[
D := \sum_{k=2}^{\infty} (e^{-(k-1)t_k} - e^{-kt_k}) \sum_{j=1}^{k} j^k e^{(1-t_k)} - \frac{1}{j}
\]
\[
= \sum_{k=2}^{\infty} e^{-k} (e^{1} - 1) \sum_{j=1}^{k} \frac{e^{(1-t_k)} - 1}{j}.
\]

For instance, letting $t_k = 1$ for $k \geq 3$, the maximum success probability is 0.730694 achieved at $t_2 = 0.450694$.

5.3 Varying the intensity of the Poisson process

The extension presented in this section constitutes a smooth transition between the above poissonised rectangular model and the full-information game from Sect. 3.2. As above, consider a homogeneous Poisson process on $[0, 1] \times [0, \infty)$, and treat values $x, y$ with $[x] = [y]$ as order-indistinguishable, but now suppose the intensity of the process is some $\lambda > 0$. Note that as $\lambda \to 0$ the ties vanish hence the best-choice probability becomes close to $\nu = 0.580164$ from the full-information game.

This process relates to a limit form of the discrete time best-choice problem, with observations $X_1, \cdots, X_n$ drawn from the uniform distribution on $\{1, \cdots, K_n\}$ where $K_n \sim n/\lambda$. See Falk et al. (2011) (Example 8.5.2) and Kolchin (1969) for the related extreme-value theory. Here, for $n$ large, the distribution of $M_n$ is close to Geometric$(1 - (1/e)^4)$. The scaling dictated by convergence to the Poisson limit is $(j/n, \lambda X_j)$. 

 Springer
Following the familiar path, we compare stopping on a record \((t, x)\), for given \([x] = k\), with stopping on the next available record. For \(k = 1\) stopping is the optimal action for all \(t\). We set \(t_1 = 0\) and \(z_1^{(\lambda)} := e^\lambda\). For \(k \geq 2\), whenever a positive solution to
\[
\sum_{j=2}^{k} \frac{z_j^{k}}{j} = h_k, \quad h_k := \sum_{j=1}^{k} \frac{1}{j} \tag{18}
\]
is smaller than \(e^\lambda\), we define \(z_k^{(\lambda)}\) to be this solution and set \(z_k^{(\lambda)} := e^\lambda(1-t_k)\). Otherwise, we set \(z_k^{(\lambda)} = e^\lambda\) (which corresponds to setting the threshold \(t_k\) to 0). The optimal stopping time is thus given by
\[
\tau = \inf \left\{ t : (t, Z_t) \text{ is a record with } t \geq 1 - \frac{\log(z_k^{(\lambda)})}{\lambda} \text{ for } k = [Z_t] \right\}.
\]
Equivalently, the stopping boundary is
\[
b(t) = \sum_{k=1}^{\infty} k(1(t_k < t \leq t_{k+1})) = \sum_{k=1}^{\infty} k(z_{k+1}^{(\lambda)} \leq e^\lambda(1-t) < z_k^{(\lambda)}).
\]
The optimal success probability decomposes into the jump and drift terms:
\[
\mathbb{P}(Z_\tau = \min_{t \in [0,1]} Z_t) = J_\lambda + D_\lambda,
\]
where
\[
J_\lambda = \sum_{k=1}^{\infty} \sum_{j=1}^{k} e^{-\lambda j} \left( (z_k^{(\lambda)})^j - (z_{k+1}^{(\lambda)})^j \right),
\]
\[
D_\lambda = \sum_{k=0}^{\infty} e^{-\lambda k}(e^\lambda - z_k^{(\lambda)}).
\]
The numerical values of the best choice probability are plotted in Fig. 9.

![Fig. 9](image-url) The success probability values for different ranges of \(\lambda\) (dotted line), in comparison with the benchmark success probability 0.580164 from the full-information game (solid line)
6 A generalisation

Finally, we mention a problem which may be approached using techniques presented in this paper. Suppose observations $X_1, \ldots, X_n$ are sampled successively from some distribution in a partially ordered space. For instance, one can consider sampling from the uniform distribution on $[0,1]^d$ endowed with the natural partial order. Define the chain records in the sample inductively, by requiring that $X_1$ be a chain record, and that $X_k$ for $k > 1$ be a chain record if the observation exceeds the last chain record from $X_1, \ldots, X_{k-1}$. The sequence of chain records is a ‘greedy’ chain in the partial order restricted to the set of sample points. A counterpart of the full-information problem in this setting is the task of maximising the probability of stopping at the last chain record (Gnedin 2007).

Acknowledgements This research was partially supported by Wrocław University of Science and Technology subsidy no. 8211104160. We would like to thank Prof. Jacek Cichoń for the suggestion to study the triangular model.

Data availability statement This article has no associated data.

Declarations

Conflicts of interest/Competing interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Baryshnikov, Y., Eisenberg, B., Stengle, G.: A necessary and sufficient condition for the existence of the limiting probability of a tie for first place. Stat. Probab. Lett. 23(3), 203–209 (1995)
Berezovsky, B.A., Gnedin, A.V.: The best choice problem. Nauka, Moscow (1984)
Bojdecki, T.: On optimal stopping of a sequence of independent random variables - probability maximizing approach. Stoch. Proc. Appl. 6, 153–163 (1978)
Campbell, G.: The maximum of a sequence with prior information. Comm. Stat. (Part C: Sequential Analysis) 1(3), 177–191 (1982)
Chow, Y.S., Robbins, H., Siegmund, D.: The theory of optimal stopping. Dover, 1st edition (1991)
Falk, M., Hüsler, J., Reiss, R.-D.: Laws of small numbers: extremes and rare events. Birkhäuser (2011)
Faller, A., Rüschendorf, L.: Approximative solutions of best choice problems. Elec. J. Probab. 17, 1–22 (2012)
Ferguson, T.S.: Who solved the secretary problem? Statist. Sci. 4(3), 282–289 (1989)
Gilbert, G.P., Mosteller, F.: Recognizing the maximum of a sequence. J. Am. Stat. Assoc. 61(313), 35–73 (1966)
Gnedin, A.: On the full-information best-choice problem. J. Appl. Probab. 33(3), 678–687 (1996)
Gnedin, A.: The representation of composition structures. Ann. Probab. 25(3), 1437–1450 (1997)
Gnedin, A.: Best choice from the planar Poisson process. Stoch. Proc. Appl. 111, 317–354 (2004)
Gnedin, A.: Recognising the last record of a sequence. Stochastics 79, 199–209 (2007)
Gnedin, A., Krengel, U.: A stochastic game of optimal stopping and order selection. Ann. Appl. Probab. 5, 310–321 (1995)
Gnedin, A., Miretskiy, D.: Winning rate in the full-information best-choice problem. J. Appl. Probab. 44, 560–565 (2007)
Goldenshluger, A., Zeevi, A.: Optimal stopping of a random sequence with unknown distribution. Math. Oper. Res. 47(1), 29–49 (2021)
Hill, T.P., Kennedy, D.P.: Sharp inequalities for optimal stopping with rewards based on ranks. Ann. Appl. Probab. 2, 503–517 (1992)
Kirschenhofer, P., Prodinger, H.: The number of winners in a discrete geometrically distributed sample. Ann. Appl. Probab. 6(2), 687–694 (1996)
Kolchin, V.F.: The limiting behavior of extreme terms of a variational series in polynomial scheme. Theory Probab. Appl. 14, 458–469 (1969)
Kuchta, M.: Iterated full-information secretary problem. Math. Meth. Oper. Res. 86, 277–292 (2017)
Nuti, P.: On the best-choice prophet secretary problem (2020) arXiv:2012.02888
Petruccelli, J.D.: On a best choice problem with partial information. Ann. Statist. 8(5), 1171–1174 (1980)
Pinsky, R.G.: The secretary problem with biased arrival order via a Mallows distribution. Adv. Appl. Math. 140, (2022)
Porosiński, Z.: The full-information best choice problem with a random number of observations. Stoch. Proc. Appl. 24(2), 293–307 (1987)
Porosiński, Z., Szajowski, K.: Full-information best choice problem with random starting point. Math. Jap. 52, 57–63 (2000)
Resnick, S.I.: Extreme values, regular variation and point processes. Springer (2008)
Sakaguchi, M.: A note on the dowry problem. Rep. Statist. Appl. Res. Un. Japan. Sci. Eng. 20, 11–17 (1973)
Sakaguchi, M.: Optimal stopping problems for randomly arriving offers. Math. Japon. 21(2), 201–217 (1976)
Sakaguchi, M.: Dowry problems and OLA policies. Rep. Statist. Appl. Res. JUSE 25 (1978)
Samuels, S.M.: Review of [Scherwentke 1989]. Math. Reviews. MR1029991 (91a:60116) (1991a)
Samuels, S.M.: Secretary problems in: Handbook of sequential analysis. Marcel Dekker, Inc., B.K. Ghosh and P.K. Sen (eds) (1991b)
Scherwentke, P.: Best choice problem for discrete random variables. Optimization 20, 517–521 (1989)
Stewart, T.J.: Optimal selection from a random sequence with learning of the underlying distribution. J. Am. Stat. Assoc. 73(364), 775–780 (1978)
Tamaki, M.: On the optimal stopping problems with monotone thresholds. J. Appl. Probab. 52(4), 926–940 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.